INFORMATION TO USERS

The Pennsylvania State University

The Graduate School

Department of Mechanical Engineering

IDENTIFICATION OF DYNAMIC NONLINEARITIES IN ROBOTS

FOR PREDICTIVE MAINTENANCE

A Thesis in

Mechanical Engineering

by

Chi-Ching Kao

Submitted in Partial Fulfillment
of the Requirements for
the Degree of

Doctor of Philosophy

May 1995

UMI Number: 9531947
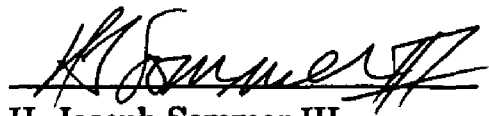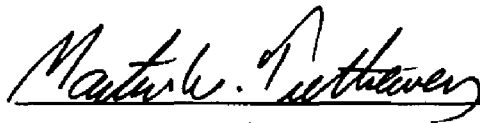
UMI

300 North Zeeb Road
Ann Arbor, MI 48103

We approve the thesis of Chi-Ching Kao.
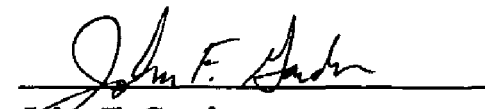
Date of Signature

H. Joseph Sommer III
Professor of Mechanical Engineering
Thesis Advisor
Chair of Committee

4/19/95

Martin W. Trethewey
Professor of Mechanical Engineering

4/19/95

John F. Gardner
Associate Professor of Mechanical Engineering

4/20/95

Kathryn W. Lilly
Assistant Professor of Mechanical Engineering

4/19/95

Joseph P. Cusumano
Associate Professor of Engineering Science and Mechanics

4/25/95

Robert J. Heinsohn
Professor of Mechanical Engineering
Acting Head of the Department of Mechanical Engineering

21 April 1995

ABSTRACT

Identification of dynamic nonlinearities in robot joint drives using the Hilbert transform and internal excitation for predictive maintenance was developed. For conventional robots, the dominant source of end effector positioning error comes from gear backlash and Coulomb friction in their joint drive mechanisms. After a long period of operation, these nonlinear effects will increase dramatically due to wear. If nonlinearity in joint drives were examined on a regular basis, maintenance can be scheduled to repair or replace the degraded component. Positioning accuracy of the end effector can then be improved.

When the Hilbert transform is applied to the frequency response function of a system, difference between the Hilbert transformed and the original frequency response functions indicates nonlinearity in the system. The Hilbert transform analysis was examined on a simulated one degree of freedom system with Coulomb friction. Indicators such as resonant peak amplitude ratio, skewness and kurtosis of the shape of the resonant mode in the Hilbert transformed function, and the RMS ratio between the transformed and original frequency response functions were proved effective to identify different levels of Coulomb friction nonlinearity.
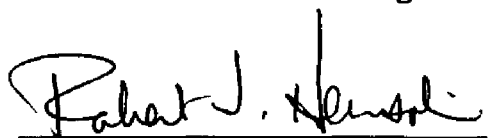
The procedure was tested on a two-link prototype manipulator using its control computer to generate internal actuator excitation and to perform data acquisition. Frequency response functions of the joint drives were measured using sine sweep excitation as the input signal and data from accelerometers attached to manipulator links as output responses. When the Hilbert transform was applied to the measured frequency response functions, the ratio between the Hilbert transform and the measured frequency response functions with respect to a set of baseline data was used to identify the presence, magnitude and location of Coulomb friction nonlinearity in the joint drive system.

TABLE OF CONTENTS

LIST OF FIGURES

Figure   Page

Figure   Page

LIST OF TABLES

LIST OF ACRONYMS

DFT	Discrete Fourier Transform
FRF	Frequency Response Function
FFT	Fast Fourier Transform
FT	Fourier Transform
HT	Hilbert Transform
HTF	Hilbert Transformed Frequency Response Function
LED	Light-Emitting Diode
MTBF	Mean Time Between Failure
MTTF	Mean Time To Failure
NPR	Noncausal Power Ratio
PID	Proportional-Integral-Derivative
PMP	Predictive Maintenance Program
RMS	Root-Mean-Square
SNR	Signal-to-Noise Ratio

ANALYTICAL NOMENCLATURE

m	mass
n	gear ratio
c	damping
k	stiffness
N	number of samples
t	time (second)
F	forcing level
F_{low}	sweep lower bound frequency (Hz)
F_{high}	sweep upper bound frequency (Hz)
R	friction level
T	sampling period (second)
Ts	sweep duration (second)
τ	time (second)
Δf	frequency resolution
ζ	damping ratio
ω	frequency (Hz or rad/sec)
ω_n	natural frequency
μ_k	central moment about mean
σ	standard deviation
α_3	skewness
α_4	kurtosis
$e(t)$	input to a nonlinear element
$c(t)$	output of a nonlinear element
$f(t)$	forcing function
$G(\omega)$	frequency response function
$H(\omega)$	Hilbert transformed frequency response function
$g(t)$	time domain causal signal
$g_{odd}(t)$	odd part of $g(t)$
$g_{even}(t)$	even part of $g(t)$
$sgn(t)$	signum function
$x(k)$	discrete time domain function
$x(t)$	time domain function, displacement
$\dot{x}(t)$	velocity

$\ddot{x}(t)$	acceleration
$x_o(t)$	odd part of $x(t)$
$x_e(t)$	even part of $x(t)$
$y(t)$	time domain function
$X(n)$	discrete frequency domain function
$X(\omega)$	frequency domain function
$X_o(\omega)$	odd part of $X(\omega)$
$X_e(\omega)$	even part of $X(\omega)$

ACKNOWLEDGMENTS

First, I would like to thank Dr. Joe Sommer for his guidance and support during the course of this project as well as my master's work. Over the years, he has always been able to give me valuable advice and provide needed financial support.

I would also like to thank Dr. Marty Tretheway for his support on testing equipment and suggestions on testing procedures. I would like to thank Dr. John Gardner, Dr. Kathryn Lilly and Dr. Joe Cusumano to serve in the committee.

I want to thank the Mechanical Engineering Department to provide me with interesting experience as a teaching assistant. I appreciate the opportunity to work with Dr. John Lamancusa in the Microcomputer Interface Laboratory. I thank Ron Gathagan of the instrument room for providing a good supply depot of materials.

Without the support of my parents, brother and sister, I wouldn't be able to sustain this long. I thank their never-ending patience and encouragement. Finally, my most sincere thank is to my wife I-Ching and my son Kevin. They make my stay at Penn State a joyful one.

Chapter 1 BACKGROUND

1.1 Robots in Industry

The study and development of robotic manipulators started in the mid-1940s when master-slave manipulators were used to handle radioactive materials at the Oak Ridge and Argonne National Laboratories. During the following decade, research on robotics was devoted to the development of a manipulator whose operation could be programmed to follow a sequence of motion steps determined by instructions in the program. The first commercial computer-controlled Unimate robot was introduced by Unimation in 1959 and was installed in a Ford factory in 1961.

Use of robots has increased dramatically in recent years. At the end of 1982, there were 1152 robots in the United Kingdom, 13,000 robots in Japan, and 6250 in the U.S. [1]. At the end of 1989, there were 5908 robots in the United Kingdom, 219,667 robots in Japan, and 36,977 robots in the U.S. [2]. Today in the U.S., the majority of robots are used in automotive, electronics, aerospace, food, and appliance industries.

As the number of robots in factories grows, robot maintenance becomes more vital. Essentially, industrial robots are designed to handle moderately heavy payloads and to operate at high speed while maintaining sufficient precision to ensure product quality. However, the basic mechanical structure of most robots is effectively composed of cantilevered beams in a serial kinematic chain connected by hinged or sliding joints. Although conventional links in robot arms can be considered rigid, robot joints may develop significant backlash and friction problems over long periods of operation. These joint errors accumulate along the serial kinematic linkage to effectively reduce accuracy and repeatability of the end effector and thus decrease product quality. Consequently, it is necessary to have a robot maintenance management program that can provide information about defects in joints such as extent of degradation and which joints are affected. This program must give the maintenance staff an early warning sign when serious defects are detected inside the robot linkage.

1.2 Maintenance Programs

Maintenance costs are a major part of the total operation cost of all manufacturing and production plants and can represent 15 to 40 percent of the costs of goods produced [3]. Recent surveys of maintenance management indicate that about one third of all maintenance costs are wasted as the result of unnecessary or improperly performed maintenance [3]. The results of ineffective maintenance management represents a loss of more than 60 billion dollars each year in the United States [3]. Loss of production time and product quality that result from inadequate maintenance dramatically reduce the ability to compete in the worldwide market with other countries that have implemented more advanced maintenance philosophies. The main reason for this drawback is the lack of predictive data that quantifies the actual need for repair or maintenance of plant machinery, equipment, and systems [3].

Traditionally, industrial and process plants utilize two types of maintenance management: either run-to-failure maintenance or preventive maintenance. Recently, many plants have additionally tried to adopted the predictive maintenance concept into their programs. Different maintenance management methods have their advantages and disadvantages. However, a good maintenance program should be able to reduce or eliminate unnecessary repairs, prevent catastrophic machine failures, and reduce the negative impact of the maintenance operation on the profitability of manufacturing and production plant.

Run-to-failure is a reactive management technique that waits for machine or equipment failure before any maintenance action is taken. The logic of run-to-failure is simple: when a machine breaks down, fix it. Therefore, it is truly a "no maintenance" approach and a plant does not spend any money on maintenance until failure occurs. However, this is the most expensive method of maintenance management. A plant must keep a large inventory of spare parts, pay overtime labor to repair sudden failures which leads to high machine downtime and low production availability. Analysis of maintenance costs indicated that a repair performed in the reactive run-to-failure mode will cost on the average about three times higher than the same repair made within a scheduled or preventive mode [3]. Experience also shows that most mechanical problems can be minimized and major repairs can usually be prevented if the problems can be detected and repaired at an early stage. No effective plant can use a run-to-failure maintenance program.

Alternatively, preventive maintenance is a time-driven program. In other words, maintenance tasks are performed based on elapsed time or hours of operation. All preventive maintenance programs assume that machines will degrade within a certain time of operation. Machine repairs or rebuilds are then scheduled based on the mean-time-to-failure (MTTF) or mean-time-between-failure (MTBF) statistics. The problem with this approach is that, when production demands are high, machines may degrade to severe levels due to heavier work loads before the scheduled maintenance time. On the other hand, when production demands are low, machines may still be in good operation condition at the scheduled time. Thus, maintenance schedules based on MTTF or MTBF data will either result in unnecessary repair or lead to run-to-failure operation.

Predictive maintenance is a condition-driven method which attempts to identify problems before machine failures occur. The main idea of predictive maintenance is regular monitoring of the actual mechanical condition, system efficiency, and other indicators of operating conditions. Maintenance activities are then scheduled on an "as-needed" basis. That is, if the monitored feature changes dramatically, then there must be something happening in the system. Therefore, necessary repair or adjustment must be taken to prevent failure. Predictive maintenance is more cost-effective than traditional maintenance programs and can minimize the number of breakdowns of all mechanical equipment in the plant and ensure that repaired equipment is in acceptable mechanical condition.

1.3 Identification of Nonlinear Mechanical Systems

Condition identification, which is an integral component of predictive maintenance, is an inverse problem of dynamics and is closely related to analyses of structural, mechanical, and control systems. The identification process attempts to develop or to improve the mathematical representation of these systems using experimental data. Physically, these systems may be inherently nonlinear. However, they are often approximated as linear systems if their operating ranges are small, and if the involved nonlinearities are smooth. Linearized models may then be used to analyze the behavior of the actual systems.

To date, identification techniques for linear systems are well developed both theoretically and experimentally. But, when conditions and/or assumptions required for the linearization are violated, significant inaccuracies arise. For instance, beams and plates

demonstrate cubic rather than linear stiffness for large deflections and the level of nonlinearity increases with the amplitude of vibration [4]. In recent years, techniques to identify nonlinear systems have been developed to provide more qualitative and quantitative information on nonlinearity of the system.

In the field of mechanical systems, identification of nonlinear systems includes detection of nonlinearities, quantification of nonlinearities, and system parameter estimation [5]. Nonlinearities can be detected by direct and indirect methods. Direct methods apply special functions, special transforms, or specific excitation and filtering techniques to obtain measurements of system nonlinearities. Typically, functions such as sig-function, high order correlation functions, bispectral functions, and probability density functions are used in the direct methods. For instance, the Hilbert transform is very effective to determine the existence of nonlinearities in mechanical systems [6]. Indirect methods often apply linearity tests to systems and check their violation. These tests are based on superposition principles, reciprocity, distortions of Nyquist plot, and parameter dependency.

Once nonlinear behavior has been detected, the question is then how to construct an accurate simple model to analyze the effect of nonlinearity. The classical approach to model a nonlinear system is using the Volterra series, which describes the input-output relationship of a system in the time domain. By taking the multi-dimensional Fourier transform of the Volterra series, higher order frequency response functions (FRFs) can be obtained to identify and to characterize a nonlinear system [7]. Common nonlinearities, such as cubic stiffness, clearance, and Coulomb friction, can be characterized by applying the moment integrals of the measured and Hilbert transformed FRFs [6]. The Noncausal Power Ratio (NPR), which is defined as the ratio of the noncausal power to the total FRF power, also provides an index to quantify the magnitude of nonlinearities [8].

For a robot manipulator, the system model should be able to describe three types of common nonlinearities:

- continuous type, such as cubic stiffness or quadratic damping,
- dead-band type, such as backlash,
- saturation type, such as Coulomb friction.

It also needs to provide measurement of the magnitude of nonlinearities in order to study the effect of nonlinearities on accuracy and repeatability of the end-effector. The location of

sources of nonlinearities along the kinematic chain should also be identified so that maintenance can be scheduled to repair the degraded joint.

For mechanical systems, physical laws and relationships, which govern the system's behavior, provide knowledge to construct a mathematical model for parameter estimation. Experience shows that better estimation can be obtained using fewer parameters. However, the resulting model should be accurate and physically interpretable to represent the behavior of the system. In system identification, the choice of a technique is always linked to the purpose of the analysis and there are no universal ways for different types of problems.

1.4 Excitation Methods for System Identification of Mechanical Systems

The procedure of system identification normally involves two stages: system input-output data acquisition, and data analysis. For mechanical systems, experimental modal testing is often used to generate data for identification routines. Different analysis routines require suitable types of excitation to the system. System properties are then characterized via analyzing the relationship between the input excitation and the output response measurement.

Basically, excitation sources can be divided into two main categories: single frequency sinusoidal and broadband signals. The most commonly used single frequency sinusoidal approach (slow sine sweep) is to scan the complete frequency range by slowly changing (sweeping) the excitation frequency continuously. High quality data are obtained because of high signal-to-noise ratio (SNR) and high energy density per spectral line. The main drawbacks for sine testing are related to the long measurement time since steady-state responses are required, and the dedicated hardware which is needed to generate the excitation. Alternatively, the stepped-sine method is based upon excitation by a sinusoidal with fixed amplitude and frequency. By stepping the sine frequency through the frequency range of interest, FRFs can be evaluated for a number of discrete frequency points [9]. Using this approach, frequency resolution can be increased when necessary such as in the region near resonance where rapid changes of system characteristics occur.

Since the evolution of Fast Fourier Transform (FFT) signal analyzers, broadband signals such as random (white noise), transient (hammer impact), and periodic chirp (fast

sine sweep) have been widely used as excitation sources. In order to improve SNR and to reduce FFT leakage problems, new signals combining periodic and random (pseudo random), transient and random (burst random), or transient and chirp (burst swept sine) have been introduced to the broadband category [9]. Although considerable measurement time can be saved by using broadband excitation, detailed system characteristics might be lost due to limited frequency resolution and dynamic range. For instance, the measured FRF of a nonlinear system can be effectively linearized if random excitation were used.

Different purposes of analysis can be achieved by applying different excitation sources in an identification routine. For instance, using sinusoidal excitation, the Hilbert transform can be used to determine the nonlinearity in a system by checking its FRF [6]. Instead, using transient or random excitation, the Hilbert transform technique in the time domain can be used to evaluate the natural frequency and damping ratio of a linear system as well as the apparent damping of a nonlinear system [10, 11].

For structures, external excitation are applied as input sources since no internal driving mechanism is available. Usually, tests are performed in a laboratory using shakers and impact hammers with dedicated measurement equipment. However, when on-site test is required on a factory floor such as robot welding assembly line, installation of test equipment becomes a major inconvenience. Shakers must be installed, calibrated, and removed for each test. Impact hammers may not provide enough energy to excite the robot. Also, the use of impact force excitation may actually degrade the robot itself.

Consequently, internal excitation, which utilizes robot joint actuators as the driving device, is an attractive approach to solve this problem. Using motor excitation, adjustment of gear backlash was developed for industrial robot joints [12]. Recently, nonlinear behavior of an experimental flexible robot link was characterized via open loop stepped sine measurement using motor generated internal excitation [13]. Modal analysis of an two link flexible manipulator was performed using both external and internal excitation [14]. Results showed that the natural frequencies and mode shapes found using internal excitation method closely agreed with those found using external excitation. Further studies are needed on the effect of internal excitation with respect to external excitation.

Conventional analysis routines utilize the single-point excitation method. Only one input excitation is applied and one output response is measured although the input and

output locations may be varied around the structure during the test. It should be noted that approaches using several simultaneously applied excitations are also possible and have been used for testing large aerospace structures. Specifically, the purpose of using multi-point excitation is to reduce the probability of introducing systematic bias errors into the measurement. Recent research on modal analysis of an aircraft frame applied both single-input and multiple-input excitation with relative phases between the simultaneous input signals [15]. Results from this analysis indicated that a multiple-input approach is more effective than single-input approach when identifying closely spaced or weakly excited symmetric and antisymmetric modes.

On the other hand, approaches applying single input excitation and several output measurements on different locations of a structure will provide information on changes of behavior between measurement points. Since nonlinearities in engineering structures are usually localized, this information may be able to help indicate the location of nonlinearities. For instance, if measurement points are assigned along links of a robot arm and waist excitation is applied, nonlinearities in robot joints will primarily affect the measured FRFs in adjacent links and more distal links. Defects which are developing in joints may then be detected by analyzing how these FRFs may vary on a daily, weekly, or monthly basis.

1.5 Research Objectives

The objective of this study is to investigate various defect identification techniques for predictive maintenance of robot manipulators. Three aspects will be investigated:

1) develop a routine to detect the presence of nonlinearities in robot manipulators using the Hilbert transform and internal excitation;

2) develop algorithms to assess the magnitude of nonlinearities detected in a robot manipulator; and,

3) develop a method to check the location of nonlinearities in robot joints along the kinematic chain.

Chapter 2 LITERATURE REVIEW

2.1 Predictive Maintenance Programs
2.1.1 Overview

The normal life-cycle of a machine can be described by the "bathtub" curve as shown in Figure 2.1. The greatest probability of failure comes at startup due to installation problems or human error. Following this initial period, the chance of failure is relatively low for an extended period of operation. Eventually after a useful life, failure rates increase as wear occurs. Following preventive maintenance philosophy, many companies may typically teardown and fix machinery on an annual or scheduled basis. However, this actually increases the chance of startup failure as shown in Figure 2.2 [16].

A predictive maintenance program (PMP) reduces this kind of failure by detecting impending breakdowns to allow planned component repairs. It also reduces the severity of many failures by reducing or preventing secondary damage. In a PMP, maintenance work can be scheduled according to collected data indicating the current condition, performance, and reliability of machines [17].

An example of a PMP in a chemical plant is presented by Maniscalco [18]. Maintenance inspectors use portable data collectors to record a variety of measurements, including vibration, temperature, pressure, flow and shaft speed, on all critical machines. Four levels of machine condition are used in the inspection reports. They are:

O (OK) - No defect. All machine functions are normal.

S (Scheduled repairs) - Not a critical defect. Repairs should be made at the earliest convenience.

H (Hot) - A significant defect was found. Secondary damage is developing and failure will occur very soon.

E (Emergency) - A very serious problem was detected and the machine should be shut down for repairs immediately.

This program successfully detected a failing bearing in a compressor and by replacing the bearing before the compressor was damaged, the company saved considerable downtime and secondary repair costs.
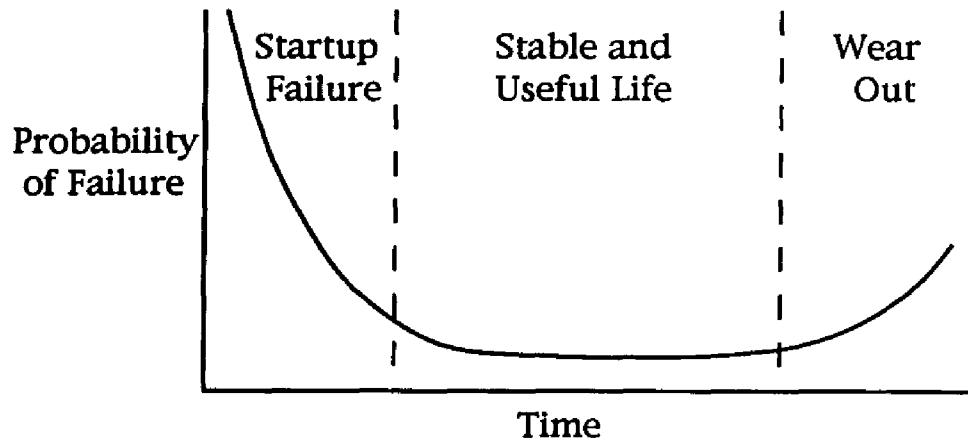
Figure 2.1 Bathtub Curve.
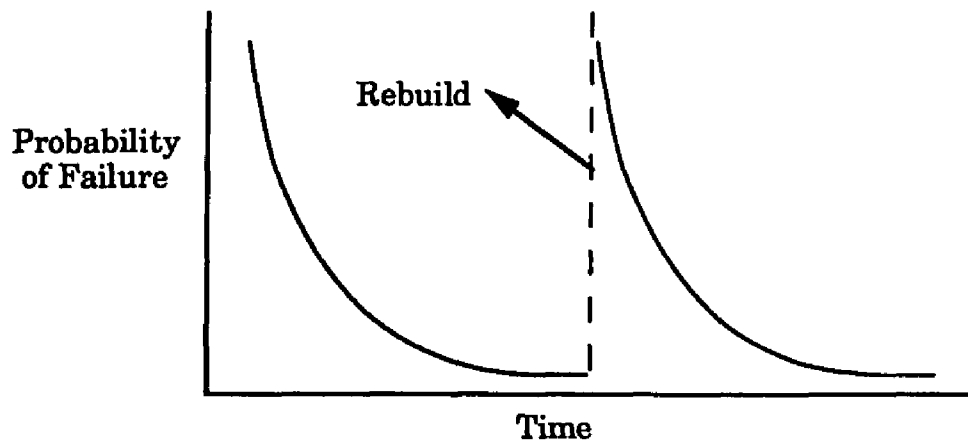
Figure 2.2 Effect of Scheduled Rebuild on Bathtub Curve.

In many cases, PMP usually leads to simpler repairs instead of costly major replacement due to failure induced damage to other machinery. If implemented successfully, maintenance costs can be greatly reduced due to fewer failures, more effective utilization of parts, and a reduction in unnecessary maintenance. Operating costs also can be reduced due to improved efficiency and productivity.

However, this is not a quick way to solve the existing problems in maintenance. The benefits of PMP must be balanced against the cost to implement it. PMP will not be effective without careful planning and full commitment from management. Experience has shown that it takes about three years before the full impact of the program is achieved. Typically, annual return is five to ten times the program's cost for an effectively established program [16].

2.1.2 Methods for Predictive Maintenance

There are five nondestructive techniques that are often used for predictive maintenance management of industrial machinery: vibration monitoring, thermography, tribology, process parameter monitoring, and manual/visual inspection [3]. Each technique provides specific information on different types of plant equipment. A total predictive maintenance program must incorporate several techniques to assist the maintenance manager in determining the actual need for maintenance and to reduce the maintenance impact on production.

Most comprehensive predictive maintenance programs use vibration analysis as the primary tool since the majority of plant equipment are mechanical [3, 17]. The vibration characteristics of a machine will change as machine condition changes. There are two underlying assumptions: (1) all common failure modes have distinct vibration frequency components that can be isolated and identified, and (2) the amplitude of each distinct vibration frequency component will remain constant unless there is a change in the operation dynamics of the machine.

Vibration analysis utilizes noise or vibration created by mechanical equipment to determine their operating condition. Recent advances in microprocessor technology provide a powerful and cost-effective tool for vibration data acquisition, management and interpretation. By regularly monitoring and evaluating the vibration of plant machinery,

any degradation of the mechanical condition can be detected before serious damage occurs. Vibration monitoring can be applied to all types of machinery and typically utilize the following techniques: broadband trending, narrowband trending, and signature analysis [19].

Broadband trending acquires broadband or overall vibration readings from selected points on a machine [3]. These data are then compared to baseline readings taken from a new machine or a vibration severity chart in order to determine the relative condition of the machine. Broadband or overall root-mean-square (RMS) data represent the total vibration of the machine at the specific measurement points. However, the limitation of this method is the lack of sensitivity in the data [20]. Unless a problem is severe, the overall measurement level may not change significantly. Also, it does not provide any information about the individual frequency components or machine dynamics that created the measured value.

Narrowband trending monitors the total energy for a specific bandwidth of vibration frequencies. Narrowband methods utilize vibration frequencies that represent specific machine components or failure modes and thus can provide the relative condition of selected machine components. This technique provides the ability to monitor the condition of gear sets, bearings, and other machine components without manual analysis of vibration signatures [19].

Signature analysis deals with the extraction of information from measured vibration signal patterns of a mechanical system [21]. It utilizes processing techniques in time domain, frequency domain, and amplitude domain as well and other probability and statistical properties. The major benefit from signature analysis is that it can provide specific and simple feature readings of the mechanical system. Therefore, not only qualitative but also quantitative studies about trending of system conditions can be made. From these trending data, a more objective criterion for plant maintenance can be established.

For example, analyzing each frequency component of the vibration spectrum generated by a machine over a period of operation provides a thorough time history on that system [17]. Any changes of intensity at specific frequencies due to damage developing in the machine can then be identified before serious failure occurs [19]. According to the predefined alert and alarm limits, a maintenance manager can decide whether minor adjustments or major repairs are necessary for that machine.

Signature analysis is neither an easy nor a cheap task. It requires powerful instrumentation systems and sophisticated analysis and processing techniques. Typically, signature analysis consists of two phases: acquisition/preprocessing and feature analysis. At the first phase, signature data, such as velocity and acceleration, is acquired by sensors on the machine. Then, preprocessing procedures such as noise reduction, data compression, or data transformation are applied to the raw data to obtain more concise data for later analysis. In the analysis phase, selected feature parameters can be extracted from the processed data. These indices describe specific properties of the current condition of the mechanical system [21].

2.1.3 Predictive Maintenance Strategy

Usually, there are four stages to establish a predictive maintenance program: [3]
- Determine the initial operating condition of the machinery
- Take data to establish a time history data base
- Choose analysis parameters
- Define alert and alarm limits.

During the initial period, machines should be monitored to establish valid baseline data sets [17]. Vibration signatures should be acquired for each machine to determine their operating conditions. For new equipment, this data can be compared to the specifications provided by manufacturers. After that, measurements are usually taken every two weeks at each measurement point to develop a sufficient statistical data base [3]. Trending and projected time-to-failure are then evaluated according to this time history data [16]. After the initial stage, the frequency of data collection will depend on the types of machines. It also can be adjusted according to the actual rate of change in data [16].

The next step is to choose analysis parameters, which will be monitored and updated routinely, for each machine based on their specific requirements. By monitoring this data, common failure modes can be identified for individual machines. For example, narrowband readings can be used as parameters to detect impending gear backlash and meshing problems. By analyzing the amplitude of vibration at the gear meshing frequency, broken teeth and cracks on gears were predicted and later found in the gearbox of a calender roll machine [19].

Alert and alarm limits are defined to provide early warning on machine degradation [17]. For example, machine condition can be evaluated using readings such as vibration velocity, shock pulse level, and vibration spike energy as severity indices. When the monitored parameter exceeds the predefined alert limit, maintenance must be scheduled to examine the degraded equipment. When the alarm limit is reached, emergency procedures must be taken to prevent machine failure. General guidelines to specify alarm limits in vibration spectrum analysis for various types of machine were developed by Berry [22].

Two basic types of alert and alarm limits, static and dynamic, are commonly used in parameter analysis to check how much and how fast the dynamic behavior of a machine has changed [3]. The static limit defines the maximum acceptable parameter variation from the baseline data. This type of limit does not consider the rate of change or historical trends of a machine and, therefore, can not be used to predict when the alarm will be reached. Alternatively, the dynamic limit uses the rate of change in vibration characteristics as an index to represent the severity of defect. It assumes that the rate of change is more important than the actual level of change. The dynamic limit can be used to predict when the alert and alarm limits will be reached. Both types of limits are applied to monitor, trend, and alarm vibration measurements.

2.1.4 Possible Predictive Maintenance Methods and Strategies for Robots

For industrial robots, vibration monitoring will be the dominant tool for PMP. Since commercial robot links can be considered effectively rigid, performance will mostly be affected by defects in the joint drive mechanisms. Vibration monitoring can detect misalignment of motor shafts, worn gears, and failing bearings. Advanced analysis routines should be able to detect expanded backlash and excessive friction due to wear in gears.

Process parameter monitoring in the form of periodic on-line end effector position repeatability testing using proximity probes, machine vision, and three-dimensional end-point sensors can be a useful trend indicator for robots, but requires special jigs, fixtures, or sensors to be located at strategic points on or near the robot workspace. This type of process monitoring has been addressed as the robot calibration problem by Bosnik, Whitney, and Mooring [23, 24, 25, 26]. Additionally, Dagalakis suggested that periodic estimation of

dynamic parameters and tuning of controller gains for individual joint drive systems can provide direct information about which joint is affected by degraded components [27, 28]. Thermography may be used to detect electrical problems.

2.2 Parameter Identification for Robot Manipulators

One of the basic criteria to evaluate the performance of a robot manipulator is its positioning accuracy. In order for the controllers to direct the end effector to a specified location, a model is required to describe the physical characteristics of the robot. The purpose of robot parameter identification is then to establish this model via extracting accurate system parameters from experimental data. According to their contribution, robot parameter identification can be divided into three categories: calibration, identification of dynamic properties, and identification of nonlinearities. Static positioning accuracy of a robot end effector can be increased through calibration. Trajectory tracking ability, which is often affected by uncertainties in dynamic properties of arm links, can be improved by using controllers with robust structures or on-line system identification scheme. While continuous nonlinearities such as quadratic damping and cubic bending stiffness can be analyzed and controlled effectively, discontinuous nonlinearities such as backlash and Coulomb friction require more study.

2.2.1 Kinematic Calibration

The purpose of robot calibration is to improve and to verify the accuracy of a robot's kinematic model as well as the electrical parameters of incorporated transducers. The kinematic model, which relates the end effector's pose (position and orientation) in the world coordinates to robot joint coordinates (individual joint displacements), describes the geometry of robot arm motion. In order to put the end effector at a desired pose, robot arms are required to adjust their joints to joint coordinates calculated from the inverse kinematic model. Position control of robot end effectors is currently done using individual joint proportional-integral-derivative (PID) controllers that completely ignore coupled dynamic effects of arm links [29]. However, due to manufacturing tolerances and errors, small variation in the geometry of the arm links will often cause significant errors in end effector positioning. Research results showed that misalignment of joints may affect the accuracy of

a typical industrial robot as much as 3.26 cm [23]. Calibration is then necessary to compensate for these inaccuracies.

Usually, individual joint displacements (translation or rotation) can be measured by internal encoders at each joint. Calibration is then used to define the relationship between the actual end effector pose and the corresponding joint encoder readings. Two approaches are typically utilized in calibration: forward and inverse. Based on the forward kinematic model, forward calibration measures the pose of end effector in world coordinates when the joint encoder readings are specified. The error between the actual pose and the commanded pose of the end effector is then used to estimate parameters in the kinematic model. Forward calibration also can estimate joint compliance, backlash, and gear transmission errors. Much of the research in robot calibration has been devoted to the forward calibration problem [23, 24, 25, 26]. Research results showed that calibration can reduce the model error of a PUMA 560 robot from 4.8 mm to 0.3 mm [25].

While the forward calibration predicts the pose of end effector from the given joint angles, the inverse calibration recommends values of the joint encoder readings necessary to drive the end effector to a desired pose in world coordinates [30]. Instead of modeling the potential sources of error, inverse calibration utilizes approximation functions to correct the encoder readings recommended by the robot's kinematic model. These functions are found by measuring the error at discrete locations throughout the robot's workspace and then fitting trivariate functions to the error samples. The trivariate functions are then used in control loops to improve the positioning accuracy of the end-effector by localized error compensation. Simulation results showed that end effector location error can be reduced from 1.2 mm average to 0.12 mm average [30].

There are three phases during the calibration process: measurement, estimation, and correction [26]. During the measurement phase, data for actual end effector pose obtained by an independent external measuring system and data for the calculated pose from the internal encoder readings is recorded over the entire work space. In the estimation phase, an optimal set of physical and electrical parameters, such as link length, joint angle, and transducer offset, is evaluated to produce the least error between the actual and calculated pose at each calibration posture. Finally, in the correction phase, modified parameters are applied in the controllers to improve accuracy of end effector positioning. These optimal

parameters also can be included in a time history database for process parameter monitoring to track the developing degradation of components in a robot predictive maintenance program.

Several pose measuring techniques have been used to calibrate industrial robots [31, 32]. The cube method applies three pairs of dial gauges to detect three mutually orthogonal faces of a cube held in the end effector of a robot [33]. The pose of the end effector is then calculated according to the measured data. For more accurate measurement, non-contact displacement transducers are used to reduce the effect of applying contact force to the robot [32].

The laser tracking method uses a laser position detector to measure the distance between stationary laser beam sources and a reflector target on the moving end effector [34]. Three-dimensional pose information can be obtained via combining position data for the end effector from two or three measurement stations. Alternatively, the three-cables method employs three tension cables attached to the robot wrist and cable-feeding device to measure the position of end effector [35]. Position data can be obtained when the robot is moving. However, orientation of end effector can not be determined by this method.

In the ball-bar method, an extensible ball-bar with an internal probe is attached between the robot end effector and a magnetic socket on a tripod [36]. The robot is programmed to follow an arc with radius equal to the nominal length of the ball-bar. Radial positioning errors can be measured from the changes in length of the ball-bar. Since the ball bar is designed to measure one-dimensional radial displacement, its application for robot testing is limited.

Using theodolite triangulation, two or more theodolites are used to measure the horizontal and vertical angles of a target on the robot wrist [25]. Three-dimensional position of the target is calculated according to these theodolite readings. Due to the improvement of image processing techniques, a video camera can be used to replace theodolites in calibration [37]. Targets with known dimensions are attached to the robot end effector. Images of these pre-arranged targets are analyzed to calculate the position of the end effector. A two-camera system with light-emitting diodes (LEDs) attached to the robot wrist can be used to determine pose of the end effector [38]. However, good accuracy is not easy to obtain in real

environments using a video camera since it is very sensitive to vibration and lighting variations.

2.2.2 Identification of Dynamic Parameters

The dynamic behavior of a robot manipulator is governed by its equations of motion that describe the relationship between the input joint torques and the output response of arm links. Using computed torque control methods, desired output motion can be obtained by solving the inverse dynamic problem to specify the necessary input torques. Since all links in a robot move simultaneously, the effective dynamic properties such as mass and inertia of the overall manipulator are configuration dependent. Gravity, centrifugal force, Coriolis force, and interaction torques caused by acceleration of adjacent links also affect the dynamic response of robot [27].

Although joint PID controllers, which utilize a linear control approach, are adequate in most position control applications, such as pick-and-place or spot-welding, undesirable overshoot problems may occur [29, 39]. Overshoot is more significant for direct drive arms where there is little friction and no gearing to attenuate dynamic overshoot effects. When trajectory tracking tasks such as plasma-welding and laser-cutting, or high speed operations in the presence of obstacles are performed, more sophisticated control algorithms, that account for the complicated system dynamics, are required.

In the computed torque method, which is a feedback linearization approach, control torques of a desired trajectory are calculated based on a dynamic model of the robot arm which attempts to compensate for inertia and nonlinear dynamic effects of the arm [29, 40]. This method is very effective when exact dynamic parameters of the analytical model of robot are available. However, in practice, model uncertainty always limits the application of this approach. More efficient control strategies are needed to minimize performance sensitivity due to model uncertainty.

Two types of uncertainty are often encountered when designing robot controllers: parametric uncertainty and high-frequency unmodeled dynamics [29, 41]. The first type corresponds to inaccuracies in terms included in the model, while the second type corresponds to inaccuracies on the system model itself. Parametric uncertainty includes inaccuracies on mass and inertia of arm links, inaccuracies on torque constants of actuators,

unknown loads, uncertain load position in the end effector, and so on. Unmodeled dynamics includes structural resonant modes, neglected actuator dynamics, and effects due to finite sampling rate.

Based on various formulations for equations of motion, several methods have been used to identify parameters of arm dynamics and load dynamics of robots. Arm link mass, center of mass position, and inertia matrices can be estimated off-line by measuring the input joint forces/torques and the resulting output joint motion (position, velocity, and acceleration) [42, 43]. Parameters of arm dynamics are often considered constant since arm links of conventional industrial robots are moderately rigid. The estimated parameters are then used in the control system. Payload mass, center of mass position, and the inertia matrix can be estimated on-line by measuring joint input torque/output motion or using wrist force/torque sensing.

Additionally, methods to estimate lumped dynamic parameters of an electrically actuated joint drive system of a robot were developed by Dagalakis and Myers [28]. FRFs of joint drive system were determined using motor generated random excitation. Coefficients of the equivalent moment of inertia, damping, and stiffness were then estimated using complex polynomial curve fitting technique. These parameters can be used to better predict the dynamic response of the robot and may also be used for precise tuning of its controllers.

Instead of trying to identify the exact dynamic properties for each manipulator, an alternative strategy is to design a controller that maintains good trajectory tracking ability while being able to handle modeling errors and disturbances. This controller design should be applicable to robots of the same model without changing its control structure. Two major design approaches have been studied in recent years to overcome uncertainty problems: robust control and adaptive control. Unlike the linear joint PID control, these methods are inherently nonlinear control techniques [41].

One typical method of robust control is sliding mode control which is based on a nominal system model and additional terms on model uncertainties which are described by their bounds [41, 44]. Complicated high order systems are first transformed into equivalent lower order systems with new state variables which define the surface of control switching. Control gains of transformed systems are then selected according to the knowledge about

parameter bounds so that control stability can be ensured. Measurements or estimates of the original states are often required, and the on-line computation load is heavy.

Although robust control can handle disturbances, quickly varying parameters, and unmodeled dynamics quite well, it requires reasonable *a priori* estimates of the parameter bounds. In order to achieve perfect tracking performance, high control activity is required along the switching surface and that may lead to control chattering [41]. High frequency structural modes of the robot may be excited due to chattering and degrade the precision of position control. Therefore, a boundary layer neighboring the switching surface is introduced to smooth out control discontinuity and thus eliminate chattering.

Adaptive control is an approach to control systems with constant or slowly varying uncertain parameters [41]. The idea of adaptive control is to estimate the uncertain parameters on-line based on the measured system signal. Estimated parameters are then used in the computation of control performance. In other words, system parameters are updated continuously during operation based on measured performance. Note that the uncertain parameters may not be estimated exactly, but the error of trajectory tracking will converge to acceptable range within finite time.

Two types of adaptive control approach are often used: model reference and self-tuning [41, 45, 46]. Model reference adaptive control uses a reference model of a joint (such as second order differential equations) to generate a reference joint output. The controller gains are adjusted according to the difference between reference and real joint outputs. Real joint outputs will converge to the desired reference joint outputs asymptotically. The self-tuning adaptive control treats each joint as a slowly varying system whose coefficients are estimated on-line based on measured performance. Control inputs are then calculated using the system model with estimated parameters.

Adaptive controllers improve their performance as the adaptation process continues and require little or no *a priori* information about uncertain parameters of the system [41]. The drawback is that system parameters are assumed to vary rather slowly. This may not be true when robots move at high speeds. In addition, computation loads for adaptive control are very heavy, too.

Although both robust and adaptive controllers are able to handle high-frequency unmodeled dynamics, control activities typically avoid that frequency range [41]. Control

laws are often designed so that their control bandwidth is lower than the frequency range of the unmodeled dynamics. Structural resonant modes can be examined by using modal analysis techniques. As new robot structures are designed to be more flexible to reduce weight, modal analysis of arm linkage is necessary to ensure that controller design is effective. Trade-offs always exist between control performance and system limitation.

2.2.3 Identification of Nonlinearities

Every mechanical system demonstrates some nonlinear effects. Nonlinearities are either in continuous forms such as quadratic damping and cubic bending stiffness, or in discontinuous forms such as backlash, stiction, and Coulomb friction. Most effective controllers can maintain performance while compensating for these nonlinear phenomena. However, when nonlinearities become very severe due to developing component failures, pure control compensation is not enough. It is necessary to identify the nonlinearities so that degraded components can be repaired or replaced.

For industrial robots, gear backlash and Coulomb friction in drive mechanisms are the dominant sources of nonlinearities and can cause actual joint response to differ dramatically from the commanded response of the joint controller. Experience has shown that up to 25 percent of actuator torque is spent to overcome friction in non-direct drive robot arms [47].

