IMPERIAL COLLEGE OF SCIENCE, TECHNOLOGY AND MEDICINE

University of London

IDENTIFICATION OF THE DYNAMIC CHARACTERISTICS OF NONLINEAR STRUCTURES

by

Rongming Lin

A thesis submitted to the University of London for
the degree of Doctor of Philosophy and for
the Diploma of Imperial College.

Dynamics Section

Department of Mechanical Engineering

Imperial College of Science, Technology and Medicine

London SW7

December 1990

ABSTRACT

Modal analysis has been extensively developed during the last two decades and has become one of the most effective means of identifying the dynamic characteristics of engineering structures. However, most of the techniques developed so far are based on the assumption that the structures to be identified are linear while, in practice, most engineering structures are *nonlinear*. It is therefore necessary to extend existing linear modal analysis techniques or develop new techniques so that structural nonlinearity can be *detected*, *quantified* and *mathematically modelled* based on the measured input-output dynamic characteristics. This thesis seeks to present complete yet new developments on the identification of dynamic characteristics of nonlinear structures.

For nonlinear structures whose modal parameters for certain modes are displacement dependent (the nonlinearity is of symmetrical type), a new nonlinear modal analysis method based on the measured first-order frequency response functions is developed. The method has been effectively applied to the data measured from practical nonlinear structures even when the modes become considerably complex. On the other hand, for structures whose nonlinearities are such that the measured first-order frequency response functions are effectively linear (nonlinearity of nonsymmetrical type), a higher-order frequency response function analysis is presented which provides opportunities for the identification of such nonlinear structures. Both the first- and higher-order frequency response function analyses are based on the classical assumption that the output of a nonlinear structure is periodic if the input is periodic. However, for some nonlinear systems (chaotic systems), this assumption is no longer valid and special techniques need to be developed in order to identify them. In this thesis, for the first time, the hidden chaotic behaviour of a mechanical backlash system with realistic system parameters has been revealed and, based on this system, qualitative as well as quantitative ways of identifying chaotic systems are presented. Both numerical studies and experimental investigations are carried out and possible engineering applications are discussed.

It is believed that nonlinearities of most engineering structures are usually localised in just a few spatial coordinates and the ability to locate these has some important engineering applications. In this thesis, location techniques based on the correlation between analytical model and measured modal parameters as well as frequency response data are developed.

Numerical study and experimental investigation demonstrate the practical applicability of these techniques. Because of the limitation of measured data available, it is essential to pinpoint where the structural nonlinearity is located before a nonlinear mathematical model can be established.

The ultimate target of the analysis of a nonlinear structure is to establish a nonlinear mathematical model (spatial model) which is a function of response amplitude. It is believed that such a target can only be achieved by combining analytical modelling (FE modelling) and modal testing techniques. In this thesis, new model updating methods are developed and extended to the mathematical modelling of nonlinear structures based on the correlation between analytical and measured modal parameters as well as frequency response data. The practical applicability of these methods is assessed based on a specific case study. Criteria on minimum data required in order to update an analytical model are established and the possibilities and limitations of analytical model improvement are discussed which make it possible for the analyst to judge whether a set of measured data will be sufficient to solve the model updating problem uniquely.

ACKNOWLEDGMENTS

The author is most grateful to his supervisor, Prof. D. J. Ewins, for his sustained encouragement, interest, stimulus and guidance given over the period in which the research work presented herein has been carried out.

Thanks are also due to the members of the Dynamics Section, both past and present, especially to Dr. J. He and Mr. D. A. Robb for their valuable help given in computing and conducting experiments.

For the friendly cooperation and useful discussions throughout the period of research, the author wishes to express his gratitude to many past and present colleagues in the Dynamics Section, in particular to Waiming To, Tom Pialucha and Ali Nobari.

Finally, the author is indebted to the British Council and the Chinese Government for providing the financial support.

NOMENCLATURE

a	constant (real)
A_r, B_r	real and imaginary parts of modal constant of r^{th} mode (real)
$A_r(\hat{x})$, $B_r(\hat{x})$	real and imaginary parts of modal constant of r^{th} mode which are functions of vibration amplitude (real)
[A]	system generalised mass matrix (real)
$_rA_{jk}$	modal constant of r^{th} mode of $\alpha_{jk}(\omega)$ (complex)
b	constant (real)
{b}	coefficient vector (real)
[B]	system generalised stiffness matrix (real)
c	constant (real)
c_r	eigenvector scaling factor (complex)
[C]	viscous damping matrix (real)
[D]	hysteretic damping matrix
$Df(\mathbf{x}^e)$	Jacobian matrix of vector field $f(\mathbf{x})$ at $\mathbf{x}^e$ (real)
$DP(\mathbf{x}^*)$	Jacobian matrix of map P at $\mathbf{x}^*$ (real)
D_c	capacity dimension of an attractor (real)
E	Young's modulus (real)
$\{e_i\}$	a vector with its i^{th} element unity and all the others zeros (real)
$[E(\omega)]$	coefficient matrix (real)
f(t)	force signal (real)
$F(\omega)$	Fourier transform of force signal f(t) (complex)
$F(x,\dot{x})$	restoring force (real)
$f(x)$	general nonlinear function (real)
$G_n[k_n, f(t)]$	n^{th} Wiener G-functional (real)

$[G(\omega)]$	coefficient matrix (real)
$H(\omega)$	frequency response function of linear systems (complex)
$h_n(\tau_1,\tau_2,...,\tau_n)$	n^{th}-order Volterra kernel (real)
$H_n(\omega_1,...,\omega_n)$	Volterra kernel transform (Volterra transfer function) (complex)
$H_n^{ij}(\omega_1,...,\omega_n)$	Volterra kernel transform corresponding to excitation at coordinate x_j and response at coordinate x_i (complex)
$H_n^R(\omega_1,...,\omega_n)$	receptance-like Volterra kernel transform (complex)
$H_n^M(\omega_1,...,\omega_n)$	mobility-like Volterra kernel transform (complex)
$H_1(\omega)$	first-order frequency response function of nonlinear systems (complex)
$H_n(\omega_1,...,\omega_n)$	n^{th}-order frequency response function defined as the measured $H_n(\omega_1,...,\omega_n)$ or $K_n(\omega_1,...,\omega_n)$ (complex)
$\mathbf{i}$	complex notation
I	current (real)
$[I]$	unit matrix (real)
k	constant (real or integer)
$k_n(\tau_1,\tau_2,...,\tau_n)$	n^{th}-order Wiener kernel (real)
$k_{n-2k(n)}(\tau_1,...,\tau_n)$	derived Wiener kernel which is of n^{th}-order (real)
$[k_r^e]$, $[k_r^b]$	expanded r^{th} extensional and bending element stiffness matrices (real)
$K_n(\omega_1,...,\omega_n)$	Wiener kernel transform (complex)
$[K]$	system stiffness matrix (real)
$[K_a]$	analytical stiffness matrix (real)
$[\Delta K]$	stiffness error matrix (real)
$[\Delta K_n]$	stiffness error matrix due to nonlinearity (real)
L	number of independent design variables (integer)
m	number of measured or calculated modes (integer)
$[m_r]$, $[k_r]$	expanded r^{th} element mass and stiffness matrices (real)

$[M]$	system mass matrix (real)
$[M_a]$	analytical mass matrix (real)
$[\Delta M]$	mass error matrix (real)
$[M_{mm}]$, $[K_{mm}]$	mass and stiffness matrices corresponding to master (measured) DOFs (real)
$[M_{ms}]$, $[K_{ms}]$	mass and stiffness matrices corresponding to master (measured) and slave (unmeasured) DOFs (real)
$[M_{ss}]$, $[K_{ss}]$	mass and stiffness matrices corresponding to slave DOFs (real)
$[M^R]$, $[K^R]$	dynamically or statically condensed mass and stiffness matrices (real)
n	number of measured coordinates (integer)
N	number of degrees of freedom specified in the analytical model (integer)
$N(\hat{x})$	describing function coefficient (complex)
$N_1(\hat{x})$, $N_2(\hat{x})$	real and imaginary part of $N(\hat{x})$ (real)
$[O]$, $[0]$	null matrices (real)
$\{p\}$, $\{\ddot{p}\}$	vector of displacement and acceleration of principal coordinates (real)
P	mapping function (real)
$P(\varepsilon)$	probability function (real)
$\{P\}$	vector of design variable change (real)
$P(\omega)$, $Q(\omega)$	polynomial function of ω (real)
$R(\omega)$, $I(\omega)$	real part and imaginary part of $\alpha_{jk}(\omega)$ (real)
ΔR, ΔI	real and imaginary part of the residue term of r^{th} mode
$\{R(\omega)\}$	coefficient vector (real)
$[R]$	residual matrix representing the contribution of unmeasured or uncalculated higher modes (complex)
s_i	first-order eigensensitivity coefficient (real)

$S_{ff}(\omega)$	auto-spectrum of force signal (real)
$S_{fx}(\omega)$	cross-spectrum of force and response signals (complex)
$S_{xx}(\omega)$	auto-spectrum of response signal (real)
[S]	eigensensitivity coefficient matrix (real)
t	time parameter (real)
$T_n[h_n, f(t)]$	n^{th} Volterra functional (real)
{T}	coefficient vector (real)
[U]	left singular vector matrix (real or complex)
$[U(\omega)]$	coefficient matrix (complex)
V	voltage (real)
$\{V(\omega)\}$	coefficient vector (complex)
[V]	right singular vector matrix (real or complex)
[V]	Vandermonde matrix (real)
x(t)	response signal (real)
$\mathbf{x}^e$	equilibrium point (real)
$\mathbf{x}^*$	fixed point of map P (real)
$X(\omega)$	Fourier transform of response signal x(t) (complex)
$\hat{x}$	vibration amplitude of certain coordinate (real)
$\{x\}$, $\{\ddot{x}\}$	vector of displacement and acceleration (real)
{z}	vector of displacement in state-space (real)
$[Z(\omega)]$	system impedance matrix (complex)
$[Z_a(\omega)]$	analytical impedance matrix (complex)
$[Z_x(\omega)]$	experimental impedance matrix (complex)
$[Z_{mm}(\omega)]$	impedance matrix corresponding to master (measured) DOFs (complex)
$[Z_{ms}(\omega)]$	impedance matrix corresponding to master (measured) and slave (unmeasured) DOFs (complex)
$[Z_{ss}(\omega)]$	impedance matrix corresponding to slave DOFs (complex)

$[\Delta Z(\omega)]$	impedance error matrix (complex)
$\alpha_{jk}(\omega)$, $[\alpha(\omega)]$	receptance element and matrix (complex)
$\{\alpha_a(\omega)\}$,$[\alpha_a(\omega)]$	analytical receptance (complex)
$\{\alpha_x(\omega)\}$,$[\alpha_x(\omega)]$	experimental receptance (complex)
$[\Delta\alpha(\omega)]$	receptance error matrix defined as $[\alpha_x(\omega)]$ - $[\alpha_a(\omega)]$ (complex)
$[\alpha_c(\omega)]$	receptance matrix corresponding to the contribution of calculated or measured modes (complex)
$[\alpha_{au}(\omega)]$	receptance matrix of updated analytical model (complex)
$\{\alpha_m(\omega)\}$	receptance corresponding to master (measured) DOFs (complex)
$\{\alpha_s(\omega)\}$	receptance corresponding to slave (unmeasured) DOFs (complex)
$[\alpha_{mm}(\omega)]$	receptance matrix corresponding to master (measured) DOFs (complex)
$[\alpha_{ms}(\omega)]$	receptance matrix corresponding to master (measured) and slave (unmeasured) DOFs (complex)
$[\alpha_{ss}(\omega)]$	receptance matrix corresponding to slave DOFs (complex)
β	constant (real)
${}_r\phi_i$	i^{th} element of r^{th} modeshape $\{\phi\}_r$ (real)
$\{\phi\}_r$, $[\phi]$	undamped r^{th} modeshape and modeshape matrix (real)
$\phi_t(\mathbf{x})$	flow which starts at $\mathbf{x}$ (real)
$\{\phi_m\}$	modeshape corresponding to master DOFs (complex)
$\{\phi_s\}$	modeshape corresponding to slave DOFs (complex)
$[\phi_a]$, $[\phi_x]$	analytical (real) and experimental (complex) modeshape matrices
γ	constant (real)
η_r	damping coefficient of r^{th} mode (real)
$\eta_r(\hat{x})$	respone-amplitude-dependent damping coefficient of r^{th} mode (real)
$\Delta\eta_r$	error on damping coefficient estimate η_r (real)

$\{\varphi\}_r$, $[\varphi]$	hysteretically damped r^{th} modeshape and modeshape matrix (complex)
$[\cdot\lambda.]$	eigenvalue matrix (complex)
$[\cdot\lambda_a.]$, $[\cdot\lambda_x.]$	analytical (real) and experimental (complex) eigenvalue matrices
λ_1, $\{\phi_1\}$	eigenvalue and eigenvector of lower response level (complex)
λ_2, $\{\phi_2\}$	eigenvalue and eigenvector of higher response level (complex)
$\frac{\partial\lambda}{\partial p}$, $\frac{\partial}{\partial p}\{\phi\}$	eigenvalue and eigenvector derivatives (real)
$[\cdot\mu.]$	coefficient matrix (complex)
ρ, $\rho(x)$	density and probability density function (real)
σ_i	i^{th} singular value of a matrix (real)
$\{\sigma\}$	constant coefficient vector (real)
$[\Sigma]$	singular value matrix (real)
τ	time parameter (real)
ω	circular frequency (real)
ω_r	natural frequency of r^{th} mode (real)
$\omega_r(\hat{x})$	respone-amplitude-dependent natural frequency of r^{th} mode (real)
$\Delta\omega_r$	error on the estimate of ω_r (real)
ζ_r	damping loss factor (real)
$\{\psi\}_r$, $[\psi]$	viscously damped r^{th} modeshape and modeshape matrix (complex)

operators

$\|A\|$	Euclidean norm of matrix [A]
$\|A\|_F$	Frobenius norm of matrix [A]
D_α^n	partial differential operator
det (A)	determinant of [A]
exp[x]	natural exponential function