For a typical joint actuator, saturation and deadband effects often exist between the input voltage and the output motion [48]. One major contribution to the deadband effect is stiction and Coulomb friction. Further, backlash and Coulomb friction may cause significant position and velocity errors, and affect system stability. To date, researchers have concentrated on how to derive models to estimate, control and thus reduce the effects of these mechanical nonlinearities.

Backlash is the lost motion error when a gear pair reverses its direction of rotation [12]. Therefore, backlash introduces a time delay between input on the drive gear and output on the driven gear. The presence of backlash in gear meshing creates a hysteretic nonlinearity. This is a two-valued nonlinearity because there are two possible output values for a given input. Due to the hysteresis loop effect, the joint servo system could exhibit a

continuous oscillation (limit cycle) or chattering condition. Excessive amount of backlash may further cause instability.

A schematic drawing of backlash between mating gear teeth is shown in Figure 2.3. The dashed line indicates the ideal tooth thickness that would eliminate backlash. The backlash can be measured as either an angular or a linear displacement quantity. θ_B is the angle measured at the center of the lower gear that corresponds to the backlash gap. B is the distance on the circumference of the pitch circle, and B_{LA} is the distance on the line of action that corresponds to the same backlash.

Most modern robots use precision gears to reduce backlash. However, complete elimination of backlash is impossible since clearance is needed to avoid tooth interference. There are three inherent backlash sources in a single gear: size variation, tooth errors, and pitch circle runout [12]. In addition, center distance variation between two meshing gears also contributes to backlash. These primary sources of backlash are due to intentional design allowance, manufacturing tolerance, or wear during operation. Other sources of backlash include bearing eccentricities, gear center shift, shaft deformation, thermal dimensional changes, deflection of teeth, and forced vibration. They are considered secondary sources because their contribution to backlash error is generally smaller.

Adjustment methods for gear backlash in industrial robot joints were developed by Dagalakis and Myers [12]. Tests using joint excitation were performed on the wrist joint of a PUMA 560 robot. Band-limited Gaussian random noise was injected into the drive system of the joint through either the joint controller or the power amplifier of the joint. Output response of the link was monitored using an accelerometer attached on the link. The coherence function of the input excitation signal and output response of the link was used as an index to measure the tightness of gear mesh such that the effect of backlash nonlinearity could be minimized. Further studies using joint velocity and position encoder feedbacks as output response were conducted and showed similar results [49]. This technique also can be applied to adjust gear meshing of a heavier robot joint such as the elbow joint [50].

Analytical methods for estimation of backlash magnitude were developed by Bae [51]. Since backlash errors are effectively joint angle errors, accurate position sensing at the joint can provide sufficient information to estimate the amount of backlash. However, sensors must be placed at both the input and output gears. Backlash was then defined as the
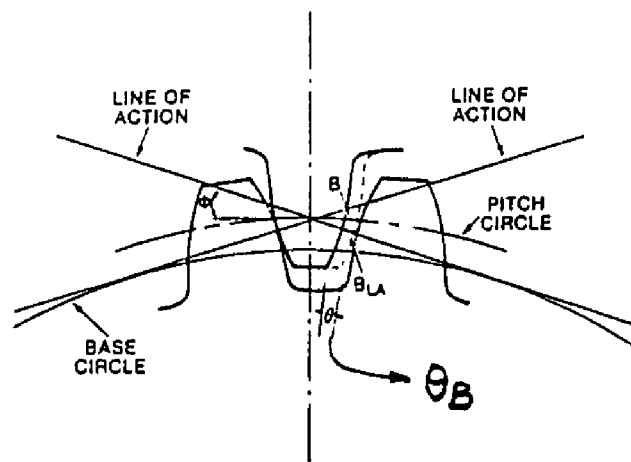
Figure 2.3 Schematic Drawing of Backlash.

difference between input of the driving gear and output of the driven gear multiplying by its gear ratio. Measurements of input and output motion were taken when the input gear changes its direction of rotation. Because backlash may be considered as a constant parameter for a limited period of time, least square estimation algorithms were reduced to a simple time average of measured backlash. Numerical simulation results showed that this identification scheme can effectively estimate the amount of backlash in gear reducers.

Friction is a dissipative force that appears at the contact surface of two bodies in contact [52]. When there is no relative motion between contacting bodies, stiction (static friction) will oppose all motion as long as the driving forces are smaller in magnitude than the stiction force. When relative motion occurs, Coulomb friction will oppose the motion with its magnitude proportional to the normal force at the contact point. In this classical dry friction model, friction force is constant and independent of the relative velocity but related to the nature of the moving surfaces. It exhibits a single-valued nonlinearity and is antisymmetric about zero velocity as shown in Figure 2.4 (a). If the contact surfaces are well lubricated, viscous friction may also be considered. In this case, the friction force is proportional to the relative velocity of moving parts.

In addition to the classical Coulomb friction model, other friction models were proposed by several researchers according to observations in their experiments. Gilbart and Winston proposed a model of bearing friction for their optical tracking telescope [53]. Instead of having constant friction torque at all velocities after break-away from stiction, there is a transition region at low velocity where friction decreases as velocity increases as shown in Figure 2.4 (b). After that region, there is a constant Coulomb friction opposing the motion. A very different model was proposed by Walrath based on experimental studies of a stabilized platform with ball bearings on the gimbals [54]. Since the transition behavior at velocity reversal is quite close to exponential, a first order differential equation is used to model friction This model introduced dynamic property of friction but did not include static friction characteristics.

Canudas de Wit, Åstrom and Braum combined Coulomb friction and viscous friction in their friction model for control compensation in DC motor drives [52]. Friction torque is considered to be a piecewise linear function of motor speed as shown in Figure 2.4 (c). This model includes both static and dynamic friction effects. Additionally, they showed that both
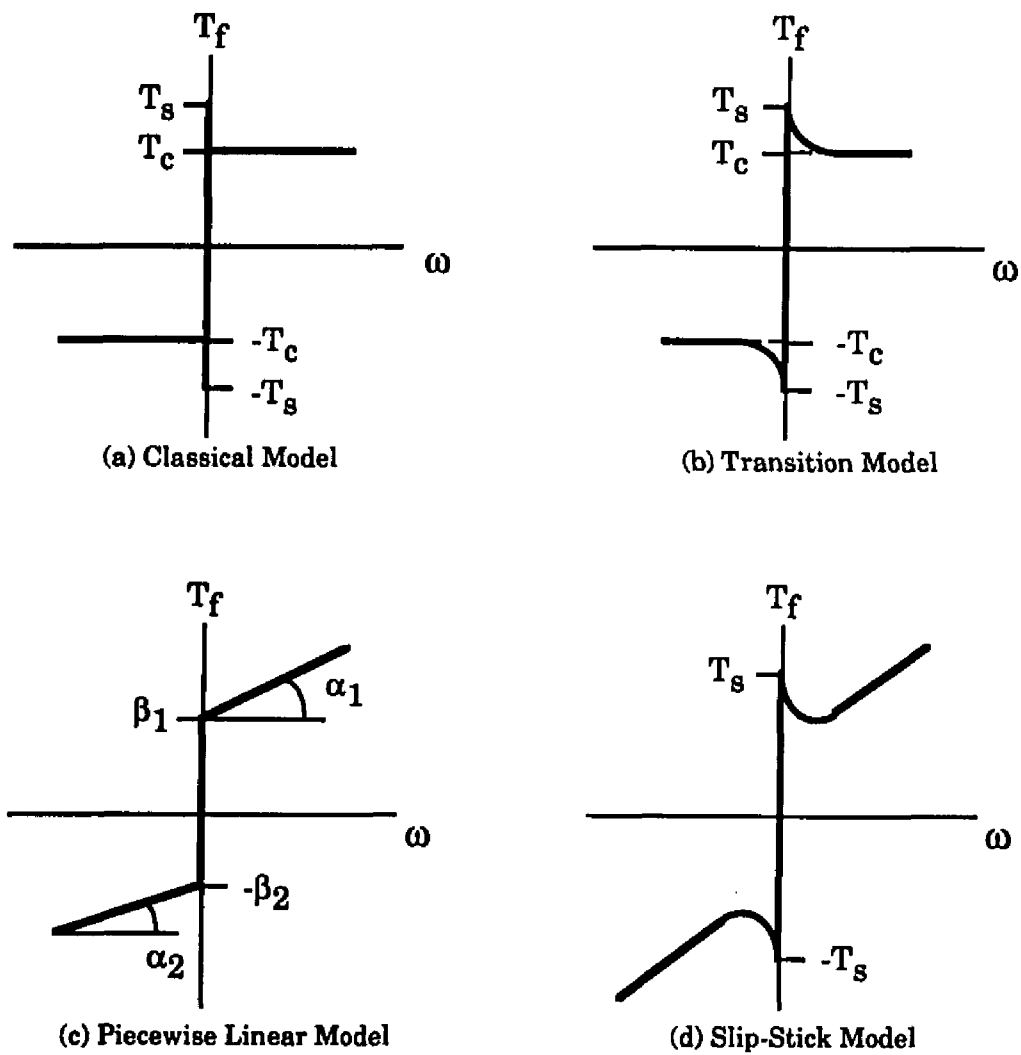
Figure 2.4 Different Models of Friction.

static and dynamic friction have nonsymmetric characteristics and depend on the direction of motion. According to tests on the servo motor drive mechanism of a PUMA 560 robot, a negative relationship between friction and velocity at low velocity after break-away was verified [55]. In this slip-stick friction model as shown in Figure 2.4 (d), friction torque decreases exponentially after stiction torque has been surmounted, reaching approximately 60 percent of the break-away torque, and then increases proportional to the velocity [56].

Estimation of friction in a motor drive system is closely related to controller design of that system. From the control point of view, friction can be treated as disturbance to the system. Different friction models can be used to estimate their magnitude either on-line or off-line. Then, controllers can be adjusted to compensate these nonlinear friction effects. The selection of adequate control techniques to compensate friction torque depends on the choice of friction model. For the dynamic first order differential equation model, it is possible to predict friction behavior and to compensate it by feedforward methods [54]. For the piecewise linear model, friction can be estimated adaptively and compensated by feedback [52]. In a slip-stick model with linear parameters, friction torque is estimated adaptively and compensated in a robot control system with computed torque scheme [56]. The major limitation of the above approaches is that in addition to position and velocity sensing, torque sensing at each joint is also required.

Backlash and Coulomb friction are effectively constant over short time periods and can be compensated. However, over a long term, these effects will change due to wear and may dramatically degrade end effector positioning accuracy. Therefore, it is necessary to perform identification testing on a regular basis to detect developing component failures.

Chapter 3
IDENTIFICATION OF NONLINEARITIES

Nonlinearities exist in all systems. When the magnitude of the nonlinear effect is small in the operating range, the system behaves linearly. However, when nonlinearity is increased or when the system is operating over a large range, the nonlinear effect becomes significant. One method to detect the nonlinearity is to examine the FRF of the system. For instance, jump phenomena occur for systems with cubic stiffness nonlinearity. The Nyquist plot, which is circular for linear system, becomes distorted if nonlinearity exists. However, for hard nonlinearities such as Coulomb friction and backlash, the nonlinear effect in the FRF is not so obvious. It is necessary to develop a method to identify them.

Tomlinson [6] indicated that the Hilbert Transform is a good candidate to detect the nonlinearity in the FRF of a system. For a linear system, the Hilbert transformed FRF will match the original FRF. Therefore, if the Hilbert transformed FRF does not match its original FRF, the system is nonlinear.

This chapter explores how to use the Hilbert Transform to detect various nonlinearities. Section 3.1 reviews the Hilbert Transform. Section 3.2 discusses models of common nonlinearities and Section 3.3 discusses the detection of nonlinearities using the Hilbert Transform.

3.1 Review of the Hilbert Transform

The Hilbert Transform (HT) which is closely related to the Fourier Transform (FT), was first applied to study the phase property of alternating current. Recently, the HT was used to check the linearity in a system's FRF. It was shown that if the Hilbert transformed FRF matches the original FRF, the system is linear. In other words, the system is nonlinear if the Hilbert transformed FRF does not match its FRF [6].

Section 3.1.1 reviews the FT and convolution theorems. Section 3.1.2 describes the HT of the FRF of a causal time signal. In Section 3.1.3, relations between continuous and discrete time system are reviewed. Implementation of the HT to measured FRF is discussed in Section 3.1.4.

3.1.1 The Fourier Transform

The Fourier Transform (FT) pair of a time function $x(t)$ is defined below, subject to the assumption of existence [57].

$$X(\omega) = \int_{-\infty}^{\infty} x(t)e^{-j\omega t}dt$$
$$x(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} X(\omega)e^{j\omega t}d\omega \tag{3.1}$$

The first equation defines the FT of $x(t)$, and the second equation defines the inverse FT. The unit of time, t, is second, and the unit of frequency, ω, is radian per second. The FT pair also can be written in the following form, where the unit of frequency, $f=\omega/2\pi$, is Hertz.

$$X(f) = \int_{-\infty}^{\infty} x(t)e^{-j2\pi ft}dt$$
$$x(t) = \int_{-\infty}^{\infty} X(f)e^{j2\pi ft}df \tag{3.2}$$

The FT pair $x(t) \leftrightarrow X(f)$ enables a mapping between time domain and frequency domain. $X(f)$ provides information about the frequency content of the time function $x(t)$. The FT convention in Equation 3.2 is used in this work. This convention, which satisfies Parseval's theorem and is consistent with the definition of the Laplace transform, is used to derive the Discrete Fourier Transform. For the simplicity of notation, $X(f)$ is rewritten as $X(\omega)$, and the FT pair is written as

$$X(\omega) = \mathcal{F}[x(t)]$$
$$x(t) = \mathcal{F}^{-1}[X(\omega)] \tag{3.3}$$

where $\mathcal{F}[\ ]$ and $\mathcal{F}^{-1}[\ ]$ denote the FT and the inverse FT.

There is a property of FT concerning the evenness and oddness of $x(t)$ and $X(\omega)$. $x(t)$ and $X(\omega)$ can be written in terms of their even and odd parts as

$$x(t) = x_e(t) + x_o(t)$$
$$X(\omega) = X_e(\omega) + X_o(\omega) \tag{3.4}$$

When FT is performed, the even part of $x(t)$ will contribute only to the even part of $X(\omega)$ while the odd part of $x(t)$ will contribute only to the odd part of $X(\omega)$, i.e.,

$$X_e(\omega) = \mathcal{F}[x_e(t)]$$
$$X_o(\omega) = \mathcal{F}[x_o(t)] \tag{3.5}$$

$X(\omega)$ also can be written in terms of its real and imaginary parts.

$$X(\omega) = \text{Re}[X(\omega)] + j\,\text{Im}[X(\omega)] \tag{3.6}$$

Then, if $x(t)$ is a function of real values, the even part of $x(t)$ will contribute only to the real part of $X(\omega)$ while the odd part of $x(t)$ will contribute only to the imaginary part of $X(\omega)$, i.e.,

$$\begin{aligned}\text{Re}[X(\omega)] &= X_e(\omega) = \mathcal{F}[x_e(t)] \\ j\,\text{Im}[X(\omega)] &= X_o(\omega) = \mathcal{F}[x_o(t)]\end{aligned} \tag{3.7}$$

The convolution integral y(t) of two functions x(t) and h(t) is given by

$$y(t) = \int_{-\infty}^{\infty} x(\tau)h(t - \tau)d\tau = x(t) * h(t) \tag{3.8}$$

The time-convolution theorem states that convolution in the time domain is equivalent to multiplication in the frequency domain, i.e., $x(t) * h(t) \Leftrightarrow X(\omega)H(\omega)$, or

$$x(t) * h(t) = \mathcal{F}^{-1}[\mathcal{F}[x(t)] \cdot \mathcal{F}[h(t)]] \tag{3.9}$$

The frequency-convolution theorem states that convolution in the frequency domain is equivalent to multiplication in the time domain, i.e., $x(t)h(t) \Leftrightarrow X(\omega) * H(\omega)$, or

$$X(\omega) * H(\omega) = \mathcal{F}[\mathcal{F}^{-1}[X(\omega)] \cdot \mathcal{F}^{-1}[H(\omega)]] \tag{3.10}$$

3.1.2 The Hilbert Transform

In continuous system theory, the real and imaginary parts, x and y, of an analytic function of complex variable z are, except for an additive constant, explicitly related to each other. Furthermore, under the assumption that the function vanishes at infinity, the following relations are obtained

$$\begin{aligned}x(t) &= \frac{-1}{\pi}\int_{-\infty}^{\infty}\frac{y(\tau)}{t-\tau}d\tau \\ y(t) &= \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{x(\tau)}{t-\tau}d\tau\end{aligned} \tag{3.11}$$

The second equation defines the Hilbert Transform (HT) of $x(t)$ and the first equation defines the inverse HT [58]. $x(t)$ and $y(t)$ form a HT pair, which enables a one-to-one mapping in the same domain. $y(t)$ is a linear functional of $x(t)$ and also can be viewed as the convolution of $x(t)$ with $1/\pi t$, i.e.,

$$\begin{aligned}x(t) &= \mathcal{H}^{-1}[y(t)] = y(t) * \frac{-1}{\pi t} \\ y(t) &= \mathcal{H}[x(t)] = x(t) * \frac{1}{\pi t}\end{aligned} \tag{3.12}$$

where $\mathcal{H}[\ ]$ and $\mathcal{H}^{-1}[\ ]$ denote the HT and the inverse HT.

For a real valued causal time signal $g(t)$, the real and imaginary parts of its FT, $G(\omega)$, can be related to HT pair under the assumption that $G(\omega)$ vanishes at infinity [6]. First, $g(t)$ is written in terms of its even and odd parts.

$$g(t) = g(t)_{even} + g(t)_{odd} \qquad (3.13)$$

where

$$g(t)_{even} = \begin{cases} g(t)/2, & t > 0 \\ g(t)/2, & t < 0 \end{cases}$$

$$g(t)_{odd} = \begin{cases} g(t)/2, & t > 0 \\ -g(t)/2, & t < 0 \end{cases} \qquad (3.14)$$

By employing the signum function $sgn(t)$, the even and odd parts of $g(t)$ can be related by

$$g(t)_{even} = g(t)_{odd} \times sgn(t)$$
$$g(t)_{odd} = g(t)_{even} \times sgn(t) \qquad (3.15)$$

where

$$sgn(t) = \begin{cases} 1, & t > 0 \\ -1, & t < 0 \end{cases} \qquad (3.16)$$

The FT of $g(t)$ is assumed to exist and can be written in terms of its real and imaginary parts.

$$G(\omega) = \text{Re}[G(\omega)] + j\,\text{Im}[G(\omega)] \qquad (3.17)$$

Applying the evenness and oddness properties described in Section 3.1.1, the real and imaginary parts of $G(\omega)$ can be related to the FT of the even and odd parts of $g(t)$.

$$\text{Re}[G(\omega)] = \mathcal{F}[g(t)_{even}] = \mathcal{F}[g(t)_{odd} \times sgn(t)]$$
$$j\,\text{Im}[G(\omega)] = \mathcal{F}[g(t)_{odd}] = \mathcal{F}[g(t)_{even} \times sgn(t)] \qquad (3.18)$$

Then, utilizing the frequency-convolution theorem that multiplication in the time domain is equivalent to convolution in the frequency domain, Equation 3.15 can be rearranged as

$$\text{Re}[G(\omega)] = \mathcal{F}[g(t)_{odd}] * \mathcal{F}[sgn(t)]$$
$$= j\,\text{Im}[G(\omega)] * \frac{-j}{\pi\omega}$$
$$j\,\text{Im}[G(\omega)] = \mathcal{F}[g(t)_{even}] * \mathcal{F}[sgn(t)] \qquad (3.19)$$
$$= \text{Re}[G(\omega)] * \frac{-j}{\pi\omega}$$

Or, more simply

$$\text{Re}[G(\omega)] = \text{Im}[G(\omega)] * \frac{1}{\pi\omega}$$
$$\text{Im}[G(\omega)] = \text{Re}[G(\omega)] * \frac{-1}{\pi\omega} \qquad (3.20)$$

Except that the variable changes from t to ω, Equation 3.20 has the same form of a HT pair as described in Equation 3.12. Note that the unit of ω is Hertz. Furthermore, terms in the left hand side of Equation 3.20 can be viewed as the real and imaginary parts of the HT of $G(\omega)$, $H(\omega)$. Therefore, if

$$H(\omega) = \text{Re}[H(\omega)] + j\,\text{Im}[H(\omega)] \tag{3.21}$$

then

$$\text{Re}[H(\omega)] = \text{Re}[G(\omega)] = \text{Im}[G(\omega)] * \frac{1}{\pi\omega}$$

$$\text{Im}[H(\omega)] = \text{Im}[G(\omega)] = \text{Re}[G(\omega)] * \frac{-1}{\pi\omega} \tag{3.22}$$

The HT of $G(\omega)$ can be written in the form of convolution,

$$\begin{aligned} H(\omega) &= \text{Re}[H(\omega)] + j\,\text{Im}[H(\omega)] \\ &= \text{Im}[G(\omega)] * \frac{1}{\pi\omega} + j\,\text{Re}[G(\omega)] * \frac{-1}{\pi\omega} \\ &= j\,\text{Im}[G(\omega)] * \frac{-j}{\pi\omega} + \text{Re}[G(\omega)] * \frac{-j}{\pi\omega} \\ &= G(\omega) * \frac{-j}{\pi\omega} \\ &= G(\omega) * \mathcal{F}[sgn(t)] \end{aligned} \tag{3.23}$$

or in the form of convolution integral.

$$H(\omega_c) = \frac{-j}{\pi} \text{PV} \int_{-\infty}^{\infty} \frac{G(\omega)}{\omega_c - \omega} d\omega \tag{3.24}$$

where PV indicates that the Cauchy Principal Value of the integral has been taken. Equation 3.23 shows that the HT of $G(\omega)$ is the convolution of $G(\omega)$ with the FT of the signum function. Equation 3.24, which is the frequency domain HT equivalent of Equation 3.11, can be used to evaluated the HT of $G(\omega)$ directly in the frequency domain. The unit of ω in Equation 3.23 is radian per second and the unit of ω in Equation 3.24 can be either Hertz or radian per second. Also, $G(\omega)$ is assumed to vanish at infinity, a condition similar to HT in the time domain [59].

3.1.3 The Discrete Fourier Transform and its Relations to Continuous System

The FRF of a system can be measured efficiently with the help of modern instruments. One important property of these measurements is that they are obtained at

discrete time intervals. Before analyzing these discrete time data, it is necessary to understand the relationship between continuous and the discrete time systems. The Discrete Fourier Transform and discrete convolution are also discussed in the following sections.

The discrete time function $x(k)$ is obtained by sampling the continuous time function $x(t)$ at a constant time interval of T seconds. Thus, $x(k)$ is composed of a sequence of data points with equal time spacing or time step. Index k specifies how many time steps have passed for the data point. The frequency contents of $x(k)$ can be studied using the Discrete Fourier Transform (DFT). Following the FT convention of Equation 3.2, the DFT pair is defined as

$$X(n) = \sum_{k=0}^{N-1} x(k)e^{-j2\pi nk/N} \qquad n = 0, 1, ..., N-1$$

$$x(k) = \frac{1}{N}\sum_{n=0}^{N-1} X(n)e^{j2\pi nk/N} \qquad k = 0, 1, ..., N-1$$

(3.25)

The first equation defines the DFT of N-point discrete time function $x(k)$ and the second equation defines the inverse DFT [60]. DFT pair enables a one-to-one mapping between discrete time domain and discrete frequency domain. $X(n)$ provides information on frequency contents in equal frequency interval of $\Delta f = 1/NT$. This constant frequency interval is commonly called frequency resolution of DFT. Note that the sampling time interval T of the sequence is assumed to be one second. If the sampling interval is not equal to one second, it is necessary to multiply the discrete time function $x(k)$ by T as a scaling factor to make $X(n)$ compatible with the continuous FT, i.e.,

$$X(n) = T\sum_{k=0}^{N-1} x(k)e^{-j2\pi nk/N} \qquad n = 0, 1, ..., N-1$$

(3.26)

Consequently, if the discrete frequency function X(n) is given, it is necessary to multiply it by Δf, the frequency resolution, as a scaling factor to make x(k) compatible with the continuous system, i.e.,

$$x(k) = \Delta f \frac{1}{N}\sum_{n=0}^{N-1} X(n)e^{j2\pi nk/N} \qquad k = 0, 1, ..., N-1$$

(3.27)

The DFT is a good approximation of the continuous FT if (1) the continuous time function $x(t)$ is periodic and band-limited, (2) the sampling rate is at least twice the highest frequency component of $x(t)$. Implicitly, the DFT assumes $x(k)$ and $X(n)$ to be periodic functions with a period of N points. $\mathcal{F}[\ ]$ and $\mathcal{F}^{-1}[\ ]$ are used to denote the DFT and the inverse DFT in the

discrete time system. The discrete convolution of two discrete functions $x(k)$ and $h(k)$ is defined by the summation

$$y(k) = \sum_{i=0}^{N-1} x(i)h(k-i) \quad k = 0, 1, ..., N-1 \tag{3.28}$$

$x(k)$ and $h(k)$ are periodic discrete functions with period of N points. Discrete convolution is commonly written as $y(k) = x(k) * h(k)$. Similarly, in order to have the same scale on the results taken from discrete and continuous convolutions, the sampling interval T is introduced as a scaling factor to modify the discrete convolution as shown in Equation 3.29.

$$y(k) = T \sum_{i=0}^{N-1} x(i)h(k-i) \quad k = 0, 1, ..., N-1 \tag{3.29}$$

If the sampling interval T is made sufficiently small, discrete convolution approximates the continuous convolution with negligible error. Note that discrete convolution yields periodic results because the functions involved are assumed to be periodic. This property is commonly referred to as the circular convolution.

Discrete convolution can be computed more efficiently if convolution theorems and Fast Fourier Transform (FFT) routines are used. When convolution between two time functions $x(k)$ and $h(k)$ is performed, the time-convolution theorem states that convolution in the time domain is equivalent to multiplication in the frequency domain, i.e.,

$$y(k) = x(k) * h(k) = \mathcal{F}^{-1}[\mathcal{F}[x(k)] \cdot \mathcal{F}[h(k)]] \tag{3.30}$$

The DFT of $x(k)$ and $h(k)$ are computed first. Next, the product $Y(n) = X(n)H(n)$ is formed. Finally the inverse DFT of $Y(n)$ is taken to yield the discrete convolution $y(k)$. Then $y(k)$ can be scaled by the sampling interval T for comparison with continuous convolution. When convolution in frequency domain is performed,

$$Y(n) = \sum_{i=0}^{N-1} X(i)H(n-i) \quad n = 0, 1, ..., N-1 \tag{3.31}$$

The frequency-convolution theorem for discrete system states that the convolution in the frequency domain is equivalent to multiplication in the time domain.

$$Y(n) = X(n) * H(n) = N\mathcal{F}[\mathcal{F}^{-1}[X(n)] \cdot \mathcal{F}^{-1}[H(n)]] \tag{3.32}$$

The inverse DFT of $X(n)$ and $H(n)$ are computed first. Next the product $y(k) = x(k)h(k)$ is formed and the DFT of $y(k)$ is taken to transform it back to discrete frequency domain. Finally, the number of samples N is multiplied to the results due to the DFT definition. $Y(n)$

can then be scaled by the frequency resolution Δf for comparison with continuous convolution.

3.1.4 Application of the Hilbert Transform to Measured Frequency Response Functions

For a measured FRF $G(\omega)$, its HT $H(\omega)$ can be evaluated using either Equation 3.23 or Equation 3.24. Because HT requires that $G(\omega)$ vanishes at infinity, it is necessary to choose an appropriate form to represent $G(\omega)$ [59]. For instance, if a system can be defined by the equation of motion as shown in Equation 3.33, its transfer function $G(\omega)$ can be measured in three forms, Receptance, Mobility, and Inertance.

$$m\ddot{x} + c\dot{x} + kx = f(t) \tag{3.33}$$

The Receptance form is obtained if the input force and the output displacement are measured. Therefore,

$$G(\omega) = \frac{\mathcal{F}[x(t)]}{\mathcal{F}[f(t)]} = \frac{1}{-m\omega^2 + jc\omega + k} \tag{3.34}$$

and $\lim_{\omega \to \infty} G(\omega) = 0$. HT can be performed with Receptance type measurement. The Mobility form is obtained if the input force and output velocity are measured. Therefore,

$$G(\omega) = \frac{\mathcal{F}[\dot{x}(t)]}{\mathcal{F}[f(t)]} = \frac{j\omega}{-m\omega^2 + jc\omega + k} \tag{3.35}$$

and $\lim_{\omega \to \infty} G(\omega) = 0$. HT also can be performed with Mobility type measurement. The Inertance form is obtained if the input force and output acceleration are measured. Therefore,

$$G(\omega) = \frac{\mathcal{F}[\ddot{x}(t)]}{\mathcal{F}[f(t)]} = \frac{-\omega^2}{-m\omega^2 + jc\omega + k} \tag{3.36}$$

and $\lim_{\omega \to \infty} G(\omega) = 1/m$. The HT pair discussed in Section 3.1.2 will not hold for Inertance type FRF. Since accelerometers are commonly used in vibration measurements, Inertance type FRF should be modified in order to apply the HT. Inertance type FRF $G(\omega)$ can be converted to mobility type FRF if $G(\omega)$ is divided by $j\omega$. The mobility type $G(\omega)$ can be used to compute its HT $H(\omega)$. $H(\omega)$ is then multiplied by $j\omega$ to convert it back to Inertance type FRF for comparison with $G(\omega)$.

The HT of a measured FRF is implemented using the convolution method as shown in Equation 3.23, i.e.,

$$H(\omega) = G(\omega) * \frac{-j}{\pi\omega} \qquad (3.37)$$

where $G(\omega)$ is the measured FRF and $H(\omega)$ is the Hilbert transformed FRF (HTF). Applying the frequency-convolution theorem and the scaling factor, HT is performed as shown in Equation 3.38.

$$H(\omega) = \Delta f\, N\, \mathcal{F}[\mathcal{F}^{-1}[G(\omega)] \cdot \mathcal{F}^{-1}[\frac{-j}{\pi\omega}]] \qquad (3.38)$$

Computations were performed using the FFT routines in MATLAB. Source code of the program is shown in Appendix A. The Inertance type FRF $G(\omega)$ is first converted to mobility type. HT is performed using frequency-convolution theorem and FFT routines. The HTF $H(\omega)$ is then converted back to Inertance type for comparison with $G(\omega)$. Note that because the measured $G(\omega)$ only covers positive frequency components, it is necessary to reconstruct the negative frequency components. The negative frequency component of $G(\omega)$ is the conjugate of the positive part of $G(\omega)$ and their magnitudes are symmetric about the zero frequency line.

The Inertance type FRF of a single degree of freedom system with equation of motion shown in Equation 3.33 was generated theoretically to examine the HT routine. The equation of motion can be rewritten as

$$m(\ddot{x} + 2\zeta\omega_n \dot{x} + \omega_n^2 x) = f(t) \qquad (3.39)$$

A reference set of system properties is used as test standard. The natural frequency ω_n is set to be 10 Hz, damping ratio is 0.1, mass is 1. The analysis frequency range is set from 0 to 25.6 Hz. Figure 3.1 (a) shows the magnitude of the original FRF $G(\omega)$ and the HTF $H(\omega)$ with a frequency resolution of 0.2 Hz. Besides the magnitude difference at the resonant peak, significant mismatch occurs when frequency is higher than 15 Hz. This is due to the overlapping effect of the FFT convolution.

The HT routine is modified to perform linear convolution by attaching a null buffer of zeros at the folding frequency of the mobility type FRF $G(\omega)$ [6]. A null buffer will not affect the FRF because the response at high frequency for a mobility type FRF will vanish. Note that when null buffer is used, the sampling frequency of the time signal is effectively increased, which is called interpolation in time domain [61]. Figure 3.1 (b) shows the magnitude of the original FRF $G(\omega)$ and the HTF $H(\omega)$ using a null buffer, which has the same size of frequency range as $G(\omega)$. It is obvious that better match is achieved between
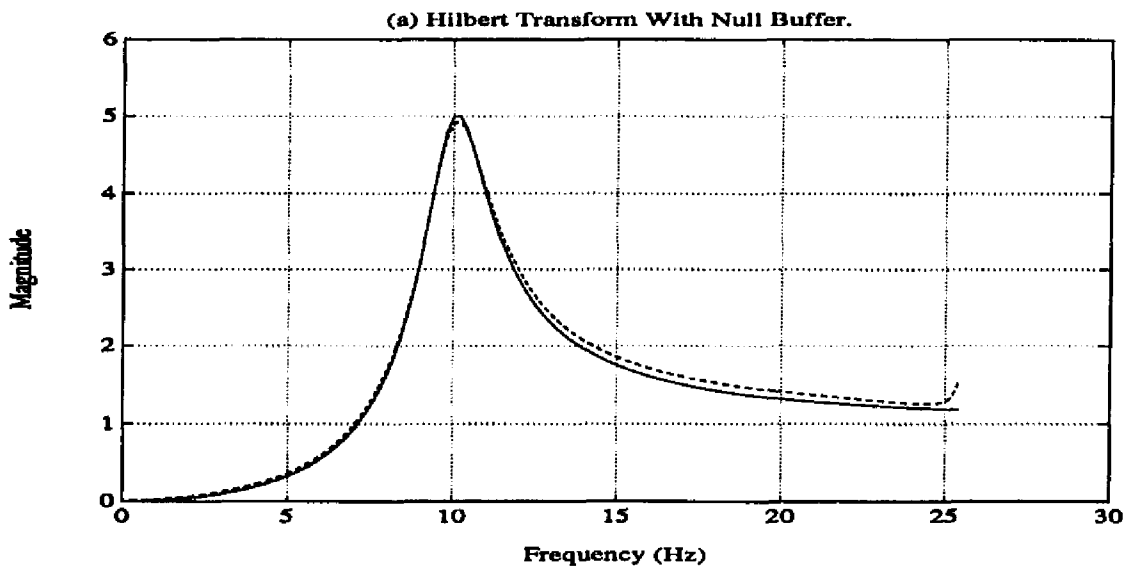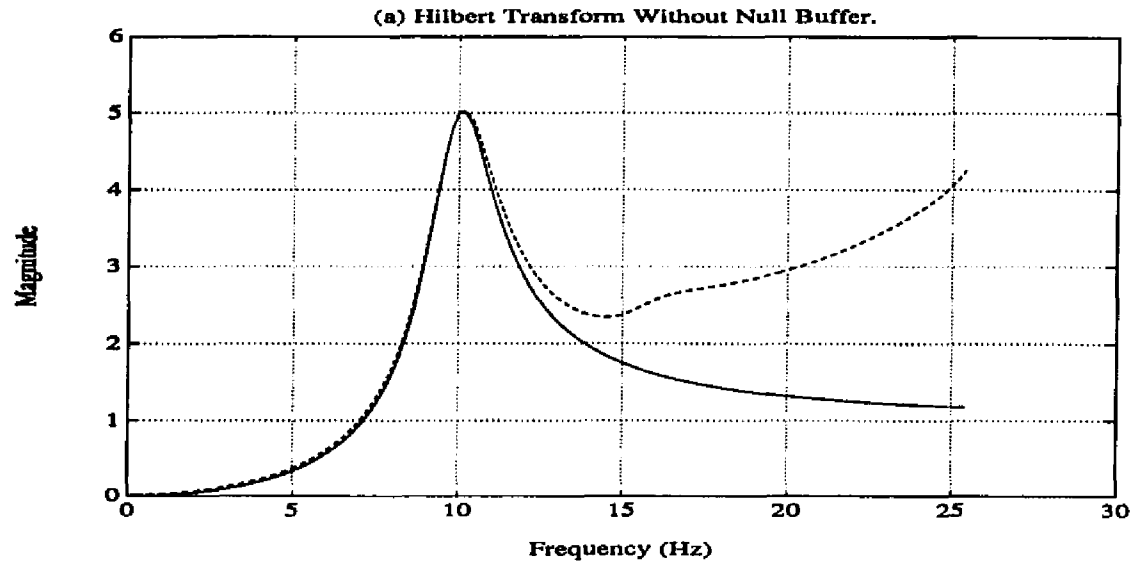
Solid Line: Original Frequency Response Function $G(\omega)$.

Dash Line: Hilbert Transformed Frequency Response Function $H(\omega)$.
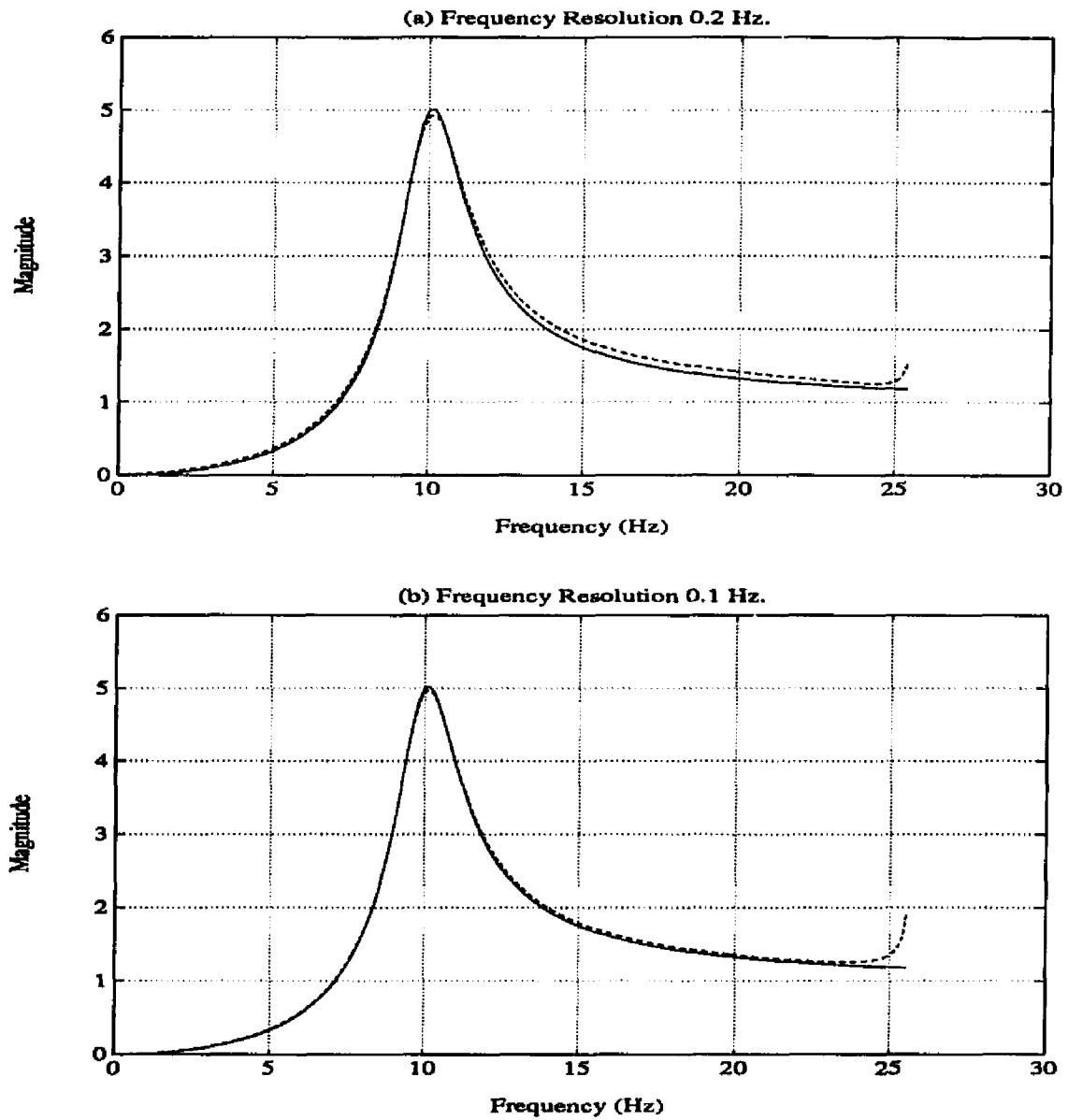
Figure 3.1 Frequency Response Functions of a Single Degree of Freedom System.

(a) Without Null Buffer. (b) With Null Buffer.

$G(\omega)$ and $H(\omega)$. One difference between $G(\omega)$ and $H(\omega)$ is the resonant peak. The peak in $H(\omega)$ is lower than that in $G(\omega)$. The reason for this mismatch is probably due to rough frequency resolution and FFT leakage.

Several examples using different frequency resolutions are shown for comparison in Figure 3.2. Comparing to Figure 3.1 (b), better match was achieved when frequency resolution was improved. In other words, the number of frequency samples has been increased. However, since the magnitude of $G(\omega)$ is not exactly equal to zero at the folding frequency, discontinuity exists between $G(\omega)$ and the null buffer. Therefore, there is a peak at the 25.6 Hz folding frequency. One way to solve this problem is to extend the frequency range of $G(\omega)$ with the same frequency resolution. Then, after HT has been taken with a larger frequency range, only the desired frequency range is retained. Figure 3.3 (a) shows that $G(\omega)$ and $H(\omega)$ match better near the folding frequency using double frequency range and 0.1 Hz resolution.

It is not necessary to double the frequency range and to use a complete buffer at the same time. An example that extends the frequency range by 50 percent and uses a half null buffer is shown in Figure 3.3 (b), which has the same result as shown in Figure 3.3 (a). By improving the frequency resolution to 0.05 Hz, Figure 3.4 shows a very good match of magnitude between the FRF and the HTF. On the other hand, the match of phase is good around the resonant mode. Large jump of phase in the HTF occurs at 0 Hz due to the reconstruction of the negative frequency components. Mismatch of phase also appears near the folding frequency due to discontinuity between $G(\omega)$ and the null buffer.

(a) Frequency Resolution 0.2 Hz.

(b) Frequency Resolution 0.1 Hz.

Solid Line: Original Frequency Response Function $G(\omega)$.

Dash Line: Hilbert Transformed Frequency Response Function $H(\omega)$.

Figure 3.2 Comparison of Different Frequency Resolutions.

(a) $\Delta f = 0.2$ Hz. (b) $\Delta f = 0.1$ Hz. (c) $\Delta f = 0.05$ Hz.

(c) Frequency Resolution 0.05 Hz.

Solid Line: Original Frequency Response Function $G(\omega)$.

Dash Line: Hilbert Transformed Frequency Response Function $H(\omega)$.

Figure 3.2 (continue)

(a) Double Frequency Range.

(b) One and Half Frequency Range.

Solid Line: Original Frequency Response Function $G(\omega)$.

Dash Line: Hilbert Transformed Frequency Response Function $H(\omega)$.

Frequency Resolution 0.1 Hz.

Figure 3.3 Comparison of Different Frequency Ranges.

(a) Double Frequency Range.

(b) One and Half Frequency Range.
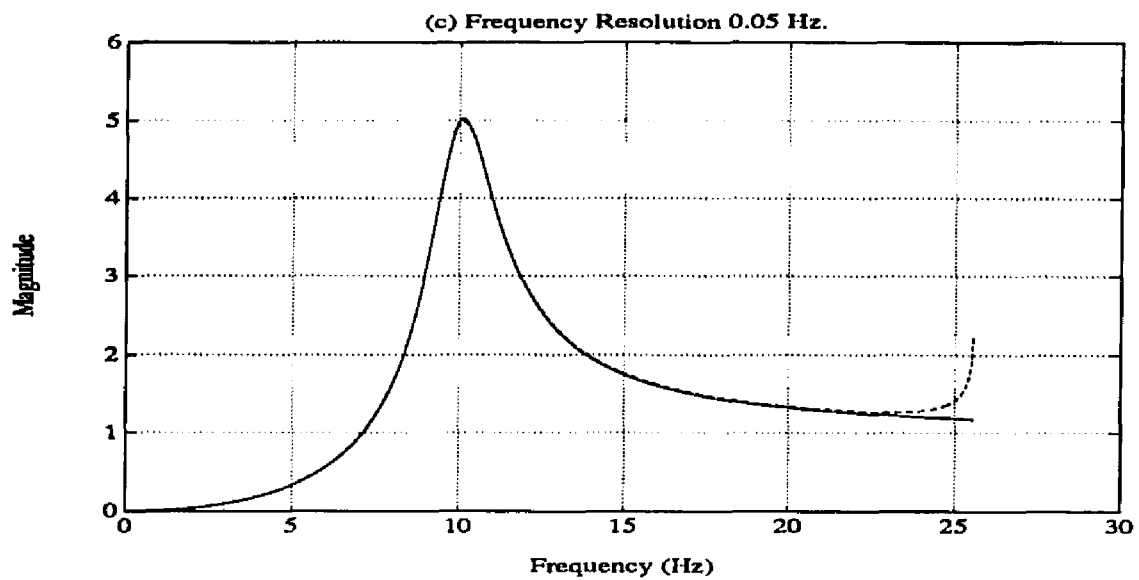
One and Half Frequency Range, Frequency Resolution 0.05 Hz.

Solid Line: Original Frequency Response Function $G(\omega)$.

Dash Line: Hilbert Transformed Frequency Response Function $H(\omega)$.

Frequency Resolution 0.05 Hz.

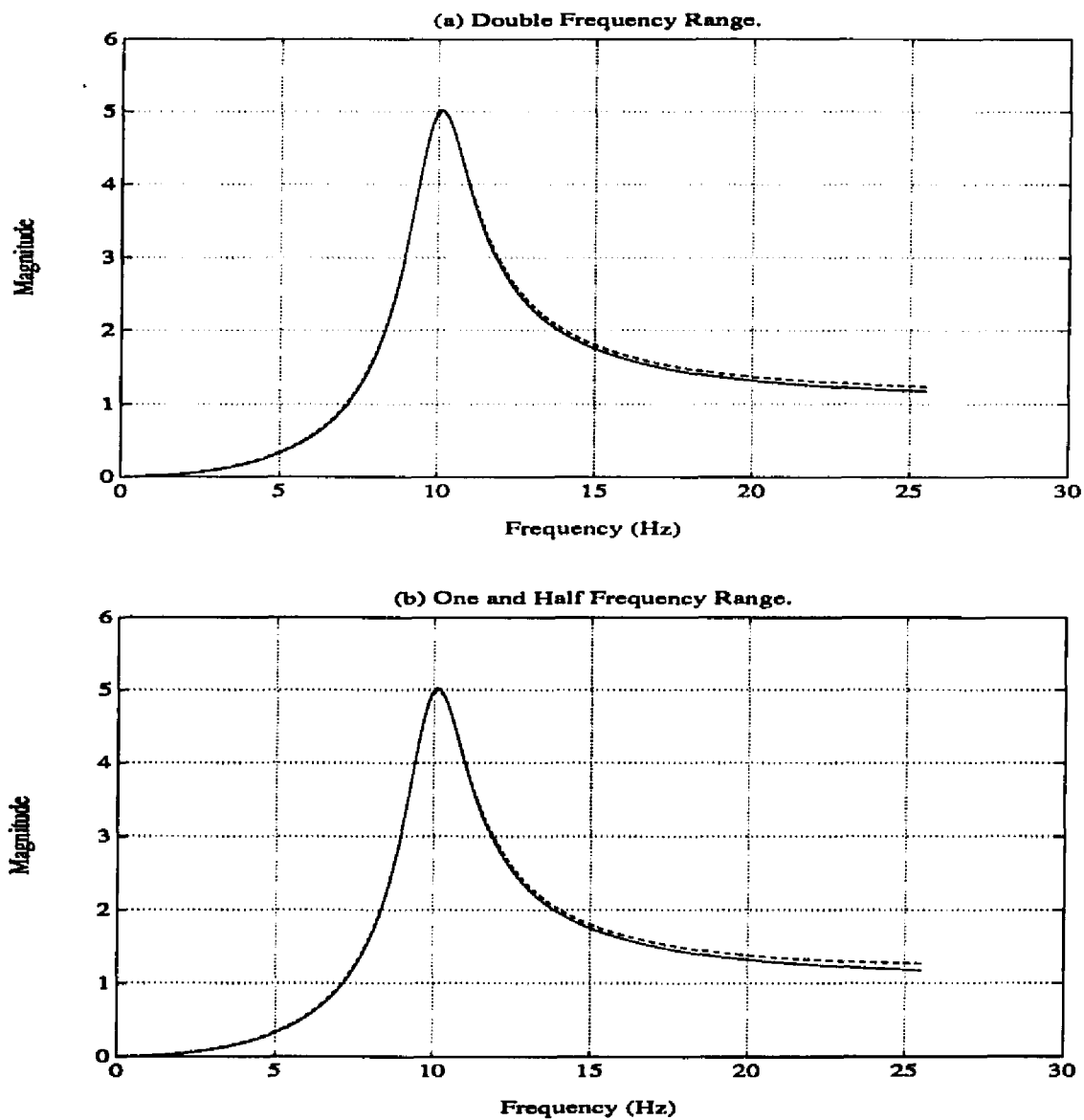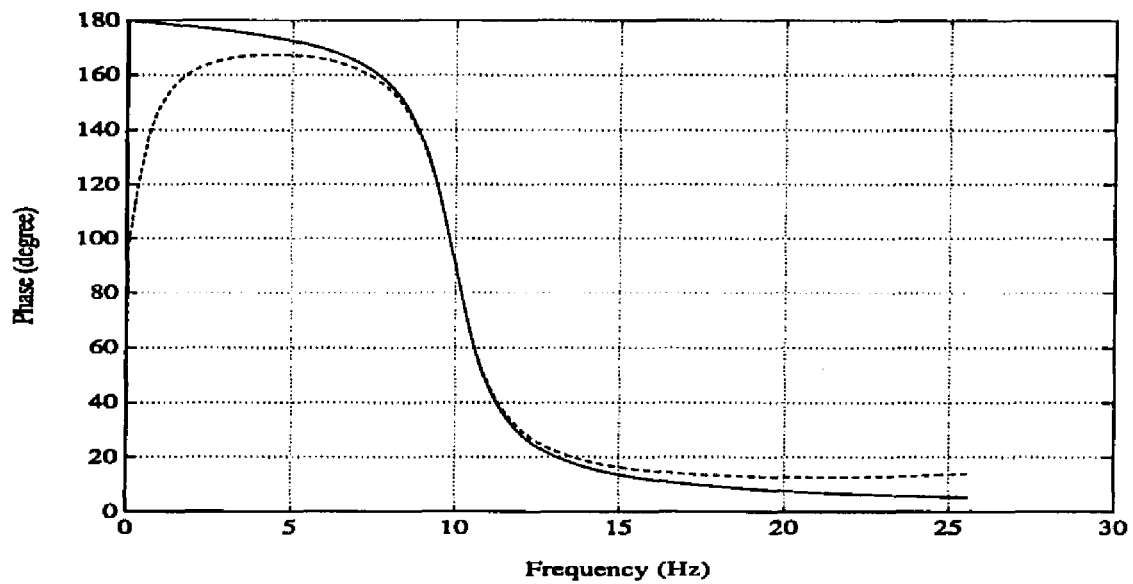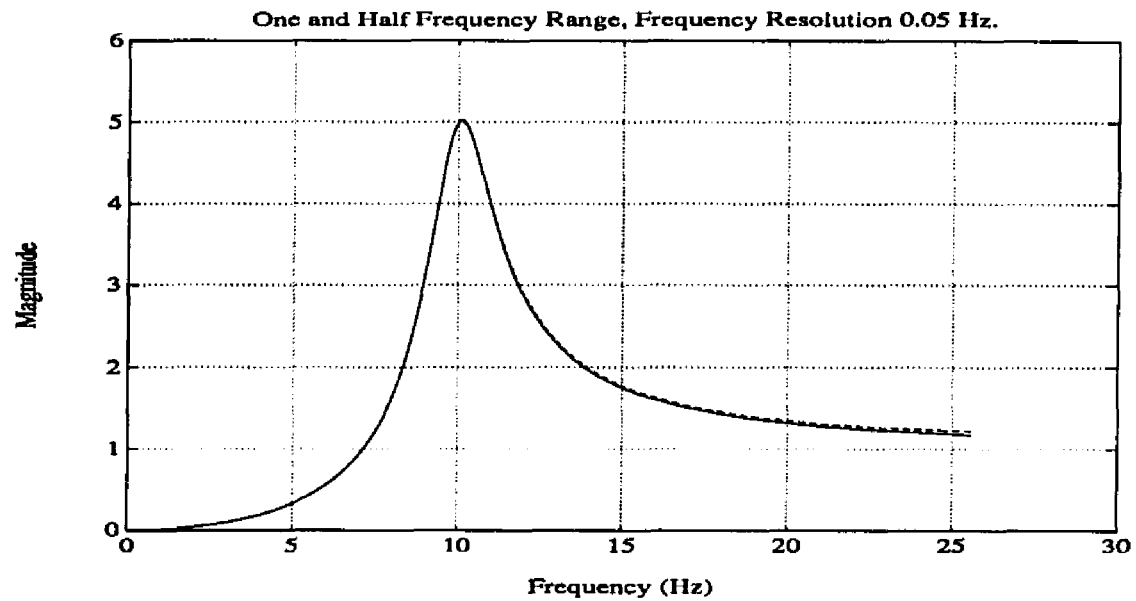Figure 3.4 Hilbert Transform of a Frequency Response Function.

3.2 Frequency Response Function Analyses of Common Nonlinearities

The frequency response method is a powerful tool for the analysis and design of linear systems. It is based on describing a linear system by a complex-valued FRF. The sinusoidal response and stability of a system can be checked by examining its FRF. However, these frequency domain analysis techniques cannot be directly applied to nonlinear systems. Fortunately, for some nonlinear phenomena and under certain conditions, an extended version of the frequency response method, the describing function method, can be used to approximately analyze and predict nonlinear behavior. Common nonlinearities such as cubic stiffness and backlash can be modeled using the describing function method. The main use of describing function analysis is to predict limit cycles in nonlinear systems.

In this section, FRFs of common nonlinearities were studied using the describing function method. Section 3.2.1 discusses the characteristics of several nonlinearities that often occur in physical systems. The describing function method is presented in Section 3.2.2. FRF analyses of a single mode system with friction are presented in Section 3.2.3 and FRF of a gear pair system with backlash is presented in Section 3.2.4.

3.2.1 Common Nonlinearities

Nonlinearities can be classified as continuous and discontinuous. Continuous nonlinearities such as cubic bending stiffness and quadratic damping can be approximated by linear functions in the range of operation. Because discontinuous nonlinearities can not be locally approximated by linear functions, they are also called "hard" nonlinearities. Hard nonlinearities are commonly found in physical systems both in small range operation and large range operation. Whether a system in small range operation should be regarded as linear or nonlinear depends on the magnitude of the hard nonlinearities and on the extend of their effects on the system performance. There are four types of common hard nonlinearities in a physical system: saturation, on-off nonlinearity, dead-zone, and backlash/hysteresis [41].

The saturation nonlinearity is usually caused by limits on component size, properties of materials, and available power. It is often associated with amplifiers and actuators in a system. The output varies linearly with the input only for small amplitude input. When the input amplitude gets out of the linear range of the amplifier, the output changes very little

and stays close to its maximum value, giving rise to the saturation phenomenon. The characteristic of saturation nonlinearity is plotted in Figure 3.5 (a), where the solid line is the real nonlinearity and the dotted line is an ideal saturation nonlinearity. The general effect of saturation is to reduce the gain of component as the input signal is increased. As a result, if a system is unstable in its linear range, its response may go into a self-sustained oscillation. On the other hand, in a linearly stable system, saturation tends to slow down the response of the system because it reduces the effective gain.

An extreme case of saturation is the on-off or relay nonlinearity. It occurs when the linear range is reduced to zero and the slope in the linearity range becomes vertical as shown in Figure 3.5 (b). On-off nonlinearity has effects similar to those of saturation. Note that Coulomb friction has a characteristic similar to on-off nonlinearity.

Dead-zone nonlinearity exists in devices such as sensors, amplifiers, and actuators. The output of the system remains zero until the magnitude of the input exceeds a certain value. Figure 3.5 (c) shows the characteristic of dead-zone nonlinearity. The direct effect of dead-zone nonlinearity is to decrease static output accuracy. They may also lead to limit cycles or system instability. In some cases however, dead-zone nonlinearity may actually stabilize a system or suppress self-oscillations.

Backlash is commonly found in gear trains and similar mechanical linkages where the coupling is not perfect. Backlash is usually caused by the small gaps that exists in transmission mechanisms. For example, there always exist small gaps between a pair of mating gears due to unavoidable errors in manufacturing and assembly. As a result, dead-zone effect occurs whenever the driving gear changes its direction of rotation. There will be no output motion of the driven gear until the driving gear contacts the driven gear. After contact has been established, the driven gear follows the rotation of the driving gear in a linear fashion. Figure 3.5 (d) shows the characteristic of backlash nonlinearity. Note the multi-valued nature of backlash. The output of the component depends on the history of the input. Hysteresis, which is frequently observed in relay components, is another case of multi-valued nonlinearity. These multi-valued nonlinearities usually lead to energy storage in the system. Energy storage is a frequent cause of instability and self-sustained oscillation in a system.
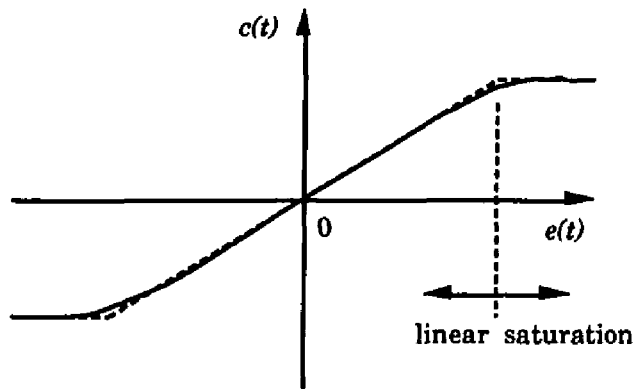
(a) Saturation Nonlinearity.
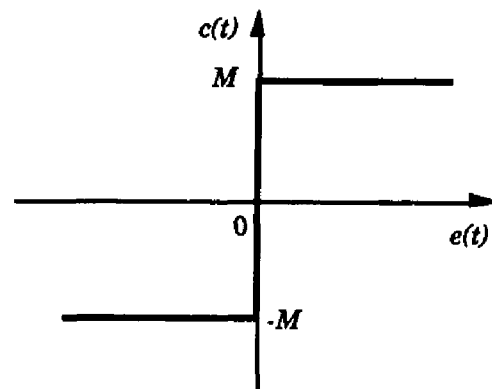
(b) On-Off Nonlinearity.
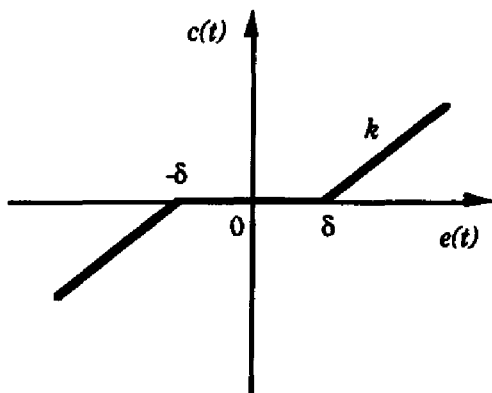
(c) Dead-Zone Nonlinearity.

(d) Backlash Nonlinearity.

Figure 3.5 Characteristics of Common Nonlinearities.

3.2.2 Describing Functions of Common Nonlinearities

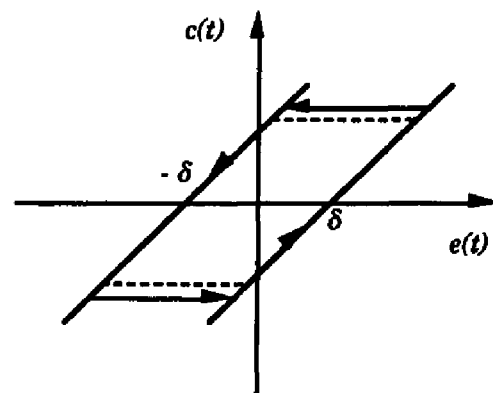The describing function method is an extended frequency response analysis of nonlinear elements in a system [41]. For the basic version of the describing function method, several assumptions are applied. There should be only one time-invariant nonlinear element in the system. If there are two or more nonlinear components in a system, they should be lumped together or retain only the primary nonlinearity and neglect the others. The output response of the nonlinear element due to sinusoidal input is approximated by a quasi-linear element. Only the fundamental harmonic term in the output response is analyzed because other higher harmonic terms are assumed to be attenuated by the linear parts of the system. This assumption is often referred to as the low-pass filtering hypothesis. The transfer function of this quasi-linear element depends on the amplitude of the input sinusoidal excitation. For simplicity, nonlinearity is assumed to have the property of an odd function. This means that the plot of the nonlinearity relation between input and output of the nonlinear element is symmetric about the origin. Note that common nonlinearities discussed in the previous section all satisfy this assumption.

Consider the sinusoidal input $e(t) = A \sin(\omega t)$ with amplitude A and frequency ω of a nonlinear element as shown in Figure 3.6. The output of the nonlinear component $c(t)$ is often a periodic, though generally non-sinusoidal, function. Using Fourier series, this periodic function can be expanded as

$$c(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} [a_n \cos(n\omega t) + b_n \sin(n\omega t)] \tag{3.40}$$

The Fourier coefficients a_i's and b_i's are functions of A and ω, determined by

$$a_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} c(t) d(\omega t)$$

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} c(t) \cos(n\omega t) d(\omega t) \tag{3.41}$$

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} c(t) \sin(n\omega t) d(\omega t)$$

Due to the assumption of an odd function, $a_0 = 0$. Due to the filtering hypothesis, only the fundamental harmonic components a_1 and b_1 are retained. Therefore, the output $c(t)$ is represented by the fundamental component $c_1(t)$.

$$c(t) \approx c_1(t) = a_1 \cos(\omega t) + b_1 \sin(\omega t) = M \sin(\omega t + \phi) \tag{3.42}$$

where $M(A, \omega) = \sqrt{a_1^2 + b_1^2}$ and $\phi(A, \omega) = \tan^{-1}(a_1 / b_1)$.

Equation 3.42 indicates that the fundamental component corresponding to a sinusoidal input is a sinusoid at the same frequency. In complex representation, this sinusoid can be written as $c_1 = Me^{j(\omega t + \phi)} = (b_1 + ja_1)e^{j\omega t}$. The describing function of the nonlinear element is defined as the complex ratio of the fundamental component of the nonlinear element by the input sinusoid.

$$M(A,\omega) = \frac{Me^{j(\omega t + \phi)}}{Ae^{j\omega t}} = \frac{M}{A}e^{j\phi} = \frac{1}{A}(b_1 + ja_1) \tag{3.43}$$

This definition is similar to the concept of frequency response function, which is the frequency domain ratio of the sinusoidal input and sinusoidal output of a system. The nonlinear element is treated as a linear element with a FRF $M(A,\omega)$ as shown in Figure 3.6. The describing function of this nonlinear element depends on the input amplitude A and the frequency ω. This representation of the nonlinear element is also called quasi-linearization.
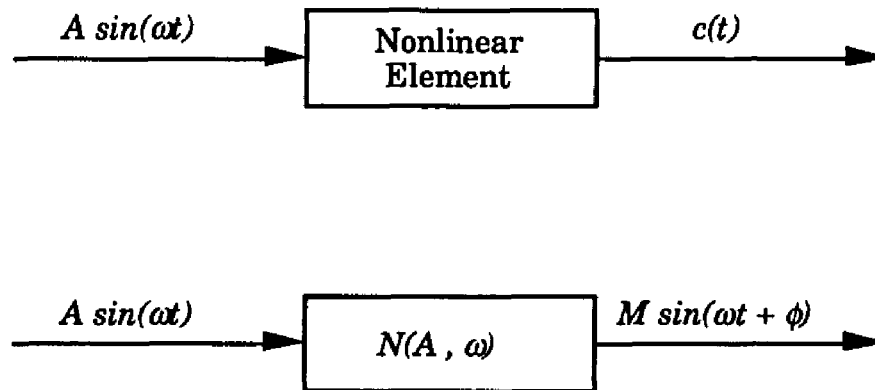
Figure 3.6 A Nonlinear Element and its Describing Function Representation.

3.2.3 Frequency Response Function of a System with Friction

The equation of motion of a single degree of freedom system with Coulomb friction can be written as

$$m(\ddot{x} + 2\zeta\omega_n\dot{x} + \omega_n^2 x) + R\frac{\dot{x}}{|\dot{x}|} = f(t) \tag{3.44}$$

or

$$\ddot{x} + 2\zeta\omega_n\dot{x} + \omega_n^2 x + \frac{R}{m}\frac{\dot{x}}{|\dot{x}|} = \frac{1}{m}f(t) \tag{3.45}$$

where the Coulomb friction term

$$\frac{R}{m}\frac{\dot{x}}{|\dot{x}|} = \frac{R}{m} \qquad \text{if } \dot{x} > 0$$
$$= -\frac{R}{m} \qquad \text{if } \dot{x} < 0 \tag{3.46}$$

The describing function of the friction term is derived according to Equation 3.40 to 3.43. Since the friction term is an even function and symmetric with respect to $\omega t = 0$ as shown in Figure 3.7, the Fourier series coefficients a_0 and b_0 are equal to 0. The coefficient of the cosine term $a_1 = \frac{4R}{m\pi}$. Therefore, the friction term can be represented as

$$\frac{R}{m}\frac{\dot{x}}{|\dot{x}|} = \frac{4R}{m\pi}\cos(\omega t) = \frac{4R}{m\pi\omega A}\omega A\cos(\omega t) = \frac{4R}{m\pi\omega A}\dot{x} \tag{3.47}$$

The equation of motion can then be written as

$$\ddot{x} + (2\zeta\omega_n + \frac{4R}{m\pi\omega A})\dot{x} + \omega_n^2 x = \frac{1}{m}f(t) \tag{3.48}$$

Assuming that the input forcing function is sinusoid with amplitude F, $f(t) = F\,sin(\omega t)$, the output displacement is a sinusoid at the same frequency with amplitude A and phase shift ψ, $x(t) = A\,sin(\omega t + \psi)$. The velocity is assumed to have a form of $\dot{x}(t) = A\omega\cos(\omega t + \psi)$. Applying the method of slowly changing amplitude and phase [62], the amplitude and phase equations can be written as

$$\dot{A} = \frac{A}{\omega}(\omega^2 - \omega_n^2)sc - (2\zeta\omega_n + \frac{4R}{m\pi\omega A})Ac^2 + \frac{F}{m\omega}c\,sin(\omega t)$$
$$\dot{\psi} = \frac{-1}{\omega}(\omega^2 - \omega_n^2)s^2 + (2\zeta\omega_n + \frac{4R}{m\pi\omega A})sc - \frac{F}{m\omega A}s\,sin(\omega t) \tag{3.49}$$

where the abbreviations $s = sin(\omega t + \psi)$ and $c = cos(\omega t + \psi)$ were used. Then, taking the mean of the right-hand sides of Equation 3.49 over a period of 2π yields

$$\dot{A} = -(\zeta\omega_n + \frac{2R}{m\pi\omega A})A - \frac{F\sin\psi}{2m\omega}$$
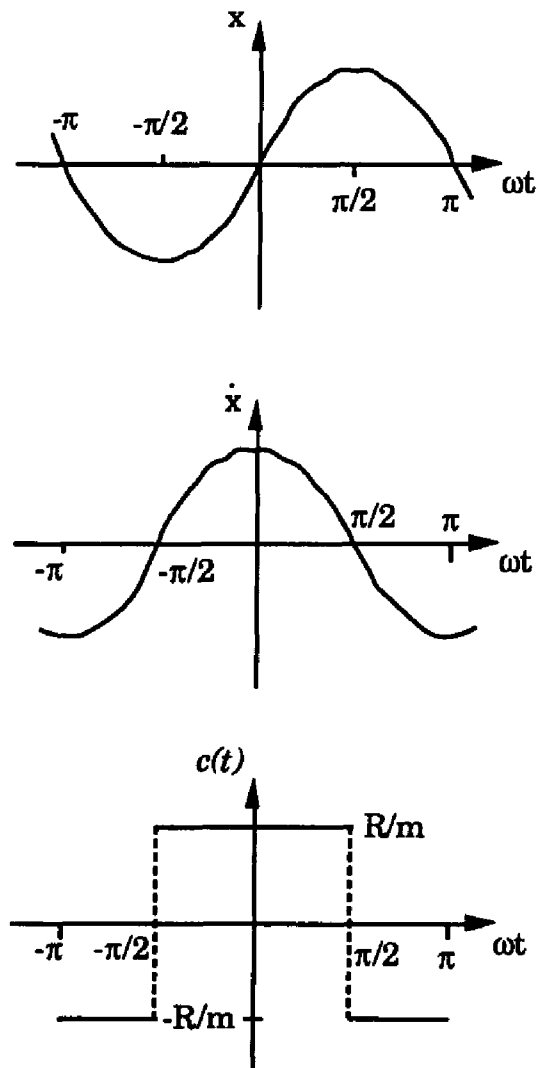$$\dot{\psi} = \frac{-1}{2\omega}(\omega^2 - \omega_n^2) - \frac{F\cos\psi}{2m\omega A} \tag{3.50}$$

Figure 3.7 Response of a Coulomb Friction Element.

The stationary solutions are obtained with $\dot{A} = 0$ and $\dot{\psi} = 0$, i.e.,

$$-(\zeta\omega_n + \frac{2R}{m\pi\omega A})A - \frac{F\sin\psi}{2m\omega} = 0$$

$$\frac{-1}{2\omega}(\omega^2 - \omega_n^2) - \frac{F\cos\psi}{2m\omega A} = 0$$

(3.51)

The amplitude A can be obtained from the following second order equation.

$$[(\omega^2 - \omega_n^2)^2 + (2\zeta\omega_n\omega)^2]A^2 + \frac{16\zeta\omega_n\omega R}{m\pi}A + [(\frac{4R}{m\pi})^2 - (\frac{F}{m})^2] = 0$$

(3.52)

One simplied solution for A is

$$A = \sqrt{\frac{(\frac{F}{m})^2 - (\frac{4R}{m\pi})^2}{(\omega^2 - \omega_n^2)^2 + (2\zeta\omega_n\omega)^2}}$$

(3.53)

The phase angle ψ cab be obtained from Equation 3.51.

$$\psi = tan^{-1}(\frac{\frac{4R}{m\pi A} + 2\zeta\omega_n\omega}{\omega^2 - \omega_n^2})$$

(3.54)

It is shown in Equation 3.53 that there is no motion until the input forcing function surpass the Coulomb friction force. The Inertance type FRF of this system can then be computed as

$$G(\omega) = \frac{-\omega^2 A\sin(\omega t + \psi)}{\frac{F}{m}\sin(\omega t)} = \frac{\omega^2 A\sin(\omega t + \pi + \psi)}{\frac{F}{m}\sin(\omega t)}$$

(3.55)

The magnitude and phase of the FRF are

$$|G(\omega)| = \frac{\omega^2 A}{F/m} = \frac{m\omega^2}{F}A$$

$$\angle G(\omega) = \pi + \psi$$

(3.56)

The FRF of this system depends on the level of Coulomb friction R, level of input force F, and frequency ω. The Inertance type FRFs with different forcing levels are shown in Figure 3.8. As the input forcing level increases, the effect of Coulomb friction is reduced so that the FRF is close to that of the system without friction. The Inertance type FRFs with different friction levels and the same forcing level are shown in Figure 3.9. As the level of Coulomb friction increases, the magnitude at the resonant peak is reduced.
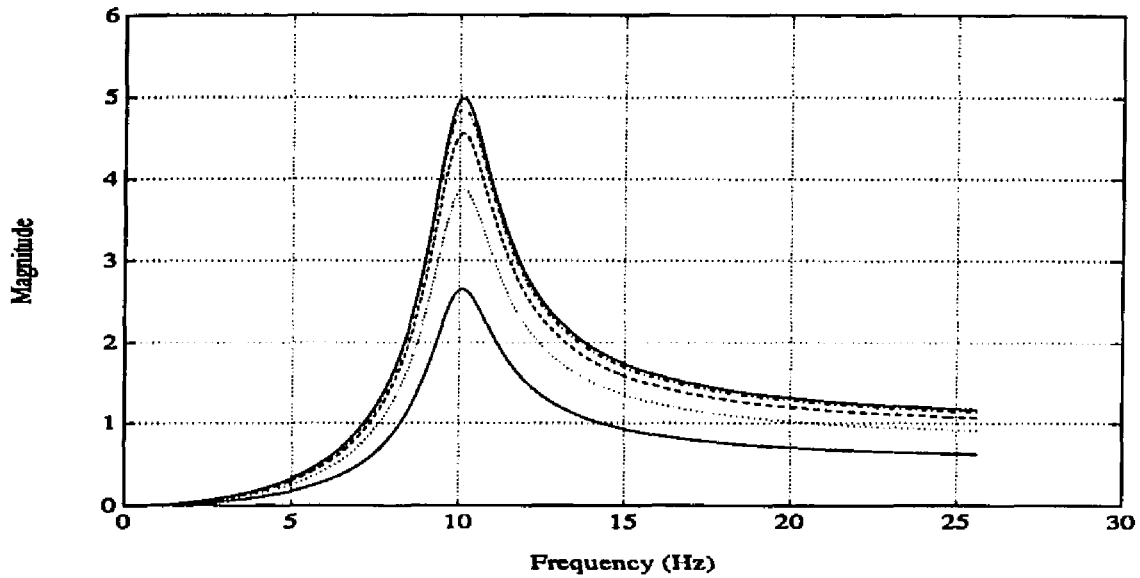
Coulomb Friction Level R = 1

Solid Line: Forcing Level F = 10

Dash-Dot Line: Forcing Level F = 5

Dash Line: Forcing Level F = 3

Dot Line: Forcing Level F = 2

Bold Line: Forcing Level F = 1.5

Figure 3.8 Frequency Response Functions of a System with Coulomb Friction
with Different Forcing Levels.
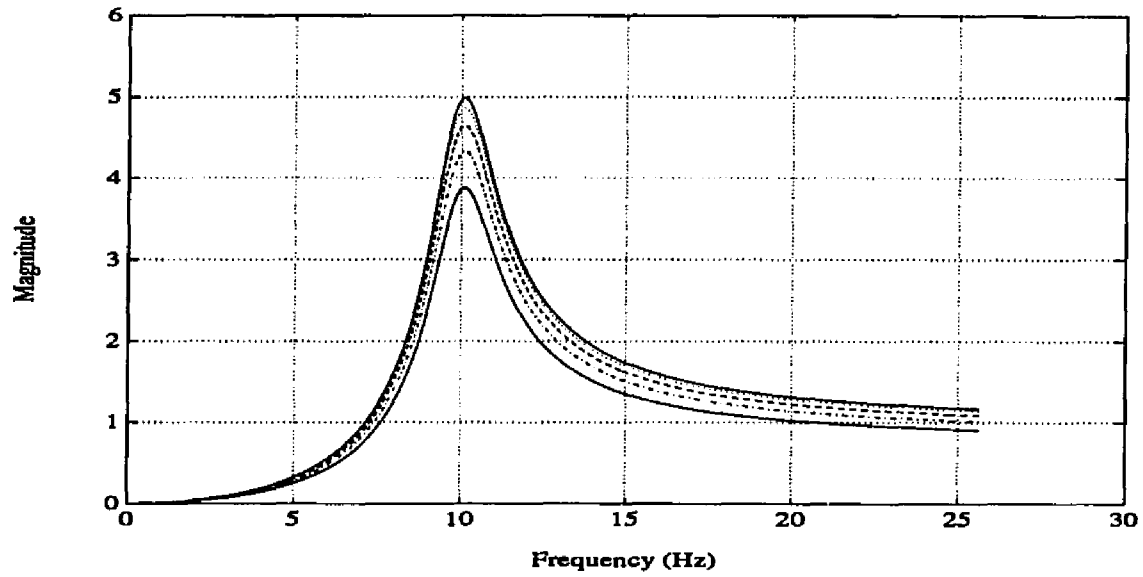
Forcing Level F = 10

Solid Line: Coulomb Friction Level R = 1

Dash-Dot Line: Coulomb Friction Level R = 2

Dash Line: Coulomb Friction Level R = 3

Dot Line: Coulomb Friction Level R = 4

Bold Line: Coulomb Friction Level R = 5

Figure 3.9 Frequency Response Functions of a System with Different Levels of Coulomb Friction.

3.2.4 Frequency Response Function of a System with Backlash

Backlash exists in all gear drives. Excess backlash may lead to undesirable inaccuracy, oscillation, and instability in control systems. In addition, it will wear out the mechanical elements. Generally, the dynamics of the gear drive with backlash depends on the relative inertia-to-friction ratio of the output driven gear. If the effective load inertia of the output gear is small compared with that of the input driving gear, the motion is dominated by friction. There will be no output motion of the driven gear when the driving gear reverses its rotation. Motion of the driven gear will be restored when the driving gear contacts the driven gear again. It is assumed that the output gear will instantaneously take on the velocity of the input gear. Figure 3.5 (d) shows a typical characteristic between input and output of the gear drive with friction controlled backlash.

A schematic plot of the gear drive system in robot joint is shown in Figure 3.10. The dynamic equations of the system are

$$L_a \frac{di_a(t)}{dt} + R_a i_a = e_a(t) - e_b(t) \tag{3.57}$$

$$T_m(t) = K_i i_a(t) \tag{3.58}$$

$$e_b(t) = K_b \frac{d\theta_m(t)}{dt} = K_b \omega_m(t) \tag{3.59}$$

$$J_m \frac{d^2\theta_m(t)}{dt^2} + B_m \frac{d\theta_m(t)}{dt} = T_m(t) - T_L(t) \tag{3.60}$$

$$T_L(t) = K(\theta_m(t) - \theta_1(t)) = J_1 \frac{d^2\theta_1(t)}{dt^2} + B_1 \frac{d\theta_1(t)}{dt} + T_1(t) \tag{3.61}$$

$$T_1(t)\theta_1(t) = T_2(t)\theta_2(t) \tag{3.62}$$

$$T_2(t) = J_2 \frac{d^2\theta_2(t)}{dt^2} + B_2 \frac{d\theta_2(t)}{dt} \tag{3.63}$$

where K_i is the torque constant, K_b is the back-emf constant, K is the torsional spring constant of the shaft, B_1, B_2 and B_m are the viscous friction constants. Rearranging Equation 3.57 to 3.60, the relationship between the control voltage input $e_a(t)$ and motor shaft angle output $\theta_m(t)$ can be written as

$$L_a J_m \frac{d^3\theta_m(t)}{dt^3} + (R_a J_m + L_a B_m)\frac{d^2\theta_m(t)}{dt^2} + (R_a B_m + K_b K_i)\frac{d\theta_m(t)}{dt} = K_i e_a(t) - \frac{dT_L}{dt} - R_a T_L$$

$$\tag{3.64}$$

or in terms of motor shaft angular velocity $\omega_m(t)$

$$L_a J_m \frac{d^2\omega_m(t)}{dt^2} + (R_a J_m + L_a B_m)\frac{d\omega_m(t)}{dt} + (R_a B_m + K_b K_i)\omega_m(t) = K_i e_a(t) - \frac{dT_L}{dt} - R_a T_L$$
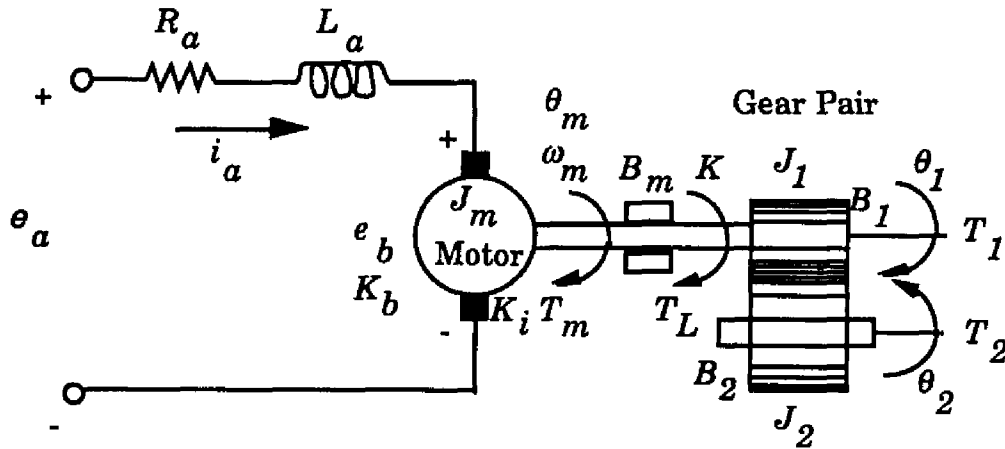
$$\tag{3.65}$$

Figure 3.10 DC Motor and Gear Drive System.

Equation 3.65 describes the dynamics of the gear drive in a position control system using velocity control. Readings from the position encoder and velocity tachometer feedback can be used to generate the proportional and derivative error signal. Because backlash exists in the gear drive, the effective load to the motor T_L is different within and out of the backlash region. For a friction controlled backlash, the effective load is

$$T_L(t) = \begin{cases} J_1 \dfrac{d^2\theta_1(t)}{dt^2} + B_1 \dfrac{d\theta_1(t)}{dt} & \text{within backlash} \\ J_1 \dfrac{d^2\theta_1(t)}{dt^2} + B_1 \dfrac{d\theta_1(t)}{dt} + \dfrac{\theta_2}{\theta_1}(J_2 \dfrac{d^2\theta_2(t)}{dt^2} + B_2 \dfrac{d\theta_2(t)}{dt}) & \text{out of backlash} \end{cases}$$ (3.66)

or in terms of angular velocity

$$T_L(t) = \begin{cases} J_1 \dfrac{d\omega_1(t)}{dt} + B_1\omega_1(t) & \text{within backlash} \\ J_1 \dfrac{d\omega_1(t)}{dt} + B_1\omega_1(t) + \dfrac{\omega_2}{\omega_1}(J_2 \dfrac{d\omega_2(t)}{dt} + B_2\omega_2(t)) & \text{out of backlash} \end{cases}$$ (3.67)

Note that the driving gear and the motor shaft have the same angular velocity $\omega_m = \omega_1$. With gear ratio n, $\omega_2 = n\omega_1$ when the gear pair is out of the backlash region. Equation 3.65 has the same form as a simple mass-damper-spring system when the effective load T_L is substituted according to Equation 3.67. The effective mass, damping and stiffness are larger

when gears are out of the backlash region than within the backlash region. The increased value is due to the effect of the driven gear.