$F(d/dt)$	function of differential operator
$\mathfrak{H}[\,]$	Hilbert transform
mod	modulo function
$n!$	factor
$\prod$	product
$\mathrm{Re}(z)$, $\mathrm{Im}(z)$	real part and imaginary part of complex number z
$\sum$	summation
$\{\,\}^T$, $[\,]^T$	transpose
$[\,]^{-1}$	standard inverse
$[\,]^{-T}$	inverse and transpose of a matrix
$[\,]^H$	Hermitian transpose (complex conjugate transpose)
$[\,]^*$	complex conjugate

- CONTENTS -

ABSTRACT ii

ACKNOWLEDGEMENTS iv

NOMENCLATURE v

Chapter 1 INTRODUCTION

1.1 GENERAL INTRODUCTION 1

1.2 STRUCTURAL NONLINEARITY 3

1.3 IDENTIFICATION OF DYNAMIC CHARACTERISTICS OF NONLINEAR STRUCTURES 4

1.3.1 Frequency Domain Techniques 4

1.3.2 Time Domain Analysis 6

1.3.3 Amplitude Domain Analysis 7

1.4 MODELLING OF NONLINEAR STRUCTURES 8

1.5 PREVIEW OF THE THESIS 9

Chapter 2 IDENTIFICATION OF NONLINEARITY USING FIRST-ORDER FREQUENCY RESPONSE FUNCTIONS

2.1 GENERAL INTRODUCTION 11

2.2 MEASUREMENT OF FIRST-ORDER FREQUENCY RESPONSE FUNCTIONS 14

2.2.1 Sinusoidal Excitation Technique 15

2.2.2 Measurement Using Random Excitation 16

2.2.3 Measurement Using Transient Excitation 18

2.2.4 Comments and Practical Considerations of Nonlinearity Measurement 19

2.3 RECENT DEVELOPMENTS ON THE IDENTIFICATION OF NONLINEARITY BASED ON FIRST-ORDER FREQUENCY RESPONSE FUNCTIONS 22
2.3.1 Detection of Nonlinearity by Directly Using FRF Data 23
2.3.2 Isometric Damping Plot Technique 26
2.3.3 Inverse Receptance Method 29
2.3.4 Comments on Current Nonlinear Modal Analysis Methods 32
2.4 COMPLEXITY OF VIBRATION MODES 34
2.4.1 Theoretical Basis 34
2.4.2 Numerical Example of Complex Modes 39
2.4.3 Complex Modes from Measurements 41
2.5 A NEW METHOD OF NONLINEAR MODAL ANALYSIS OF COMPLEX MODES 42
2.5.1 Harmonic Balance Theory 42
2.5.2 Description of the Method 47
2.5.3 Application of the New Method to the Analysis of Nonlinear Systems Simulated Using Analogue Circuits 48
2.5.4 Extension of the Method to Nonlinear MDOF Systems 54
2.5.5 Application of the Method to Practical Nonlinear Structures 58
2.5.6 Identification of Nonlinear Physical characteristics 62
2.6 CONCLUSIONS 63

Chapter 3 IDENTIFICATION OF NONLINEARITY USING HIGHER-ORDER FREQUENCY RESPONSE FUNCTIONS

3.1 INTRODUCTION 66
3.2 THE VOLTERRA SERIES REPRESENTATION OF NONLINEAR SYSTEMS 70
3.2.1 Basic Characteristics of Nonlinear Systems 70
3.2.2 The Volterra Series Representation 72
3.3 HIGHER-ORDER FREQUENCY RESPONSE FUNCTIONS 73
3.3.1 Theory of Higher-Order Frequency Response Functions 73
3.3.2 Analytical Calculation of Higher-Order Frequency Response Functions ... 78
3.3.3 Measurement of Higher-Order Frequency Response Functions Using Harmonic Probing Method 87

3.4 CORRELATION ANALYSIS USING RANDOM INPUT 92
3.4.1 Theory of Wiener Series 92
3.4.2 Determination of Wiener Kernels by Cross-correlation 94
3.5 MEASUREMENT OF WIENER KERNELS 97
3.5.1 Measurement of Wiener Kernels Using Correlation Analysis with Random Input 97
3.5.2 Removal of Linear Contributions 101
3.6 IDENTIFICATION OF NONLINEARITY USING HIGHER-ORDER FREQUENCY RESPONSE FUNCTIONS 102
3.7 CONCLUSIONS 104

Chapter 4 IDENTIFICATION OF CHAOTIC VIBRATIONAL SYSTEMS

4.1 PRELIMINARIES 106
4.2 INTRODUCTION TO CHAOTIC VIBRATION 108
4.2.1 Dynamic System Theory 109
4.2.2 Steady-state Behaviour of Dynamic Systems 112
4.3.3 Chaotic Attractors 116
4.3.4 The Poincaré Map 118
4.3.5 Stability of Limit Sets and Lyapunov Exponents 121
4.2.6 The Dimension of an Attractor 124
4.3 CHAOTIC VIBRATION OF MECHANICAL SYSTEM WITH BACKLASH 126
4.3.1 Introduction 126
4.3.2 General System 127
4.3.3 Chaotic Motions and Strange Attracting Sets 129
4.3.4 Fractal Dimension of Strange Attractors 132
4.3.5 Sensitivity on Initial Conditions and Lyaponov Exponents 133
4.3.6 Forcing Parameter Field and Damping Effect on Chaos 135
4.3.7 Two-degree-of-freedom Backlash System 139
4.3.8 Experimental Investigation 140
4.4 CONCLUSION 142

Chapter 5 LOCATION OF STRUCTURAL NONLINEARITIES

5.1 PRELIMINARIES **144**

5.2 NECESSITIES AND REQUIREMENTS OF NONLINEARITY LOCATION .. **145**

5.3 TECHNIQUES FOR THE LOCATION OF NONLINEARITY **147**

5.3.1 Location Using Measured Modal Data 147

5.3.2 Expansion of Unmeasured Coordinates 149

5.3.3 Sensitivity of Modal Properties to Localised Nonlinearity 151

5.3.4 Numerical Case Studies 152

5.4 EXPERIMENTAL INVESTIGATIONS **155**

5.4.1 Simulation of Structural System with Localised Nonlinearity 155

5.4.2 Measurement Results 160

5.4.3 Location of Nonlinearity 162

5.5 EXTENSION OF THE METHOD TO MEASURED FRF DATA **164**

5.6 LOCATION OF NONLINEARITY AND MODELLING ERRORS IN ANALYTICAL MODEL IMPROVEMENT **167**

5.7 CONCLUSIONS **168**

Chapter 6 IDENTIFICATION OF MATHEMATICAL MODELS OF DYNAMIC STRUCTURES

6.1 PRELIMINARIES **173**

6.2 MODELLING OF LINEAR AND NONLINEAR STRUCTURES **174**

6.3 REVIEW OF ANALYTICAL MODEL IMPROVEMENT METHODS **176**

6.4 MODEL UPDATING USING FRF DATA - A NEW GENERALISED METHOD **178**

6.4.1 Description of the Method 178

6.4.2 Numerical Case Studies 183

6.4.3 Extension of the Method to the Modelling of Nonlinear Structures 190

6.4.4 Application of the Method to the GARTEUR Structure 192

6.4.5 Some Considerations on Computational Aspects 196

6.5 GENERALISATION OF MODEL UPDATING METHODS 198
6.6 UNIQUENESS OF UPDATING PROBLEM 200
6.7 BALANCE OF COEFFICIENT MATRICES 202
6.8 CONCLUSIONS 204

Chapter 7 POSSIBILITIES AND LIMITATIONS OF ANALYTICAL MODEL IMPROVEMENT

7.1 INTRODUCTION 206
7.2 REVIEW OF FULL MATRIX UPDATING PRACTICE 207
7.2.1 Berman's Method 208
7.2.2 The Error Matrix Method 209
7.2.3 Mathematical Illustrations 210
7.3 CONDENSED MODEL BASED ON GUYAN REDUCTION 210
7.3.1 The GARTEUR Structure 211
7.3.2 The Guyan Reduction Technique 211
7.3.3 Location of Modelling Errors 215
7.3.4 Updating of Analytical Model Based on Location Results 221
7.4 STRUCTURAL CONNECTIVITIES IN AN ANALYTICAL MODEL 221
7.5 MINIMAL DATA REQUIRED TO UPDATE AN ANALYTICAL MODEL 224
7.5.1 The Eigendynamic Constraint Method 225
7.5.2 The Inverse Eigensensitivity Analysis 227
7.5.3 Number of Modes and Coordinates Required to Update an Analytical Model 230
7.5.4 Numerical Case Studies 231
7.5.5 Generalisation of the Criteria 235
7.5.6 Application of the Methods to the GARTEUR Structures 236
7.6 CONCLUSIONS 248

Chapter 8 CONCLUSIONS

8.1 IDENTIFICATION OF STRUCTURAL NONLINEARITY.....250

8.2 LOCATION OF STRUCTURAL NONLINEARITY 252

8.3 MODELLING OF LINEAR AND NONLINEAR STRUCTURES 253

8.4 SUGGESTION FOR FURTHER STUDIES.....255

APPENDICES

Appendix I: Singular Value Decomposition (SVD) 257

Appendix II: Derivation of Eigenderivatives 261

REFERENCES

CHAPTER 1

INTRODUCTION

1.1 GENERAL INTRODUCTION

In current engineering practice, the emphasis placed on safety, performance and reliability of structural systems is becoming more and more demanding due to the continuous challenges from real life. For example, any design inadequacy in an aircraft might lead to huge loss of human life. In order to design a structural system which, after being manufactured, will satisfy the prescribed safety performance and reliability criteria, it is essential that dynamic analysis be carried out at the design stage as well as at the prototype stage and, subsequently, a mathematical model which can accurately represent the dynamic characteristics of the structure be established. Such a mathematical model is needed for response and load prediction, stress and stability analysis, structural modification and optimisation etc.

For simple structural components, such as uniform beams and plates, mathematical models (and analytical solutions) which accurately describe their dynamic characteristics are readily available. However, due to the complexity of most engineering structures, analytical solutions are often impossible to obtain (if they exist at all) and numerical approximations have to be pursued. In structural mechanics, the most commonly employed numerical method is the so-called Finite Element Analysis (FEA) method. In

FEA, a continuous structure is discretised into many 'small' elements (the size of the element depends on the analysis accuracy required) called 'finite elements' and then, based on the theory of dynamics (e.g., Newton's law, Lagrange equation) and mechanics of deformable bodies (e.g., stress-displacement equations, stress-strain relations), a mathematical model of the structure is derived. This is often referred as the 'analytical model' as compared with the 'experimental model' which is derived from dynamic testing. Once a mathematical model (spatial model in terms of mass, stiffness and possibly damping matrices of the structure) has been formulated, the next step of the analysis is to solve the differential equations to obtain the dynamic response. However, due to the approximations and idealisation involved, lack of knowledge about the structure and even sometimes mismodelling of structural joints and boundary conditions, it is inevitable that the mathematical model thus established will not always adequately represent the actual behaviour of the real structure.

Apart from the analytical approach to achieve a mathematical model for the study of vibration characteristics of a dynamic system, another major approach is to establish an experimental model for the system by performing a vibration test and subsequent analysis on the measured data. This process, including the data acquisition and the subsequent analysis, is now known as 'Modal Testing'. The theoretical basis and practical applications of modal testing have been discussed in detail in [1]. The most significant application of modal testing is perhaps to compare and eventually to validate an analytical model using measured vibration test data. Apart from this, mathematical models derived from measured data (referred to as 'experimental models' which can be in the form of response, modal or spatial models) are frequently used in structural modification analysis, structural coupling, force determination etc. It is usually believed that provided sufficient care is given to the experimental procedures, the results from measurement are those that should be regarded as the most correct.

In a typical engineering design process, both analytical prediction and experimental modal testing procedures are involved in an iterative way. They have complementary roles for the complete description and understanding of the dynamic behaviour of a structure and one cannot be substituted for the other. In the present work, we shall be dealing mainly with the experimental side of the problem of evaluating the dynamic characteristics of mechanical structures (mainly nonlinear structures), although analytical models are often needed and are assumed to be available in the studies of location of structural nonlinearities and mathematical modelling of nonlinear structures.

1.2 STRUCTURAL NONLINEARITY

Most of the theories upon which structural dynamic analysis is founded rely heavily on the assumption that the dynamic behaviour of the structure to be analysed is linear. By this is meant that (i) if a given loading is doubled, the resulting deflections are doubled and (ii) the deflection due to two (or more) simultaneously applied loads is equal to the sum of the deflections caused when the loads are applied one at each time. This superposition principle of linear systems can be expressed mathematically as

$$x[\alpha f(t)] = \alpha x[f(t)] \tag{1-1}$$

$$x[\sum_{k=1}^{n} f_k(t)] = \sum_{k=1}^{n} x[f_k(t)] \tag{1-2}$$

where x is the deflection, $f(t)$ is the loading force and α is a constant. Linear mathematical models of engineering structures based on this superposition principle have proven to be very useful in numerous engineering applications. From general theoretical considerations based on the superposition principle, successful methods have been developed and applied to the dynamic analysis of linear structures.