When FRF of the gear drive system is measured, several input/output combinations can be used to examine the effect of gear backlash. Possible inputs include the control voltage to the motor, position encoder and velocity tachometer readings of the motor shaft. Note that these parameters are not affected by backlash. Possible outputs include the joint angle, velocity of the driven gear and readings of an accelerometer attached on the link. Note that these parameters are affected by backlash. The FRF of the motor gear drive system will probably lie between FRFs of the system within and out of the backlash region.

The input/output relationship of a friction controlled backlash nonlinearity is plotted in Figure 3.11, with slope n and width 2δ. For a sinusoidal input $e(t) = A\sin(\omega t)$, $A \geq \delta$, the output $c(t)$ of the nonlinearity can be represented in one cycle as

$$c(t) = \begin{cases} (A-\delta)n & \pi/2 < \omega t \leq \pi - \omega t_1 \\ (A\sin(\omega t) + \delta)n & \pi - \omega t_1 < \omega t \leq 3\pi/2 \\ -(A-\delta)n & 3\pi/2 < \omega t \leq 2\pi - \omega t_1 \\ (A\sin(\omega t) - \delta)n & 2\pi - \omega t_1 < \omega t \leq 5\pi/2 \end{cases} \quad for \qquad (3.68)$$

where ωt_1 is the time required for the driving gear to rotate 2δ after reversal and to make contact with the driven gear. Because the output $c(t)$ is symmetric, $a_0 = 0$. However, $c(t)$ is neither odd nor even function. According to Equation 3.41, the coefficients of the fundamental frequency can be obtained by integrating from $\pi/2$ to $5\pi/2$ for one cycle.

$$a_1 = \frac{4n\delta}{\pi}(\frac{\delta}{A} - 1)$$

$$b_1 = \frac{4n}{\pi}[\frac{\pi}{2} - \sin^{-1}(\frac{2\delta}{A} - 1) - (\frac{2\delta}{A} - 1)\sqrt{1 - (\frac{2\delta}{A} - 1)^2}] \qquad (3.69)$$

The describing function of the backlash can then be calculated by

$$|N(A)| = \frac{1}{A}\sqrt{a_1^2 + b_1^2}$$

$$\angle N(A) = \tan^{-1}(\frac{a_1}{b_1}) \qquad (3.70)$$

Note that for perfect friction controlled backlash, the driven gear output only depends on the input amplitude A.
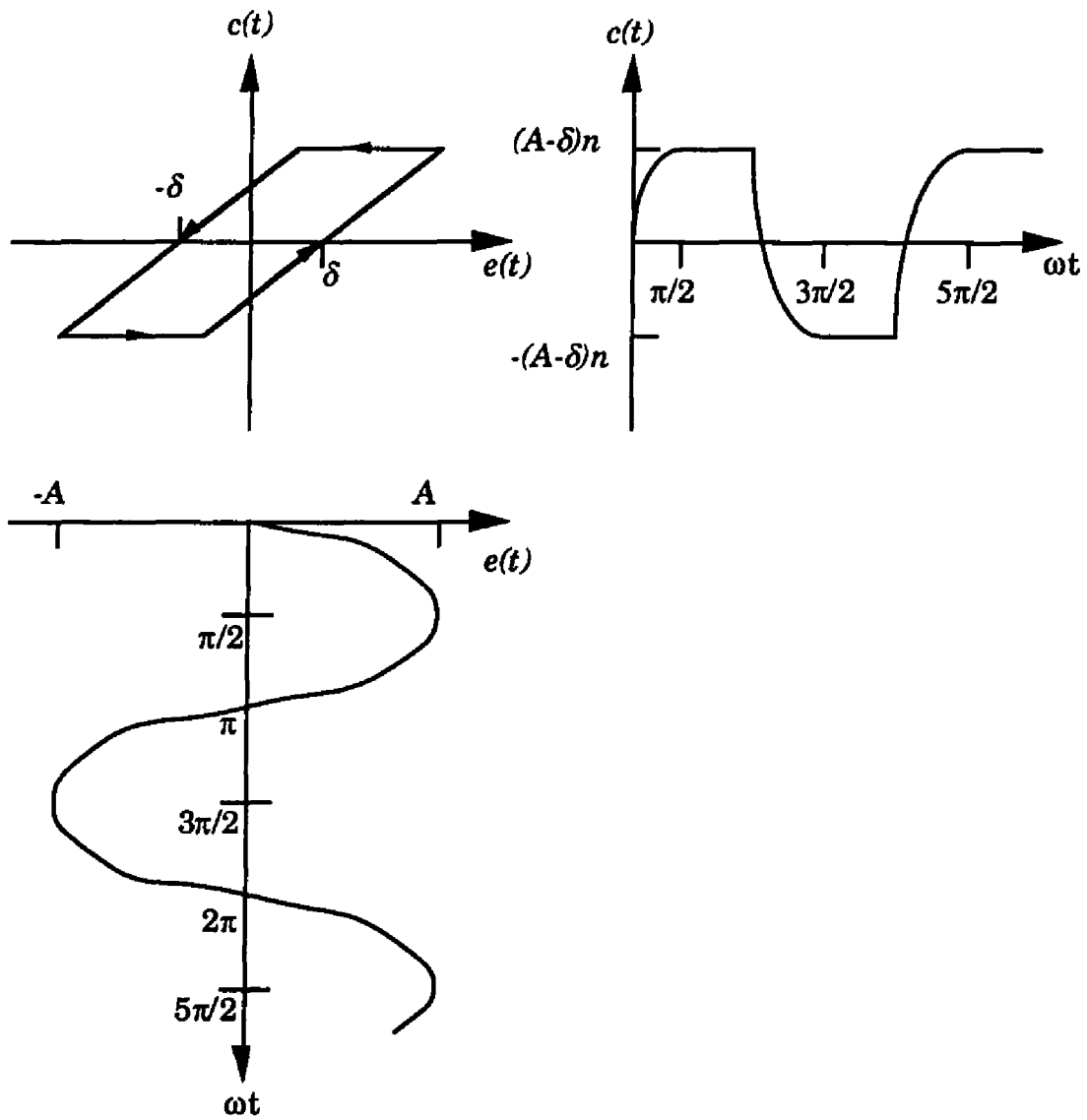
Figure 3.11 Input Output Relationship of Backlash Nonlinearity.

3.3 Detection of Nonlinearities Using the Hilbert Transform

For a system that has a FRF $G(\omega)$, its HT $H(\omega)$ can be calculated following the procedure discussed in Section 3.1.4. Examining the HT in convolution form as shown in Equation 3.24, the HT can be considered as the evaluation of an analytic function $G(\omega)$ inside and on a closed contour on the complex plane [59]. It was shown that for the FRF of a linear system, the complex integral of the HT will hold [59]. On the contrary, for the FRF of a system with common nonlinearities, the complex integral will not hold. Therefore, these relations can be used to detect nonlinearity in the measured FRF following the criterion [6]

$$
\begin{aligned}
H(\omega) &= G(\omega) \Rightarrow \text{System is linear} \\
H(\omega) &\neq G(\omega) \Rightarrow \text{System is nonlinear}
\end{aligned}
\tag{3.71}
$$

HT works more effectively if the FRF data have been measured using a sinusoidal excitation. Care must be taken to avoid misleading results which may arise if the data relate only to a limited frequency range, thereby restricting the effectiveness of the HT [63]. Accurate results can be expected for HT when the original FRF $G(\omega)$ vanishes near the end of the frequency range. However, it is not important which type of measurement, Inertance, Mobility or Receptance, is made because it can be converted to an appropriate type that $G(\omega)$ vanishes near the end of the frequency range.

The HT is applied to the FRF of the system with Coulomb friction nonlinearity discussed in Section 3.2. The analysis frequency range is 25.6 Hz with a frequency resolution of 0.05 Hz. In order to have better HT accuracy, the frequency range of 38.4 Hz is used to perform the FRF simulation and the HT. The transformed FRF was then truncated to the original frequency range of 25.6 Hz for analysis. Several measuring indices were applied to examine the level of nonlinearity in the FRF and HTF.

3.3.1 Test on a System with Coulomb Friction

In the friction test, HT was first applied to the FRF of the system with the same friction level and different forcing levels as shown in Figure 3.12. In the HTF, when the forcing level decreases, the resonant peak also decreases and the mismatch at the folding frequency increases. Then HT was applied to the FRF of the system with different friction level and the same forcing level as shown in Figure 3.13. Similar results were observed in the HTF. When the friction level increases, the resonant peak decreases and the mismatch

(a) R = 1, F = 3.

(b) R = 1, F = 5.

(c) R = 1, F = 10.

(d) R = 1, F = 20.

R: Friction Level

F: Forcing Level

Solid Line: FRF

Dash-Dot Line: HTF

Figure 3.12 Hilbert Transform of the Frequency Response Function of a System with Coulomb Friction (constant friction, various forcing levels).
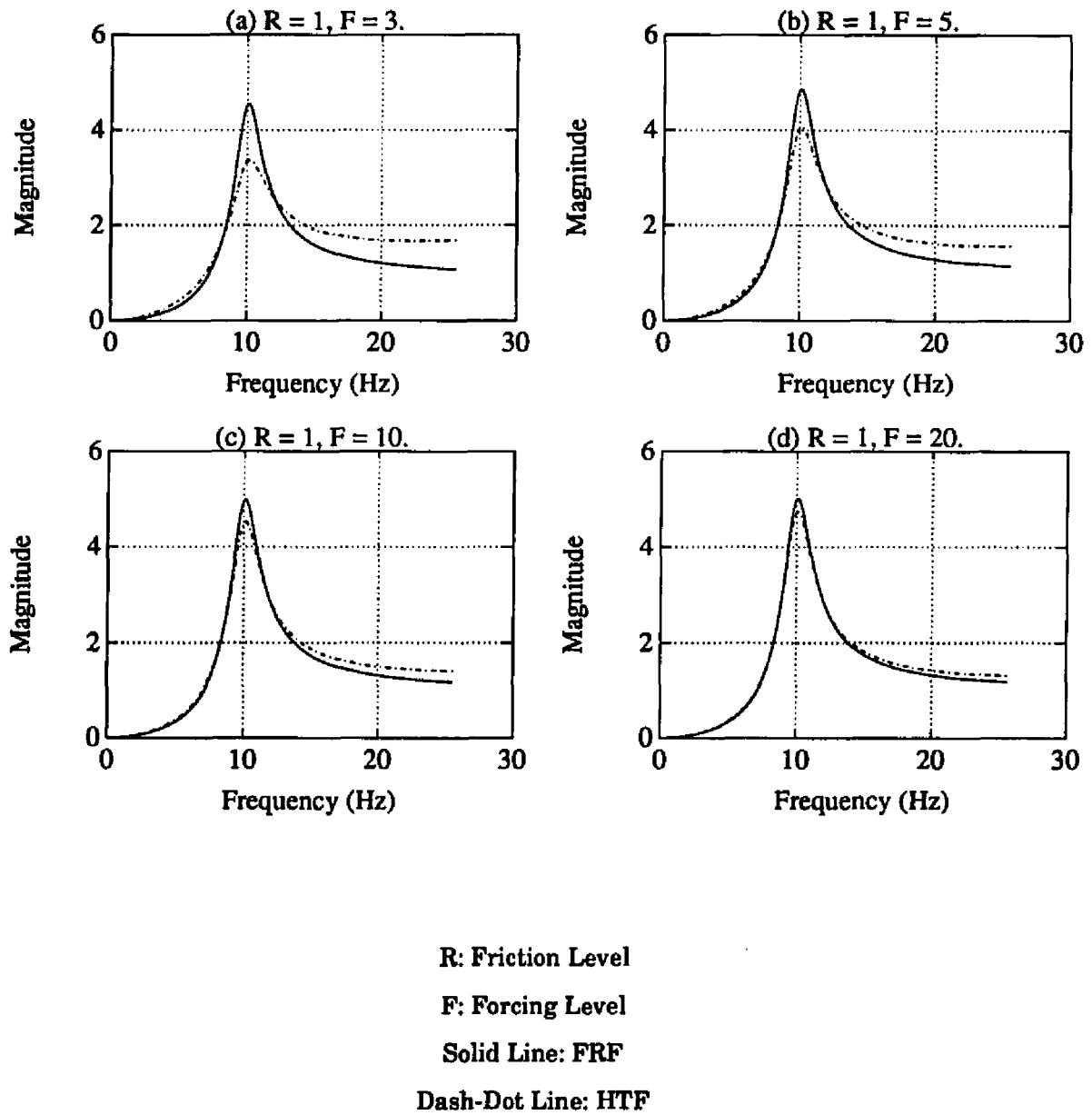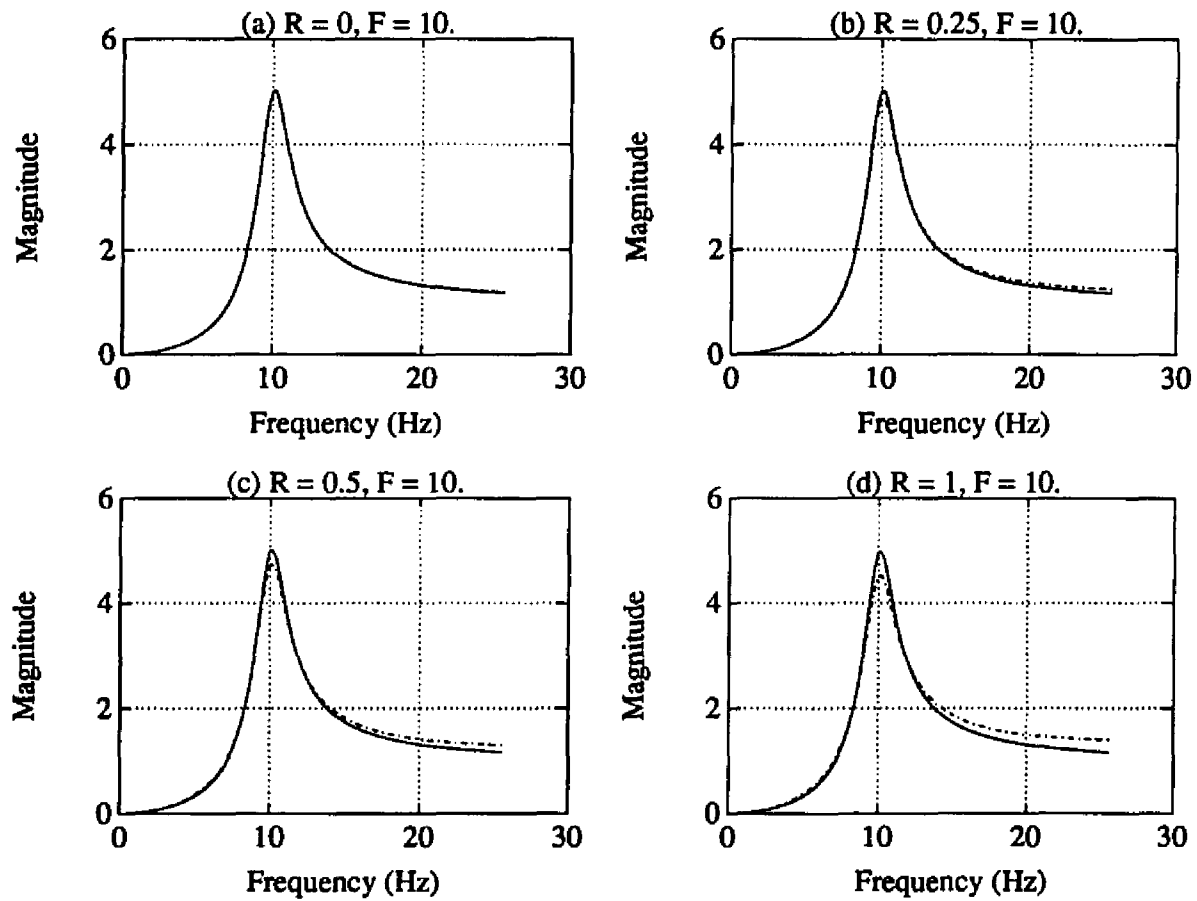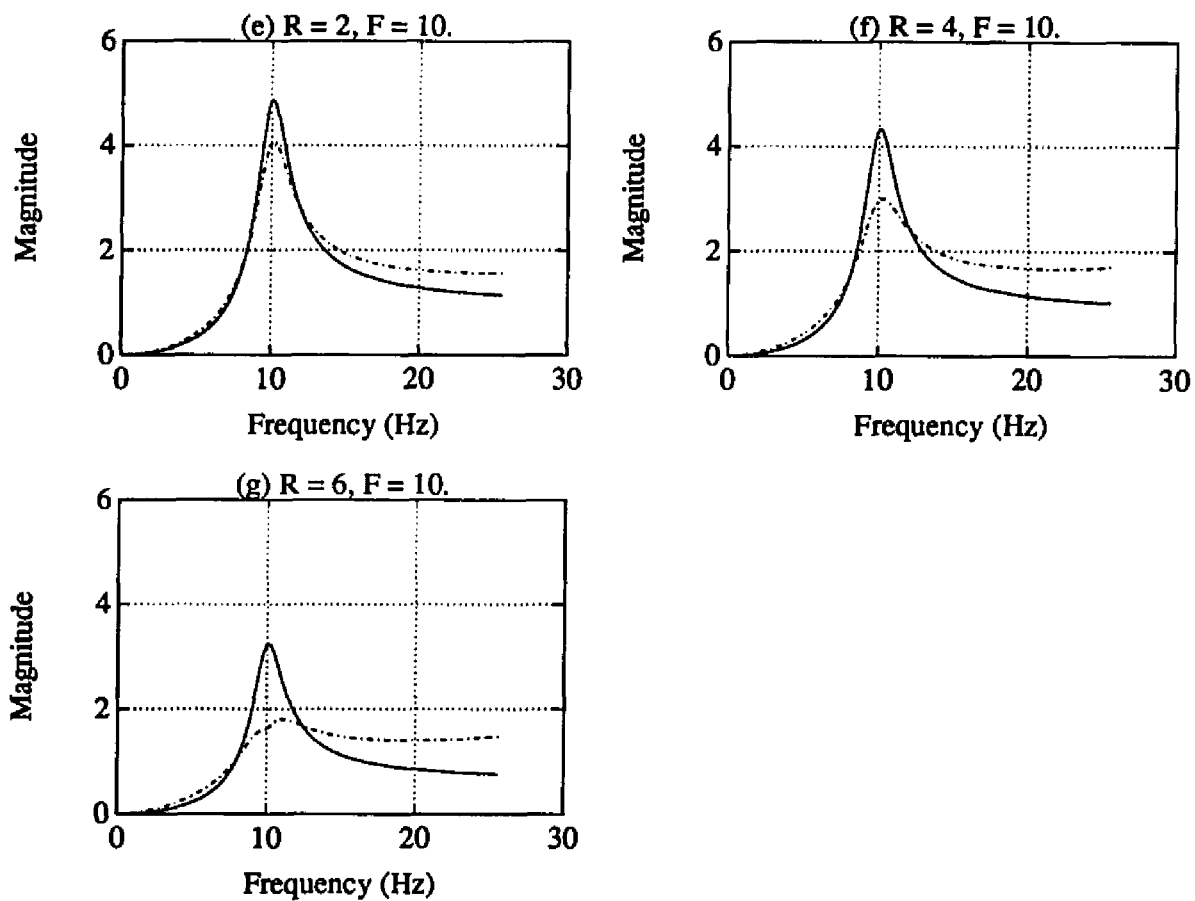
Figure 3.13 Hilbert Transform of the Frequency Response Function of a System with Coulomb Friction (constant forcing, various friction levels).

(e) R = 2, F = 10.

(f) R = 4, F = 10.

(g) R = 6, F = 10.

R : Friction Level

F : Forcing Level

Solid Line: FRF

Dash-Dot Line: HTF

Figure 3.13 (continue)

at the folding frequency increases. This result shows that HT can detect the significance of friction nonlinearity by examine the magnitude of the resonant peak in the HTF. The magnitude of the resonant peak decreases when relative friction level increases with respect to the forcing level.

### 3.3.2 Indices to Examine Significance of Nonlinearity

From the friction test, the HT is effective in identifying a system with Coulomb friction nonlinearity. The next task is how to determine the significance of nonlinearity. Several indices based on examining the resonant mode in the HTF or the ratio between the HTF and the FRF were utilized. These indices measures either the specific properties (peak amplitude ratio and geometric shape descriptors) of the resonant mode or the overall property (RMS ratio) of the HTF.

As discussed in the previous section, the magnitude of the transformed resonant peak decreases when nonlinearity increases. The relation of the amplitude between the transformed resonant peak and the original peak was analyzed. Ratio of peak amplitude was obtained when the transformed peak was divided by the original peak in the FRF. Figure 3.14 (a) shows the ratio of peak amplitude with respect to the relative level of friction. Figure 3.14 (b) shows an alternative view of the same data. The peak amplitude decreases proportionally to the increase in relative friction level. Note that in this simulation, the highest possible relative friction level is about 0.785 according to Equation 3.53. At this critical ratio, the input forcing will not surpass the friction so that there is no motion in the system.

Another way to assess the level of nonlinearity is to follow the half-power point method from conventional modal analysis. Instead of checking the resonant frequency and damping ratio of the resonant mode, the shape of the resonant peak may be examined. If the actual meaning of variables of both axes is disregarded, the FRF is generally similar to a Gaussian distribution function in the vicinity of resonance. Therefore, it is possible to examine the geometric shape properties of the resonant mode using descriptors borrowed from statistics.

For a distribution function $f(x)$, the mean is defined as

$$\overline{X} = \sum f(x)x \qquad (3.72)$$
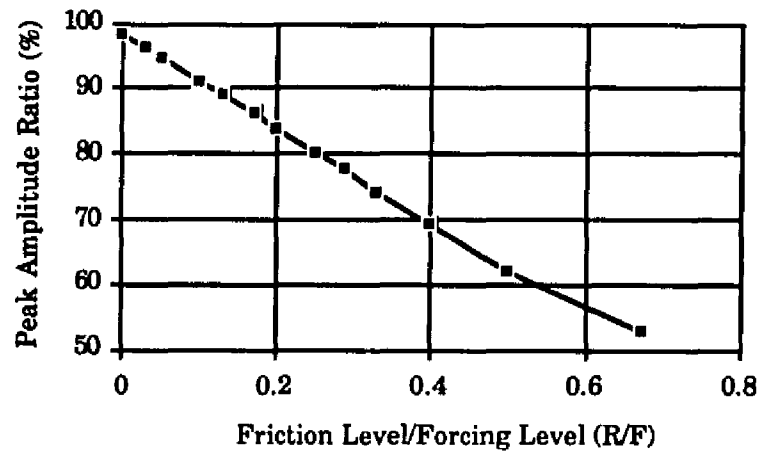
(a) Friction with Respect to Forcing.
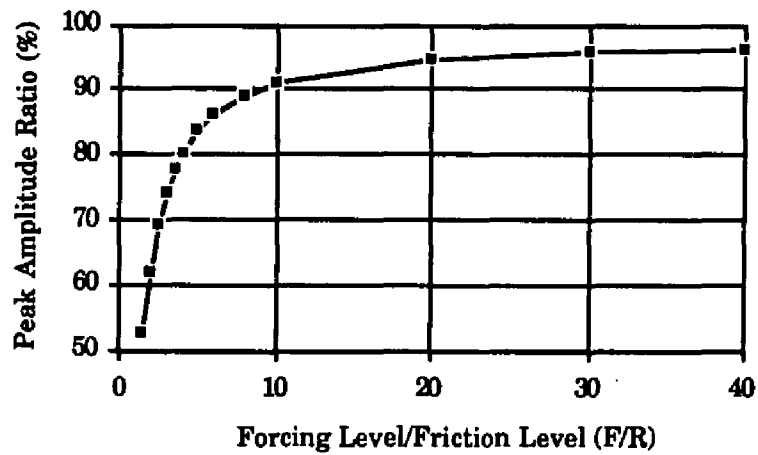
(b Forcing with Respect to Friction.

Figure 3.14 Peak Amplitude Ratio to Relative Friction Level.

Note that the area under the curve is equal to one, i.e., $\sum f(x) = 1$. The variance is defined as

$$VAR = \sigma^2 = \sum f(x)(x - \overline{X})^2 \qquad (3.73)$$

The standard deviation is defined as

$$\sigma = \sqrt{\sum f(x)(x - \overline{X})^2} \qquad (3.74)$$

The kth central moment about the mean is defined as

$$\mu_k = \sum f(x)(x - \overline{X})^k \qquad (3.75)$$

Note that $\mu_0 = 1, \mu_1 = 0, \mu_2 = \sigma^2$. The coefficient of skewness is defined as

$$\alpha_3 = \frac{\mu_3}{\sigma^3} \qquad (3.76)$$

The coefficient of kurtosis is defined as

$$\alpha_4 = \frac{\mu_4}{\sigma^4} \qquad (3.77)$$

Both skewness and kurtosis are dimensionless quantities. Skewness measures the degree of asymmetry of the distribution curve about the mean value. Skewness will be positive or negative according as the curve is skewed to the right or left respectively. For a symmetric curve, skewness is equal to zero. Kurtosis measures the degree of peakedness of the curve at the mean value. A large peak at the mean will have a larger kurtosis value than a flat peak. Note that kurtosis is equal to three for a normal distribution.

Using these measures applied to FRF peaks, some changes are necessary to interpret the meaning of variables. The frequency variable ω in the FRF is represented by x. The amplitude of the frequency response is viewed as the distribution function $f(x)$. Skewness and kurtosis were then used as geometric shape descriptors to analyze the shape of resonant peak between the half-power points (3 dB points), the frequencies where amplitude is equal to 0.707 times the peak amplitude. For proper comparison, the amplitude of the frequency response was normalized so that the area above the 3dB point is equal to one.

The geometric shape descriptors were utilized to examine the resonant peak in the FRF of the one degree of freedom system described in Section 3.1.4. The skewness and kurtosis were calculated with respect to different damping ratios in the system. Both the skewness and kurtosis increase as the damping ratio increases as shown in Figure 3.15. These trends show that the shape of the resonant peak above the 3dB point is more asymmetric and less peaked when damping in the system is increased.
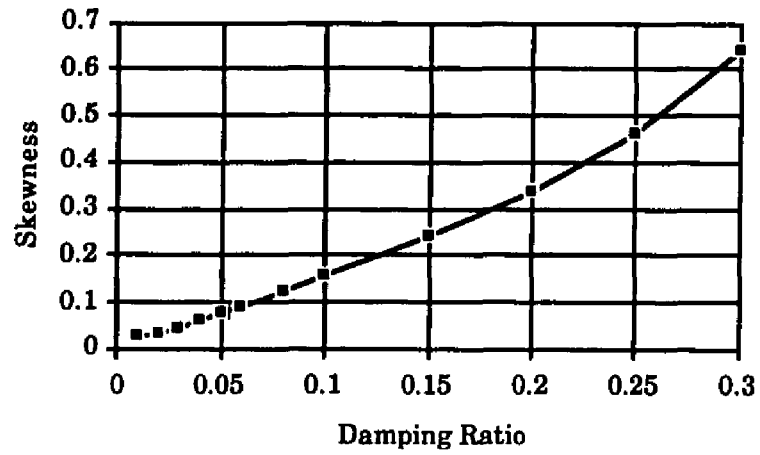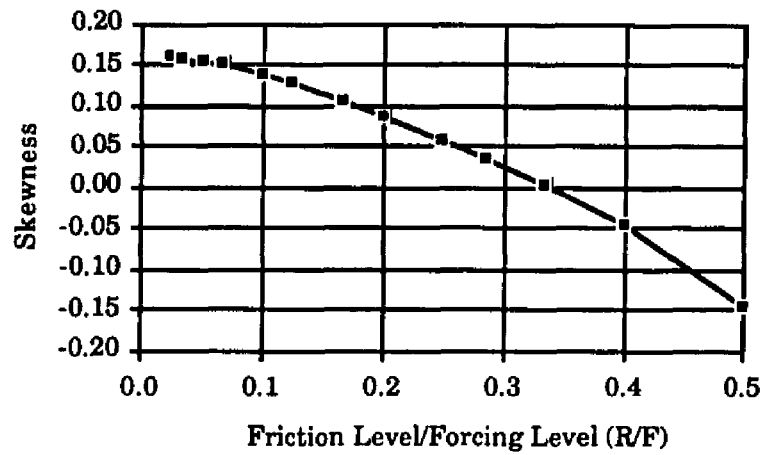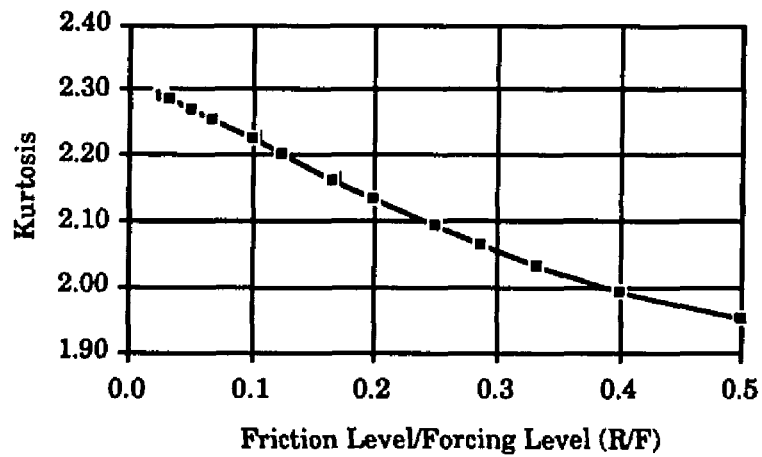
(a) Skewness.
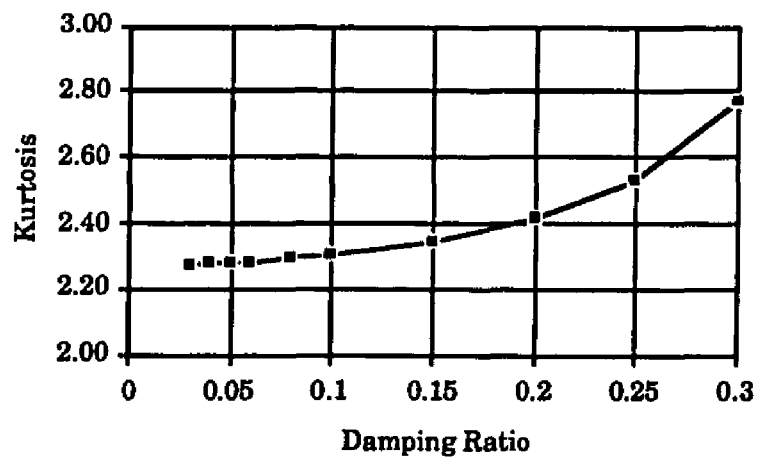
(b) Kurtosis.

Figure 3.15 Skewness and Kurtosis of the Resonant Peak of a Linear System.

The FRF and the HTF of the one degree of freedom system with Coulomb friction as discussed in Section 3.3.1 was examined with the shape descriptors. Using the 3dB point limit, both skewness and kurtosis values did not change for the resonant peak in the FRF with different friction/forcing ratio. When they were applied to the HTF, both skewness and kurtosis values decrease sharply when the friction/forcing ratio increases as shown in Figure 3.16. It means that as the relative friction nonlinearity increased, the shape of the resonant peak above 3dB point in the HT changed from skewed to the right to skewed to the left. Meanwhile, the shape of the resonance in the HTF is less peaked as the nonlinearity increased. Note that, with the same damping ratio, the skewness and kurtosis values of the linear system are upper bound limits to the system with Coulomb friction.

Instead of examining the resonant mode in the FRF and HTF, an alternative approach is to check the RMS ratio between the HTF and the FRF. This RMS index will show the overall level of mismatch across certain frequency range. This method was applied to the system with Coulomb friction. The ratio between the HTF and the FRF was calculated by first dividing the HTF by the FRF in complex domain, and taking the absolute magnitude. After subtracting unity from the magnitude, the RMS value was calculated. Either across a broad band or a narrow band frequency range, the RMS ratio between the HTF and the FRF increases sharply when the friction/forcing ratio increases as shown in Figure 3.17. In other words, mismatch between the HTF and the FRF increases as nonlinearity increases.
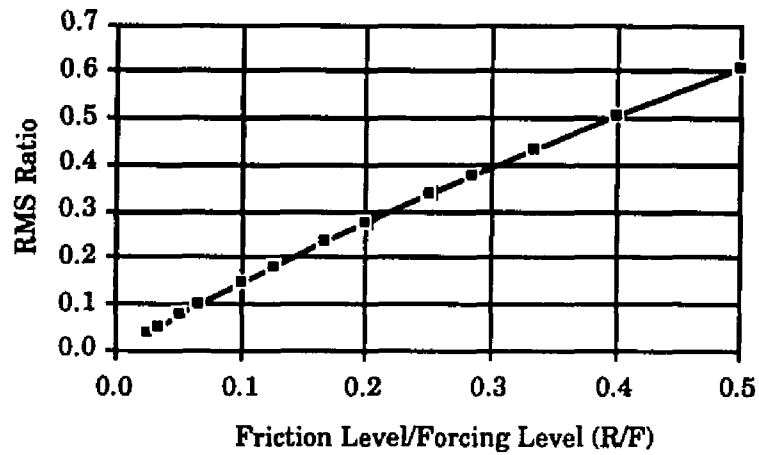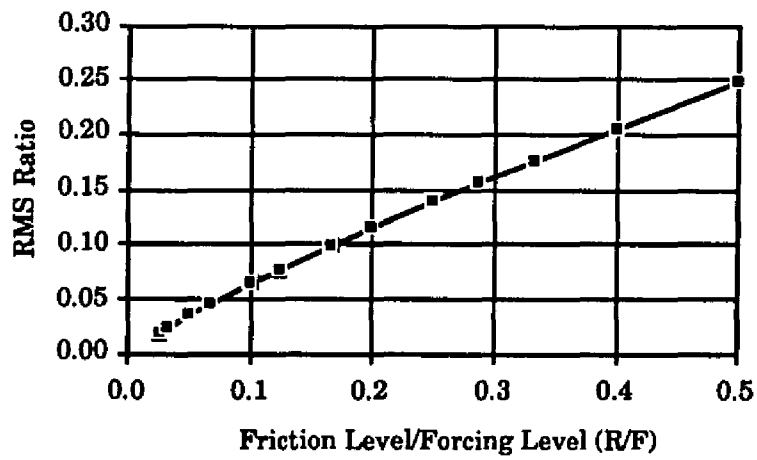
(a) Skewness.

(b) Kurtosis.

Figure 3.16 Skewness and Kurtosis of the Resonant Peak in the HTF of a System with Coulomb Friction.

(a) Broad Band (0 to 25 Hz).

(b) Narrow Band (5 to 15 Hz).

Figure 3.17 RMS Ratio between the HTF and the FRF of a System with Coulomb Friction.

Chapter 4
EVALUATION ON PROTOTYPE MANIPULATOR

For conventional robots, the dominant source of end effector positioning error comes from gear backlash and Coulomb friction in their joint drive mechanisms. After a long period of operation, these nonlinear effects will increase dramatically due to wear. In typical velocity control used by robots, joint PID controllers can not effectively compensate this nonlinear degradation. Therefore, it is necessary to identify these nonlinearities so that proper maintenance can be arranged to reduce or remove them.

For mechanical equipment such as robots, vibration analysis is the preferred tool to examine their operating condition because the vibration characteristics of a machine will change as machine condition changes. When the Hilbert transform is applied to analyze the FRF of robot joint drive systems, nonlinearity due to increasing gear backlash and Coulomb friction can be detected. Instead of using external excitation sources, robot drive actuators can be used as an internal excitation source to generate desired excitation signal for FRF measurements.

Section 4.1 describes FRF measurement using robot internal excitations. Section 4.2 discusses excitation tests on a two-link prototype manipulator using externally generated excitation signal. A programmed excitation test and data acquisition procedure are described in Sections 4.3 and 4.4. Section 4.5 shows the baseline FRF of the manipulator using shoulder and elbow excitations. Friction tests on the shoulder joint and elbow joint are then discussed in Sections 4.6 and 4.7.

4.1 Frequency Response Function Analysis of a Robot Using Internal Excitation

The characteristic of a structure can be described by its FRF, which acts like the signature of the structure. In conventional modal analysis routines, system properties such as natural frequencies and damping ratios are then extracted by analyzing the FRF or the relation between the input excitation and the output response. Single frequency sinusoidal or broadband signals, such as white noise and transient impact, are commonly used as excitation. Note that external excitation sources such as shakers and impact hammers are

used to generate input excitations since internal actuators are not available for most structures. The properties of the excitation signals can then be controlled.

On the contrary, without the knowledge of the excitation signals, simple vibration analysis may be preferred on a machine under actual operating conditions. It is assumed that the vibration characteristic of the machine will change as machine condition changes. Broadband, narrowband and signature analyses are used to analyze the vibration spectrum. Changes of intensity at specific vibration frequency due to damage developing in the machine can then be identified. In essence, vibration analysis acts like an open loop FRF examination while modal analysis acts like a closed loop FRF examination of the machine.

When defect identification is considered, shortcomings exist in both techniques. When actual machine operating conditions are examined by vibration analysis, there is no control on the excitation sources. When external excitation is used in modal analysis, real operating condition of the machine cannot be exactly simulated. Furthermore, inconvenience to install the external shakers practically eliminate the chance to monitor the machine on a regular basis. Consequently, FRF analysis of a machine using its own internal excitation sources is an attractive approach to solve the problem. FRF can be measured using controlled internal excitation.

For robots, the joint actuators can be used to generate internal excitation signals for FRF analysis. FRF can be measured for excitation generated by different joint actuators along the kinematic chain. Different input excitation signals such as noise, sine sweep, or single frequency sinusoid can be injected into the joint drive amplifier. Note that conventional robots are operated in velocity control mode and there are typically no torque sensors in robot joints. Consequently, the excitation signal is actually the velocity command to the joint actuators. Although motor current may be measured and related to motor torque, this is extremely difficult for motors using pulse width modulated control. Therefore, the FRF input excitation is in the form of an excitation velocity command instead of an actual input force or torque.

Readings from build-in joint encoders, velocity tachometers, or accelerometers attached to the arm linkage can be used as the output responses of the robot. Note that these readings are also valuable feedback information to the robot control system. Various FRFs can be generated by using different combinations of input excitation and output

response readings. The FRF measurements can be performed using commercial frequency analyzer or implemented within the robot's controller. These programs can be written to perform FRF analysis on a regular basis, as an integral part in the predictive maintenance management program. Following this approach of using robot joint excitation and accelerometer readings, Dagalakis and Myers successfully applied the coherence function of the measured FRF to adjust robot gear backlash [12].

The purpose of the FRF analyses is to form a time history database on the characteristic of the robot. Routinely measured FRFs are then compared to the baseline FRF. Any change in the new FRF indicates change in the characteristic of the robot. The first goal of the analysis is to identify the presence of significant nonlinearity, that may indicate developing component defects. Then, try to identify the magnitude and location of that nonlinearity to localize degraded components.

4.2 Excitation Tests on Prototype Manipulator

FRF analyses were implemented on a prototype two-link manipulator developed by Whalen [14]. The shoulder link (proximal link) is driven by a Megatorque motor which accepts velocity command control. The elbow link (distal link) is driven by a servo DC motor which accepts voltage velocity commands. Available sensor information includes shoulder joint encoder, elbow joint encoder, shoulder motor velocity tachometer, and accelerometers attached along robot links. Previous modal analyses of the robot using both external and internal excitations show that the resulting estimates of natural frequencies and mode shapes match each other very well [14]. Seven major resonant modes of the manipulator occur below 40 Hz at approximately 1.25, 2, 11.8, 18.7, 22.8 , 25.3, and 28.3 Hz.

Initial FRF measurements of the robot were performed using shoulder motor excitation. Excitation signals were generated by signal generators and injected into the motor through its velocity command input terminal. An accelerometer was attached on the proximal link near the elbow joint to perform output measurement. Input and output data were recorded using Wavepak signal analyzer. These preliminary measurements can be adopted as the initial baseline data for the manipulator.

Three types of excitation signal (white noise, sine sweep and stepped sine) were used. Due to their different characteristics, care must be taken when interpreting the resulting

FRFs. A schematic plot of the data analysis system is shown in Figure 4.1. The analysis frequency range was set from 0 to 40 Hz and the number of spectrum lines was set to 800 to achieve a frequency resolution of 0.05 Hz. Ensemble average of five frames was used for white noise and sine sweep excitations. With these specifications, the acquisition time for each frame is 20 seconds. Therefore, 100 seconds is needed to generate the FRF.

In the noise excitation test, band limited random noise was generated by a noise generator. The frequency range of the noise was restricted from 0 to 40 Hz. The FRF of the robot using noise excitation was shown in Figure 4.2. Note that the FRF has effectively been linearized by using random excitation. Consequently, some nonlinear phenomena in the system may not appear in the FRF. The quality of the measurement was examined by checking the coherence function. A low coherence usually indicates imperfect measurement but may also indicate existence of nonlinearity in the system. The quality of the FRF from noise excitation was poor.

In the sine sweep excitation test, a sweep generator was used to generate the sine sweep signal. The sweep range was set from 0 to 40 Hz with a sweep time of 5 seconds, i.e., the output frequency varies automatically from 0 to 40 Hz in 5 seconds. The sweep generator was set in continuous mode such that the end of a sweep cycle was connected to the start of a new sweep cycle. With an acquisition time of 20 seconds, at least three complete sweeps could be captured during each ensemble measurement. Figure 4.3 shows the FRF of the robot using sweep excitation. One significant flaw is that the FRF curve is very ragged due to the sweep time being too fast and the frequency resolution was too fine.

Tests using 400 spectrum lines were performed to verify the above question. The frequency resolution was then 0.1 Hz and the acquisition time was reduced to 10 seconds. Only one complete sweep cycle was captured in each ensemble measurement. Figure 4.4 shows the resulting FRF of the robot with the same input excitation level. Compared to Figure 4.3, it is obvious that the FRF curve is smoother with lower frequency resolution.

With the same frequency resolution, another set of tests were performed by adjusting the sweep frequency manually. The sweep frequency was adjusted to increase slowly from 0 to 40 Hz and then decrease slowly from 40 to 0 Hz. Up to 10 ensembles were used to measure the FRF of the robot. Figure 4.5 shows a typical FRF of manually adjusted sweep excitation with 400 spectrum lines and 0.1 Hz frequency resolution. Figure 4.6 shows
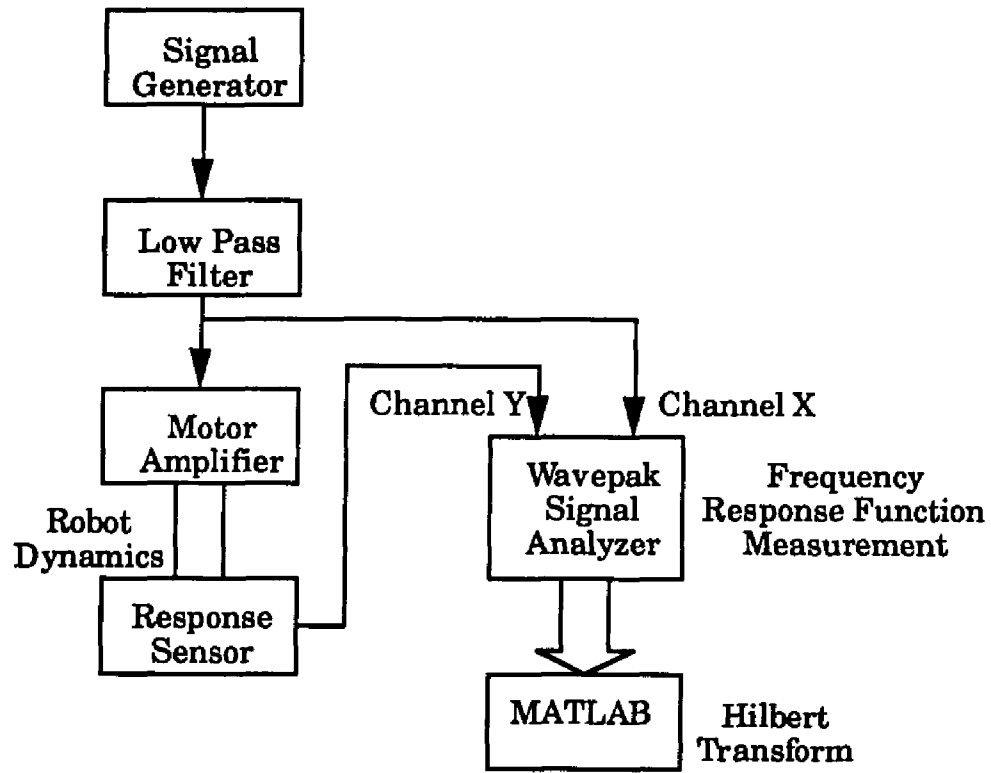
Figure 4.1 Schematic Plot of Data Analysis System.

Figure 4.2 Frequency Response Function from Noise Excitation.

(400 spectrum lines)

Figure 4.3 Frequency Response Function from Sweep Excitation.

(800 spectrum lines)
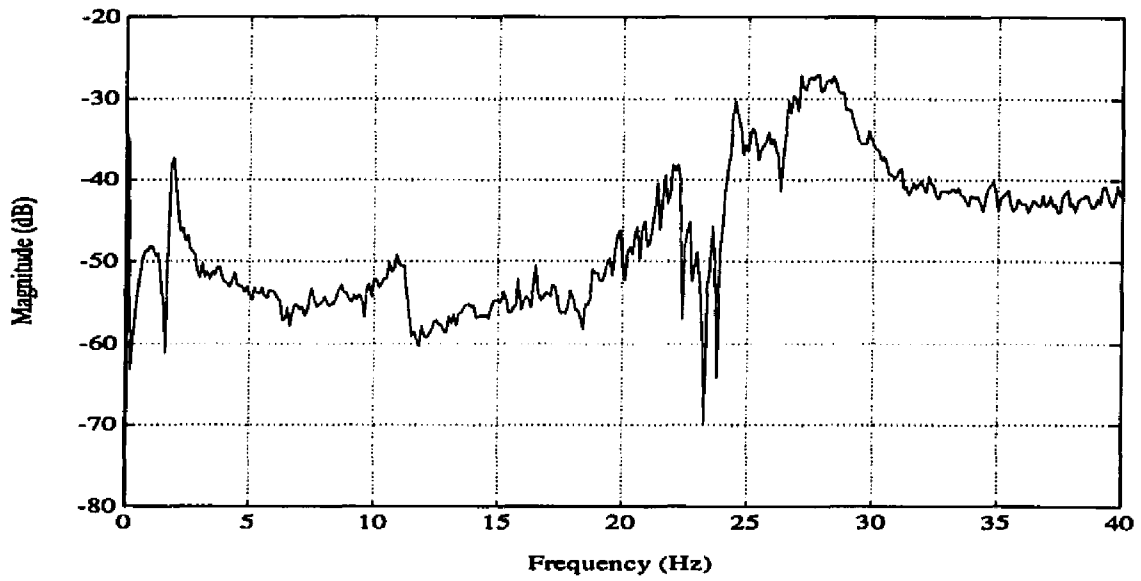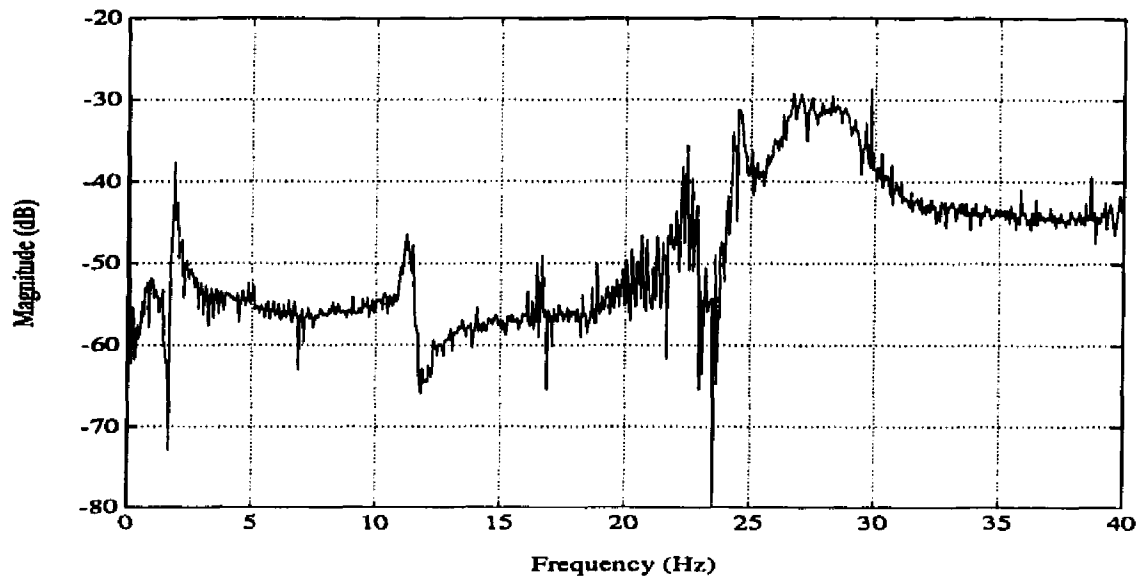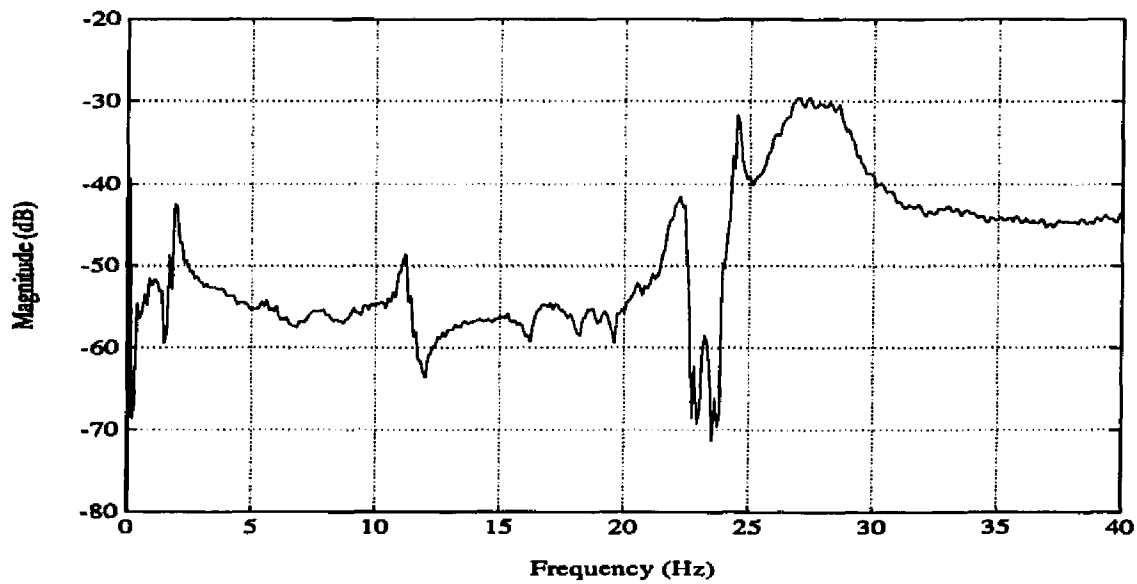

Figure 4.4 Frequency Response Function from Sweep Excitation.
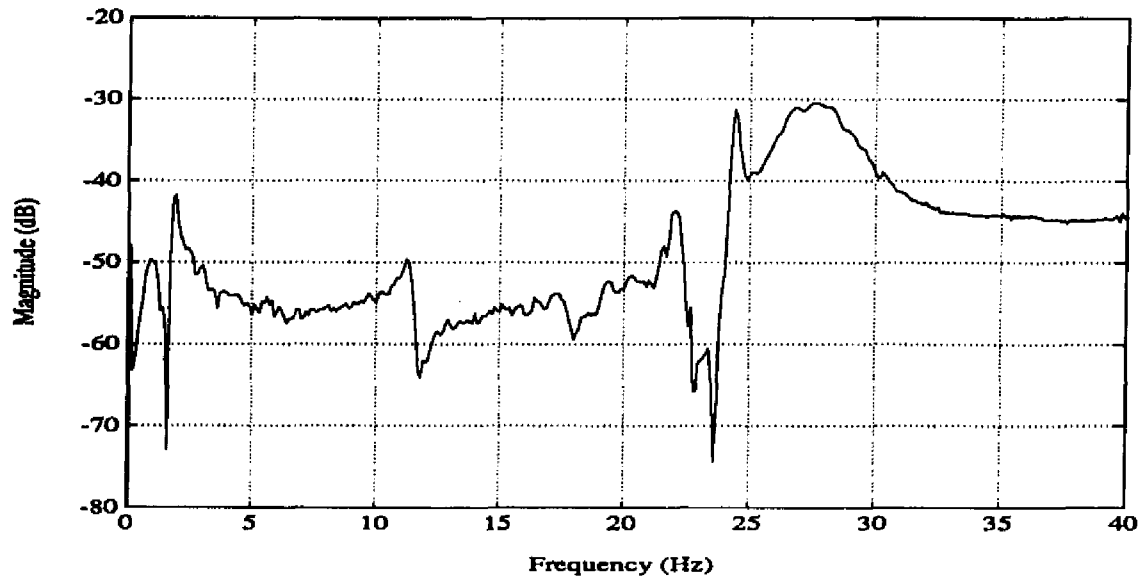
(400 spectrum lines)

Figure 4.5 Frequency Response Function from Manual Sweep Excitation.
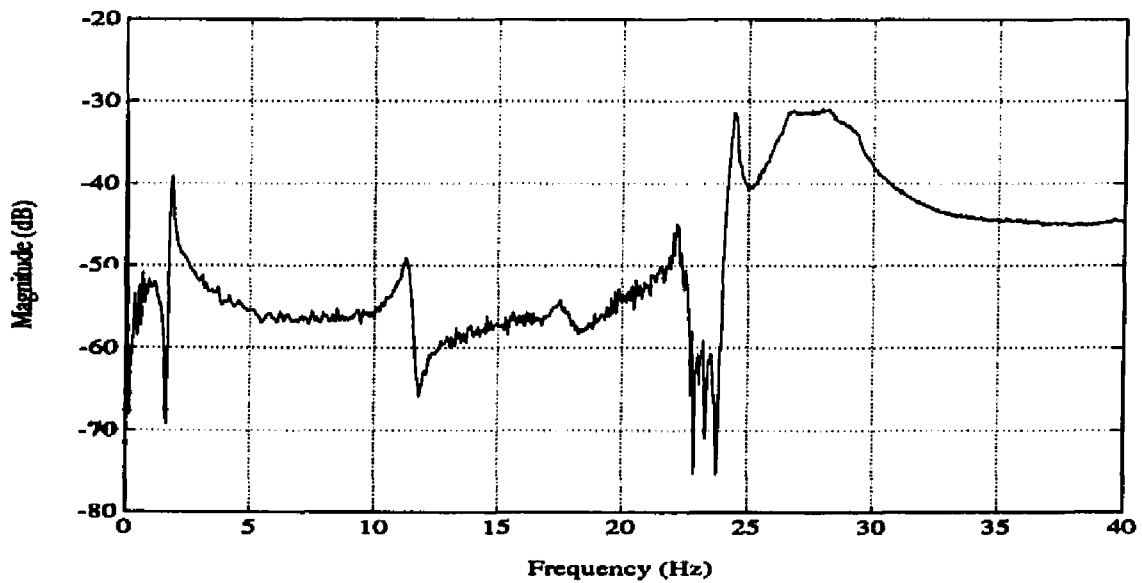
(400 spectrum lines)

Figure 4.6 Frequency Response Function from Manual Sweep Excitation.

(800 spectrum lines)

another FRF of manual sweep test with 800 spectrum lines and 0.05 Hz frequency resolution. This test was performed by sweeping the frequency range very slow so that steady state can be achieved. Generally, the quality of these two FRFs is better comparing to the FRFs shown in Figure 4.3 and 4.4. A few ragged sections in the frequency response curve were attributed to nonuniformly adjusted sweep frequency. Essentially, to have a successful sweep measurement, an appropriate sweep time had to be chosen. For a high resolution measurement, the sweep time must be sufficiently long to get a clean FRF.

The sine sweep excitation can be characterized as transient or steady state methods depending on the choice of sweep time. For a short sweep time, fast sine sweeps or chirp behave like transient excitation. When longer sweep time is used, the resulting slow sine sweep behaves like sinusoidal excitation. One way to explore the proper sweep time is to measure the FRF twice, the first time sweeping up and the second time sweeping down through the frequency range. However, because the sweep generator does not have sweep down capability, this test could only be performed manually.

Stepped sine testing is the traditional method to do FRF measurement. The signal generator was used to generate a sinusoid wave with a fixed amplitude and frequency. This sinusoidal excitation signal was then injected to the motor control through its velocity command input. To cover the entire frequency range, the excitation signal was stepped from one discrete value to another to provide enough spectral line density in the FRF. That means finer frequency steps were used near the resonance. It was necessary to ensure that steady state condition was achieved before the measurement was made. Due to this constraint, the stepped sine testing is time consuming and therefore is not used in routine FRF measurement on robot. However, since sinusoidal excitation was used, the measured FRF can be related to the theoretical frequency response analysis of the system [63]. If significant nonlinearities were detected in the FRF of robot using other excitation methods, a stepped sine test may be useful to provide more detailed information on the characteristics of the robot. A typical FRF of the robot using stepped sine test was shown in Figure 4.7.

Besides the disadvantage of being time consuming, another problem occurs during the measurement process. Because finer frequency resolution is required to characterize the response near the resonance, the robot was effectively excited near its natural frequencies for a long time. This will potentially increase the chance to damage the structure and the joint
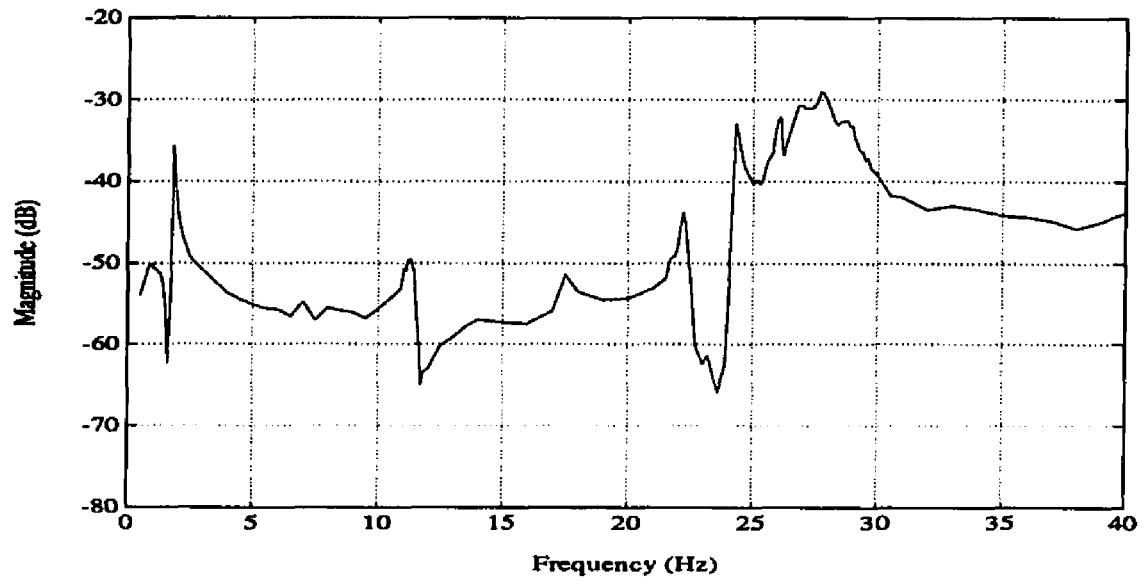
Figure 4.7 Frequency Response Function from Stepped Sine Excitation.

drive system of the robot. For this reason, stepped sine testing is not recommended for routine FRF measurement of robot.

Since all the input excitation amplitudes for different tests were close to 0.25 Volt RMS, the FRFs were practically measured in similar conditions. Comparing the obtained FRFs, sweep sine excitation with appropriate sweep time provides both good measurement quality and time saving advantages over other excitation methods. Basically, these FRFs are very similar. Seven major resonant mode were observed at approximately 1, 1.8, 11.3, 17.5, 22.2, 24.3 and 27.7 Hz. One minor resonant mode was also observed at 23.2 Hz between two major resonant modes.

Using sine sweep excitation, FRFs were measured with different excitation levels. The RMS amplitude of the excitation signal was varied from 0.25 Volt to 0.5 Volt. The resulting FRFs were shown in Figure 4.8. There is no significant change at the first four modes. The fifth mode at 22.2 Hz shifts to lower frequency. The sixth mode at 24.3 Hz becomes wider. The seventh mode at 27.7 Hz becomes thinner. One significant change occurs at the minor mode between the fifth and the sixth modes. As excitation amplitude increases, the response at this mode increases. The corresponding coherence functions also show that nonlinear effects increase in this section.

The HT of these FRFs were shown in Figure 4.9. Mismatch at peaks increases when excitation level increases as predicted by simulation. When applying the HT, an extended frequency range of 40 Hz was used and a null buffer was attached to the FRF from 40 to 51.2 Hz because the Wavepak analyzer truncates 20 percent at half the sampling frequency to reduce the transient effect of built-in low pass filters.
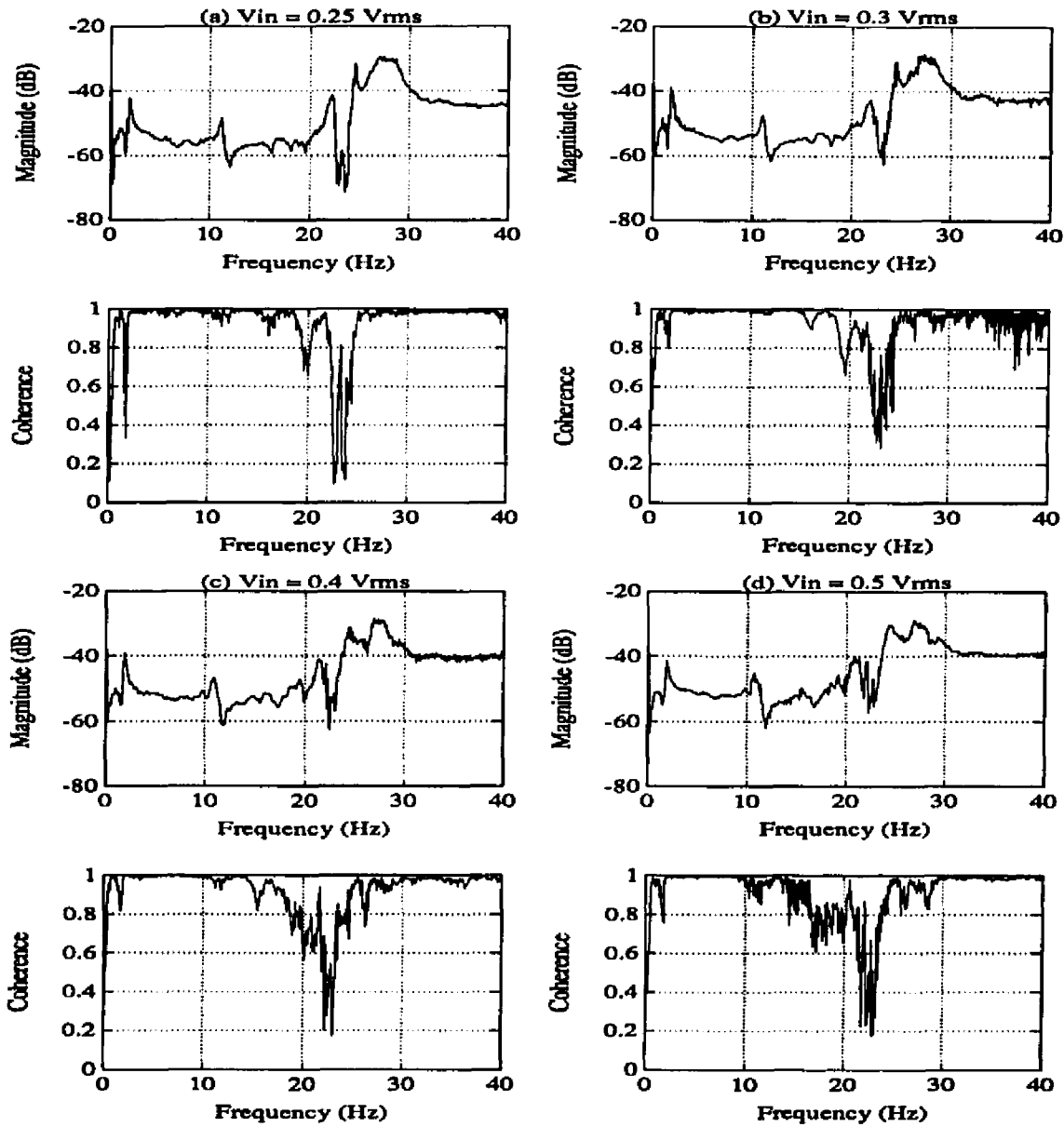
Figure 4.8 Frequency Response Functions from Sweep Excitation with Different Input Excitation Level.
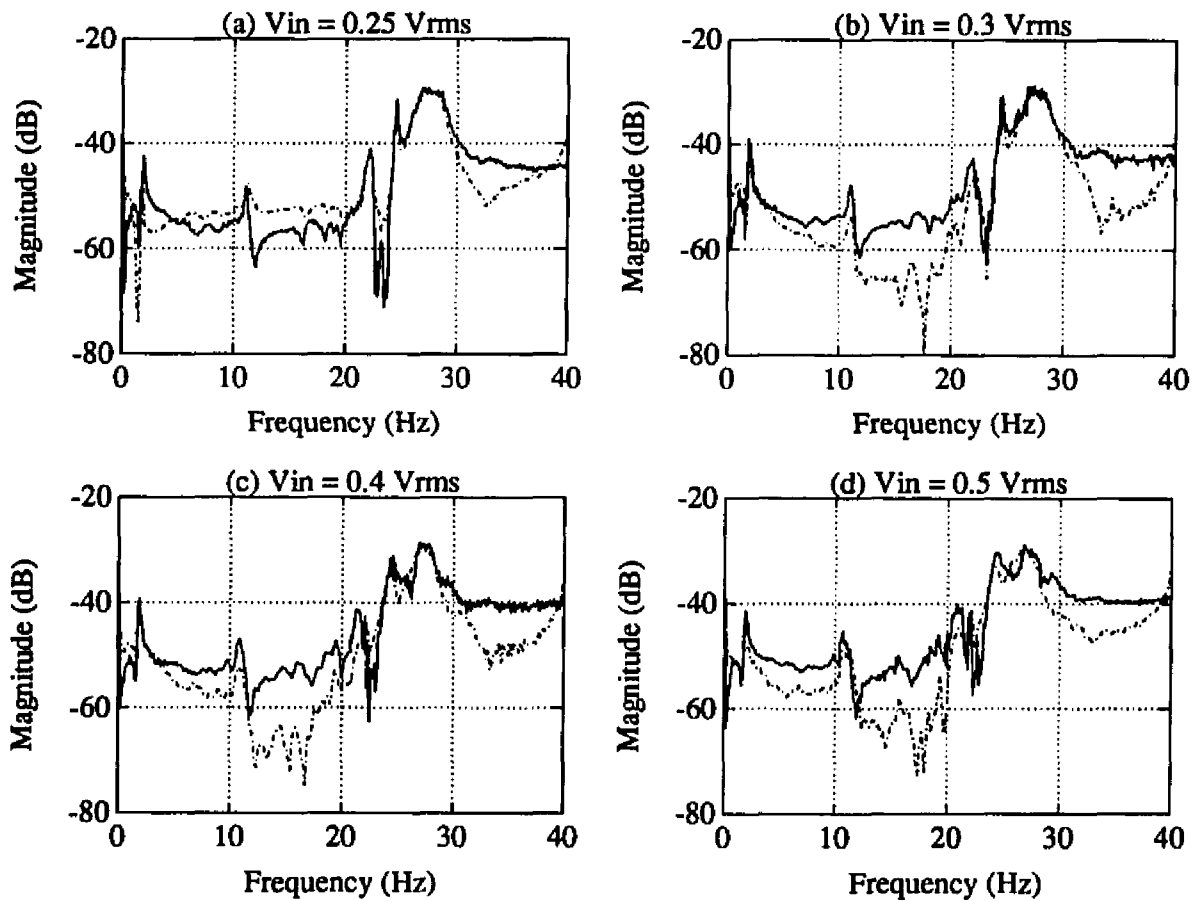
Solid Line: Measured FRF

Dash-Dot line: HT of the FRF

Figure 4.9 Hilbert Transform of the Frequency Response Function from Sweep Excitation.

4.3 Programmed Excitation Test on Manipulator

The control program of the manipulator was modified to perform internal both shoulder and elbow excitation by programmed sine sweep. The recorded sensor data was then processed to generate the FRF for the HT analysis. Using this program, FRF measurement can be performed automatically by higher level robot control routines rather than injecting excitation signals into the motor amplifier. In a robot PMP, robots could be programmed to performed the FRF measurement routinely and the resultant data analyzed by the host computer and compared to the FRF database for possible changes in the characteristics of the robots.

A block diagram of the excitation/data acquisition system is shown in Figure 4.10. The internal excitation signal was generated by the control program and sent to the motor amplifier via the D/A output channel on the DAC02 board. Sensor signals were low-pass filtered to avoid aliasing and then sampled via the A/D input channels on the DAS16 board. The basic program architecture developed by Whalen was adopted [14]. Because the same computer was used to generate excitation and to collect response data, proper excitation/ response sequencing must be observed to get correct samples. Modifications to the timing method of the original program were made.

Figure 4.10 Block Diagram of Programmed Excitation and Data Acquisition System.

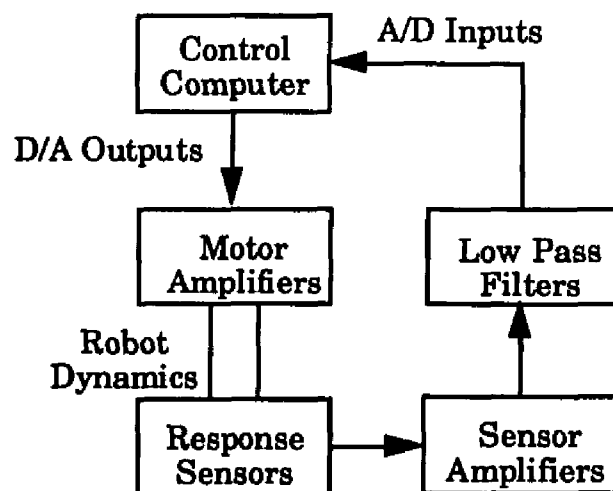### 4.3.1 Programmed Sine Sweep Signal

The sine sweep excitation signal was generated according to Equation 4.1 [64] where A defines the sweep amplitude, F_{low} and F_{high} define the lower and upper bound frequencies of sweep, T_s defines the duration for one sweep cycle. Note that this equation generates one sweep cycle. The next sweep cycle starts when t is reset to 0. One advantage on using programmed excitation is that signal parameters can be adjusted easily and accurately in the program for different test conditions. For instance, a specific sweep range in the neighborhood of a resonant peak can be examined by changing the lower and upper bound frequencies.

$$v = A \, sin\left(\pi[2F_{low}t + \frac{F_{high} - F_{low}}{T_s} t^2] \right) \tag{4.1}$$

Different excitation forcing levels can be obtained by adjusting the sweep amplitude A. To cover the frequency range for analysis, the lower bound frequency was set to 0 Hz and the upper bound frequency was set close to half of the sampling frequency. The frequency of the sweep signal was increased linearly from 0 to the upper bound frequency during one sweep cycle. For this up-only sweep, frequency goes back to 0 Hz to start the next sweep cycle. An alternative approach is to decrease the frequency from the upper bound linearly to 0 Hz in the second sweep cycle. This up-then-down sweep avoids sudden change in the control routine, and provides smoother transition between two sweep cycles. The duration of sweep was chosen so that near steady-state response could be measured by the sensors.

4.3.2 Excitation and Data Acquisition

The manipulator control program was modified to perform excitation and data acquisition on the same computer. Due to the limitation of the primary interface board used in the project, simultaneous excitation and data acquisition is not possible on the same board. Within each sampling period, the control loop must perform multi-channel data acquisition just after excitation takes place. A flow chart of the control loop was shown in Figure 4.11 to explain the timing method. Note that while one motor was used to generate internal excitation, the other motor is still regulated by PID control. Therefore, the control loop must incorporate both functions.
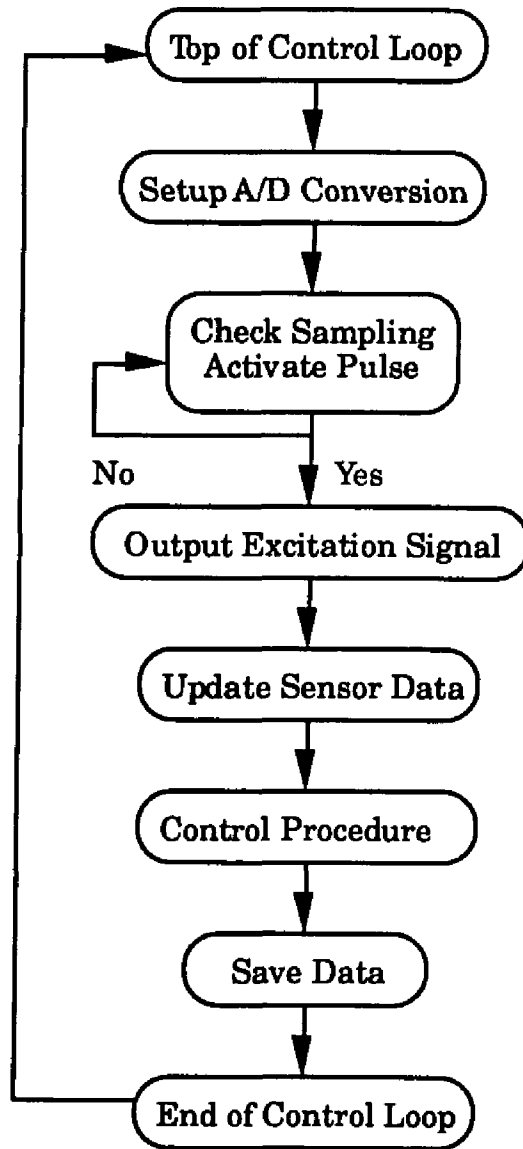
Figure 4.11 Flow Chart of the Excitation and Data Acquisition Control Loop.

In the system sampling routine, a sampling activate pulse was generated at required sampling frequency by counter zero on the CTM05 counter/timer board. Because the CTM05 board used a 1K Hz clock to generate the sampling activate pulse, sampling frequency is available on integer divisions of the clock frequency. A sampling frequency of 125 Hz was used, which provided a 62.5 Hz range for analysis.

The sampling activate pulse was fed to the gate input of counter one on the CTM05 board to generate the sampling trigger pulses at the maximum conversion frequency of A/D converter. These pulses were then fed into the trigger input on the DAS16 board to trigger successive A/D conversions. In addition, the activate pulse was connected to the digital input port on the CTM05 board to provide a monitoring channel for the control program. When the leading edge of the sampling activate pulse was detected by the control loop, the corresponding excitation signal was sent out via the D/A output on the DAC02 board.

The DAS16 board was configured to make A/D conversions on all eight channels using the DMA mode. However, because some overhead time is required for excitation output, the first five channels (channel zero to four) were not used to collect sensor data. Experiments showed that correct response samples could be obtained from channel five to channel seven. This method provided three-channel sampling with minimum time lag between data channels, and between excitation and response.

Unnecessary program steps in the DMA acquisition subroutine of the DAS16 board were modified to reduce overhead time. The excitation signal was calculated in advance and stored in a data array to reduce computation in the control loop. Note that in the excitation test while one motor generates excitation, the other motor is still under PID control. Due to this limitation on computer speed, higher sampling frequency was not used to ensure correct sampling in the test. The excitation signal was resampled at A/D channel seven. The velocity monitoring signal of the shoulder motor was sampled at A/D channel four.

4.3.3 Data Analysis

FRF and HT analyses were performed using routines written in MATLAB as shown in Appendix B. The recorded data from the excitation test was first separated into individual variables and converted to appropriate units. Sensor data from A/D conversion was also

checked for possible saturation. The MATLAB function *spectrum* was used to calculate the FRF using the overlapped averaging method. The HT was then performed on the FRF.

Using the averaging method, data was divided into blocks to calculate its frequency spectrum. The resulting frequency resolution in the spectrum was calculated according to Equation 4.2, where Δf is the frequency resolution, N is the number of data points per block, and T is the sampling period. To utilize a conventional FFT routine, the number of data points must be of two's power. For sufficient frequency resolution with sampling frequency of 125 Hz, possible choices are 512 points with 0.2 Hz resolution, 1024 points with 0.1 Hz resolution and 2048 points with 0.05 Hz resolution. The calculated spectrums for each data block were then averaged to obtain the final spectrum. To get a smoother spectrum, data blocks were overlapped with one another.

$$\Delta f = \frac{1}{NT} \tag{4.2}$$

The parameters for data averaging were decided according to the properties of the sine sweep signal. The sine sweep signal was set to sweep in the up-then-down option with sweep frequency range from 0 to 60 Hz and sweep duration of 30 seconds for each complete up-then-down cycle. Each data block contained 512 points. Ten data blocks were used for averaging with 20 percent data point overlap between blocks. A total of 4250 data points were taken from each excitation test. Note that the total data acquisition time was 34 seconds which covered at least one complete sweep cycle. The frequency spectrum of the sine sweep signal is shown in Figure 4.12. The magnitude of the sweep signal remains nominally at the same level below the cutoff frequency of 60 Hz. The 1024 points per block option was not used because there was significant drop-out in the frequency spectrum within the frequency range of interest.

FRF of the excitation test was obtained using accelerometer data as output response and the sine sweep signal as input excitation. The calculated sweep signal instead of the sampled sweep signal was used after no significant difference was found. The HT routine discussed in Section 3.1.4 was applied to the measured FRF. The ratio between these two spectrums was also checked.
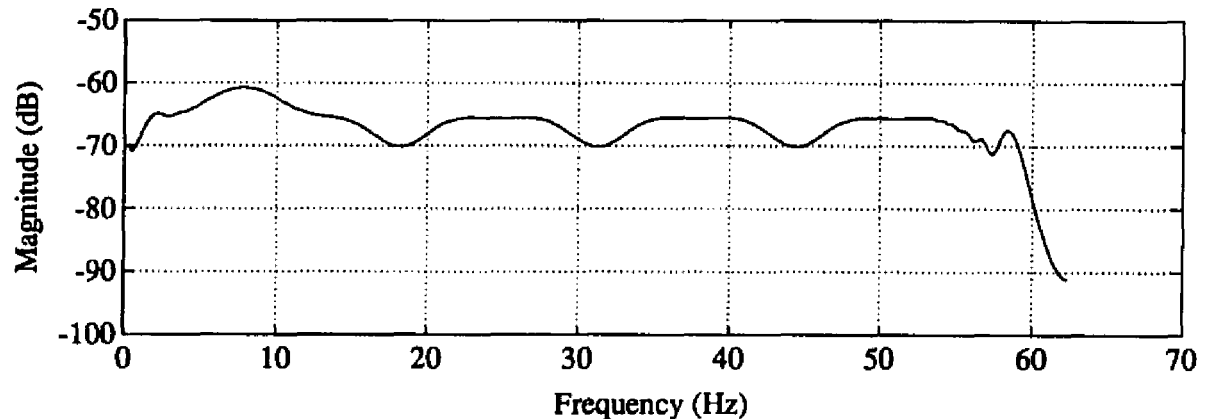
Figure 4.12 Power Spectrum of the Sine Sweep Signal.

4.4 Test Setup

Standard hardware configurations and software parameters were chosen to provide consistent test conditions for FRF measurements. The D/A output channel zero on the DAC02 board was used to control the shoulder motor, while channel one was used for the elbow motor. In shoulder excitation tests, the sine sweep signal was fed to the shoulder motor to generate internal excitation, and accelerometer data was sampled at the DAS16 A/D input channels. Meanwhile, the elbow motor was controlled to hold the elbow link at a constant joint angle. In elbow excitation tests, the sine sweep signal was fed to the elbow motor, and accelerometer data was sampled at the DAS16 A/D inputs. To increase link rigidity, the proximal link was cross-braced using steel wires while the distal link was reinforced by plastic plates.

Accelerometers (PCB series 302A) were mounted at four locations along the manipulator links as shown in Figure 4.13. Location A was at the proximal link plate near the shoulder joint at approximately 1/7 its length. Location B was at the shoulder end on the elbow joint. Note that accelerometers were mounted outside the proximal link plate to simulate mounting on real robots. Location C was at the elbow end on the distal link. Location B and C were at about the same distance from the elbow shaft. Location D was at the distal end of the distal link.

Elbow Joint   Shoulder Joint

D   C   B   A

Distal Link

Proximal Link

Elbow Motor   Shoulder Motor

Test Bench
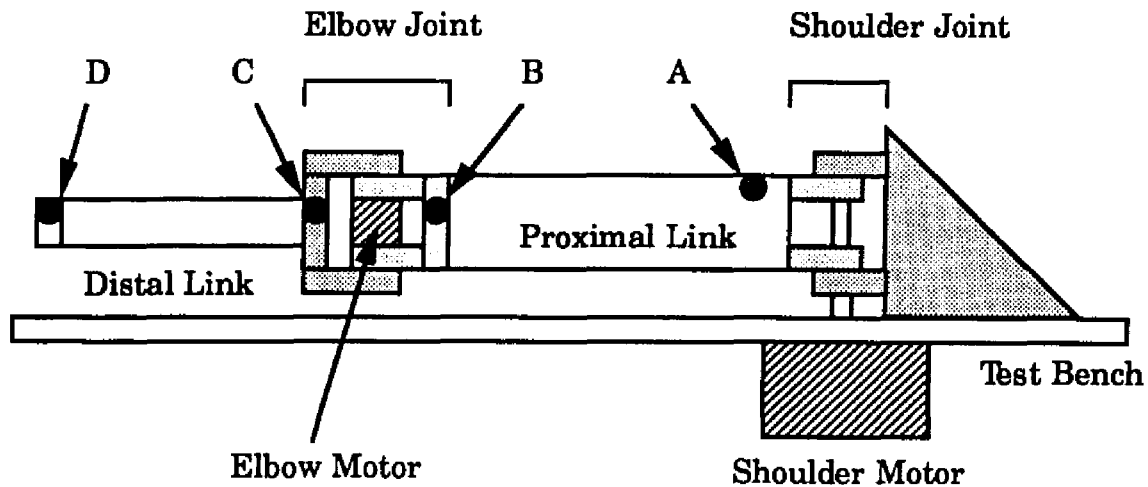
● Accelerometer Mounting Location

Figure 4.13 Schematic Plot of Accelerometer Mounting Locations.