Failure to obey the superposition principle implies that the structure is nonlinear. In fact, most practical engineering structures exhibit a certain degree of nonlinearity due to nonlinear dynamic characteristics of structural joints, nonlinear boundary conditions and nonlinear material properties. For practical purposes, they are in many cases regarded as linear structures because the degree of nonlinearity is small and therefore insignificant in the response range of interest. For other cases, the effect of nonlinearity may become so significant that it has to be taken into account in the analysis of dynamic characteristics of the structure.

It is often supposed that unless a real measurement is taken, the existence of a nonlinearity in a practical structure cannot be foreseen based on analytical prediction nor the degree of nonlinearity can be analytically quantified. Experimental investigation becomes essential in the identification of dynamic characteristics of nonlinear structures.

The present research focuses on the identification, location and mathematical modelling of practical nonlinearities based on measurement of the input/output dynamic characteristics of nonlinear structures. Although there have been several efforts directed towards the

identification of nonlinearity, there has not been a complete and systematic development of identification techniques which are required in order to analyse the numerous different nonlinear phenomena that occur in engineering practice. Further, the location of nonlinearity (information about where the nonlinearity is located in a structure), which has important engineering applications has not been investigated to date. Also, as an ultimate target of nonlinear system analysis, the establishment of a nonlinear spatial mathematical model which is a function of the response amplitude, has not been investigated in spite of its practical relevance to numerous engineering applications.

1.3 IDENTIFICATION OF DYNAMIC CHARACTERISTICS OF NONLINEAR STRUCTURES

System identification, which is generally considered as the inverse problem of system dynamics, is in the scope of various fields such as structural and control engineering. Although mathematicians and engineers have developed a number of approaches to address the identification problem, most of the work to date has been restricted to linear systems. Nonlinear systems are, however, often assumed to be linearisable in some manner, and the resulting linear model is then used to analyse the behaviour of the system. Significant inaccuracy arises when conditions and/or assumptions required for the linearisation are violated.

The identification of linear time-invariant system is relatively well understood and theoretically well developed. The same is not true for the case of a nonlinear system. Nevertheless, over the past years, some progress has been made in the development of both theories and techniques in the identification of nonlinear systems. A very brief review is presented here in terms of frequency, time and amplitude domains and a more detailed discussion will be given in some later chapters when specific topics are described or referred.

1.3.1 FREQUENCY DOMAIN TECHNIQUES

Techniques developed for the identification of nonlineariries in the frequency domain are, in general, based on the comparison of different characteristics of the measured frequency response functions of linear and nonlinear structures (nonlinearity detection) and extension/modification of classical linear analysis methods to nonlinear structures (nonlinearity quantification). As the first task of nonlinearity analysis, the detection of the existence of nonlinearity is believed to be relatively easy. For most practical nonlinear structures, frequency response functions (FRFs) measured using sinusoidal excitation

with constant forcing amplitude will show certain form of distortion as compared with those of a linear structure. Distortions of measured FRFs when they are displayed in the form of a Bode plot [1] or reciprocal receptance [2] have been employed to detect the existence of nonlinearities. Also, as discussed in [1], when a structure is nonlinear, the isometric damping plot calculated based on the measured FRF data will show systematic variation (surface distortion) and this variation is an indication of nonlinearity. These detection techniques are simple and easy to implement in practice.

As a more sophisticated method, the Hilbert transform technique can be used to detect, and to some degree, to quantify structural nonlinearities. The theory of the Hilbert transform, which is an integral transform, is described in detail in [3]. The basis that the Hilbert transform technique can be used to identify nonlinearity is due to the fact that for a linear structure, the real and imaginary parts of a measured FRF constitute a Hilbert transform pair (that is: $\mathcal{H}[\mathrm{Re}(\alpha(\omega))]=\mathrm{Im}(\alpha(\omega))$ and vice versa), while for the FRF of a nonlinear structure, these Hilbert transform relationships do not hold. By calculating the Hilbert transform of the real part (or the imaginary part) of a measured FRF and comparing it with the corresponding imaginary part (or real part), the existence of nonlinearity can be identified based on the difference of the transform pair [4].

For most practical applications, not only does the nonlinearity need to be detected, but more importantly, it needs to be quantified. The Hilbert transform approach seeks to quantify the nonlinearity by measuring the degree to which the Hilbert transform pair differ from each other. As a more practical way of quantifying structural nonlinearity, the Inverse Receptance method was developed [5] which aims to establish the relationship between the natural frequency and the vibration amplitude of a nonlinear structure. However, the method is restricted to the case in which the mode to be analysed is real.

All the above-mentioned techniques are formulated for the identification of nonlinearity based on the measured first-order FRFs (FRFs which are obtained by considering only the fundamental frequency component of the response signal, as will be defined later). For some nonlinear structures, the measured first-order FRFs are effectively linear and for some practical vibration problems in which the harmonic components of the response become as important as the fundamental component, the measurement and analysis of higher-order FRFs becomes necessary. The theoretical basis of higher-order FRFs is the Volterra series and its extended Wiener series theory [6]. However, research activities had been restricted in electrical and control engineering since Wiener's early work [7] and until recently that the theory has been applied to the identification of nonlinear mechanical structures [8] and found to be quite useful.

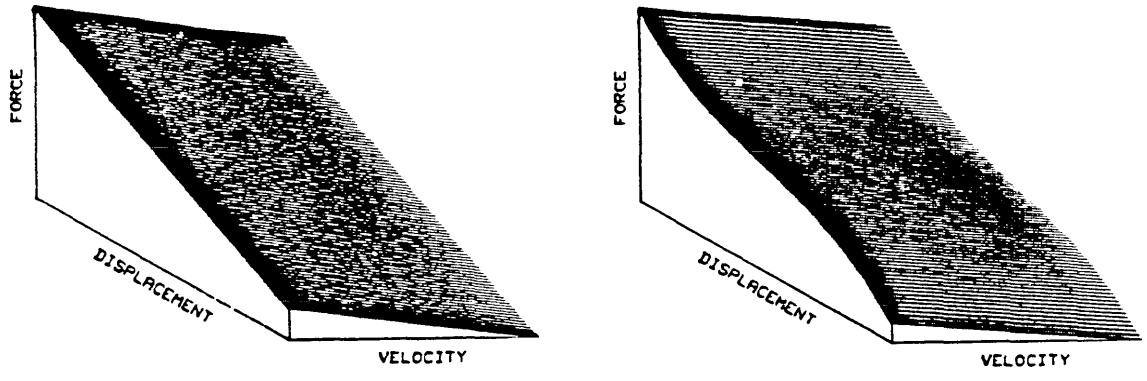
1.3.2 TIME DOMAIN ANALYSIS

As in the identification of linear structures for which time domain methods such as Ibrahim's method (ITD) [9] can be used to obtain modal parameters of a linear structure, structural nonlinearities can be identified by analysing the measured time force and response data directly. The simplest method of nonlinearity analysis in time domain is the phase-space or, more specifically, the phase-plane approach [10] which gives both local and global behaviour of a nonlinear system and provides an exact topological account of all possible system motions under various operating conditions. In the case of sinusoidal excitation, the Poincaré map which is the discrete phase-space trajectory of the motion, can be calculated to detect the existence of harmonic components and so the existence of nonlinearity.

The force state mapping technique for the identification of nonlinearity was initiated by Masri and Caughey [11] and independently developed by Crawley and O'Donnel [12-13]. The technique aims to establish the relationship between the restoring force and the vibration displacement and velocity $F(x,\dot{x})$. For an SDOF nonlinear system described by

$$m\ddot{x} + F(x, \dot{x}) = f(t) \tag{1-3}$$

where $F(x, \dot{x})$ is the restoring force, if, by some means, the mass m of the system is known and quantities x, $\dot{x}$, $\ddot{x}$ and $f(t)$ are measured, then the restoring force $F(x, \dot{x})$ at given state $(x, \dot{x})$ can be calculated. If the system is linear, then the restoring force surface is a plane as shown in Fig.1.1(a). If the system is nonlinear, then some surface distortion is expected, as shown in Fig.1.1(b) for the case of cubic stiffness nonlinearity. The extension of this technique to certain nonlinear MDOF systems has been investigated [11] and the practical application of the technique has been discussed [14].



(a) linear SDOF system (b) nonlinear SDOF system

Fig.1.1 Restoring Force Surfaces of Linear and Nonlinear Systems

Time series analysis techniques [15] have been widely used in the modelling of linear systems and have recently been extended to the identification of nonlinear systems [16-17]. In general, a linear time-invariant system can be represented by higher-order (more than the second) differential equation with constant coefficients. Such a differential equation can, in theory, be approximated by a difference equation whose accuracy depends on the time interval of sampled data points. The time series analysis seeks to calculate the coefficients of the difference equation model based on the measured input-output time series data. In the case of a nonlinear system, extra coefficients have to be identified which represent the effect of the nonlinear behaviour [16]. After the difference equation model, which describes the dynamic characteristics of a nonlinear system, has been identified, the first-order and higher-order frequency response functions of the system can be calculated [18].

1.3.3 AMPLITUDE DOMAIN ANALYSIS

Nonlinearities can also be identified by calculating the amplitude probability density function of the response due to random excitation. For a linear system, if the input force is a random signal with its amplitude probability density function (pdf) being Gaussian, then the amplitude pdf of the response will also be Gaussian. For a nonlinear system however, this simple relationship no longer holds and some distortion in the response amplitude pdf from Gaussian distribution is expected and from this distortion, the existence of nonlinearity can be identified [19]. The method is developed based on the Fokker-Planck-Kolmogorov (FPK) equation of a nonlinear system which is described in detail in [20]. To illustrate the idea, consider an SDOF nonlinear system as

$$\ddot{x} + \beta\dot{x} + \varphi(x) = f(t) \tag{1-5}$$

where $f(t)$ is a Gaussian noise signal. The corresponding Fokker-Planck-Kolmogorov equation of (1-5) is described as [20]:

$$\frac{S_0}{2}\frac{\partial^2}{\partial\dot{x}\partial\dot{x}}[\rho_1(x)\rho_2(\dot{x})] - \frac{\partial}{\partial x}[\dot{x}\rho_1(x)\rho_2(\dot{x})] + \frac{\partial}{\partial\dot{x}}[(\beta\dot{x} + \varphi(x))\rho_1(x)\rho_2(\dot{x})] = 0 \tag{1-6}$$

where S_0 is the power spectrum of the input force and $\rho_1(x)$ and $\rho_2(\dot{x})$ are the probability density functions of the displacement and velocity, respectively. Solving (1-6), $\rho_1(x)$ can be obtained as [20]:

$$\rho_1(x) = \gamma \exp\left[-\frac{2\beta}{S_0}\int_0^x \varphi(\xi)d\xi \right] \tag{1-7}$$

where γ is a constant which can be determined by the normalisation condition. From (1-7), it can be seen that only when $\varphi(x)$ is linear does $\rho_1(x)$ have a Gaussian distribution. On the other hand, $\rho_1(x)$ can be calculated based on the measured time response data and, therefore, nonlinearity can be easily identified experimentally based on the distortion of measured $\rho_1(x)$ from the standard Gaussian distribution.

In the research described in this thesis, we will concentrate mainly on the development of analysis techniques in the frequency domain although time domain techniques such as phase-plane and Poincaré map approaches will be used in the characterisation of chaotic vibrational systems.

1.4 MODELLING OF NONLINEAR STRUCTURES

For many engineering applications, accurate mathematical models (spatial models in terms of mass and stiffness matrices) of nonlinear structures are required. So far, much progress has been made in the mathematical modelling of linear structures [21]. As mentioned above, a mathematical model of a linear structure can be established either using analytical FE analysis (an analytical model) or based on measured dynamic test data (an experimental model). Due to the existence of modelling errors in most practical cases, the analytical model needs to be validated using measured test data so that an accurate mathematical model can be established. In the case when a structure to be modelled is nonlinear, its mathematical model becomes a function of response amplitude ([K]=[K(x̂)]

for the case of stiffness nonlinearity) and what needs to be established is a series of linearised models corresponding to different vibration amplitudes by correlating the analytical model and measured first-order FRF data.

In fact, since structural nonlinearity cannot be foreseen and so cannot generally be analytically predicted, measurement is crucially important in the modelling of nonlinearity. However, measurement alone cannot, in general, establish a practically realistic model because measured data are usually very limited (as will be shown, this is especially true for the case of nonlinear structures). It is believed that a reasonably accurate linear model of a nonlinear structure (corresponding to very low vibration amplitude) and location information of the localised nonlinearity are necessary in order to establish the mathematical model of a nonlinear structure.

In the present research, we shall focus on the development of techniques for both the location and the mathematical modelling of structural nonlinearities. The procedure is as proposed below. First, an analytical model is updated using vibration test data measured at very low response amplitude to obtain an accurate linear model of the nonlinear structure. Then, the nonlinearity is located, based on this linear model and measured data at higher response amplitudes. With this location information available, modelling of the nonlinearity can be concentrated on the region where the structural nonlinearity is and then by correlating the linear model and measured FRF data at different response amplitudes, a mathematical model of the nonlinear structure can be established.

1.5 PREVIEW OF THE THESIS

Despite rapid developments in the identification of dynamic characteristics of linear structures in recent decades, structural nonlinearity presents a major difficulty to the majority of applications to practical cases. The research presented in this thesis is intended to seek new developments on the identification, location and modelling of structural nonlinearities in the pursuit of better understanding of the dynamic characteristics of practical nonlinear structures.