Two accelerometers (Acc-One and Acc-Two) were used to take measurements from two locations for each test. Accelerometer signals were amplified with a gain of 10 by a signal amplifier (PCB power unit model 480D06). The amplified signal was then low-pass filtered with a cutoff frequency of 50 Hz and a gain of 10 by a filter (ITHACO model 4302) which has a 24 dB per octave rolloff rate. Note that with the moderate rolloff rate at the cutoff frequency, small amount of aliasing might affect measurement above 40 Hz. The signal from Acc-One was sampled on DAS16 A/D channel five while signal from Acc-Two was sampled on DAS16 A/D channel six. Note that the resulting AC-coupled signal has a gain of 100. To measure the signal at highest resolution without saturation, the DAS16 A/D input voltage range was configured to ± 2.5 volts.

An up-then-down sine sweep was used for testing. The signal swept from 0 to 60 Hz and then back to 0 Hz in 30 seconds with 4250 samples taken for each test. For the shoulder excitation test, the standard sweep amplitude was 0.5 volt. For the elbow excitation test, the standard sweep amplitude was 0.75 volt. The proportional, integral, and derivative gains to control the shoulder motor were 0.035, 0.0004 and 1.2 respectively [14]. The proportional, integral, and derivative gains to control the elbow motor are 0.4, 0.002 and 0.225 respectively [14].

The standard link posture for the excitation test held both links at zero degree joint angles. Two additional postures were used for the elbow excitation tests. In this additional posture, the elbow link was held at +30 and -30 degrees joint angle to prevent elbow motor drift.

4.5 Baseline Data

A set of baseline FRFs was taken as a standard for comparing shoulder and elbow excitation tests. These FRFs measured the dynamic characteristics of the manipulator links and joints at the four accelerometer locations without additional friction nonlinearities. When additional friction was introduced, variations in both the FRF and the HTF with respect to the baseline data were inspected to identify the nonlinearity.

Table 4.1 shows pairing information of accelerometers in different tests. The HT of the measured FRF was calculated to check the existing nonlinearities of the manipulator. Note that the FRF represents the combined dynamics of the flexible links and joint drives.

4.5.1 Baseline Data Using Shoulder Excitation

Figures 4.14 (a) through (d) show the HTFs superimposed on the FRFs at location A, B, C, and D using shoulder excitation. A list of major resonance and anti-resonance frequencies at all four locations is shown in Table 4.2. Note that several resonant peaks occur for all four accelerometer locations. Figures 4.15 (a) through (c) show the ratio between the HTFs and the FRFs at location A, B, and C.

At location A, the magnitude of the HTF generally matches the FRF. The coherence of the FRF at A as shown in Figure 4.14 (a) was quite low near 20 Hz. Note that resonant peaks in the HTF are usually equal to or lower than those in the FRF. In contrast, the anti-resonance at 18 Hz is much smaller for the HTF than the FRF. The phase of the HTF generally matches the phase of the FRF.

At location B, the FRF has a resonant peak at 4 Hz and drops in magnitude until the anti-resonance at 28 Hz. A minor peak appears at 32 Hz and the magnitude of the FRF rises to peaks at 37 Hz and 41 Hz. The HTF at location B matches the FRF with good coherence.

Test	Excitation Source	Elbow Angle	Accelerometer Location	FRF Figure
1	Shoulder	0	Acc-One at A	4.14 (a)
2			Acc-Two at B	4.14 (b)
			Acc-One at C	4.14 (c)
3			Acc-One at D	4.14 (d)
4	Elbow	30	Acc-One at A	4.16 (a)
5			Acc-Two at B	4.16 (b)
			Acc-One at C	4.16 (c)
6			Acc-Two at D	4.16 (d)

Table 4.1 Test Configuration for Baseline FRF Data.

Accelerometer Location			
A	B	C	D
4 Hz(R)	4 Hz(R)	4 Hz(R)	
12 Hz(AR)			
14 Hz(R)			
18 Hz(AR)			
	29 Hz(AR)	29 Hz(AR)	28 Hz(R)
	32 Hz(R)		34 Hz(AR)
37 Hz(R)	37 Hz(R)	37 Hz(R)	37 Hz(R)
39 Hz(AR)			
41 Hz(R)	41 Hz(R)	41 Hz(R)	42 Hz(R)
54 Hz(R)			

Note: (R) indicates resonance, (AR) indicates anti-resonance.

Table 4.2 Resonant and Anti-Resonant Frequencies in Baseline FRFs Using Shoulder Excitation.

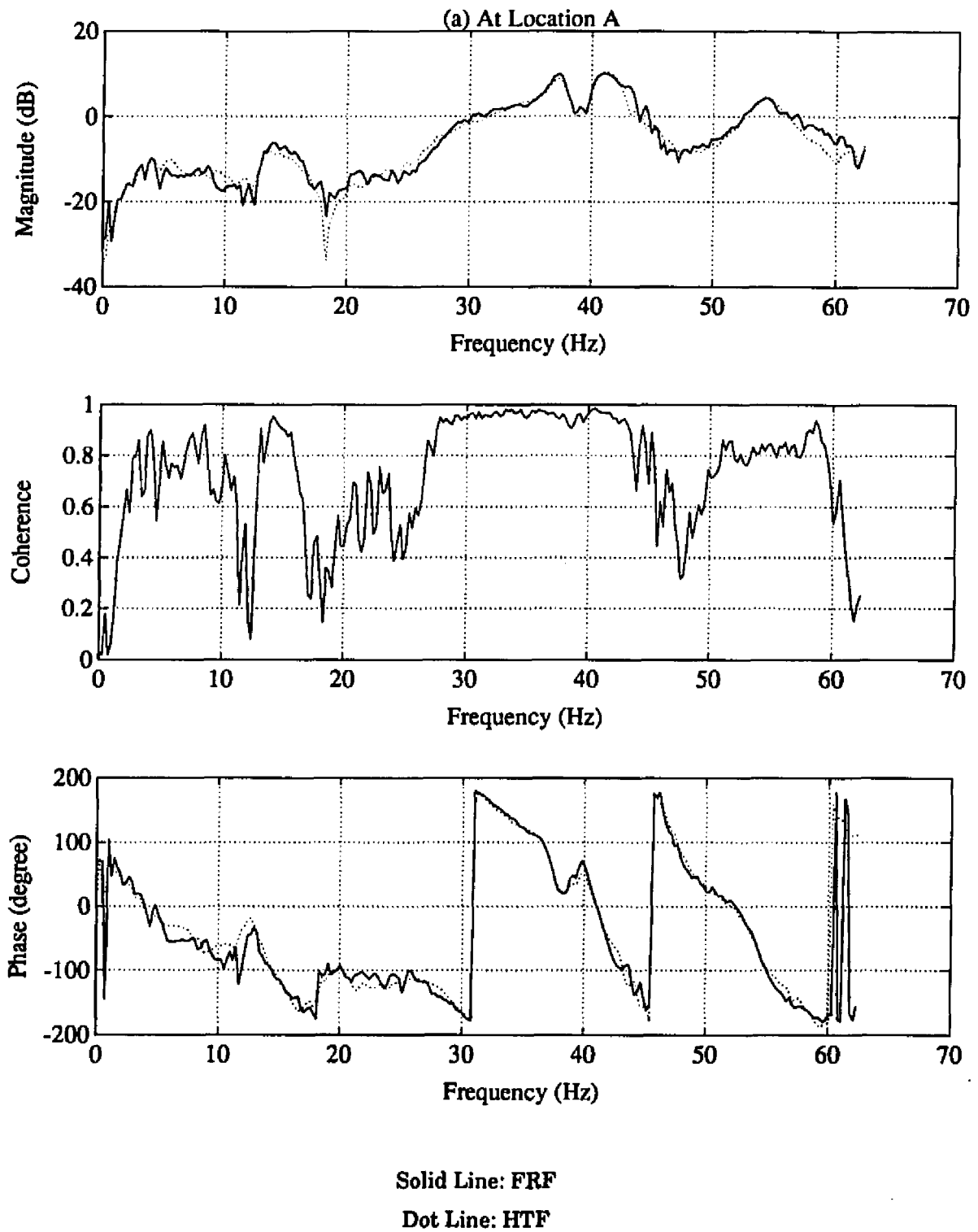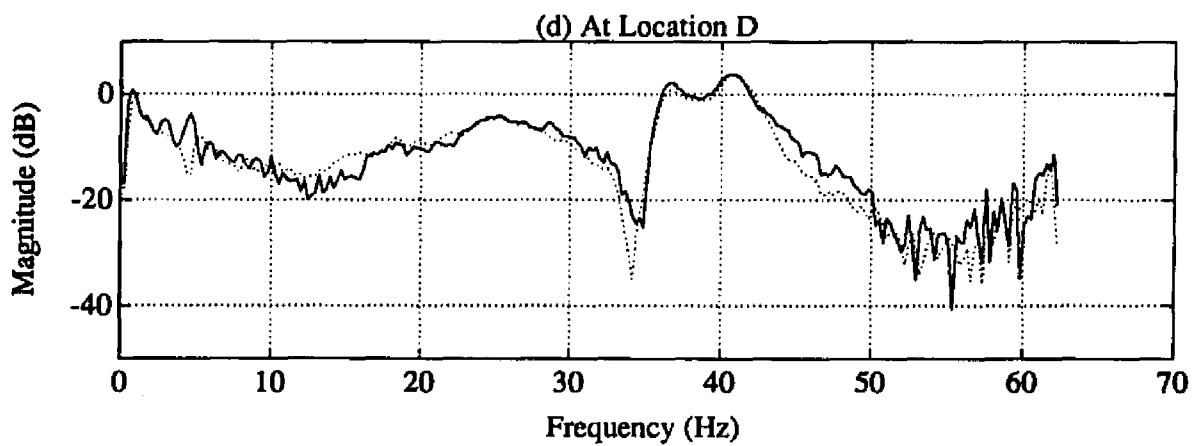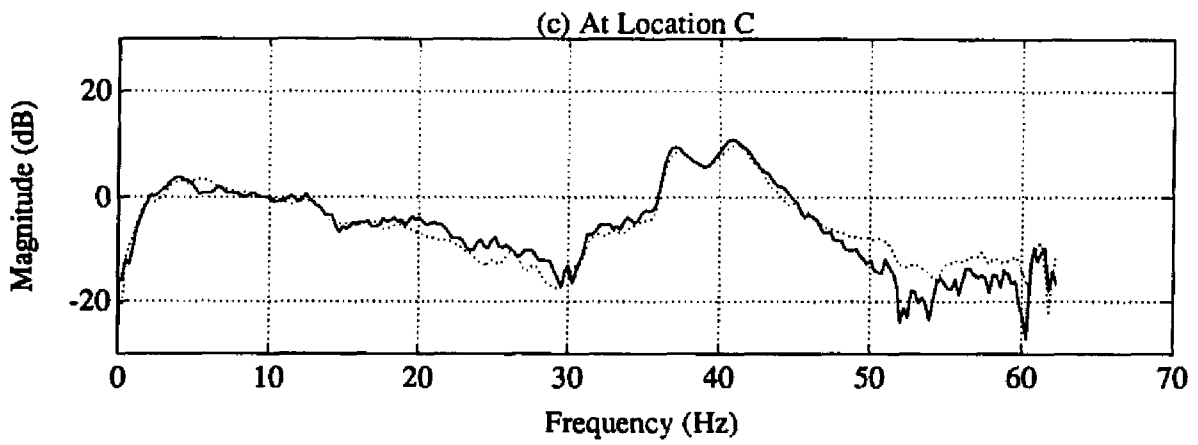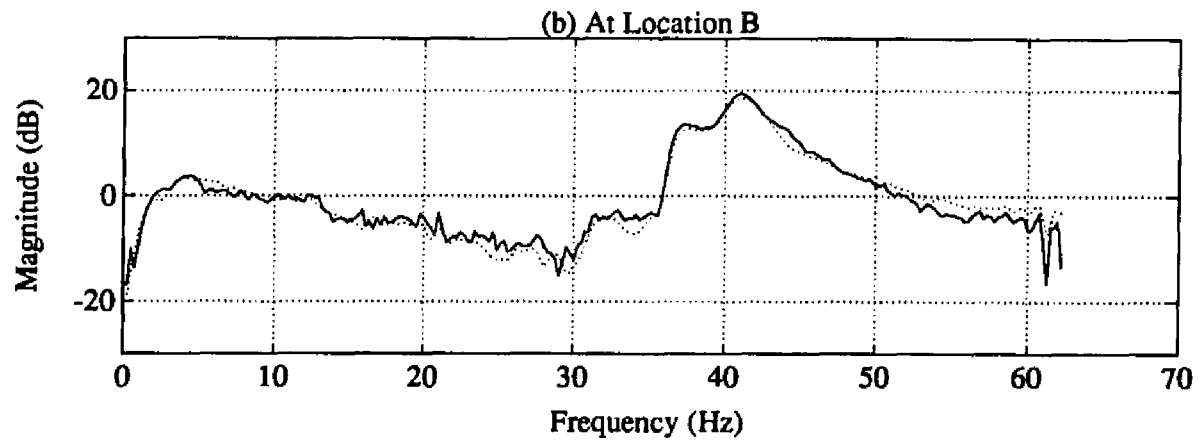Figure 4.14 Baseline FRFs and HTFs Using Shoulder Excitation.

(b) At Location B

(c) At Location C

(d) At Location D

Solid Line: FRF
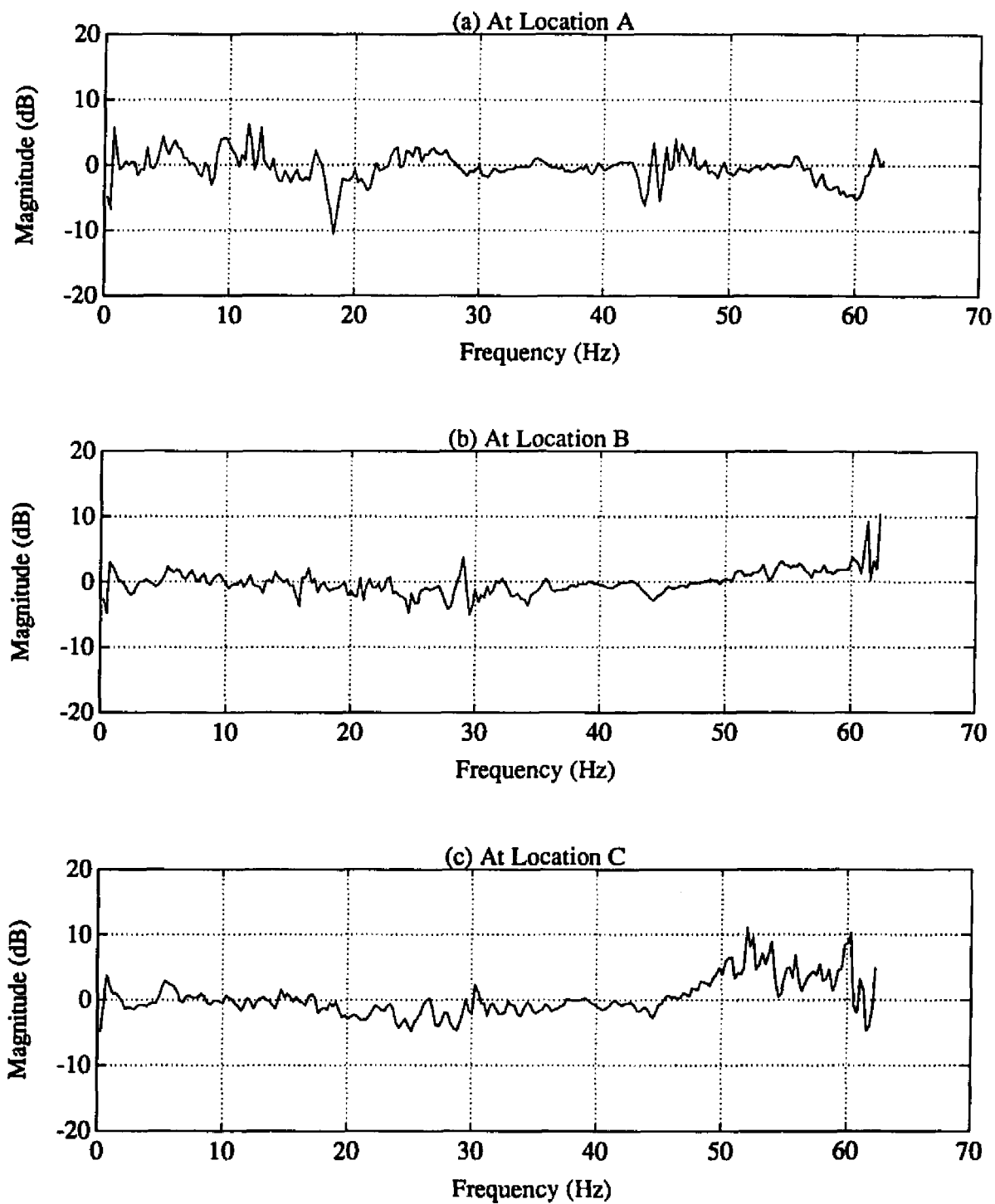Dot Line: HTF

Figure 4.14 (continue)

Figure 4.15 Difference between Baseline FRFs and HTFs Using Shoulder Excitation.

The FRF at location C is generally similar to that at location B except that the resonant peak at 37 Hz is lower at location C. Again, the HTF generally matches the FRF with fair coherence. This result indicates that the elbow joint does not affect the FRF across it. At location D, the coherence of the FRF is not good below 20 Hz although resonant peaks appear at 28, 37 and 42 Hz. The HTF does not match the FRF as well at location D.

Note that due to the limited rolloff rate of 24 dB per octave above the 50 Hz cutoff frequency, small amount of aliasing might affect measurement above 40 Hz. This effect can explain the relatively low coherence in the FRF just below 50 Hz at location A. The quality of the FRFs above 50 Hz at locations C and D were also greatly affected. One way to minimize the aliasing effect is to use a filter with sharper rolloff rate. Another approach is to use a higher sampling frequency, which provides a wider analysis bandwidth, to measure the FRF. When further analysis of the FRF is performed, the high frequency components of the FRF can be disregarded. As discussed in Section 4.3.2, higher sampling frequency was not used in the excitation test. Care must be taken to interpret analysis results above 50 Hz.

4.5.2 Baseline Data Using Elbow Excitation

Figures 4.16 (a) through (d) show the HT superimposed on the FRFs at location A, B, C, and D using elbow excitation. Note that the elbow joint angle was held at +30 degrees for the elbow excitation test to prevent elbow motor drift. Because the elbow motor is not powerful enough to generate large excitation, the response signal from the accelerometers was small compared to the shoulder excitation test. A list of major resonance and anti-resonance frequencies at all four locations is shown in Table 4.3.

At location A, the HTF generally matched the FRF at resonant peaks but varied between them. The quality of the FRF below 15 Hz was bad due to low accelerometer response levels over this frequency range. Generally, the magnitude of the FRF at location A is small compared to other accelerometer locations because location A is very proximal to the excitation source. At location B, the FRF had poor coherence for the entire frequency range. Thus, FRF measured at location B will not be used for further analysis.

The FRF at location C was much better than at location B. The HTF matches the FRF better at higher frequencies than at lower frequencies. The FRF has an anti-resonance at 2 Hz and increases in magnitude until 5 Hz. It stays at the same level to 10 Hz and drops

Accelerometer Location			
A	B	C	D
		2 Hz(AR)	2 Hz(AR)
5 Hz(R)	5 Hz(R)		6 Hz(R)
			10 Hz(AR)
20 Hz(R)		20 Hz(AR)	
31 Hz(AR)			32 Hz(AR)
33 Hz(R)	33 Hz(R)	32 Hz(R)	33 Hz(R)
41 Hz(R)			
		50 Hz(R)	
55 Hz(R)			

Note: (R) indicates resonance, (AR) indicates anti-resonance.

Table 4.3 Resonant and Anti-Resonant Frequencies in Baseline FRFs Using Elbow Excitation.

to an anti-resonance at 20 Hz. The magnitude of the FRF then rises to a peak at 32 Hz and rises again to a peak at 50 Hz. The ratio between the HTF and the FRF as shown in Figure 4.17 is quite large below 25 Hz. Results from location B and C shows that better response can be measured at more distal locations than more proximal locations with respect to the excitation source.

At location D, the quality of the FRF was not good above 40 Hz due to small response at higher frequencies. While only one clear resonant peak occurred at 33 Hz, anti-resonances appeared at 2, 10 and 32 Hz. The HTF does not match the FRF well at location D below 30 Hz.

4.5.3 Strategy to Identify Joint Nonlinearity

For a linear system, the HTF should match the FRF. Therefore, by comparing the HTF and FRF, differences would be indicators of nonlinearities in system characteristics. This information can be used to identify the presence, magnitude and location of nonlinear effects.
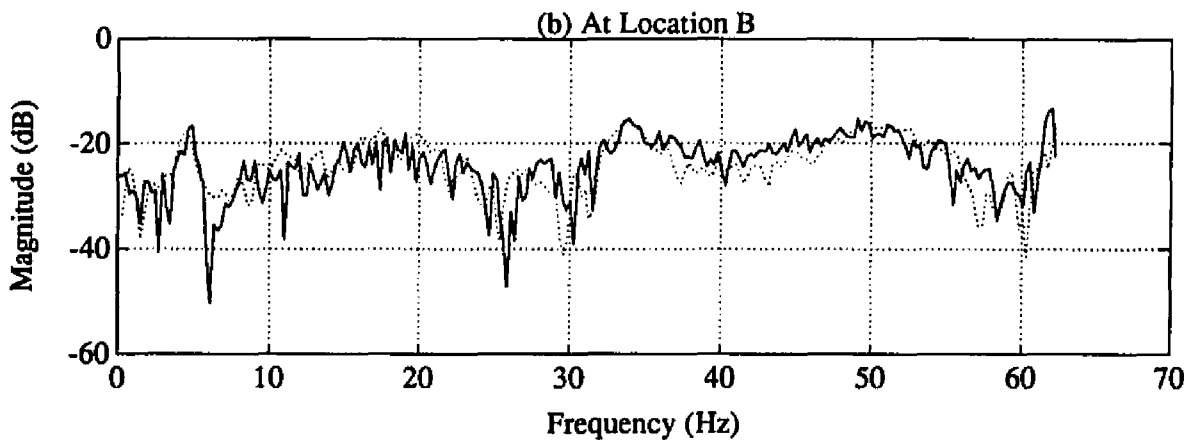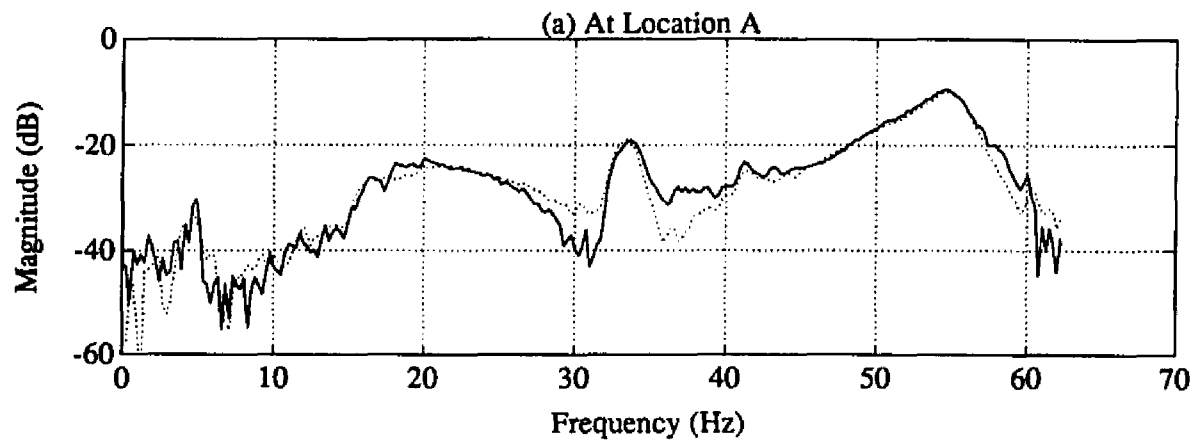
Figure 4.16 Baseline FRFs and HTFs Using Elbow Excitation.

(c) At Location C
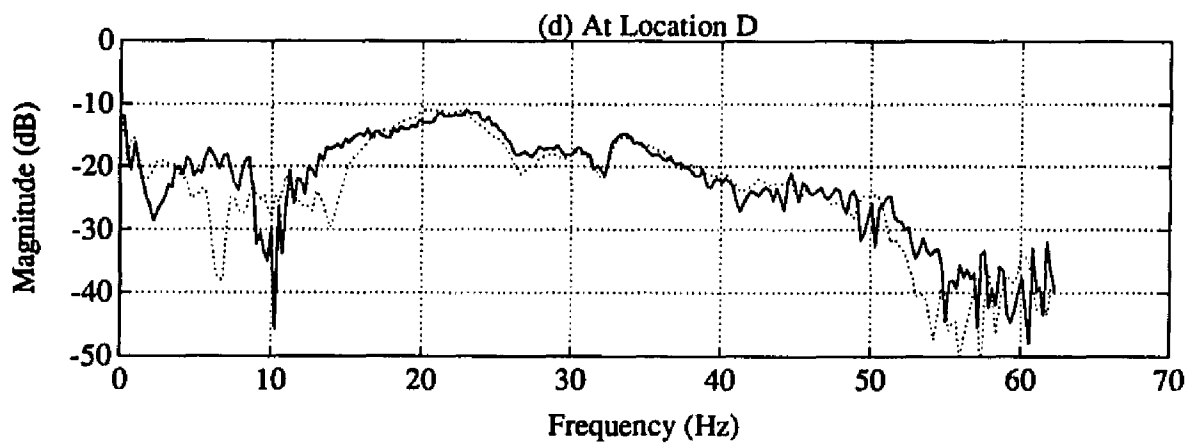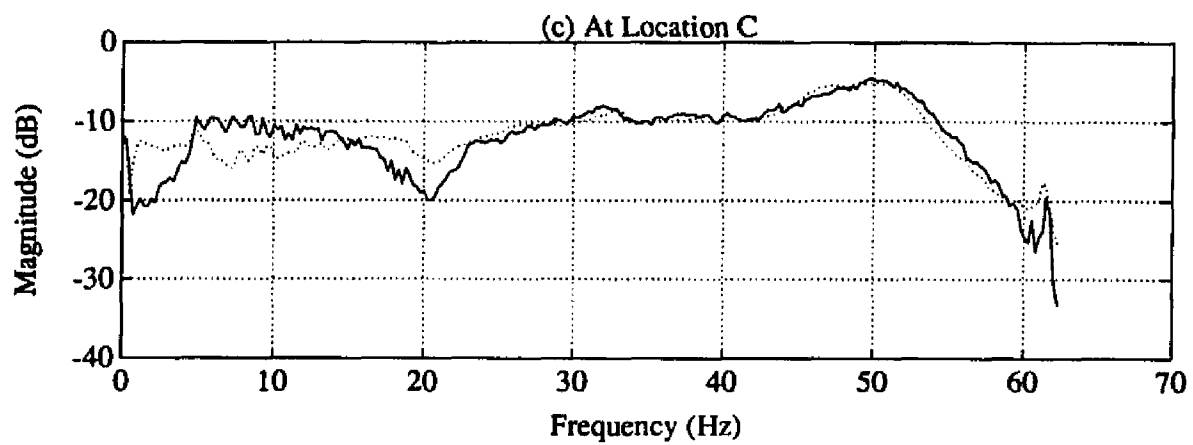
(d) At Location D

Solid Line: FRF
Dot Line: HTF

Figure 4.16 (continue)
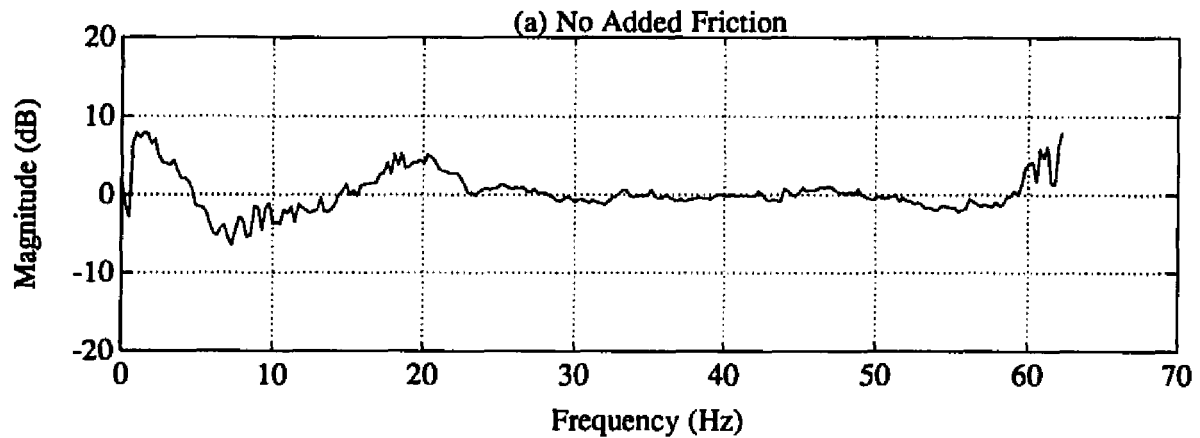
Figure 4.17 Ratio between Baseline FRF and HTF at Location C Using Elbow Excitation.

One important subject in robot PMP is the magnitude of component degradation. This information directly affects maintenance scheduling. While variation between the HTF and FRF of a system indicates the presence of defect conditions, the magnitude of variation provides a direct indication of the severity of the defect. Once the presence and magnitude of nonlinearities are detected, the next task is to identify the possible locations of these nonlinearities. This information is useful to identify the source of nonlinearities so that the degraded component can be identified and repaired.

When output response sensors are installed along the links of robot, a series of FRFs can be measured along the linkages. Since friction and backlash in joint drives are localized properties, it is possible to detect changes in the FRF across these nonlinear elements. This is based on the assumption that nonlinear effects will affect more distal measurements to a greater extent than more proximal measurements relative to the excitation source.

To test this concept on the prototype manipulator, two accelerometers were installed on each sides of the elbow joint at locations B and C. Shoulder motor excitation was used and additional friction nonlinearities were introduced. These nonlinearities should then cause significant changes across the FRF and HTF measured at Location B and C. Ratio between the HTF and FRF will indicate the presence and significance of nonlinearity. If the distal FRF were classified as more nonlinear than the proximal FRF, it can be concluded that a significant nonlinear effect exists across the elbow joint drive.

4.6 Friction Test on Shoulder Joint

Using shoulder excitation, friction tests were performed with Coulomb friction added to the shoulder joint shaft. The purpose of these tests was to verify that variation between the HT and the FRF can be related to the difference in friction levels applied at the shoulder joint. Note that the shoulder joint utilizes a direct drive brushless motor with minimum friction and no backlash in the motor. To add friction to the shaft, a hinged lever loaded by a pulley and a dead weight was used as shown in Figure 4.18. A piece of rubber was attached on the lever to contact the shaft. Because Coulomb friction is proportional to the normal force at the contact point, different levels of friction could be induced by dead weights. Table 4.4 shows test conditions for the shoulder friction test.

Figures 4.19 (a) through (c) show the HTF superimposed on the FRF at location A with three different friction levels. Compared with the baseline data as shown in Figure 4.14 (a), there is increased mismatch between the HTF and the FRF when friction is applied to the joint. Figures 4.20 (a) through (c) show the ratio between the HTF and the FRF at location A with different friction levels. The ratio between the HTF and the FRF increased as friction levels increased. Table 4.5 shows the RMS ratio and maximum difference across the frequency range. Note that the overall RMS ratio increased when friction level increased. The RMS level below 30 Hz also increased as friction level increased. This trend also happened at the maximum and minimum peak difference between 10 and 20 Hz, where one resonant and two anti-resonant peaks exist in the FRF. Because the HTF generally did not have well defined resonant peaks for all tests, skewness and kurtosis analyses were not used.

Figures 4.21 (a) through (c) show the HTF superimposed on the FRF at location B with three different friction levels. The ratio between the HTF and the FRF at location B is shown in Figures 4.22 (a) through (c). Figures 4.23 (a) through (c) show the HTF superimposed on the FRF at location C with three different friction levels. The ratio between the HTF and the FRF at location C is shown in Figures 4.24 (a) through (c). Table 4.6 summarizes the RMS ratio over several frequency ranges. The RMS control effort to hold the elbow stationary is also included in Table 4.6.

At both location B and C, the overall RMS level increased when friction increased to moderate level. However, the RMS level dropped when high level friction was added. The
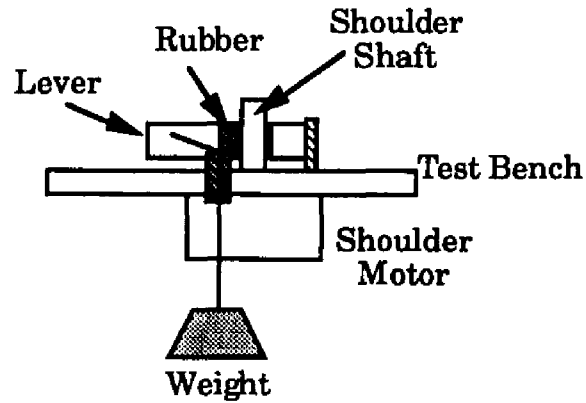
Figure 4.18 Mechanism to Generate Shoulder Friction.

Test	Excitation Source	Elbow Angle	Accelerometer Location	Friction Level	FRF Figure
1				Low (2lb)	4.19 (a)
2	Shoulder	0	Acc-One at A Acc-Two at B	Moderate (4lb)	4.19 (b)
3				High (8lb)	4.19 (c)
4				Low (2lb)	4.22 (a) 4.21 (a)
5	Shoulder	0	Acc-One at C Acc-Two at B	Moderate (4lb)	4.22 (b) 4.21 (b)
6				High (8lb)	4.22 (c) 4.21 (c)

Note: value of pounds in the friction level indicates dead weight used.

Table 4.4 Test Conditions for Shoulder Friction Test.

Figure 4.19 FRFs and HTFs for Shoulder Friction Test at Location A.
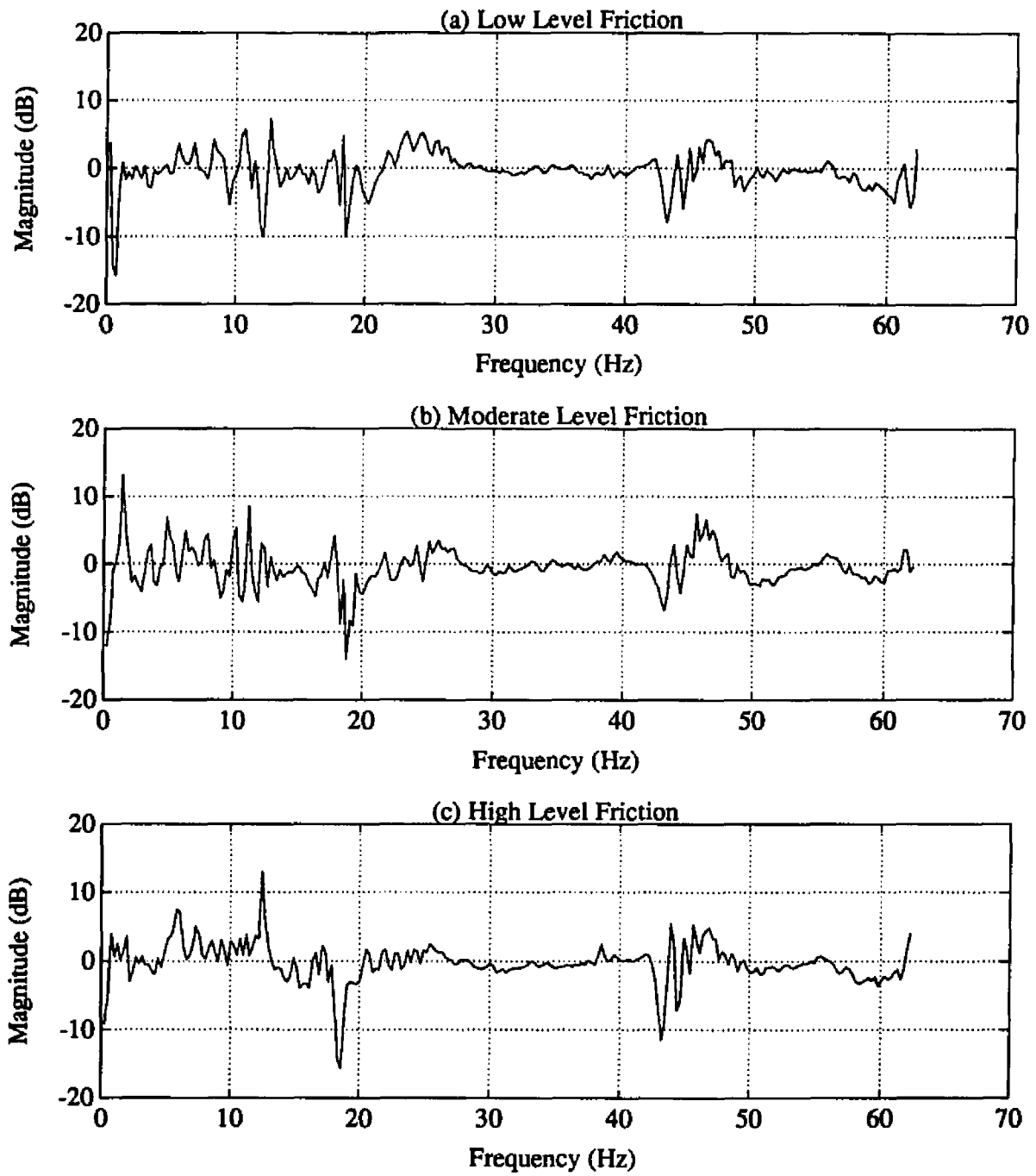
Figure 4.20 Ratio between FRF and HTF for Shoulder Friction Test at Location A.

Friction Level	RMS Ratio (0-62.5 Hz)	RMS Ratio (0-30 Hz)	Max (dB) (10-20 Hz)	Min (dB) (10-20 Hz)	Span (dB) (10-20 Hz)
None	0.249	0.295	6.2	-10.5	16.7
Low	0.289	0.392	7.3	-10.0	17.3
Moderate	0.383	0.477	8.5	-14.0	22.5
High	0.365	0.455	13.1	-15.6	28.7

Table 4.5 Ratio between FRF and HTF at Location A for Shoulder Friction Test.

Friction Level	Location B		Location C		Elbow Control Effort (RMS)
	RMS Ratio (0-62.5 Hz)	RMS Ratio (0-30 Hz)	RMS Ratio (0-62.5 Hz)	RMS Ratio (0-30 Hz)	
None	0.367	0.176	0.462	0.194	0.576
Low	0.244	0.215	0.596	0.306	0.575
Moderate	0.240	0.293	0.731	0.327	0.569
High	0.238	0.253	0.448	0.295	0.447

Table 4.6 Ratio between FRF and HTF at Location B and C for Shoulder Friction Test.

(a) Low Level Friction

(b) Moderate Level Friction
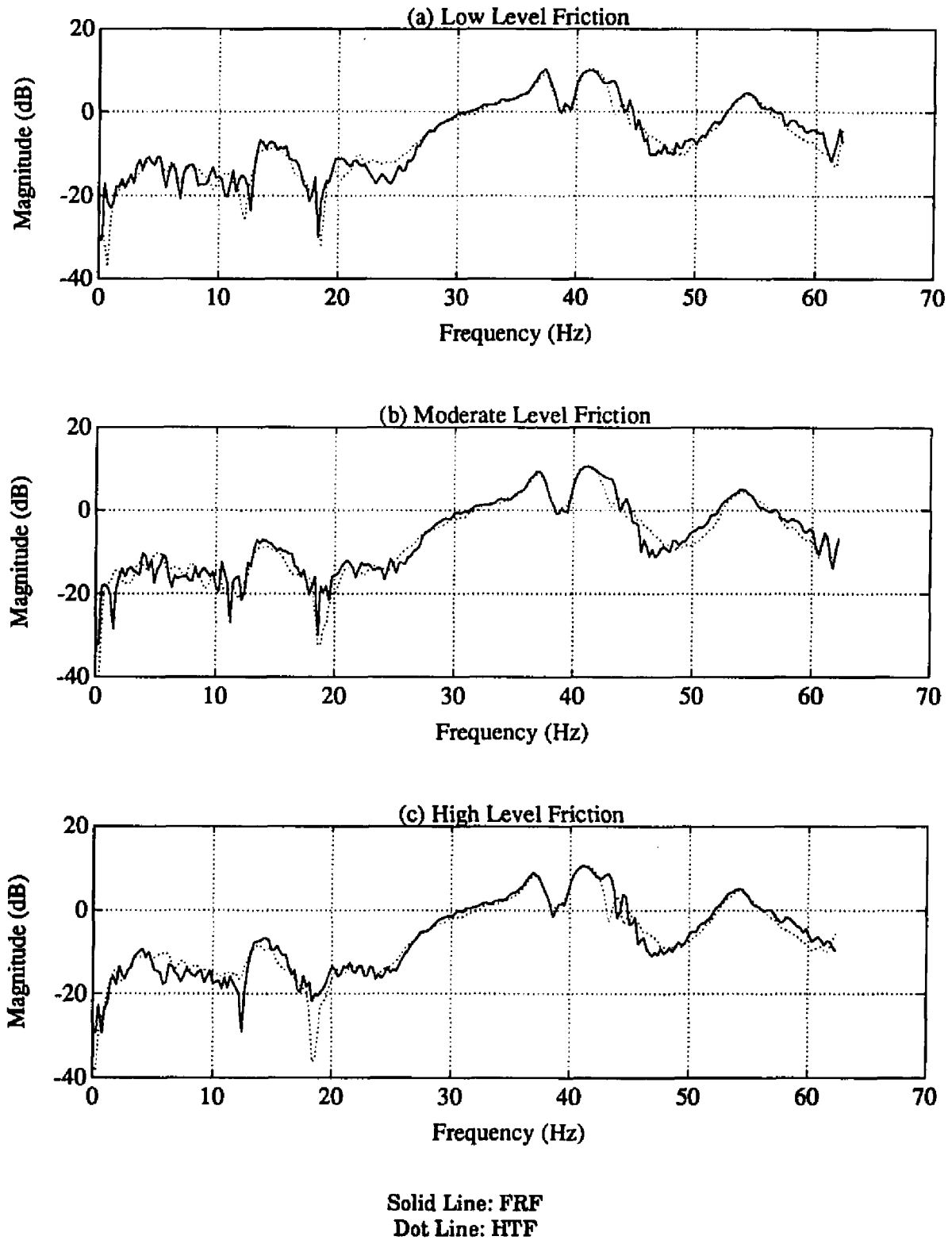
(c) High Level Friction

Solid Line: FRF
Dot Line: HTF

Figure 4.21 FRFs and HTFs for Shoulder Friction Test at Location B.

Figure 4.22 Ratio between FRF and HTF for Shoulder Friction Test at Location B.

(a) Low Level Friction

(b) Moderate Level Friction

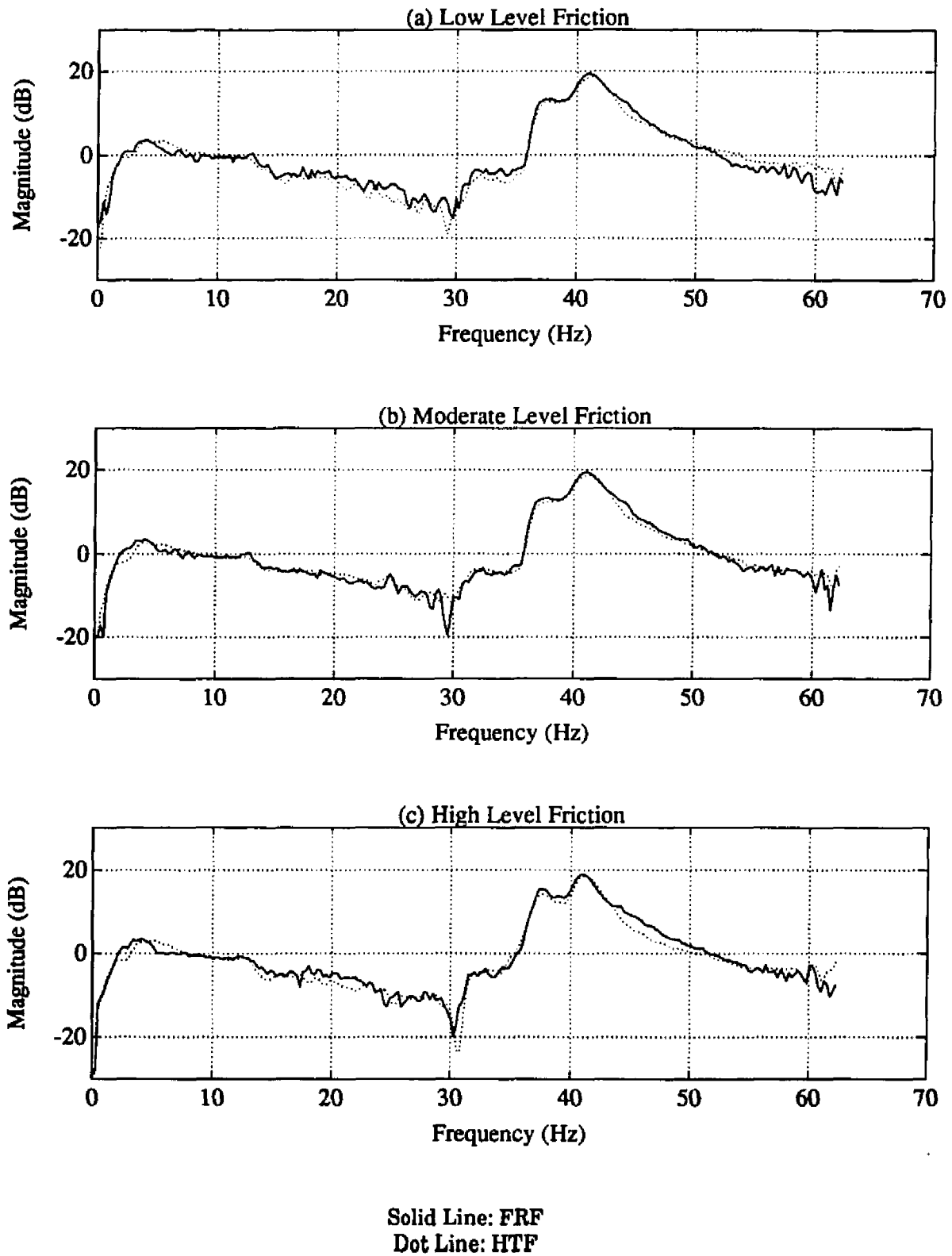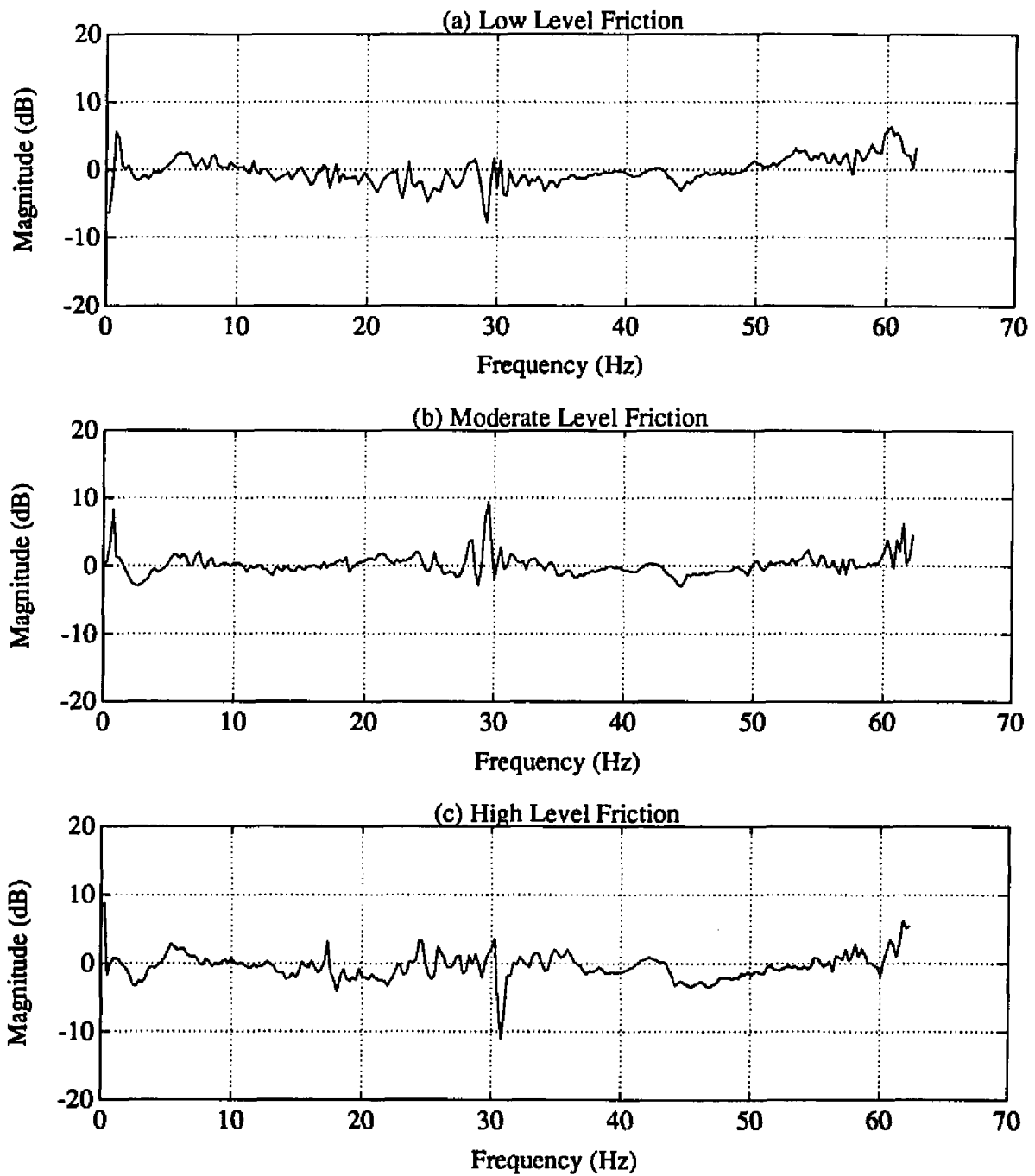(c) High Level Friction
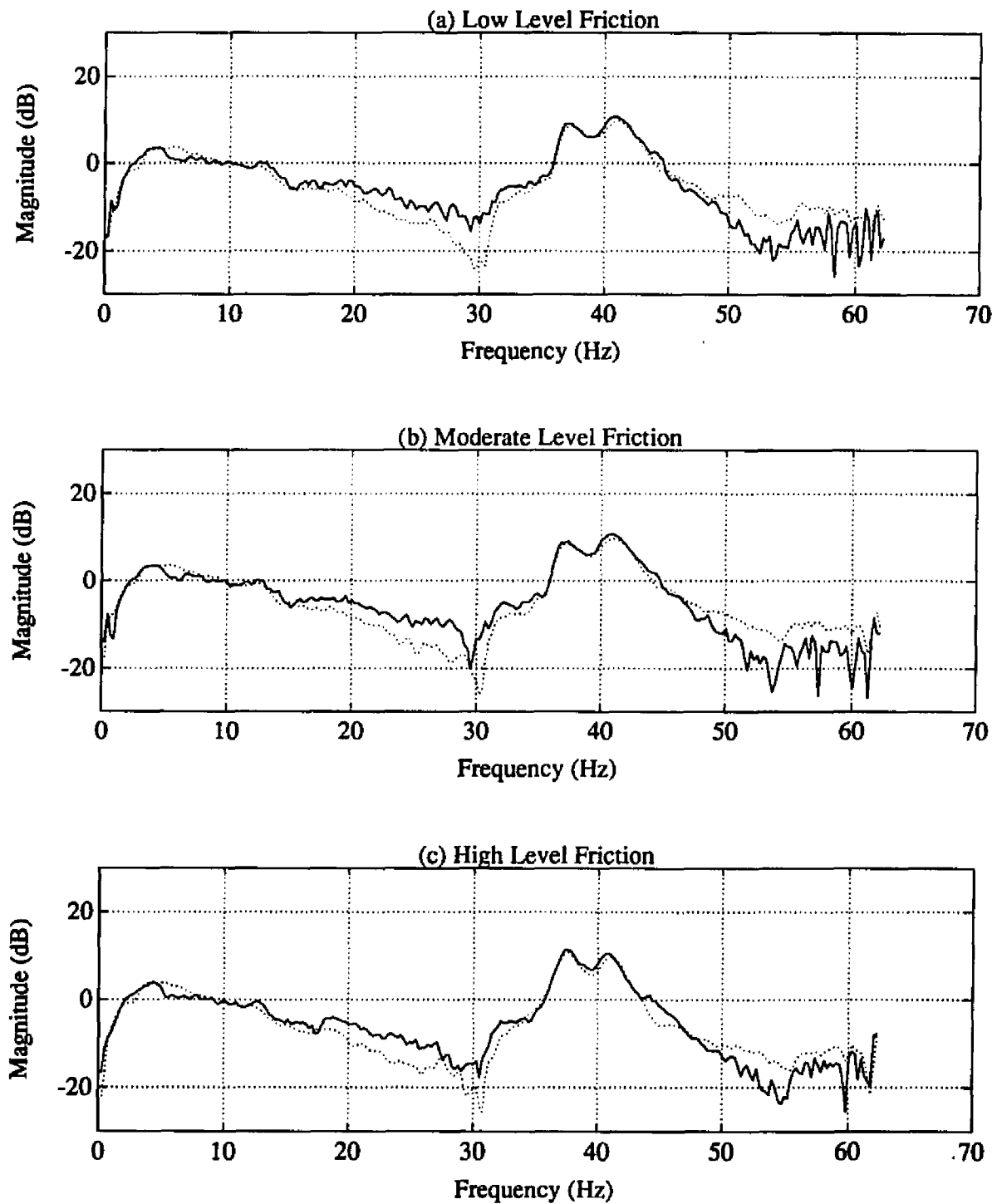
Solid Line: FRF
Dot Line: HTF

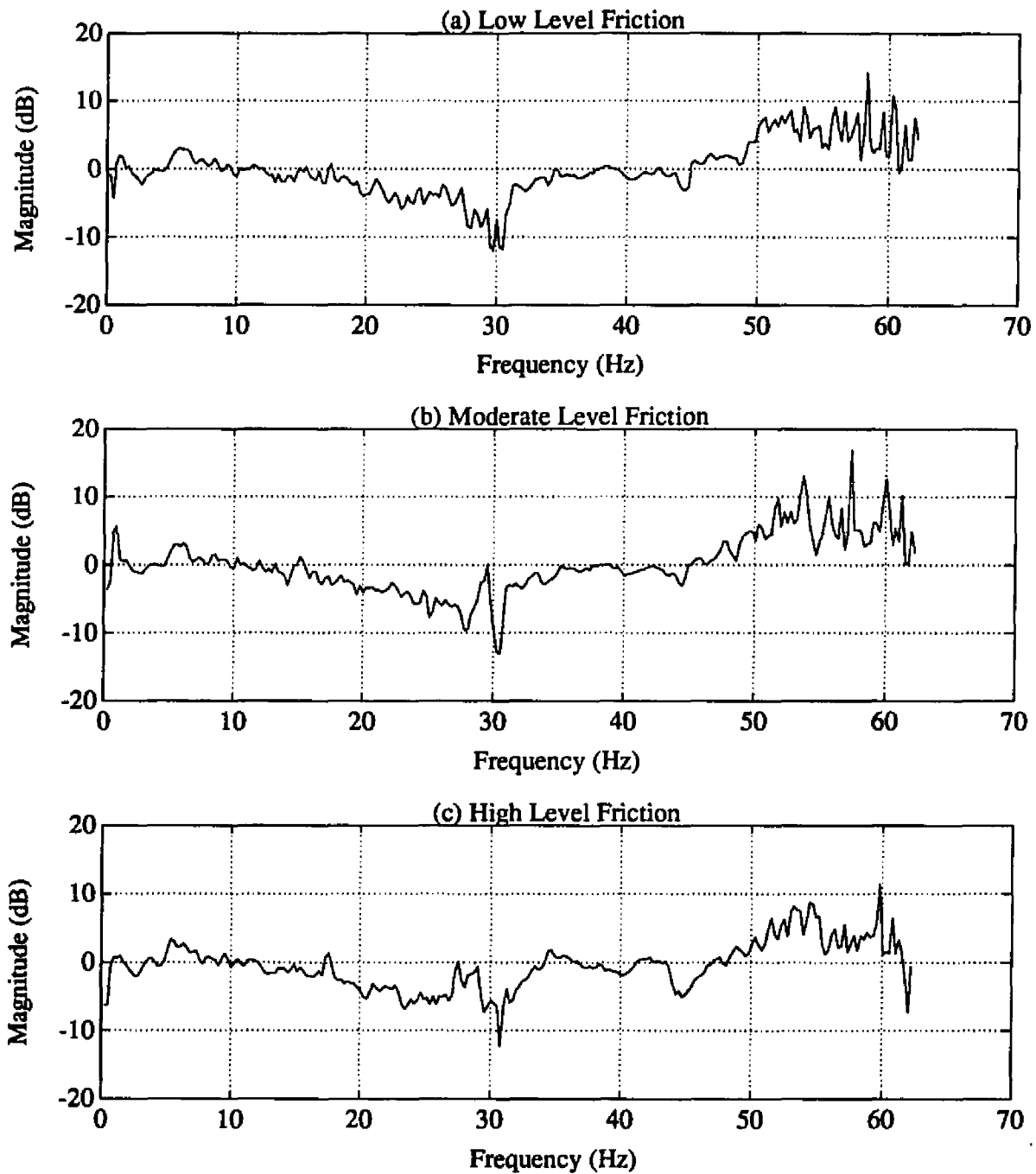Figure 4.23 FRFs and HTFs for Shoulder Friction Test at Location C.

Figure 4.24 Ratio between FRF and HTF for Shoulder Friction Test at Location C.

RMS level below 30 Hz did not show any trend at both location B and C. The increased RMS ratio at location C is due to significant mismatch at higher frequencies. The overall control effort for the elbow motor stayed at nominally the same level for different friction levels, but dropped significantly when high level friction was added to the shoulder.

In summary, the ratio between the HTF and the FRF is relatively proportional to the friction effect at location A but not at locations B and C. This result agreed with predictions because location A is the closest measurement point to the nonlinear effect. Note that RMS ratios below 30 Hz at location C were almost the same. This indicates that the friction effect at location A did not propagate across the elbow joint.

4.7 Friction Test on Elbow Joint

Using shoulder excitation, friction tests were performed with added friction at the elbow joint shaft. The purpose of these tests was to verify firstly that variation between the HT and the FRF can be related to the difference in friction level applied at the elbow joint. Secondly, the difference is larger distal to the measuring location (location C) than more proximal locations (location A and B) with respect to the excitation source. The elbow joint utilizes a DC motor with a planetary gear head that already contains significant friction and backlash effects. To add friction, a string was wrapped around the elbow shaft and attached to the end of the proximal link as shown in Figure 4.25. Different levels of friction were added by adjusting the number of wraps around the shaft as shown in Table 4.7.

Figure 4.26 shows the HTFs superimposed on the FRFs at location A with low and high elbow friction levels. Compared with the baseline FRF, the peak at 41 Hz is higher when elbow friction is added. The ratio between the HTF and the FRF at location A is shown in Figure 4.27. Table 4.8 shows the RMS ratio and maximum difference across three frequency ranges. Note that the RMS ratio level at location A stayed the same for different levels of elbow friction. The maximum and minimum peak difference as well as the span between 10 and 20 Hz also stayed at about the same level. This means friction effects at the elbow joint have minimum effect on the measurement at location A.

Figure 4.28 shows the HTFs superimposed on the FRFs at location B with low and high friction levels. The FRFs are generally the same as the baseline FRF. Figure 4.29 shows the ratio between the HTF and the FRF at location B. Unlike the up-down style
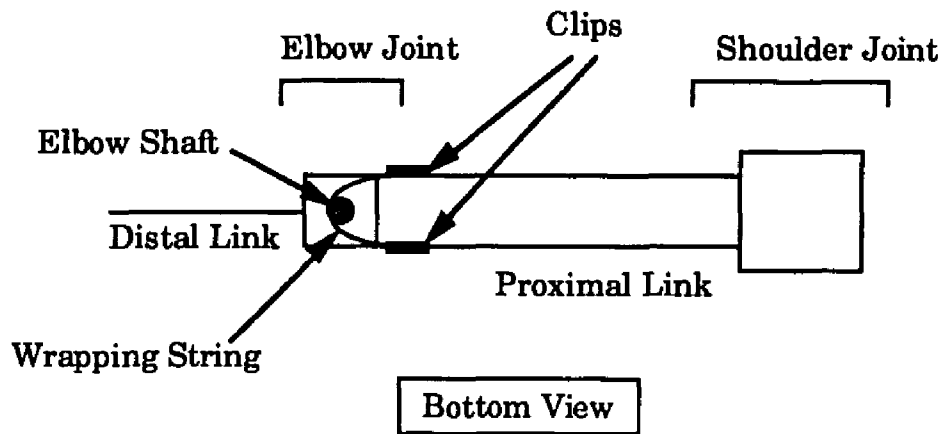
Figure 4.25 Wrapping to Generate Elbow Friction.

Test	Excitation Source	Elbow Angle	Accelerometer Location	Friction Level	FRF Figure
1	Shoulder	0	Acc-One at A	low (2 turns)	4.26 (a)
2				high (4 turns)	4.26 (b)
3			Acc-Two at B Acc-One at C	low (2 turns)	4.28 (a) 4.30 (a)
4				high (4 turns)	4.28 (b) 4.30 (b)
5	Elbow	30	Acc-One at C	low (2 turns)	4.32 (a)
6				high (4 turns)	4.32 (b)

Note: value of turns in the friction level indicates number of string wraps.

Table 4.7 Test Conditions for Elbow Friction Test.
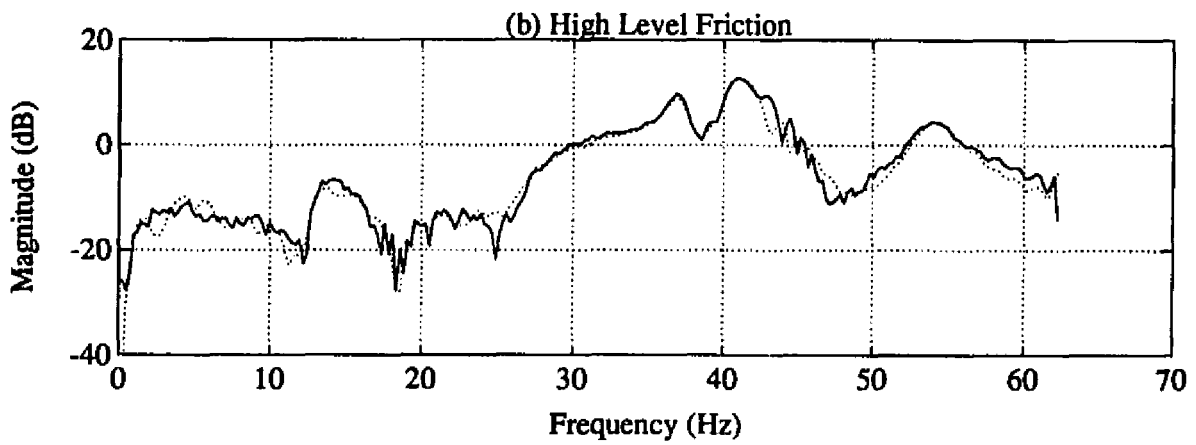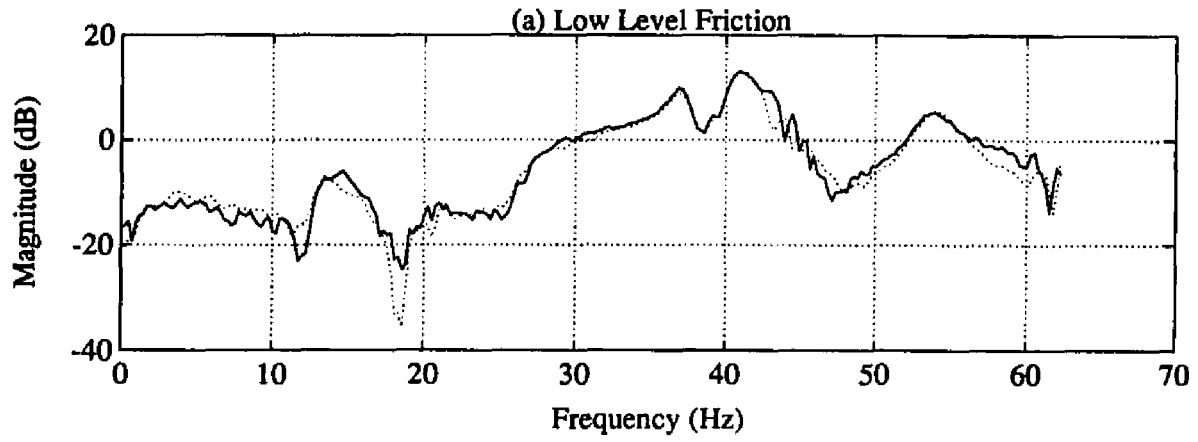
Figure 4.26 FRFs and HTFs for Elbow Friction Test at Location A
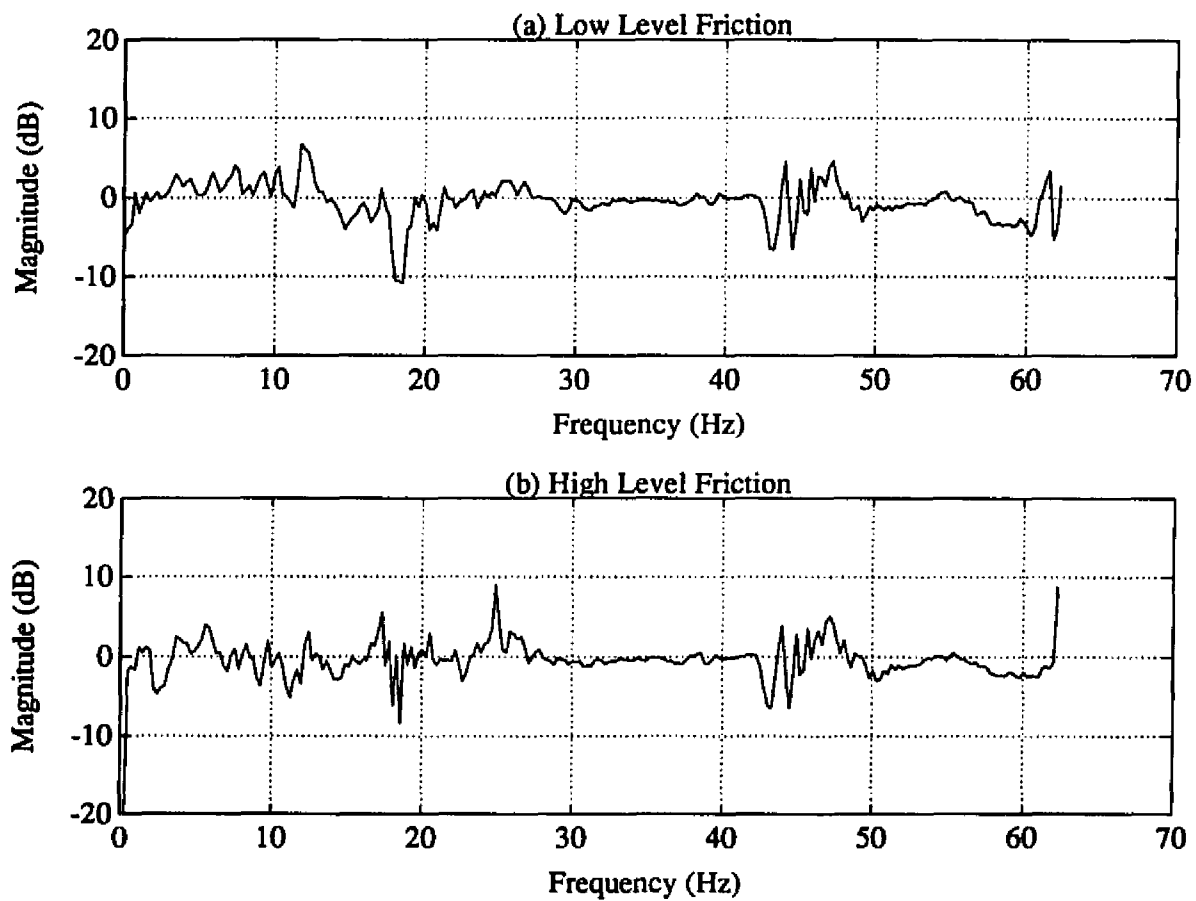
Figure 4.27 Ratio between FRF and HTF for Elbow Friction Test at Location A.

Friction Level	RMS Ratio (0-62.5 Hz)	RMS Ratio (0-30 Hz)	Max (dB) (10-20 Hz)	Min (dB) (10-20 Hz)	Span (dB) (10-20 Hz)
None	0.249	0.295	6.2	-10.5	16.7
Low	0.256	0.288	6.7	-10.9	17.6
High	0.278	0.291	5.5	-8.3	13.8

Table 4.8 Ratio between FRF and HTF at Location A for Elbow Friction Test.

Friction Level	Location B		Location C	
	RMS Ratio (0-62.5 Hz)	RMS Ratio (0-30 Hz)	RMS Ratio (0-62.5 Hz)	RMS Ratio (0-30 Hz)
None	0.267	0.176	0.462	0.194
Low	0.285	0.315	0.255	0.168
High	0.399	0.300	0.733	0.360

Table 4.9 Ratio between FRF and HTF at Location B and C for Elbow Friction Test.

(a) Low Level Friction

(b) High Level Friction
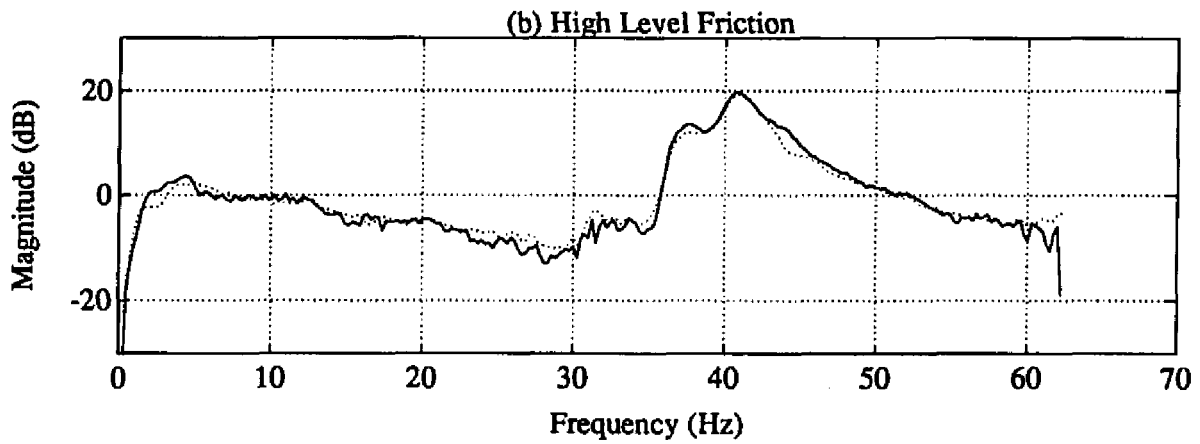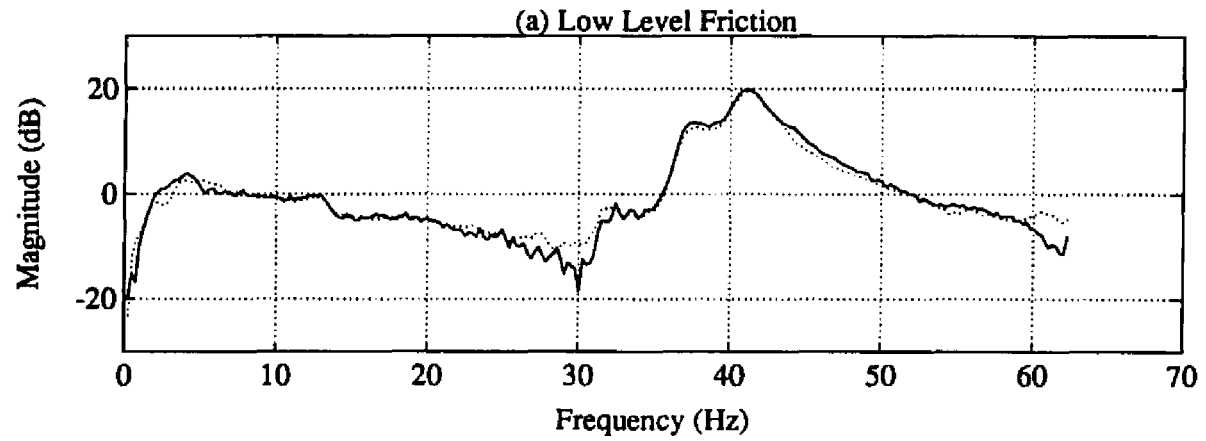
Solid Line: FRF

Dashed Line: HTF

Figure 4.28 FRFs and HTFs for Elbow Friction Test at Location B.
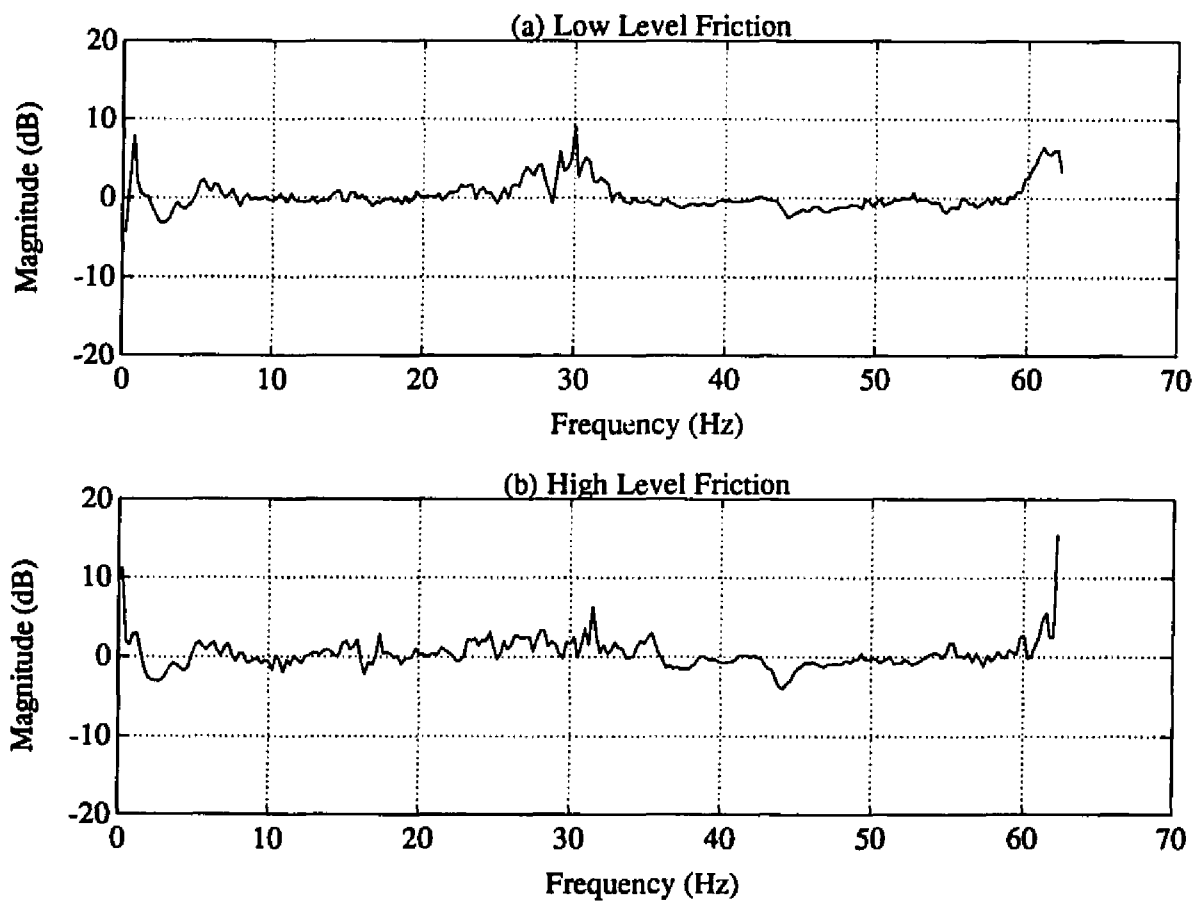
Figure 4.29 Ratio between FRF and HTF for Elbow Friction Test at Location B.

peaks around 30 Hz in the baseline data, only positive peaks occur when elbow friction was added. The RMS ratio at location B increased to a higher level as shown in Table 4.9.

Figure 4.30 shows the HTFs superimposed on the FRFs at location C with low and high friction levels. Figure 4.31 shows the ratio between the HTF and the FRF at location C. For low level friction, the FRF and HTF are similar to the baseline data. Compared to the baseline data as shown in Figure 4.15 (c), the ratio between FRF and HTF actually decreased for most of the frequency range except near 30 Hz. When higher levels of friction were added, the anti-resonance at 30 Hz widened significantly. Ratio between HTF and FRF also increased significantly between 15 Hz and 35 Hz. The overall RMS ratio also shows this trend.

Additionally, friction tests were performed with added friction at the elbow joint shaft but using elbow excitation rather than using shoulder excitation. As discussed in the baseline data measurement using elbow excitation, only FRF at location C was measured to examine the effect of friction. Figure 4.32 shows the HTFs superimposed on the FRFs at location C with low and high friction levels. Figure 4.33 shows the ratio between the HTF and the FRF at location C. As the friction level increased, the anti-resonance at 2 Hz became smaller, the magnitude around 10 Hz was also smaller, and the anti-resonance at 20 Hz was not well defined. The overall quality of the FRF was degraded. The ratio between the HTF and the FRF was quite large below 25 Hz.

(a) Low Level Friction

(b) High Level Friction
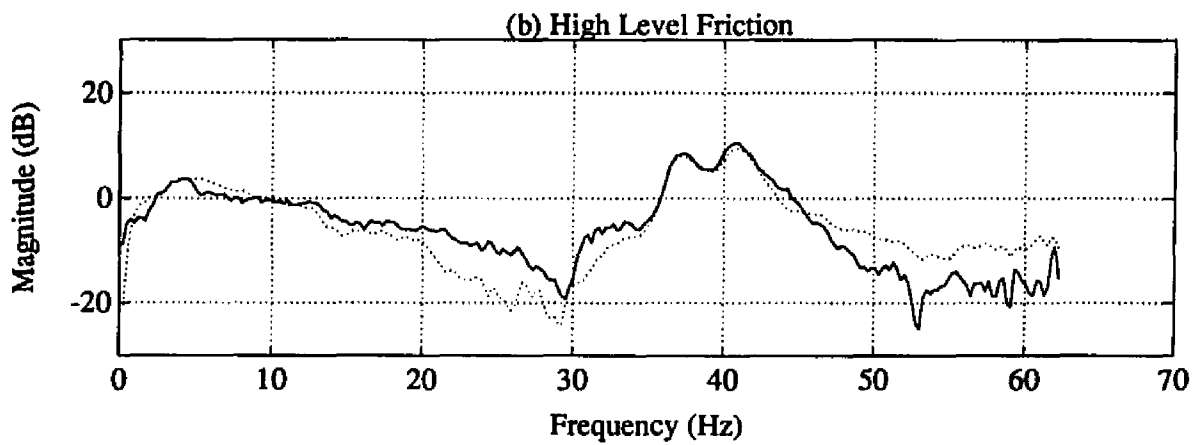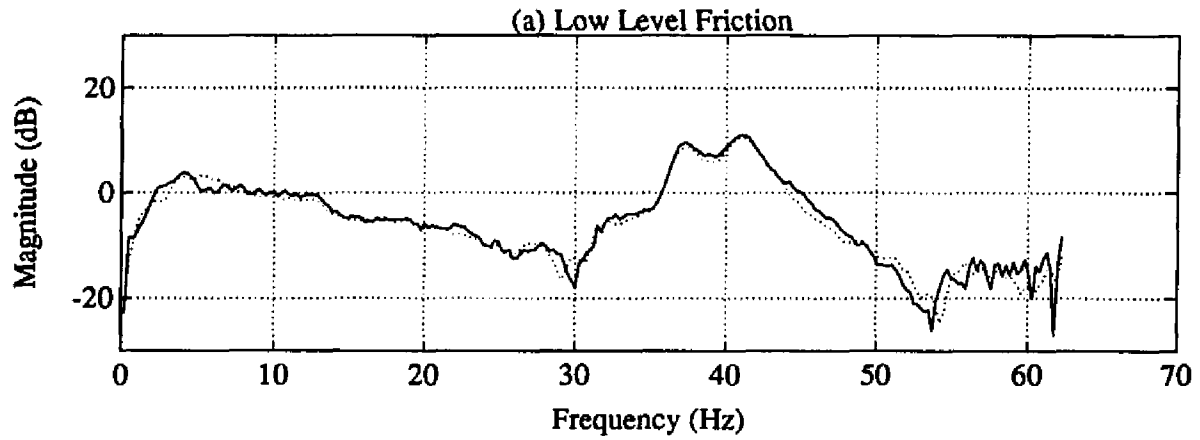
Solid Line: FRF

Dashed Line: HTF

Figure 4.30 FRFs and HTFs for Elbow Friction Test at Location C.