Based on the analysis of measured first-order FRFs, some of the recently-developed techniques for the identification of nonlinearity are reviewed in Chapter 2 and their advantages and disadvantages when applied to practical problems are discussed. Then, a new improved method for the nonlinear modal analysis of complex modes is developed to cope with the practical situations in which measured modes become complex. The method has been successfully applied to the data measured from practical nonlinear structures

even when the modes are considerably complex. On the other hand, for structures whose nonlinearities are such that the measured first-order FRFs are effectively linear (the nonlinearity being of nonsymmetrical type), a higher-order frequency response function analysis becomes necessary for the identification of such nonlinear structures. The theoretical basis of higher-order FRF analysis is presented in Chapter 3 with special attention given to the numerical assessment of the practical applicability the technique. Both first- and higher-order FRF analysis techniques are largely based on the classical assumption that the output of a nonlinear structure is periodic if the input is periodic. For some nonlinear systems however, this assumption is no longer valid (chaotic systems, in which a periodic input will result in an output of "random" nature) and special techniques need to be developed in order to identify them. In Chapter 4, for the first time, the hidden chaotic behaviour of a mechanical backlash system with realistic system parameters has been revealed and, based on this system, qualitative as well as quantitative ways of identifying chaotic structures are presented. Both numerical studies and experimental investigations are carried out and possible engineering applications are discussed.

It is believed that nonlinearity in most engineering structures is usually localised in certain spatial coordinates and the ability to locate these has some important engineering applications. In Chapter 5, nonlinearity location techniques based on the correlation between an analytical model and measured modal parameters and/or measured frequency response function data are developed. Numerical studies and experimental investigation demonstrate the practical applicability of these techniques.

The ultimate target of the analysis of a nonlinear structure is to establish a nonlinear mathematical model which is a function of response amplitude. It is believed that such a goal can only be achieved by combining analytical modelling (FE modelling) and experimental modal testing techniques. In Chapter 6, a new model updating method is developed and extended to the mathematical modelling of nonlinear structures based on the correlation between an analytical model (of a linear system) and measured frequency response data. As compared with existing methods, the new method shows marked advantages. The practical applicability of the method is assessed based a special case study. In Chapter 7, criteria on minimum data required in order to update an analytical model are established and the possibilities and limitations of analytical model improvement are discussed which make it possible for the analyst to judge whether a set of measured data will be adequate to solve the model updating problem uniquely.

Finally, Chapter 8 reviews all the new developments presented in this thesis and indicates the direction for possible further studies.

CHAPTER 2

IDENTIFICATION OF NONLINEARITY USING FIRST-ORDER FREQUENCY RESPONSE FUNCTIONS

2.1 GENERAL INTRODUCTION

As mentioned in Chapter 1, it is believed that all practical engineering structures are nonlinear to some extent, due to nonlinearities in structural joints, boundary conditions and material properties. Some structures may be only slightly nonlinear so that they can be analysed based on a linear theory with satisfactory accuracy. For others whose nonlinearities are such that their dynamic characteristics deviate considerably from linear behaviour, nonlinear analysis techniques have to be employed.

As far as analyses of nonlinear systems are considered, there are two types of activity which are loosely termed here as "theoretical analysis" and "experimental analysis". The theoretical analysis methods assume that a mathematical model of the nonlinear system to be analysed is known (usually in the form of differential equations) and what is of interest is the prediction of the response of the system due to a certain input. In contrast, the experimental analysis methods seek ways of identifying mathematical models of nonlinear systems based on measured input/output dynamic characteristics.

Since, in general, nonlinear problems do not possess closed form solutions, both theoretical and experimental analyses are approximate and the accuracy of each analysis depends on the mathematical nature of the problem and the specific methods employed. What is of major interest in this thesis is the experimental identification of nonlinear

structures based on structural modal testing, but since the theoretical analysis provides the basis for the experimental identification of nonlinear structures, it is necessary to give a brief introduction to the theoretical analysis of nonlinear systems.

Nonlinear systems with either inherent nonlinear characteristics or nonlinearities deliberately introduced into the system to improve their dynamic characteristics have found wide applications in most diverse fields of engineering. The principal task of nonlinear system analysis is obtain a comprehensive picture, quantitative if possible, but at least qualitative, of what happens to the system if it is driven into its nonlinear regime. According to whether the system variables such as vibration displacement in the mechanical structure are perturbed only slightly or largely from their operating points (for most nonlinear mechanical structures, the nonlinear effect becomes more severe as the vibration amplitude increases, but there are some exceptions such as friction nonlinearity), the nonlinear characteristics can be divided into *local* or *global* behaviour. Local behaviour can be investigated by rather general and efficient linear methods that are based on the powerful superposition principle as explained in Chapter 1 because, in this case, the dynamic characteristics of the system are completely dominated by linear behaviour. However, if these linear methods are extended to describe the global behaviour of a nonlinear system, the results can be erroneous both quantitatively and qualitatively since, in this case, the nonlinear characteristics may be essential but the linear methods may fail to reveal it. Therefore, there is a strong emphasis on the development of methods and techniques for the analysis and design of nonlinear systems.

However, it has to be mentioned that the development of nonlinear methods faces real difficulties for a variety of reasons. There are no universal mathematical methods for the solution of nonlinear differential equations which are the mathematical models of nonlinear systems. The methods which exist deal with specific classes of nonlinear equations and therefore have limited applicability to system analysis. The classification of a given system and the choice of an appropriate method of analysis is not at all an easy task. Furthermore, even in simple nonlinear problems, there are numerous new phenomena qualitatively different from those expected in linear system behaviour, and it is impossible to encompass all these phenomena in a single and unique method of analysis.

Although there is no universal approach to the analysis of nonlinear systems, nonlinear methods generally fall into one of the three following approaches: (i) the phase-space topological method, (i) the stability analysis method, or (iii) the approximate method of nonlinear analysis. These are summarised below.

(i) The phase-space or, more specifically, the phase-plane approach has been used for solving problems in mathematics and engineering at least since Poincaré in 1880's. The approach gives both local and global behaviour of the nonlinear system and provides an exact topological account of all possible system motions under various operating conditions. It is a powerful concept underlying the entire theory of ordinary differential equations (linear, nonlinear, time-varying and time-invariant). However, it is limited to the case of second-order equations. For higher-order systems, this approach is very cumbersome to use.

(ii) The stability analysis of nonlinear systems, which is heavily based on the work of Lyapunov, is a powerful approach to the qualitative analysis of system global behaviour. By this approach, the global behaviour of the system is investigated utilizing a given form of nonlinear differential equations without explicit knowledge of their solution. Stability is an inherent feature of a wide class of systems such as aerospace structures.

(iii) Approximate methods for solving problems in mathematical physics were first developed at the beginning of this century. They have been received with much interest by engineers and have promptly obtained wide application in diverse fields of engineering. The basic merit of approximate methods lies in their being direct and efficient and they permit a simple evaluation of the solution for a wide class of problems arising in the analysis of nonlinear oscillations.

In the theoretical analysis of nonlinear systems whose equations of motion can be formulated analytically, there are quite a number of approximate methods available to examine their nonlinear vibration behaviour. According to different input signals, methods in general can be categorised into *deterministic* methods, in which the excitation signals are deterministic such as sinusoids and *statistical* methods, in which the input signal is of a random nature. Statistical analysis methods include the method(s) of random linearisation [22-23] and the amplitude domain analysis based on the FPK equations [20] as discussed in Chapter 1. On the other hand, in deterministic analysis, the most commonly used methods are the Linstedt-Poincaré method [24], the method of multiple scale [24-26] and the harmonic balance method [27]. What is of particular interest here is the harmonic balance method (often called describing function method) because this harmonic balance analysis provides the mathematical basis for a new nonlinear modal analysis method developed in this chapter. The harmonic balance method is heavily based on the Krylov-Bogoliubov approach [28] and is applicable to nonlinear systems described by higher-order differential equations. The mathematical basis of the method and the

applicability conditions when it is used to identify the nonlinearities of mechanical structures will be discussed in detail later in this chapter when the new nonlinear modal analysis method is introduced. However, it will be appropriate here to introduce the measurement techniques available in the dynamic characterisation of nonlinear structures since what is of primary interest is the identification of nonlinear structures from the actual test data.

2.2 MEASUREMENT OF FIRST-ORDER FREQUENCY RESPONSE FUNCTIONS

A brief introduction has been made so far of the theoretical analysis of nonlinear systems based on the known differential equations. However, the primary target which is sought in this study is the identification of the unknown mathematical models of nonlinear structures based on measured input/output dynamic characteristics. Therefore, it becomes necessary before the introduction of any identification techniques to discuss how the dynamic characteristics of a structure (linear or nonlinear) can be measured.

First of all, it is necessary to explain what is meant by the so-called *first-order frequency response functions* of a nonlinear structure. In concept, first-order frequency response functions (first-order FRFs) are the extension of frequency response functions (FRFs) of linear systems to nonlinear systems. Similar to the measurement of FRFs of a linear structure, in the case of sinusoidal excitation (the excitation is a pure sinusoid), the first-order FRF $H_1(\omega)$ of a nonlinear structure is defined as the spectral ratio of the response $X(\omega)$ and the force $F(\omega)$ at the excitation frequency: $H_1(\omega)=X(\omega)/F(\omega)$. During the estimation of $H_1(\omega)$, all the harmonic components (subharmonics, superharmonics and combinational resonances) are ignored and only the fundamental frequency component of the response is retained. Similarly, in the case of random excitation (the excitation is wide-band random signal), first-order frequency response function is defined as the spectral ratio of cross-spectrum of the force and response and the auto-spectrum of the force: $H_1(\omega)=S_{fx}(\omega)/S_{ff}(\omega)$ (or its equivalent form $H_1(\omega) =S_{xx}(\omega)/S_{fx}(\omega)$). The measured first-order FRFs of a nonlinear structure based on sinusoidal and random excitations are in general different and their relationship will be discussed in Chapter 3.

For linear structures, the first-order frequency response functions (often referred simply as frequency response functions) are unique and, therefore, will not vary according to different excitation techniques and conditions. For nonlinear structures, however, the measured first-order frequency response functions are, in general, not unique. They depend not only on the excitation conditions (input force levels), but also on the different

excitation signals used to measure them. Therefore, the first problem of nonlinearity investigation will necessarily be to decide a proper means of excitation so that nonlinearity can easily be revealed and then identified. There are three types of excitation method widely used in vibration study practice - sinusoidal, random and transient - and each of them is discussed below.

2.2.1 SINUSOIDAL EXCITATION TECHNIQUE

In testing a linear structure, if the input is a sinusoid, the response will also be a sinusoid with the same frequency as that of the excitation and the frequency response function at this excitation frequency is simply the ratio of the amplitudes (usually complex) between the response and the input signals. This observation of its special characteristics naturally made the sinusoidal excitation to be the first choice of excitation signal at the very beginning of structural dynamic testing and it still remains one of the most favourable excitation techniques in today's modal testing practice because of its uniqueness and precision, although other techniques such as random, transient etc. have also been developed.

The main advantages of sinusoidal excitation are: (i) the input force level can be accurately controlled and hence it becomes possible to excite the structure at specified response levels required and (ii) since all the input energy is concentrated at one frequency each time, and the noise and harmonic components in the response signal are averaged out through an integration process, the signal-to-noise ratio is generally good as compared with other excitation methods. As in most cases the study of nonlinearity requires either response or force controls, the characteristics of (i) become important in the successful identification of structural nonlinearity.

When the response level is set to be constant during the measurement (response amplitude is constant at different excitation frequencies), a nonlinear structure is said to be linearised and the measured first-order frequency response functions can be analysed using standard linear modal analysis methods in exactly the same way as for the frequency response functions measured on a linear structure. On the other hand, when the input force is kept constant during the measurement (the amplitude of the input force is constant at different excitation frequencies), the measured first-order frequency response functions are nonlinear (they are characteristically different from FRFs measured from linear structures) and in this case, special nonlinear modal analysis methods have to be used to analyse them in order to identify the existing nonlinearity.

However, it is worth mentioning that in some cases, the accurate control of force or response can be a problem due to the electro-dynamic characteristics of the shaker. A feedback control process is required in order to achieve either response or force control as will be discussed in some detail in Chapter 5.

The main drawback of the sinusoidal excitation technique is that it is relatively slow when compared with many of the other techniques used in measurement. The reason for this is that the excitation is performed based on frequency by frequency basis and, at each frequency, time is needed for the transient response components to decay and the system to settle to its steady-state vibration. However, it is believed that in many applications, correct measurement of the dynamic characteristics of a structure becomes more important than the measurement time involved. As will be discussed in chapter 6 on analytical model updating practice, the accuracy of measured frequency response functions becomes vitally important for a successful correlation to be achieved.

2.2.2 MEASUREMENT USING RANDOM EXCITATION

The term 'random' applies to the amplitudes of the excitation force which, in statistical terms, have a Gaussian or Gaussian-like probability distribution. Wide-band random excitation is widely used in structural dynamic testing because it approximates more closely the statistical characteristics of vibration service environments than does a pure sinusoidal excitation.

With this type of excitation, individual time records in the analyser contain data with random amplitude and phase for each frequency component. On average, however, the spectrum is flat and continuous, containing energy approximately the same level for every frequency in the range of interest. The spectrum distribution is easy to control in a random test, and it can be limited to cover the same range as the analysis.

The excitation is random and continuous in time, but the record length is finite, and so the recorded signals (force and response) are, in general, nonperiodic. However, during the signal analysis, these nonperiodic signals are assumed to be periodic and as a result, leakage errors occur in the estimation of frequency response functions. These errors can be minimised by using window functions, or weighting, which act as a soft entry and exit for the data in each record. A suitable weighting function to use with random data is the Hanning window.