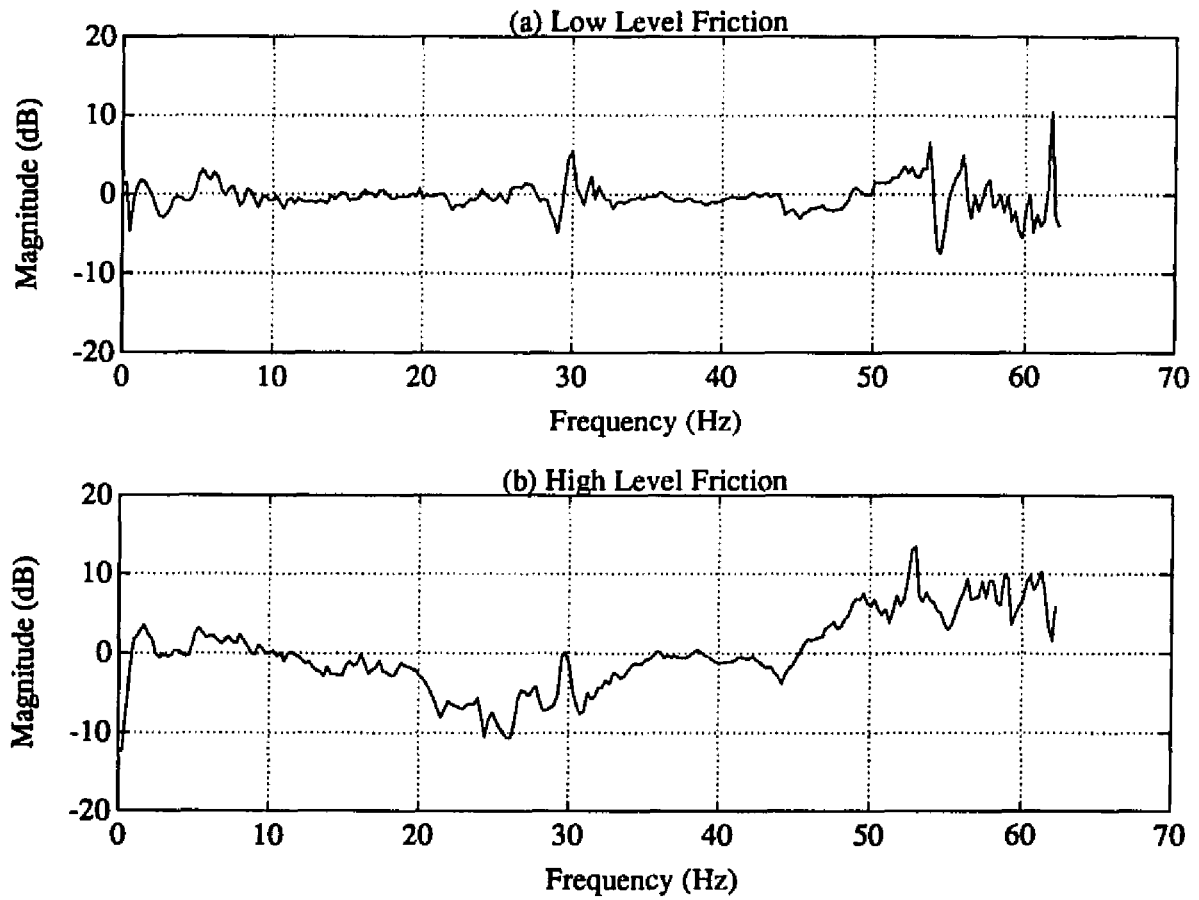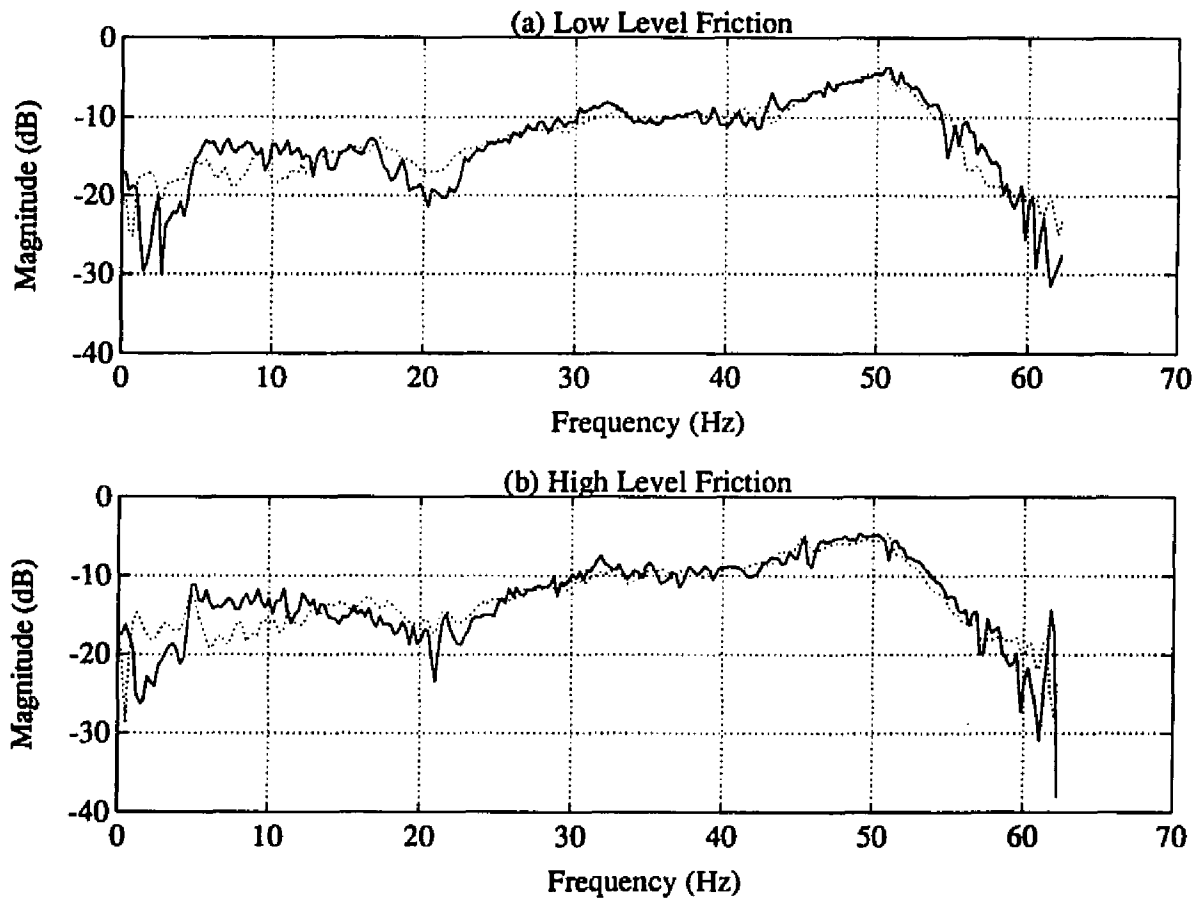Figure 4.31 Ratio between FRF and HTF for Elbow Friction Test at Location C.

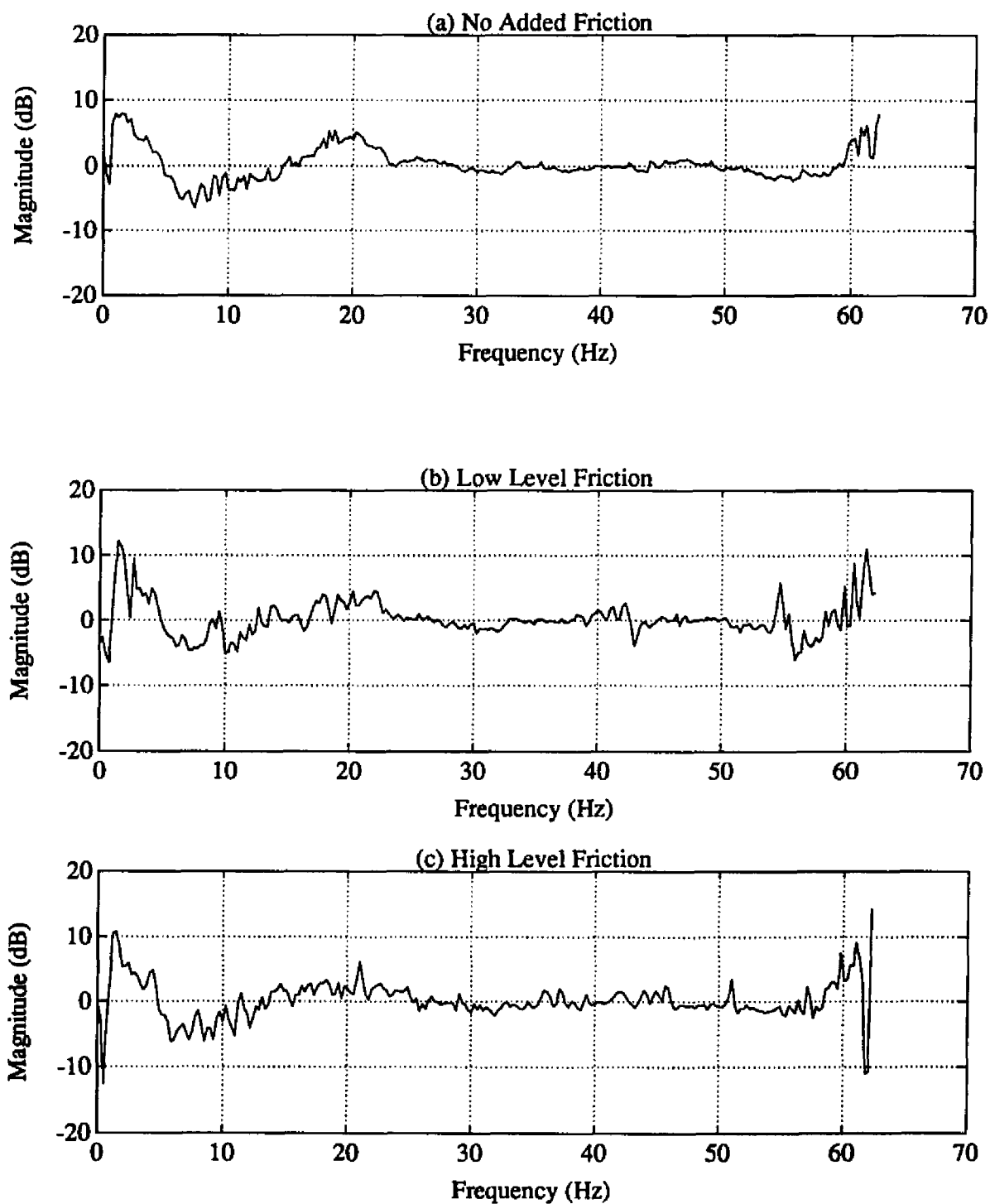**Figure 4.32 FRFs and HTFs for Elbow Friction Test at Location C Using Elbow Excitation.**

Figure 4.33 Ratio between FRF and HTF for Elbow Friction Test at Location C Using Elbow Excitation.

Chapter 5

CONCLUSIONS AND RECOMMENDATIONS

5.1 Conclusions

In this thesis, a procedure to examine nonlinearity in robot joint drives was developed. FRF of the joint drive was measured using controlled robot actuator excitation. Sine sweep excitation was used as the input signal. Output responses were measured using accelerometers attached to robot links. When the HT was applied to the measured FRF, the ratio between the HTF and FRF with respect to the baseline data was used to identify the presence, magnitude and location of nonlinearity in the joint drive system. This procedure may be used for predictive maintenance of robot manipulators to detect defects developing in robot joint drives. The procedure was tested on a two-link prototype manipulator to identify Coulomb friction nonlinearity.

Several changes from conventional vibration diagnoses were investigated. The use of internal excitation is particular valuable in robot maintenance because internal actuators are available and can be controlled. While the FRF can be examined for changes in the characteristics of a system, the application of the HT to the FRF provides an additional tool to identify the changes more effectively. This study also showed that sensor locations are very important to detect variation in areas where nonlinear effects occur.

Controlled internal excitation, which is a closed loop approach, is a valuable tool to examine the characteristics of a system. Conventional vibration diagnoses only examine response from the system, which is an open loop approach, while modal analyses generally utilize external excitation. By using internal excitation, techniques from both fields can be combined to examine a system more effectively. This is especially important where external excitation is not available. Internal excitation also helps to identify problems in the system actuators. For robot maintenance, the control computer can be programmed to generate necessary excitation signals and to command the actuators to perform the diagnosis on a regular basis.

Variation of the FRF and HTF from baseline data was used to identify presence of nonlinearity in the system. In the friction tests conducted at both the shoulder joint and the elbow joint, small amount of externally applied friction clearly changed the ratio between the

HTF and FRF relative to baseline data. This was also evident in the RMS ratio and the peak difference span at location A for the shoulder friction test and at location C for the elbow friction test. For robot PMP, the HTF and FRF should be checked against the baseline data to record long term trends with respect to the robot performance.

The level of mismatch between the FRF and HTF indicates magnitude of friction nonlinearity. In the shoulder friction test, the peak difference span at location A increased when the applied friction increased relatively as shown in Figure 5.1. Although the RMS ratio actually decreased a little from moderate to high level friction as shown in Figure 5.2, the RMS values showed significant variation from baseline data when no additional friction was applied. Note that the ratio between HTF and FRF also showed higher positive-negative peaks when friction level increased as shown in Figure 4.20. In the elbow friction test, RMS ratio between FRF and HTF at location C decreased for low friction level but increased significantly when high friction levels were applied as shown in Table 4.8. This is due to the fact that moderate backlash exists in the gear drive of the elbow joint. Small amounts of additional friction reduced the backlash nonlinearity while large amounts of additional friction increased the overall nonlinear effect. For robot PMP, nonlinearity indicators such as RMS ratio can be used to check long term variations from baseline data. The magnitude of these indicators can then be related to the performance of the robot. The alert and alarm limits can be set accordingly to provide advance warning of impending component failure.

The location of nonlinear effects can be identified by checking FRF and HTF at strategic sensor locations along the robot structure. In the elbow friction test using shoulder excitation, both the RMS ratio and peak difference span at location A stayed at about the same level for higher friction. But at locations B and C, RMS ratio changed dramatically when different levels of friction were applied. This means that some nonlinear effect occurred between location B and C, i.e., the elbow joint drive. Although RMS ratio at location B and C in the shoulder friction test also showed increased value, relatively larger increase at location A suggested that nonlinear effect occurred at the shoulder joint. Signature shape can also be observed in the mismatch between HTF and FRF at location B. While positive-negative peaks occurs in the shoulder friction test as shown in Figure 4.22, generally only positive peaks occurs in the elbow friction test as shown in Figure
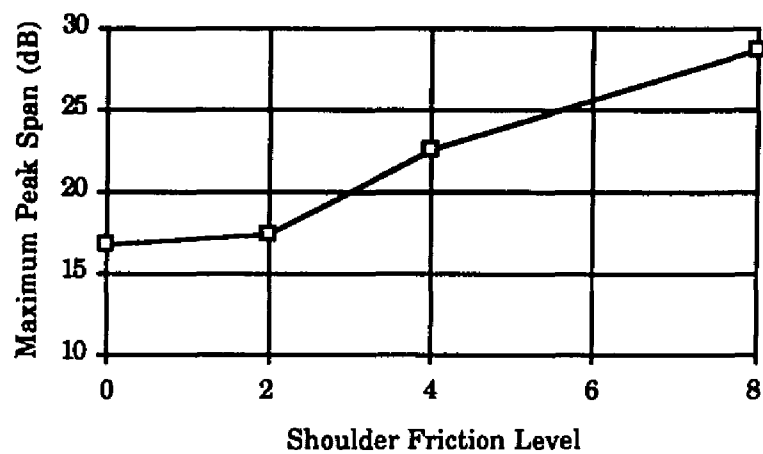
Figure 5.1 Maximum Peak Span at Location A for Shoulder Friction Test.
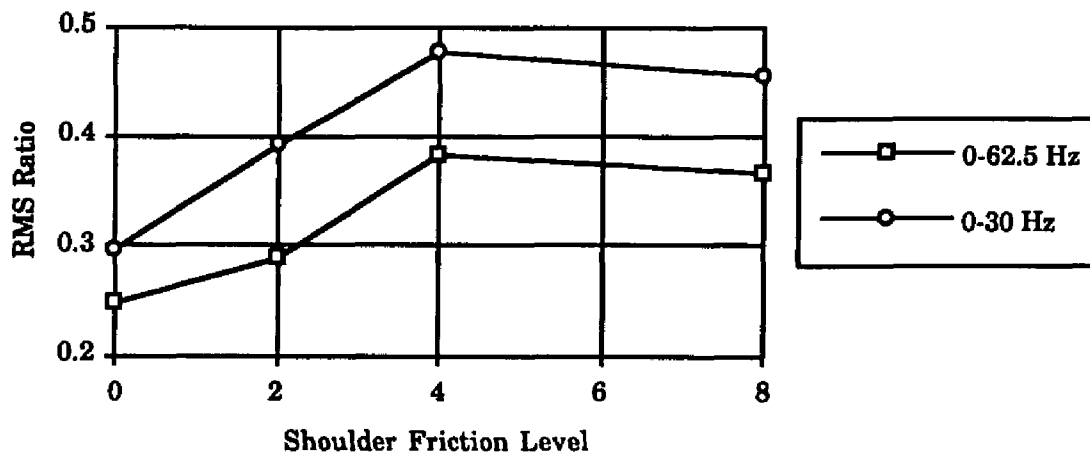

Figure 5.2 RMS Ratio at Location A for Shoulder Friction Test.

4.29. In robot PMP, if the location of significant nonlinear effects are identified, maintenance can be scheduled to repair or replace the degraded component.

Identification of nonlinearity is more effective when sensors are placed at strategic locations in a system. Two facts were observed. Firstly, the sensor closest to the nonlinear effect provides more accurate information on the magnitude of the effect. For instance in the shoulder friction test, data from location A was a better indicator of friction levels than data from location B. Secondly, identification is more effective when the excitation source, nonlinear effect, and response sensor are in sequence along the kinematic chain. In other words, nonlinear effects typically do not propagate proximally toward the fixed link. This effect can be seen in the elbow friction test using elbow excitation. While data from location C was a good indicator of friction levels, data from location B, which is more proximal relative to the elbow motor, cannot be used. The same effect was observed in the elbow friction test using shoulder excitation. Data at location C showed large variations, while small variation occurred at location B and minimum change existed at location A.

The HT was very effective in detecting Coulomb friction nonlinearity in the two-link flexible manipulator. The HTF and FRF analyses provided qualitative information about nonlinearity in the system. Detailed comparison between the HTF and the FRF can be performed to examine the effect of nonlinearity within certain frequency ranges. By using various indicators to extract quantitative measures from these analyses, nonlinear analysis became more objective. These data can then be used to establish the alert and alarm limits for the PMP. Note that some indicators may be more effective than others in identifying certain type of nonlinear effect or in certain frequency ranges. Care must be taken to prevent incorrect interpretation.

This thesis established techniques for the identification of nonlinearities in robot joint drive using internal robot actuator excitation. By monitoring the characteristic of robot joint drives on a regular basis, these techniques will enhance the predictive maintenance program of robot manipulator. Positioning accuracy of end effectors can be improved when degraded components in joint drives are identified and repaired.

5.2 Recommendations

When conducting tests and nonlinear analyses discussed in this thesis, several notes are suggested. In the HT procedure, convergence of the HT must be checked when the FRF

is converted from inertance to mobility. Note that in some cases the FRF might not need to be converted to obtain good results.

Although various nonlinear indicators were developed in simulation, only the RMS ratio was used in tests on the real manipulator. This is due to the flexible nature of the system and experimental noise in measured FRF compared to the simulated smooth FRF. Care must be taken when choosing nonlinear indicators for different systems and measuring conditions. The nominal excitation level for the test must be chosen so that the nonlinear effect can be observed effectively.

As described in Section 4.3.1, the sine sweep excitation signal can be programmed to sweep certain frequency ranges with certain durations. However, since the power density spectrum of the sine sweep signal is highly dependent on the sampling frequency, number of data blocks, number of samples per block and extent of data block overlap, combinations of these parameters must be checked before use.

In this investigation, accelerometers were the only sensor type used. It is possible to use other output sensors in the FRF measurement. Velocity feedback on the shoulder motor and the position encoders on both shoulder and elbow joints are possible candidates. Note that for real systems any sensors can be used in the identification procedure providing they are effective.

As more direct drive robots are used, identification of friction effects will become more important in robot performance. To perform the nonlinear identification procedure described in this thesis on a regular basis helps monitor the condition of robots, minimize the cost of maintenance, and guarantee the quality of manufactured products.

REFERENCES

[1]   "Robot Population Reports," *Robotica*, Vol. 1, Part 1, 1983, pp. 67-71.

[2]   "Robots Worldwide," *Robotica*, Vol. 9, Part 1, 1991, pp. 1-5.

[3]   Mobley, R. K., *An Introduction to Predictive Maintenance*, New York, NY: Van Nostrand Reinhold, 1990.

[4]   White, R. G., "Effects of Non-Linearity due to Large Deflections in the Resonance Testing of Structures," *Journal of Sound and Vibration*, Vol. 16, No. 2, 1971, pp. 255-267.

[5]   Natke, H. G., Juang, J.-N., and Gawronski, W., "A Brief Review on the Identification of Nonlinear Mechanical Systems," *Proceedings of the 6th International Modal Analysis Conference*, Vol. 2, 1988, pp. 1569-1574.

[6]   Tomlinson, G. R., "Developments in the Use of the Hilbert Transform for Detecting and Quantifying Non-Linearity Associated with Frequency Response Functions," *Mechanical Systems and Signal Processing*, Vol. 1, No. 2, 1987, pp. 151-171.

[7]   Gifford, S. J., and Tomlinson, G. R., "Understanding Multi Degree of Freedom Nonlinear Systems via Higher Order Frequency Response Functions," *Proceedings of the 7th International Modal Analysis Conference*, Vol. 2, 1989, pp. 1356-1364.

[8]   Park, Y.-S., and Kim, W.-J., "Nonlinearity Identification and Quantification Using Inverse Fourier Transform," *Proceedings of the 9th International Modal Analysis Conference*, Vol. 2, 1991, pp. 1018-1024.

[9]   Van der Auweraer, H., Vanherch, P., Sas, P., and Snoeys, R., "Accurate Modal Analysis Measurement with Programmed Sine Wave Excitation," *Mechanical Systems and Signal Processing*, Vol. 1, No. 3, 1987, pp. 301-313.

[10] Agneni, A., and Balis-Crema, L., "Damping Measurement from Truncated Signals via Hilbert Transform," *Mechanical Systems and Signal Processing*, Vol. 3, No. 1, 1989, pp. 1-13.

[11] Mohammad, K. S., and Tomlinson, G. R., "A Simple Method of Accurately Determining the Apparent Damping in Non-Linear Structures," *Proceedings of the 7th International Modal Analysis Conference*, Vol. 2, 1989, pp. 1336-1346.

[12] Dagalakis, N. G., and Myers, D. R., "Adjustment of Robot Joint Gear Backlash Using the Robot Joint Test Excitation Techniques," *The International Journal of Robotics Research*, Vol. 4, No. 2, 1985, pp. 65-79.

[13] Wyckaert, K. Vanherch, P., Sas, P., and Van Brussel, H., "The Identification of the Nonlinear Dynamic Behavior of a Flexible Robot Link," *Proceedings of the 8th International Modal Analysis Conference*, Vol. 1, 1990, pp. 530-539.

[14] Whalen, M. D., "Initial Experiments in End-Point Control of a Flexible, Two-Link Manipulator," Ph.D Dissertation, Department of Mechanical Engineering, Pennsylvania State University, 1991.

[15] Lembregts, F. Lipkens, J., Leuridan, J. and Van der Auweraer, H., "Comparison of Stepped-Sine and Broad Band Excitation to an Aircraft Frame," *Proceedings of the 6th International Modal Analysis Conference*, Vol. 2, 1988, pp. 1706-1709.

[16] Sachs, N. W., "Predictive Maintenance Cuts Costs," *Power Transmission Design*, Vol. 28, No. 11, 1986, pp. 45-48.

[17] Tranter, J. T., "The Application of Computers to Machinery Predictive Maintenance," *Sound and Vibration*, Vol. 24, No. 12, 1990, pp. 14-19.

[18] Maniscalco, M. A., "Microlog System Adds Accuracy to Predictive Maintenance Program," *Power Transmission design*, Vol. 29, No. 6, 1987, pp. 32-43.

[19] Berry, J. E., "Diagnostic Evaluation of Machinery Using Vibration Signature Analysis," *Sound and Vibration*, Vol. 20, No. 6, 1986, pp. 10-17.

[20] Chandler, J., "Meaningful Vibration Measurements for Predictive Maintenance," *Sound and Vibration*, Vol. 25, No. 11, 1991, pp. 18-22.

[21] Braun, S., *Mechanical Signature Analysis: Theory and Application*, Orlando, FL: Academic Press, 1986.

[22] Berry, J. E., "How to Specify Machinery Vibration Spectral Alarm Bands," *Sound and Vibration*, Vol. 24, No. 9, 1990, pp. 16-28.

[23] Mooring, B. W., "The Effect of Joint Axis Misalignment on Robot Positioning Accuracy," *Proceedings of the 1983 ASME International Conference on Computers in Engineering*, Vol. 2, 1983, pp. 151-155.

[24] Mooring, B. W., and Tang, G.-G., "An Improved Method for Identifying the Kinematic Parameters in a Six Axis Robot," *Proceedings of the 1984 ASME International Conference on Computers in Engineering*, Vol. 1, 1984, pp. 79-84.

[25] Whitney, D. E., Lozinski, C. A., and Rourke, J. M., "Industrial Robot Forward Calibration Method and Results," *Journal of Dynamic Systems, Measurement, and Control*, Vol. 108, No. 1, 1986, pp. 1-8.

[26] Bosnik, J. R., "Static and Vibrational Kinematic Parameter Estimation for Calibration of Robotic Manipulators," Ph.D Dissertation, Department of Mechanical Engineering, Pennsylvania State University, 1986.

[27] Dagalakis, N. G., and Myers, D. R., "Self-Adjustment of a Robot Joint Controller," *Proceedings of the 1986 IEEE International Conference on Industrial Electronics, Control, and Instrumentation*, Vol. 2, 1986, pp. 530-537.

[28] Dagalakis, N. G., and Myers, D. R., "Estimation of the Dynamic Parameters of a Robot Joint Drive System," *Proceedings of the 1986 IEEE International Conference on System, Man, and Cybernetics*, Vol. 1, 1986, pp. 655-659.

[29] Asada, H., and Slotine, J.-J. E., *Robot Analysis and Control*, New York, NY: John Wiley and Sons, 1986.

[30] Shamma, J. S., and Whitney, D. E., "A Method for Inverse Robot Calibration," *Journal of Dynamic Systems, Measurement, and Control*, Vol. 109, No. 1, 1987, pp. 36-43.

[31] Vira, N., "Robot's End Point Sensing: Hardware and Software Techniques," *Computers in Industry*, Vol. 12, No. 1, 1989, pp. 1-12.

[32] Nowrouzi, A., Kavina, Y. B., Kochekali, H., and Whitaker, R. A., "An Overview of Robot Calibration Techniques," *The Industrial Robot*, Vol. 15, No. 4, 1988, pp. 229-232.

[33] McEntire, R. H., "Three Dimension Accuracy Measurement Methods for Robots," *The Industrial Robot*, Vol. 3, No. 3, 1976, pp. 105-112.

[34] Gilby, J. H., and Parker, G. A., "Laser Tracking System to Measure Robot Arm Performance," *Sensor Review*, Vol. 2, No. 4, 1982, pp. 180-184.

[35] Fohanno, T., "Assessment of the Mechanical Performance of Industrial Robots," *Proceedings of the 12th International Symposium on Industrial Robots and 6th International Conference on Industrial Robot Technology*, 1982, pp. 349-358.

[36] Vira, N., and Lau, K., "An Extensible Ball Bar for Evaluation of Robot's Positioning Performance," *Journal of Robotic Systems*, Vol. 4, No. 6, 1987, pp. 799-814.

[37] Podoloff, R. M., and Seering W. P., "An Accuracy Test Procedure for Robotic Manipulators Utilizing a Vision Based, 3-D Position Sensing System," *Proceedings of the 1984 American Control Conference*, Vol. 1, 1984, pp. 19-22.

[38] Dainias, A., and Juberts, M. "Accuracy Remote Measurement of Robot Trajectory Motion," *Proceedings of the 1985 IEEE International Conference on Robotics and Automation*, 1985, pp. 92-99.

[39] Arimoto, S., and Miyazaki, F., "Stability and robustness of PID Feedback Control for Robot Manipulators of Sensory Capability," *Proceedings of the 1st International Symposium on Robotics Research*, 1984, pp. 783-799.

[40] Walker, M. W., and Orin, D. E., "Efficient Dynamic Computer simulation of Robotic Mechanisms," *Journal of Dynamic Systems, Measurement, and Control*, Vol. 104, No. 3, 1982, pp. 205-211.

[41] Slotine, J.-J. E., and Li, W., *Applied Nonlinear Control*, Englewood Cliffs, NJ: Prentice-Hall, 1991.

[42] Mukerjee, A., and Ballard, D. H., "Self-Calibration in Robot Manipulators," *Proceedings of the 1985 IEEE International Conference on Robotics and Automation*, 1985, pp. 1050-1057.

[43] Olsen, H. B., and Bekey, G. A., "Identification of Robot Dynamics," *Proceedings of the 1986 IEEE International Conference on Robotics and Automation*, Vol. 2, 1986, pp. 1004-1010.

[44] Slotine, J.-J. E., "The Robust Control of Robot Manipulators," *The International Journal of Robotics Research*, Vol. 4, No. 2, 1985, pp. 49-64.

[45] Slotine, J.-J. E., and Li, W., "On the Adaptive Control of Robot Manipulators," *The International Journal of Robotics Research*, Vol. 6, No. 3, 1987, pp. 49-59.

[46] Hsia, T. C., "Adaptive Control of Robot Manipulators - A Review," *Proceedings of the 1986 IEEE International Conference on Robotics and Automation*, Vol. 1, 1986, pp. 183-189.

[47] Craig, J. J., *Introduction to Robotics: Mechanics and Control*, Reading, MA: Addison-Wesley, 1986.

[48] Kubo, T., Anwar, G., and Tomizuka, M., "Application of Nonlinear Friction Compensation to Robot Arm Control," *Proceedings of the 1986 IEEE International Conference on Robotics and Automation*, Vol. 2, 1986, pp. 722-727.

[49] Dagalakis, N. G., and Myers, D. R., "Adjustment of Robot Joint gears Using Encoder Velocity and Position Information," *Journal of Robotic Systems*, Vol. 2, No. 2, 1985, pp. 229-234.

[50] Dagalakis, N. G., and Myers, D. R., "Use of Coherence Analysis for the Evaluation and Adjustment of Robot Gear Performance," *Robotics and Computer Integrated Manufacturing*, Vol. 2, No. 3/4, 1985, pp. 237-245.

[51] Bae, I., "On-Line Estimation of Mechanical Nonlinearities in Compliant Two-Link Robot Manipulators," Ph.D Dissertation, Department of Mechanical Engineering, Pennsylvania State University, 1989.

[52] Canudas de Wit, C., Åstrom, K. J., and Braun, K., "Adaptive Friction Compensation in DC-Motor Drives," *Journal of Robotics and Automation*, Vol. RA-3, No. 6, 1987, pp. 681-685.

[53] Gilbart, J. W., and Winston, G. C., "Adaptive Compensation for An Optical Tracking Telescope," *Automatica*, Vol. 10, No. 2, 1974, pp. 125-131.

[54] Walrath, C. D., "Adaptive Bearing Friction Compensation Based on Recent Knowledge of Dynamic Friction," *Automatica*, Vol. 20, No. 6, 1984, pp. 717-727.

[55] Armstrong, B. "Friction: Experimental Determination, Modeling and Compensation," *Proceedings of the 1988 IEEE International Conference on Robotics and Automation*, Vol. 3, 1988, pp. 1422-1427.

[56] Canudas de Wit, C., Noel, P., Aubin, A., and Brogliato, B., "Adaptive Friction Compensation in Robot Manipulators: Low Velocity," *The International Journal of Robotics Research*, Vol. 10, No. 3, 1991, pp. 189-199.

[57] Bracewell, R. N., *The Fourier Transform and Its Applications*, 2nd Ed. Revised, New York, NY: McGraw-Hill, 1986.

[58] Bose, N. K., *Digital Filters: Theory and Applications*, New York, NY: North-Holland, 1985.

[59] Worden, K., and Tomlinson, G. R., "The High-Frequency Behavior of Frequency Response Functions and its Effect on their Hilbert Transforms," *Proceedings of the 8th International Modal Analysis Conference*, Vol. 1, 1990, pp. 121-130.

[60] Brigham, E. O., *The Fast Fourier Transform and Its Applications*, Englewood Cliffs, NJ: Prentice-Hall, 1988.

[61] Hagedorm, P., *Non-Linear Oscillations*, 2nd Ed., New York, NY: Oxford University Press, 1988.

[62] Lim, J. S., Oppenheim, A. V., *Advanced Topics in Signal Processing*, Englewood Cliffs, NJ: Prentice-Hall, 1988.

[63] Ewins, D. J., *Modal Testing: Theory and Practice*, Taunton, Somerset, England: Research Studies Press, 1989.

[64] *Waveform Equations: LeCroy LW400 Series Arbitrary Waveform Generator*, Chestnut Ridge, NY: LeCroy Corporation, 1994.

APPENDIX A

This MATLAB program calculates the Hilbert Transform of a Inertance type FRF from the Wavepak data using (1) all spectrum data lines in Wavepak output file, (2) partial null buffer.

```
% Variables
% Frange : Analysis frequency range
%      N : # of spectrum lines
%     TN : Total # of spectrum lines used in calculation
%     df : Frequency resolution
%      g : Original FRF
%      h : Hilbert transform of FRF

% Assign parameters
Frange = input('Enter the Analysis Frequency Range (Hz)[40]: ');
if isempty(Frange), Frange = 40; end;
N = input('Enter number of spectrum lines [400]: ');
if isempty(N), N = 400; end;

if N == 200,
   TN = 256;
elseif N == 400,
   TN = 512;
elseif N == 800,
   TN = 1024;
end
df = Frange / N;
w = 0:df:df*TN;

% Calculate the Hilbert Transform in Form of
% 1: Receptance
% 2: Mobility
% 3: Inertance

choice = input('Choice [2]: ');
if isempty(choice) choice = 2; end;
if choice <= 2,
   g = g./(sqrt(-1)*2*pi*w(1,1:N));    % Transform to Mobility
   g(1,1) = 0;
   if choice == 1,
      g = g./(sqrt(-1)*2*pi*w);         % Transform to Receptance
      g(1,1) = g(1,2);
   end;
end;

% Hilbert Transform using null buffer
g1 = conj(g(:,N:-1:2));
gg = [g,zeros(1,TN-N),0,zeros(1,TN-N),g1];
gg = ifft(gg);
```

```
% Generate Sgn(w) using F[sgn(t)]
sgn = -sqrt(-1) / pi ./ w(:,2:TN);
ss = [0., sgn, 0, -sgn(:,TN-1:-1:1)];

% Hilbert transform convolution
ss = ifft(ss);
hh = gg.*ss;
hh = fft(hh) * 2 * TN * df;
ss = fft(ss);
gg = fft(gg);

% Transform back to inertance
w = w(1,1:N);
g = gg(1,1:N);
h = hh(1,1:N);
if choice <= 2,
  g = g.*(sqrt(-1)*2*pi*w);        % From Mobility
  h = h.*(sqrt(-1)*2*pi*w);
  if choice == 1,
    g = g.*(sqrt(-1)*2*pi*w);      % From Receptance
    h = h.*(sqrt(-1)*2*pi*w);
  end;
end;

% Program End
```

APPENDIX B

This program takes sample data from the sweep excitation tests to calculate the FRF using MATLAB function spectrum(). Note that the samples must be transferred to the data[] array. FRF can be calculated for various excitation source and response sensors. The Hilbert transform is then performed on the FRF.

```
% Variables
%        N:   # of data points per block
%        Fs:  sampling frequency in Hz
%        df:  frequency resolution
% FRFtype:    type of FRF
%        w    frequency (Hz)
%        g:   original FRF
%        h:   Hilbert transformed FRF


% Read sampling frequency from keyboard
Fs = input('Enter sampling frequency (Hz) = [125] ');
if isempty(Fs) Fs = 125; end

[TotalN,i] = size(data);

% Set up FFT & FRF parameters
flag = 1;
while (flag == 1)
  NP = input('Enter # of data points discarded at the beginning = [20] ');
  if isempty(NP) NP = 20; end
  NB = input('Enter # of data blocks for averaging = [10] ');
  if isempty(NB) NB = 10; end
  N = input('Enter # of samples per block = [512] ');
  if isempty(N) N = 512; end
  choice = input('Overlapping (Y/N) = [Y] ');
  if isempty(choice) choice = 'Y'; end
  if all(choice == 'Y' | choice == 'y')
    Nperc = input('Enter percentage of overlapping (%) = [20] ');
    if isempty(Nperc) Nperc = 20; end
    Nperc = fix(N * Nperc / 100);
  else
    Nperc = 0;
  end
  if ( (N+(NB-1)*(N-Nperc)) <= (TotalN-NP) ) flag = 0; end
end
Idata = NP+1:1:NP+N+(NB-1)*(N-Nperc);

% Separate variables
link = menu('Enter FRF Source','ACCs/Shoulder Excitation',...
        'ACCs/Elbow Excitation','ACC1/ACC2: Shoulder Excitation');
if link == 1
  sout = data(Idata,4);
elseif link == 2
  eout = data(Idata,5);
else
```

```
    sout = data(Idata,4);
    eout = data(Idata,5);
  end
acc1 = data(Idata,2);
acc2 = data(Idata,3);

% Convert data counts to real volts
offset = -11;
acc1 = 5 * (acc1 - offset) / 2048;        % Use appropriate A/D gain
acc2 = 5 * (acc2 - offset) / 2048;
if link == 1
    sout = 10 * (sout - 2048) / 2048;
elseif link == 2
    eout =  -5 * (eout - 2048) / 2048;
else
    sout = -10 * (sout - 2048) / 2048;
    eout = -5 * (eout - 2048) / 2048;
end

% Spectrum analysis
if link == 1
    p1 = spectrum(sout, acc1, N, Nperc);
    p2 = spectrum(sout, acc2, N, Nperc);
    p3 = p2 ./ p1;
elseif link == 2
    p1 = spectrum(eout, acc1, N, Nperc);
    p2 = spectrum(eout, acc2, N, Nperc);
    p3= p2 ./ p1;
else
    p1 = spectrum(sout, acc1, N, Nperc);
    p2 = spectrum(sout, acc2, N, Nperc);
    p3 = spectrum(acc2, acc1, N, Nperc);
end

df = Fs / N;          % Frequency resolution
index = 1 : N/2;
choice1 = menu('Enter FRF choice','ACC1/OUT','ACC2/OUT','ACC1/ACC2');
% Transfer function
if choice1 == 1,
    g = p1(index,4);
elseif choice1 == 2,
    g = p2(index,4);
else
    g = p3(index,4);
end
g = g(:);          % Make sure g is column vector

% Null buffer selection
choice = input('Do you want to use null buffer? Y/N [Y]:','s');
if isempty(choice) choice = 'Y'; end
if choice == 'Y'
  TN = N;
```

```
else
  TN = N / 2;
end
w = (0 : TN) * df;      % w in Hz
w = w(:);

% Convert FRF g
FRFtype = menu('Enter FRF type','g = g','g = g / w','g = g / w / w');
if FRFtype == 1
   g0 = g;
elseif FRFtype == 2
   g0 = g ./ (sqrt(-1) * 2 * pi * w(index));
   g0(1) = g0(2);
elseif FRFtype == 3
   g0 = g ./ (sqrt(-1) * 2 * pi * w(index));
   g0(1) = 0.;
   g0 = g0 ./ (sqrt(-1) * 2 * pi * w(index));
   g0(1) = g0(2);
end

% Generate G(w) and sgn(w)
g1 = conj(g0(N/2:-1:2));
g0(1) = real(g0(1));
if choice == 'Y'
   addzeros = zeros(index);
   addzeros = addzeros(:);
   gg = [g0; addzeros; 0; addzeros; g1];
else
   gg = [g0; real(g0(N/2))/2; g1];
end

sgn = -sqrt(-1)/pi ./w(2:TN);
sgn = sgn(:);
ss = [0; sgn; 0; -sgn(TN-1:-1:1)];

% Hilbert transform
gg = ifft(gg);
ss = real(ifft(ss));
hh = gg .* ss;
hh = fft(hh) * 2 * TN * df;

h = hh(index);
w = w(index);      % w in Hz
if FRFtype == 2
   h = h .* (sqrt(-1) * w * 2 * pi);
elseif FRFtype == 3
   h = h .* (sqrt(-1) * w * 2 * pi);
   h = h .* (sqrt(-1) * w * 2 * pi);
end

%Program End
```

VITA

Chi-Ching Kao was born on November 12, 1963 to Jin-Cherng Kao and Ming-Shu Lee in Taipei, Taiwan, the Republic of China. After graduating from the High School of the National Taiwan Normal University in Taipei, he was admitted to the Mechanical Engineering Department of the National Taiwan University in 1981. He earned a Bachelor of Science degree in 1985.

In 1987, He was admitted to the Mechanical Engineering Department of the Pennsylvania State University, where he earned a Master of Science degree in 1989. The master thesis topic was "Image Processing for Turbulent Flame Morphology." Later, he was admitted to the Ph.D program. His research involves identification of nonlinearities for robot manipulators.

In addition to his research activity, He served as a teaching assistant to Machine Element Design, Industrial Robotics and Microcomputer Interface Application. His fields of interest include System Dynamics, Control, Robotics, Digital Signal Processing, and Computer Vision.