In order to eliminate the leakage problem, a pseudo-random excitation signal can be used instead of a true random signal. The pseudo-random signal is periodic and repeats itself with every record of analysis. A single time record of a pseudo-random signal resembles a true random wave form, with a Gaussian-like amplitude distribution. However, the spectral properties are quite different from those of a random signal because of its periodicity. First, the periodic nature of the pseudo-random signal removes the leakage error entirely so that a rectangular window must be used in the analysis and secondly, the spectrum becomes discrete, only containing energy at the frequencies sampled in the analysis.

For the FRF measurement of linear systems, random and pseudo-random excitations are attractive to analysts and researchers because of their potential time-saving in obtaining frequency response functions. In random and pseudo-random excitation measurement, the structure is excited simultaneously at every frequency within the range of interest. It is this wideband excitation characteristic that makes the random and pseudo-random excitation faster than sinusoidal excitation. As compared with true random excitation, in addition to the advantage of being leakage error free, pseudo-random excitation is much faster because as the random source is true noise, it must be averaged for several time records before an accurate FRF can be determined.

As for nonlinear structures, from the measured first-order frequency response function point of view, a random test in general linearises nonlinear structures due to the randomness of the amplitude and phase of the input force signal and the averaging effects, therefore, the measured first-order frequency response functions using random test are linear. The theoretical aspects of this linearisation process will be discussed in Chapter 3. However, the linearisation of a nonlinear structure when using random excitation does not mean that it is impossible to identify nonlinearity using random excitation. Corresponding to different input excitation levels (power spectra), the measured linearised first-order frequency response functions are different and if a set of these frequency response functions are measured at different excitation levels, the identification of nonlinearity could, in some cases, become possible. On the other hand, as will be shown in Chapter 3, this conventional random test technique can be extended to measure the higher-order frequency response functions of a nonlinear structure and these provide valuable information concerning the nature of the nonlinearity and can be used to serve the purpose of nonlinearity identification. Pseudo-random excitation on the other hand, is in general not suitable for the first-order FRF measurement of nonlinear structures. This is because a pseudo-random signal is periodic and so contains limited discrete frequency components. When such a input signal is applied to a nonlinear

structure, modulation and intermodulation distortion will be generated due to nonlinearity and, unfortunately, these distortion products (e.g. $2f_1$, $3f_1$, ... due to modulation of input component f_1) will fall exactly on the other frequency components of the input signal (e.g. $2f_1$, $3f_1$, ...). So the distortion products add to the output and therefore interfere with the measurement of frequency response functions [29]. Unlike random excitation, in which these distortions can be averaged out, pseudo-random is periodic and so the averaging has no effect on the measured FRF.

Although the first-order frequency response functions measured using random excitation are different when the input force spectrum levels are different, these differences could be very small when practical nonlinear structural tests are considered. One reason for this is the dropout of the input force spectrum around resonance frequencies due to the impedance mismatch between the test structure and the electro-dynamic shaker. Since the energy input around structural resonance(s) is mainly responsible for the vibration level of a structure, dropout of the input force spectrum around resonance(s) means that the structure cannot easily be driven into its very nonlinear regime and the measured frequency response functions corresponding to different input force levels will not, in general, be very different from one another. With sinusoidal excitation, this impedance mismatch can be compensated using a feedback control system, but for random excitation, such compensation seems to be difficult and this is a practical problem for the identification of nonlinearity using random test.

2.2.3 MEASUREMENT USING TRANSIENT EXCITATION

One of the most popular excitation techniques used in structural dynamic testing is transient excitation, sometime referred as 'impact testing'. This popularity is because transient excitation has some unique characteristics as compared with shaker-based excitation techniques. The main advantages of using transient excitation can be summarised as:

(i) transient excitation does not require a dynamic shaker to generate the input excitation force; this is usually produced using an impactor such as a hand-held hammer and therefore the test structure remains unmodified during the test,

(ii) because there is no attachment required in the test, transient excitation provides easier access to the measurement points of the structure and,

(iii) transient excitation requires less equipment (no shaker and its related power amplifier involved) and measurement time, therefore, it is ideal for mobile experiments.

As in the case of random excitation, the derivation of input and output relationship under transient excitation relies on Fourier transform theory and is based on the Duhamel Convolution Integral. The measured frequency response function depends on estimates of the auto-spectrum $S_{ff}(\omega)$ of the force signal and the cross-spectrum $S_{xf}(\omega)$ of the force and response and it is calculated as $H(\omega)=X(\omega)/F(\omega)=S_{xf}(\omega)/S_{ff}(\omega)$ (or its equivalent form $H(\omega)=S_{xx}(\omega)/S_{xf}(\omega)$.

An ideal impulse is the delta function $\delta(t)$ which, after being Fourier transformed, produces a force spectrum with equal amplitude at all frequencies. Unfortunately, this ideal impulse is practically impossible to achieve. The waveform which can be produced by an impact is a transient (short time duration) energy transfer event whose spectrum is continuous, with a maximum amplitude at zero frequency and amplitude decaying with increasing frequency. The spectrum shape of the transient signal is mainly determined by the time duration of the signal. The shorter the time duration of the signal, the broader the range of energy distribution in the frequency domain. On the other hand, the time duration of an impact is determined by the mass and stiffness of both the impactor and the structure. Therefore, by properly choosing the material and of the hammer tip and its mass, it is possible to generate the required transient signal with desired spectrum characteristics. However, the spectrum can only be controlled at the upper frequency limit, which means the technique is not suitable for zoom analysis.

Although it has been suggested that the high crest factor of transient excitation makes it possible for the nonlinear behaviour of the structure to be provoked and then possibly identified, there has not been much evidence so far which seems to support the advantage of using transient excitation to identify structural nonlinearity based on the measured first-order FRFs. However, a special hammer has been designed to measure second-order FRFs of nonlinear structures [30] as will be discussed in Chapter 3.

2.2.4 COMMENTS AND PRACTICAL CONSIDERATIONS OF NONLINEARITY MEASUREMENT

As discussed above, there are three main types of excitation technique available for the vibration testing of a structure. Each of them has its advantages and disadvantages and a proper choice of excitation technique depends, in general, on the measurement accuracy required and time available for the test. For linear structures, since the measured frequency response functions are, in theory, unique and independent of the excitation, all techniques should be equally applicable. For nonlinear structures, however, the choice of excitation becomes important for the hidden nonlinearities to be revealed and then

identified because, in this case, the measured dynamic properties are excitation-dependent.

Transient excitation is one of the most often used techniques in structural dynamic testing because of its simplicity and speed in obtaining frequency response functions. It requires less equipment and is therefore suitable for mobile experiments. Since there is no shaker involved, the structure remains unmodified during the test. The coherence functions obtained from transient tests, being an indication of the measurement quality, are usually better than those from random tests in the sense that low coherence only occurs at anti-resonances due to the low signal-to-noise ratio of the response signal while in the random excitation case, low coherence occurs not only at anti-resonances, but also at resonances due to the dropout of input force spectrum around resonances caused by the impedance mismatch between the test structure and shaker. As for the identification of nonlinearity, although it is believed that it might be possible to use the transient excitation because of its high crest factor which provokes the structural nonlinearity, there have not been many studies carried out to support this idea.

When random excitation is used, the measured first-order FRFs are always linear, whether the structure is linear or not. In the case where the test structure is linear, the measured FRFs are unique and will not vary according to different excitation levels, while on the other hand, if the test structure is nonlinear, a series of linearised first-order FRFs will be obtained corresponding to different excitation levels. These measured first-order FRFs can be used to detect whether a structure is linear or not by comparing their values for different excitation levels and in cases where only an approximate linear model of a nonlinear structure is of interest, regardless of the type of nonlinearity the structure possesses, these linearised FRFs can often provide an accurate linear approximation of the nonlinear structure from a response prediction point of view.

On the other hand, the conventional random excitation technique can be extended to the case of higher-order frequency response function measurement based on the Wiener theory of nonlinear systems [7]. As will be shown in the next chapter, higher-order FRFs can be used in some cases to identify the type of structural nonlinearity and, together with the measured first-order FRFs, to predict the response due to certain inputs more accurately than those obtained using the measured first-order FRFs alone.

In the case where accurate quantification of structural nonlinearity is required, e.g. how the modal and/or spatial model of a nonlinear structure will change for different vibration response levels, sinusoidal excitation is generally regarded as the best choice because of

its flexibility of input force level control. There are two different types of controlled sinusoidal measurement technique commonly used for testing a nonlinear structure, referred as the 'constant response' and 'constant force'* measurement procedures. In constant response measurements, the response amplitude of test nonlinear structure at a certain point is kept constant at different excitation frequencies by adjusting the input force levels and, as a result, the measured first-order FRFs are linear. However, corresponding to different response levels, the measured first-order FRFs are different and by analysing them using linear modal analysis methods, a relationship between modal model and response levels can be established. The problem here is that the measurement is extremely time-consuming and therefore expensive. Furthermore, the measured range of response amplitude, which is important in nonlinearity analysis, could be limited because of the dramatic changes of receptance amplitude around resonances, especially when the structure is very lightly damped. In the case of constant force measurements, the amplitude of the input force is constant at each of the different excitation frequencies. Due to the varying receptance amplitudes, the response amplitudes are different at different measurement frequencies and, therefore, the measured first-order FRFs are nonlinear and contain information of a series of linearised FRFs measured at constant response amplitudes. Such nonlinear first-order FRFs are used from time to time in nonlinearity investigations and it will be shown in this chapter that they can be analysed based on the nonlinear modal analysis method developed to establish the relationship between the modal model and response levels of a nonlinear structure.

In practical measurements, because of the existence of different types of nonlinearity, care must be taken in determining the necessary excitation range so that the nonlinearity(ies) can be exposed to a satisfactory extent. In general, nonlinearities can be categorised into four different types. For the majority of nonlinearities commonly encountered in practice, either stiffness nonlinearities or damping nonlinearities, increasing the excitation force level will be similar to increasing the degree of nonlinearity. Examples of such nonlinearities are cubic stiffness and quadratic damping. For some nonlinearities such as backlash, the structure will remain linear until its response exceeds a certain limit. For frictional damping, on the other hand, increasing the excitation level will decrease the degree of nonlinearity and for some nonsymmetric nonlinearities such as bilinear and quadratic stiffness, the nonlinearity will have no effect on the measured first-order FRFs and in order to identify such special types of nonlinearity, the introduction of higher-order FRFs becomes necessary.

Following these observations, it is clearly important to choose properly the response range and thus the excitation range required so that the nonlinearity can be exposed and

* See Appendix III for more details

then identified satisfactorily. For most nonlinear structures in practice, relatively high excitation levels are recommended. For some others, such as structures with frictional damping, the situation can be the other way round and in order to identify such nonlinearities, the excitation levels should be set as low as possible. Lastly, for structures whose nonlinearities have no effect on the measured first-order FRFs, the measurement of higher-order FRFs is recommended.

2.3 RECENT DEVELOPMENTS ON THE IDENTIFICATION OF NONLINEARITY BASED ON THE FIRST-ORDER FREQUENCY RESPONSE FUNCTIONS

Once a structure is suspected of being nonlinear and its first-order FRFs are measured as discussed above, it becomes necessary to take nonlinearity into account in the subsequent modal analysis of the data. In practical nonlinearity analysis, in order to understand the nature and extent of nonlinearity, three requirements must be achieved by appropriate application of modal analysis methods. First, the existence of nonlinearity needs to be detected. Second, the extent of the nonlinearity needs to be quantified. Finally, physical characteristics of nonlinearity need to be established.

It is believed that the first requirement (detection) is comparatively easy to achieve. In fact, simply by comparing the difference between measured frequency response functions at different excitation levels, the existence of nonlinearity can be discovered. The quantification of nonlinearity, that is the establishment of the relationship between modal parameters and response levels based on the measured first-order FRFs, is the main topic of this chapter and will be discussed in detail. The last objective and the most difficult task in nonlinear modal analysis is the identification of physical characteristics of nonlinearity - the relationship between the structural spatial properties (such as stiffness) and response amplitudes. As will be shown in chapter 6, the establishment of such a nonlinear spatial model can only be achieved based on the correlation between the analytical model and vibration test data.

In efforts to achieve these various objectives of nonlinearity analysis, a large number of papers have been published in recent years. Methods which are commonly used in practical nonlinear modal analysis are to be reviewed and their advantages and disadvantages will be discussed.

2.3.1 DETECTION OF NONLINEARITY BY DIRECTLY USING FRF DATA

As far as nonlinearity detection is concerned, many techniques are available. The simplest methods are those which are based on the use of raw measured frequency response data without any post-measurement data processing. Detection techniques of this type are the Bode plot [1] and the reciprocal receptance plot [2] methods.

The basis of using Bode and or reciprocal receptance plot methods to detect nonlinearity is that, due to the existence of nonlinearity, the measured first-order FRFs will be distorted systematically from the corresponding linear frequency response functions. Since linear FRFs are very well recognised, the existence of nonlinearity can be revealed by examining the abnormal behaviour of the measured first-order FRFs.

To illustrate use of the Bode plot to detect nonlinearity, frequency response functions measured from a practical Beam/Absorber nonlinear structure are used. As shown in Fig.2.1, when the excitation force level increases, the distortion in the measured FRF data also increases and the apparent resonance frequency (the frequency of maximum FRF value) drops. The existence of softening stiffness nonlinearity of the structure is clearly demonstrated.

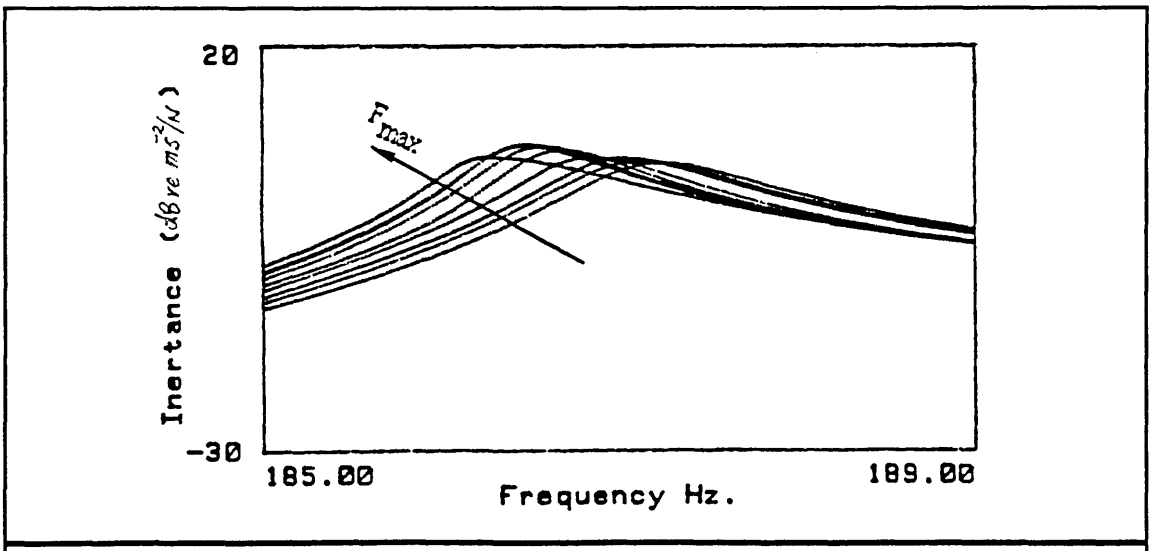


Fig. 2.1 - FRF Data Measured from Practical Beam-absorber Structure

As an alternative but more versatile technique, FRF data can be displayed in their reciprocal form to detect the existence of nonlinearity. The advantage of displaying data in

this format is that, in the case when the modal constant of the mode to be analysed is effectively real, this technique can not only detect the existence of nonlinearity, but can also give an indication as to whether the nonlinearity exists in the stiffness or damping. Suppose the residual contribution of other modes has been subtracted or can be neglected, the receptance data around r^{th} mode of the structure can be expressed as :

$$\alpha_{jk}(\omega) = \frac{{}_rA_{jk}}{\omega_r^2 - \omega^2 + i\eta_r \omega_r^2} \tag{2-1}$$

Rewriting equation (2-1) in its reciprocal form and assume modal constant ${}_rA_{jk}$ to be real:

$$\frac{1}{\alpha_{jk}(\omega)} = \frac{\omega_r^2 - \omega^2}{{}_rA_{jk}} + i\,\frac{\eta_r \omega_r^2}{{}_rA_{jk}} = R(\omega) + i\,I(\omega) \tag{2-2}$$

From equation (2-2), it can be seen that if FRF data are expressed in their reciprocal form, ω_r (related to stiffness nonlinearity) and η_r (related to damping nonlinearity) can be identified separately from the real and imaginary parts of the reciprocal FRF.

When FRF data are obtained for linear structures, the relationships of $R(\omega)$ *vs* ω^2 and $I(\omega)$ *vs* ω^2 are straight lines for the case of hysteretic damping as shown in Fig.2.2. Any distortion from a straight line gives indication of the existence of nonlinearity. When FRF data from nonlinear structures are to be analysed, the effect of a stiffness nonlinearity will show up in the real part of the reciprocal of receptance data while the effect of a damping nonlinearity will appear only in the imaginary part of the data. To illustrate this point, FRF data with stiffness and damping nonlinearity measured from simulated analogue circuits are analysed and the real part ($R(\omega)$ *vs* ω^2) and imaginary part ($I(\omega)$ *vs* ω^2) of the reciprocal FRF are shown in Figs.2.3 & 2.4. In the case of stiffness nonlinearity, distortion of the reciprocal FRF data only appears in the real part, as shown in Fig.2.3. On the other hand, as shown in Fig.2.4, when damping nonlinearity is considered, its effect is clearly observed to be confined to the imaginary part of the reciprocal FRF data.

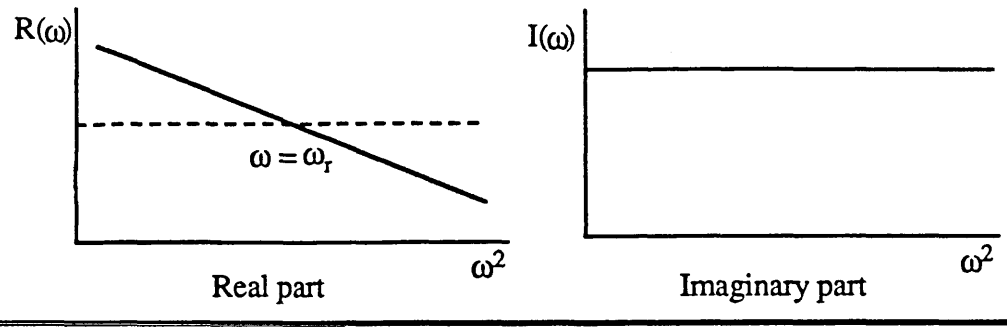


Fig.2.2 - Real and Imaginary Parts of Linear Reciprocal FRF Data

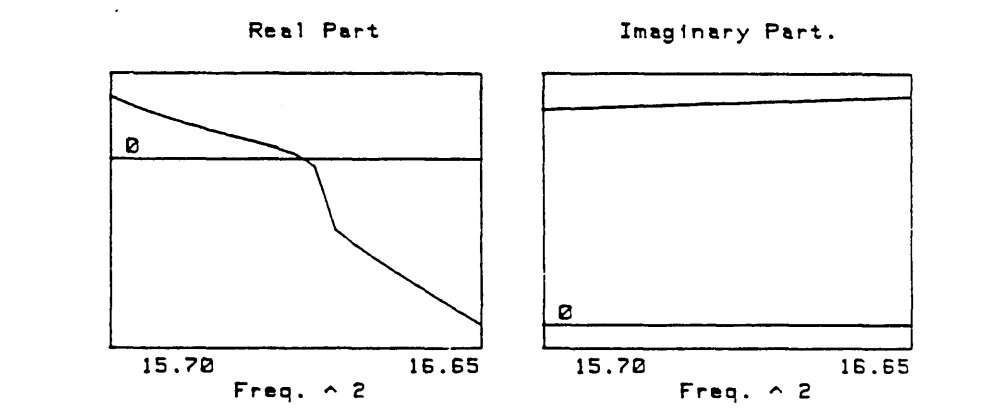


Fig.2.3 - Real and Imaginary Parts of Reciprocal FRF Data with Stiffness Nonlinearity

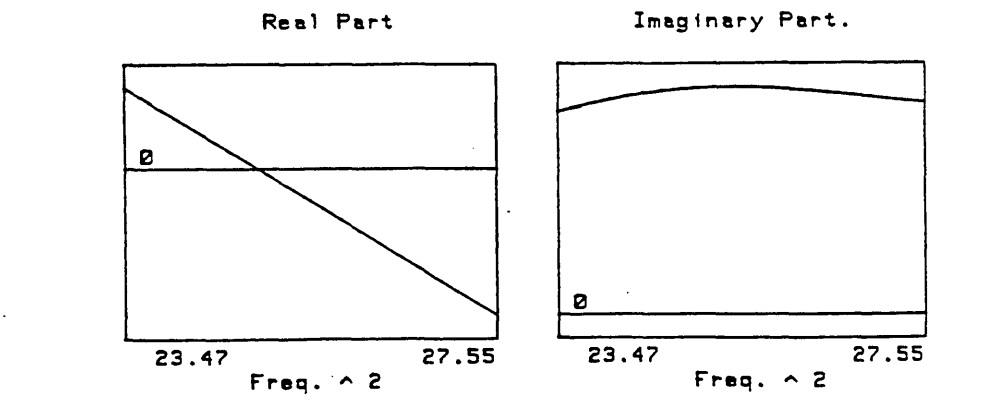


Fig.2.4 - Real and Imaginary Parts of Reciprocal FRF Data with Damping Nonlinearity

So far, it has been demonstrated that both Bode and reciprocal receptance plots can be used to detect the existence of nonlinearity, and that the latter technique can tell whether the nonlinearity is of stiffness or damping type in the case where the measured modes are effectively real. However, these methods can only be used to provide rough and basic demonstration of the existence of nonlinearity. It is not possible to quantify the extent of the nonlinearity based on these methods. In what follows, an alternative method of nonlinearity detection - the isometric damping plot technique - will also be discussed.

2.3.2 ISOMETRIC DAMPING PLOT TECHNIQUE

It has been established [1] that structural nonlinearity can be detected by inspection of the isometric damping plots which can be calculated from measured FRF data. The argument which supports this technique is generally believed to be that structural nonlinearity (usually stiffness nonlinearity) distorts the spacing of frequency response data around the Nyquist 'circle' from their positions when no nonlinearity exists. Since the distortion caused by nonlinearity is systematic, the consequent distortion of the damping estimate plot using different pairs of points around the Nyquist circle will display a specific pattern depending on the type of nonlinearity. These patterns in the damping plot can then be recognised and compared to detect and possibly to identify the nonlinearity. The mathematical basis of this technique will be discussed next with the new explanation for the reason why damping estimates vary when different pairs of frequency points are used.

Again, suppose that the residual effect of other modes can be neglected and that the modal constant is effectively real for the mode to be analysed, then the receptance $\alpha_{jk}(\omega)$ around r^{th} mode can be expressed as that of equation (2-1). When $\alpha_{jk}(\omega)$ is plotted in the Nyquist plane, a circle as shown in Fig.2.5 can be obtained. If the data are measured on a linear structure, then the damping loss factor of r^{th} mode η_r can be calculated as follows:

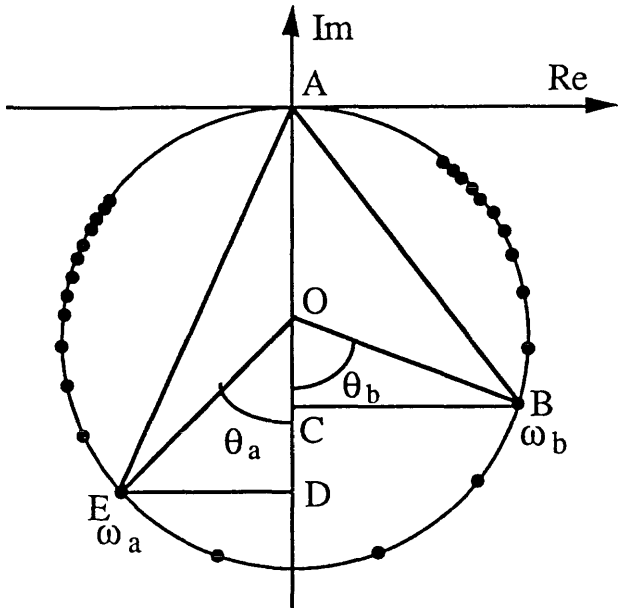


Fig.2.5 - Nyquist Circle of Receptance Data

$$tg\frac{\theta_b}{2} = \frac{|BC|}{|AC|} = \frac{|Re(\alpha_{jk}(\omega_b))|}{|Im(\alpha_{jk}(\omega_b))|} = \frac{\omega_r^2 - \omega_b^2}{\eta_r \omega_r^2} \tag{2-3}$$

$$tg\frac{\theta_a}{2} = \frac{|DE|}{|AD|} = \frac{|Re(\alpha_{jk}(\omega_a))|}{|Im(\alpha_{jk}(\omega_a))|} = -\frac{\omega_r^2 - \omega_a^2}{\eta_r \omega_r^2} \tag{2-4}$$

Adding equations (2-3) and (2-4), the damping loss factor η_r is given by:

$$\eta_r = \frac{\omega_a^2 - \omega_b^2}{\omega_r^2 \left(tg\frac{\theta_b}{2} + tg\frac{\theta_a}{2}\right)} \tag{2-5}$$

When different combinations of points (ω_b,ω_a) are used, a flat plane which is the surface plot of the estimated damping ratio $\eta_r(\omega_b,\omega_a)$ against its two variables ω_b and ω_a, can be obtained in the case of linear FRF data.

On the other hand, if the measured FRF data are from a nonlinear structure, distortion of the isometric damping plot (no longer a flat plane) calculated based on equation (2-5) will, in general, be expected as shown in Fig.2.6 for the data measured from a Beam/Absorber structure shown in figure 2.1.

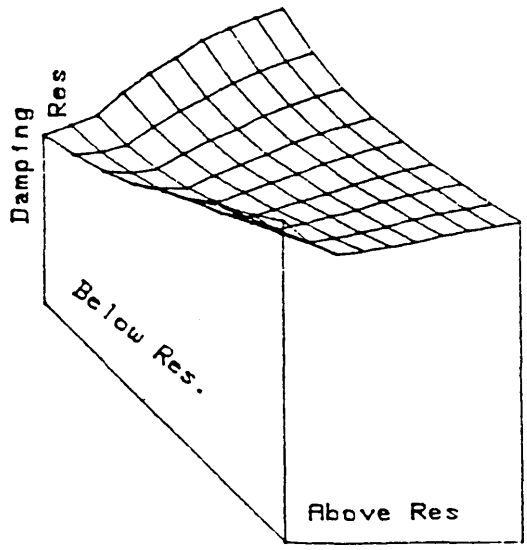


Fig.2.6 - Isometric Damping Plot for Data Measured from Beam-absorber Structure

The reason why this distortion occurs is discussed below. Suppose the FRF data are measured from a system with stiffness nonlinearity, then the natural frequency ω_r of r^{th} mode which is sensitive to nonlinearity becomes a function of response amplitudes $\omega_r=\omega_r(\hat{X})$ (where $\hat{X}$ is the response amplitude of certain coordinate) and since different frequency points have different response amplitudes for data measured with constant force, the effective natural frequencies at different data points are therefore, different. With this in mind, equation (2-3) and (2-4) become:

$$tg\frac{\theta_b}{2} = \frac{|BC|}{|AC|} = \frac{|Re(\alpha_{jk}(\omega_b))|}{|Im(\alpha_{jk}(\omega_b))|} = \frac{\omega_r^2(\hat{X}_b) - \omega_b^2}{\eta_r\,\omega_r^2(\hat{X}_b)} \tag{2-6}$$

$$tg\frac{\theta_a}{2} = \frac{|DE|}{|AD|} = \frac{|Re(\alpha_{jk}(\omega_a))|}{|Im(\alpha_{jk}(\omega_a))|} = -\frac{\omega_r^2(\hat{X}_a) - \omega_a^2}{\eta_r\,\omega_r^2(\hat{X}_a)} \tag{2-7}$$

The exact damping loss factor η_r can be calculated as:

$$\eta_r = \frac{\omega_a^2 - \omega_b^2 + \omega_r^2(\hat{X}_b) - \omega_r^2(\hat{X}_a)}{\omega_r^2(\hat{X}_b)\,tg\frac{\theta_b}{2} + \omega_r^2(\hat{X}_a)\,tg\frac{\theta_a}{2}} \tag{2-8}$$

In the case when damping is linear, the calculated damping loss factors based on equation (2-8) will be constant and therefore, it is not difficult to see that the distortion of the isometric damping plot obtained based on equation (2-5) is determined by the the

difference between the estimates of equation (2-5) and (2-8) which, to the first-order approximation, becomes:

$$\Delta\eta_r \approx \frac{\omega_r^2(\hat{X}_b) - \omega_r^2(\hat{X}_a)}{\omega_r^2(\hat{X}_b)\ \mathrm{tg}\ \frac{\theta_b}{2} + \omega_r^2(\hat{X}_a)\ \mathrm{tg}\ \frac{\theta_a}{2}} \tag{2-9}$$

From equation (2-9), it can be seen clearly that in the case of stiffness nonlinearity, the distortion of the damping plot calculated based on equation (2-5) is caused by the different response amplitudes and therefore different natural frequencies of various frequency points chosen.

The isometric damping plot technique, as has been demonstrated, can be used to detect the existence of nonlinearity. However, as in the case of nonlinearity detection based on Bode and Nyquist plot techniques, anything beyond detection in the identification of nonlinearity will be truly difficult because the method is of a qualitative nature rather than quantitative, although it has been suggested that, by comparing the different distortion patterns of damping plot of commonly encountered nonlinearities, the identification of the type of nonlinearity may become possible. With a more ambitious objective of quantifying nonlinearity, the Inverse Receptance method was developed [5]. The method will be presented next and its limitations when applied to FRF data measured from practical nonlinear structures will be discussed.

2.3.3 INVERSE RECEPTANCE METHOD

As discussed before, nonlinearities in FRF data will cause distortion of a plot of the inverse receptance data and such characteristics as plot distortion have been employed for the detection of the existence of nonlinearity. These inverse receptance data can be further employed for the purpose of quantifying nonlinearity. In this section, the Inverse Receptance method [5] is introduced. The limitations of the method for the analysis of practical nonlinear structures will be discussed and further possible improvements will be pointed out.

For a nonlinear SDOF system, the natural frequency $\omega_n(\hat{X})$ and damping loss factor $\eta(\hat{X})$ are, in general, response amplitude $\hat{X}$ dependent. With this in mind, the reciprocal of the receptance can be expressed as:

$$\frac{1}{\alpha(\omega)} = \frac{\omega_n^2(\hat{X}) - \omega^2}{A} + i\frac{\eta(\hat{X})\,\omega_n^2(\hat{X})}{A} \tag{2-10}$$

where A is the modal constant which is assumed to be real. Separate equation (2-10) into its real and imaginary parts, so that

$$\mathrm{Re}(1/\alpha(\omega)) = \frac{\omega_n^2(\hat{X}) - \omega^2}{A} \tag{2-11}$$

$$\mathrm{Im}(1/\alpha(\omega)) = \frac{\eta(\hat{X})\,\omega_n^2(\hat{X})}{A} \tag{2-12}$$

Suppose that the input force signal $F(\omega)$ is also recorded during the measurement, so that the response amplitude at each frequency can be easily calculated as:

$$|\hat{X}(\omega)| = |F(\omega)|\,|\alpha(\omega)| \tag{2-13}$$

It becomes clear that if the modal constant A can be estimated by some means, then the relationships of $\omega_n(\hat{X})$ *vs* $\hat{X}$ and $\eta(\hat{X})$ *vs* $\hat{X}$ can be calculated based on equations (2-11) and (2-12) as:

$$\omega_n^2(\hat{X}) = \omega^2 + A\,\mathrm{Re}(1/\alpha(\omega)) \tag{2-14}$$

$$\eta(\hat{X}) = A\frac{\mathrm{Im}(1/\alpha(\omega))}{\omega_n^2(\hat{X})} \tag{2-15}$$

where $\hat{X}$ is the response amplitude at frequency ω, and can be calculated from equation (2-13). The calculation of the modal constant A in this Inverse Receptance method is based on a trial-and-error approach and the criterion for the correct estimation of A is based on the fact in which satisfactory results have been obtained.

Based on this method, FRF data measured from analogue computer circuit with cubic hardening stiffness nonlinearity have been analysed and the results are shown in Fig.2.7. The effect of a hardening stiffness nonlinearity is clearly demonstrated.

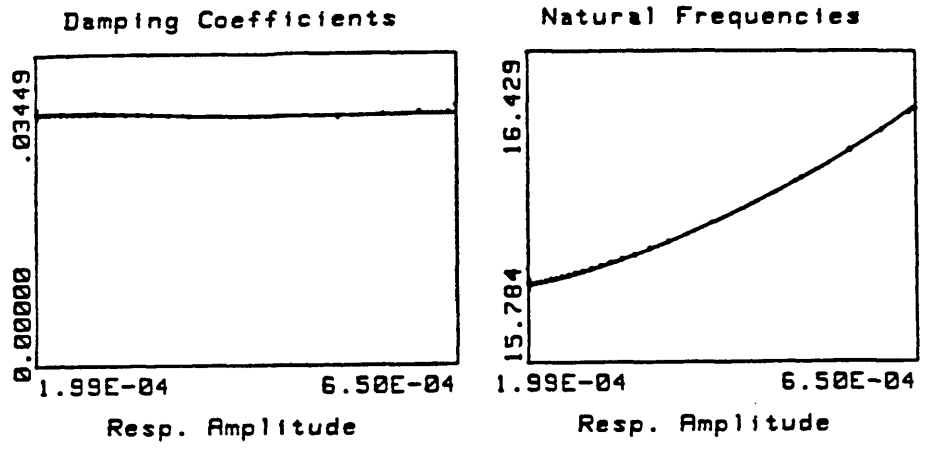


Fig.2.7 - Analysis of Stiffness Nonlinearity Using Inverse Receptance Method

However, an assumption that the modal constant of the mode to be analysed should be *real* and *constant* was made during the development of the Inverse Receptance method. The validity of this assumption when analysing FRF data measured from practical nonlinear structures will be discussed later on but, the influence of mode complexity on the estimation of $\omega_n(\hat{x})$ and $\eta(\hat{x})$ based on this method will be discussed here. Suppose the modal constant is complex and can be expressed as $A(\cos\theta+i\sin\theta)$, then equations (2-14) and (2-15) becomes:

$$\omega_n^2(\hat{x}) = \omega^2 + (\mathrm{Re}(1/\alpha(\omega))A\cos\theta + \mathrm{Im}(1/\alpha(\omega))A\sin\theta) / \cos 2\theta \qquad (2\text{-}16)$$

$$\eta(\hat{x}) = \frac{\mathrm{Im}(1/\alpha(\omega))A - (\omega_n^2(\hat{x}) - \omega^2)\sin\theta}{\omega_n^2(\hat{x})\cos\theta} \qquad (2\text{–}17)$$

Comparing equations (2-16) and (2-17) with those of (2-14) and (2-15), and bearing in mind that modal parameter changes due to nonlinearity are usually of second order, a small degree of complexity could seriously impair the results obtained based on the Inverse Receptance method just outlined. To illustrate this point, the same FRF data as those shown in Fig.2.7, but with 10° modal constant complexity artificially added are analysed using the method and the results are shown in Fig.2.8. In this case, not only is the damping value seriously in error (it should be constant for data points on either side of the resonance since the damping is linear), the calculated natural frequencies are incorrect as well (compare with figure 2.7). This demonstrates the limitation of the Inverse Receptance method and recommends the necessity of further development so that nonlinearity can be analysed accurately when measured modes become complex.

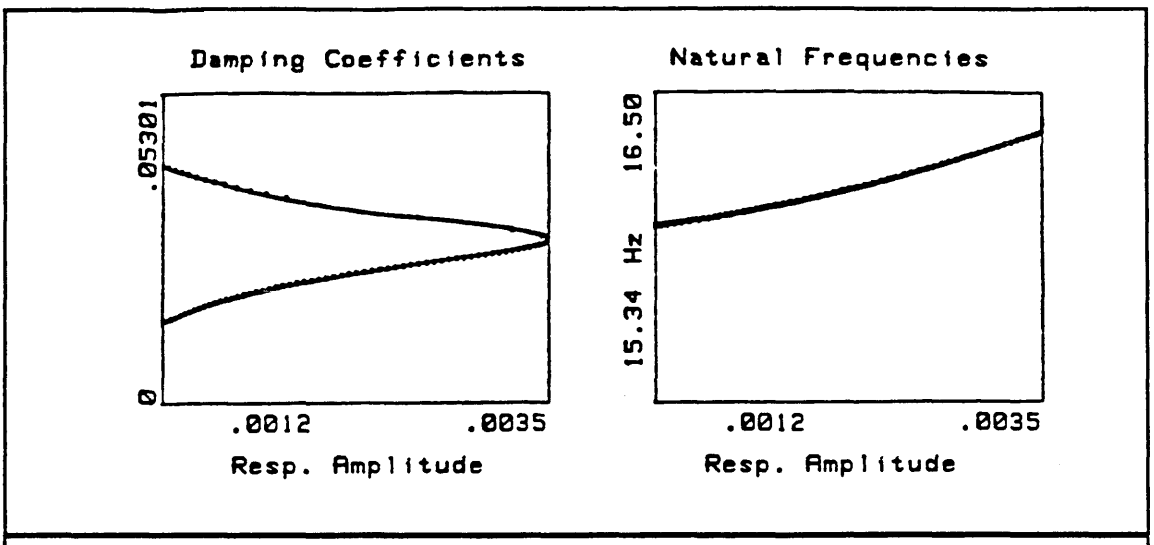


Fig.2.8 - Analysis Results of Complex Mode Using Inverse Receptance Method

2.3.4 COMMENTS ON CURRENT NONLINEAR MODAL ANALYSIS METHODS

So far, some of the currently available nonlinear modal analysis methods (based on the extension of linear modal analysis) have been reviewed. Based on the systematically abnormal behaviour of measured nonlinear FRF data, qualitative methods as discussed above can be used to detect the existence of nonlinearity. By comparing FRF data measured at different input force levels, the overlaid Bode plots can be used to check whether or not the measured data display nonlinearity. On the other hand, when FRF data are presented in their reciprocal form, the effect of stiffness and damping nonlinearities can be separated into the real and imaginary parts of the data and, by examining the plots of both parts, the type of existing nonlinearity (stiffness or damping) can be revealed. Also, the isometric damping plot technique can be used to detect the existence of nonlinearity by investigating the variation of damping ratios calculated using different pairs of frequency points on the Nyquist circle. The reason for this damping variation in the case of stiffness nonlinearity, as demonstrated, is due to the different response amplitudes, and so different effective natural frequencies, of different receptance data points.

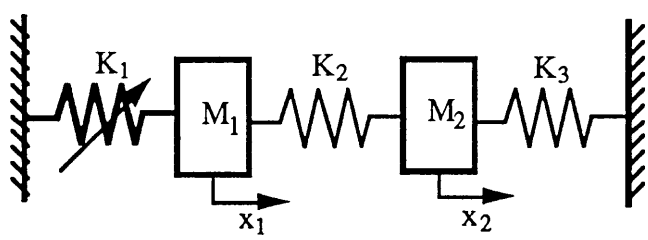
With these methods available, the task of detecting the existence of nonlinearity can be accomplished reasonably successfully if the structural nonlinearity has some contribution to the measured first-order FRFs. However, since all these methods are qualitative in

nature, it is difficult for them to establish the extent of nonlinearity which a structure possesses.

The Inverse Receptance method seeks to quantify nonlinearity by establishing the relationship between modal parameters and response amplitudes: $\omega_r(\hat{x})$ *vs* $(\hat{x})$ and $\eta_r(\hat{x})$ *vs* $(\hat{x})$. The method was developed based on an assumption that the modal constant of the mode to be analysed is real and constant. Although valid for FRF data measured from nonlinear SDOF systems, for those measured on practical nonlinear structures, this assumption is, in general, no longer valid for following reasons:

(i) measured data may contain mode complexity;
(ii) the modal constant of a mode is, in theory, a function of response amplitude.

The effect of mode complexity on the analysis results based on Inverse Receptance method has been demonstrated and detailed discussions on the existence of genuine complex modes will be presented later on. Here, only the second point (the modal constant of a nonlinear system is a function of response amplitude) will be illustrated based on a 2DOF system with cubic stiffness nonlinearity as shown in Fig.2.9. Assuming K_1 is nonlinear and can be expressed as $K_1=400000(1+\beta\hat{x}^2)$N/m (where $\hat{x}$ is the vibration amplitude of mass m_1 when it vibrates sinusoidally) and solving the eigenvalue problem of this system, for the first mode, the natural frequency $\omega_1(\hat{x})$ and modal constant $A_{11}(\hat{x})$ of $\alpha_{11}(\omega)$ can be expressed as:



$K_2 = K_3 = 400000$N/m $\quad M_1 = M_2 = 1.0000$ kg

$K_1 = 400000\,(1 + \beta\hat{x}^2)$

Fig.2.9 - A 2DOF Nonlinear System

$$\omega_1^2(\hat{x}) = \frac{4 + \beta\hat{x}^2 - \sqrt{4 + \beta^2\hat{x}^4}}{2} \times 400000 \qquad (2\text{-}18)$$

$$A_{11}(\hat{X}) = \frac{\beta\,\hat{X}^2 + \sqrt{4 + \beta^2\hat{X}^4}}{4 + \beta^2\hat{X}^4 + \beta\,\hat{X}^2\sqrt{4 + \beta^2\hat{X}^4}} \tag{219}$$

From equations (2-18) and (2-19), it is clear that if cubic stiffness is introduced as shown in figure 2.9, both natural frequency and modal constant of mode 1 are functions of response amplitude. The relationships $\omega_1(\hat{X})$ *vs* $\hat{X}$ and $A_{11}(\hat{X})$ *vs* $\hat{X}$ are illustrated in Fig.2.10 with β=1.0 and nondimensionalised response $\hat{X}$ = 0. - 1.0.

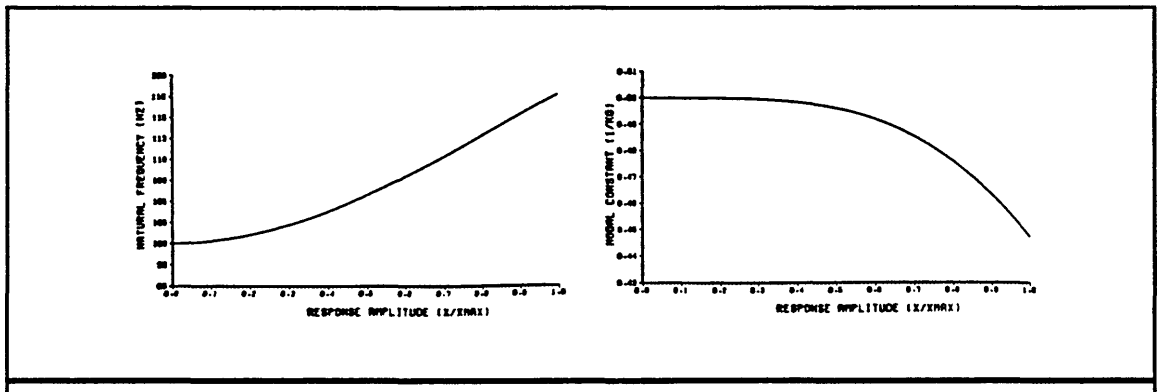


Fig. 2.10 - Natural Frequencies and Modal Constants of a Nonlinear System

As far as the quantification of nonlinearity is concerned, since FRF data measured from practical nonlinear structures usually contain mode complexity and the modal constant of the mode to be analysed cannot be assumed to be constant, the analysis results obtained based on the Inverse Receptance method can be erroneous and sometime misleading. Therefore, it becomes necessary to develop more realistic techniques so that the complexity of the mode and the variation of modal constant can be taken into account and more accurate modal parameters of nonlinear structures can be obtained.

2.4 COMPLEXITY OF VIBRATION MODES

2.4.1 THEORETICAL BASIS

There exist two different types of mode known as real modes and complex modes in structural vibration analysis. In real mode vibration, individual elements of a system move exactly in or out of phase with each other while in the case of complex modes, individual elements vibrate with different phase angles (relative to each other). The reason for the existence of complex modes is known to be a nonproportional distribution of the structure's damping. However, the degree of complexity of modes when the damping is nonproportional is largely determined by the closeness of the natural frequencies of the

system. In what follows, the necessary and sufficient conditions for the existence of complex modes, the influence of mode spacing on the complexity of a mode and the relationship between viscous and hysteretic damping models from mode complexity point of view, will be discussed.

It is well known that an undamped linear dynamic system described by

$$[M]\{\ddot{x}\} + [K]\{x\} = \{0\} \tag{2-20}$$

possesses real modes when [M] is nonsingular and $([M]^{-1}[K])$ has a full set of eigenvectors, [31]. Such real modes can be used to find the principal coordinates in which the equations of motion of the system are decoupled. Suppose $[\phi]$ is the mass normalised modeshape matrix and let $\{x\}=[\phi]\{p\}$, then equation (2-20) becomes:

$$[M][\phi]\{\ddot{p}\} + [K][\phi]\{p\} = \{0\} \tag{2-21}$$

Pre-multiply equation (2-21) by $[\phi]^T$ and since $[\phi]^T[M][\phi]=[I]$ and $[\phi]^T[K][\phi]=[\cdot\lambda.]$, then equation (2-21) can be decoupled in terms of principal coordinates $\{p\}$ as:

$$\{\ddot{p}\} + [\cdot\lambda.]\{p\} = \{0\} \tag{2-22}$$

In the presence of damping (assuming viscous damping for the convenience of analysis), equation (2-20) is modified to become

$$[M]\{\ddot{x}\} + [C]\{\dot{x}\} + [K]\{x\} = \{0\} \tag{2-23}$$

In this case, the criterion for the existence of real modes of the damped system is that the real modes of the corresponding conservative system (without damping) can be used to decouple the equations of motion of the damped system. For damped systems, in general, the decoupling property is violated and the modes become complex. However, certain conditions on the form of the damping matrix have been found under which a damped system can still possess real modes. Such damping condition have been discussed in detail by Caughey [32-33] who pointed out that the sufficient condition for the existence of real modes in a damped system is that the damping matrix of the system can be expressed as:

$$[C] = \sum_{s=1}^{N} \beta_s [M] ([M]^{-1}[K])^{s-1} \tag{2-24}$$

where N is the dimension of the system. In the case when $\beta_1=\beta_2=1$ and $\beta_s=0$ (s=3,N), equation (2-24) becomes the familiar Rayleigh damping which is

$$[C] = \beta_1 [M] + \beta_2 [K] \tag{2-25}$$

To prove the sufficiency of equation (2-24), pre-multiply both sides of equation (2-24) by $[\phi]^T$ and post-multiply by $[\phi]$, then

$$[\phi]^T [C] [\phi] = \sum_{s=1}^{N} \beta_s [\cdot\lambda.]^{s-1} = [\cdot\mu.] \tag{2-26}$$

From equation (2-26), since the damping matrix is diagonalised by the real modes of the corresponding conservative system, these real modes are also the real modes of the damped system. On the other hand, if the corresponding conservative system has no repeated eigenvalues, then condition described in equation (2-24) is also the necessary condition. To illustrate this point, rewrite $\sum_{s=1}^{N} \beta_s[\cdot\lambda.]^{s-1} = [\cdot\mu.]$ into linear algebraic equations in terms of unknowns $\{\beta\}$ as:

$$\begin{bmatrix} 1 & \lambda_1 & \cdots & \lambda_1^{N-1} \\ 1 & \lambda_2 & \cdots & \lambda_2^{N-1} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ 1 & \lambda_N & \cdots & \lambda_N^{N-1} \end{bmatrix} \begin{Bmatrix} \beta_1 \\ \beta_2 \\ \cdot \\ \cdot \\ \cdot \\ \beta_N \end{Bmatrix} = \begin{Bmatrix} \mu_1 \\ \mu_2 \\ \cdot \\ \cdot \\ \cdot \\ \mu_N \end{Bmatrix} \tag{2-27}$$

The coefficient matrix is a Vandermonde matrix [V] whose determinant is given as:

$$\det(V) = \prod_{i>j\geq 1}^{N} (\lambda_i - \lambda_j) \tag{2-28}$$

Therefore, if all eigenvalues are distinct, then $[V]^{-1}$ exists and the unknown coefficients $\{\beta\}$ can be uniquely determined.

When the damping matrix of the system does not satisfy the condition set in equation (2-24), then the damping matrix cannot be diagonalised using the real modes of the corresponding conservative system and the modes of the damped system will in general become complex. Since the extent of departure of a given damping matrix from that of equation (2-24) (often referred as nonproportionality) determines how complex the modes of the system will be, the quantification of such departure becomes necessary in order to study complex modes. Some research work on the quantification of nonproportionality of a given damping matrix and the complexity of modes has been reported and the relationship between complexity of modes and nonproportionality of damping matrix has been investigated [34].

The degree of complexity of a certain mode depends on the closeness of the natural frequencies of the system. In the case when all the modes of the system are well separated, even though the damping matrix is nonproportional (localised damping for example), the modes will not be substantially complex. Theoretically, considerably complex modes can only occur when modes become close. This effect of mode spacing on the complexity of modes is to be discussed below based on the perturbation theory.

In order to illustrate the effect of mode spacing on the complexity of modes, a hysteretic damping model is assumed in the analysis although the relationship between the hysteretic and viscous damping models will be discussed later. Also, assume that the structural damping matrix [D] is of second order in its Euclidean norm sense when compared with the system's stiffness matrix [K], then to first order approximation, the r^{th} modeshape of the damped system $\{\varphi\}_r$ can be expressed in terms of the modal parameters of the corresponding conservative system and damping matrix [D] as:

$$\{\varphi\}_r = \{\phi\}_r + i \sum_{s=1;\ s\neq r}^{N} \frac{\{\phi\}_r^T [D] \{\phi\}_s}{\lambda_r - \lambda_s} \{\phi\}_s \qquad (2\text{-}29)$$

In the case when the r^{th} mode is well isolated, then $\{\phi\}_r^T[D]\{\phi\}_s$ (which is a scalar), will be of second order compared with $(\lambda_r-\lambda_s)$ and, therefore, $\{\varphi\}_r$ will be effectively real. However, if there are close modes, say mode r and mode (r+1), then when s=(r+1),

$\{\phi\}_r^T[D]\{\phi\}_{r+1}$ will no longer be of second order of $(\lambda_r-\lambda_{r+1})$ and $\{\varphi\}_r$ will become considerably complex.

In the above discussion, both viscous and hysteretic damping models have been used. The relationship between these two damping models and the complexity of modes need to be discussed.

In the case of viscous damping, the eigenvalue problem of the system becomes quadratic as:

$$([M]\, s^2 + [C]\, s + [K])\, \{x\} = \{0\} \tag{2-30}$$

While the standard eigenvalue problem is in the form of

$$([A]\, \lambda + [B])\, \{z\} = \{0\} \tag{2-31}$$

where [A] and [B] can be complex matrices in general. In order to solve the quadratic eigenvalue problem given in equation (2-30), some mathematical transformations are required, namely:

$$[A] = \begin{bmatrix} [C] & [M] \\ [M] & [0] \end{bmatrix},\ [B] = \begin{bmatrix} [K] & [0] \\ [0] & -[M] \end{bmatrix},\ \{z\} = \begin{Bmatrix} x \\ \dot{x} \end{Bmatrix} \text{ and } \lambda = s$$

By solving equation (2-31), the eigenvalues and the so-called 'A-normalised' eigenvectors (normalised to the system's generalised mass matrix [A]) of the system can be obtained.

In the case of hysteretic damping, the eigenvalue problem becomes:

$$([M]\, \lambda + [K] + \mathbf{i}\, [D])\, \{x\} = \{0\} \tag{2-32}$$

Compared with equation (2-31), the solution to equation (2-32) is standard and since $[A] \equiv [M]$ and $[B] \equiv ([K] + \mathbf{i}[D])$ in this case, the eigenvectors for the hysteretic damping case are therefore mass-normalised (normalised to the mass matrix [M] of the system).

Because of the different normalisation procedures used when the different damping models are considered, the corresponding eigenvectors are apparently quite different, even for the case of proportionally damped systems although, in fact, they differ only by

a complex scaling factor. These differences in amplitude as well as phase angles of the corresponding eigenvectors often cause confusion to analysts and it is therefore necessary to establish the relationship between the 'A-normalised' and the mass-normalised eigenvectors. In the case of proportional damping, the corresponding r^{th} mode eigenvectors of hysteretically- (normalised to [M]) and viscously-damped (normalised to [A]) systems can be expressed, in theory, as $\{\psi\}_r = \chi_r\{\phi\}_r$ (where χ_r is a complex scaling factor). Substitute $\{\psi\}_r=\chi_r\{\phi\}_r$ into the 'A-normalisation' condition for the r^{th} mode together with $\{\phi\}_r^T[M]\{\phi\}_r=1$ and $\{\phi\}_r^T[C]\{\phi\}_r=2\omega_r\zeta_r$,

$$\begin{Bmatrix} \{\psi\}_r \\ i\omega_r\sqrt{1-\zeta_r^2}\,\{\psi\}_r \end{Bmatrix}^T \begin{bmatrix} [C] & [M] \\ [M] & [0] \end{bmatrix} \begin{Bmatrix} \{\psi\}_r \\ i\omega_r\sqrt{1-\zeta_r^2}\,\{\psi\}_r \end{Bmatrix} = 1 \qquad (2\text{-}33)$$

χ_r can be calculated as:

$$\chi_r = \frac{e^{\pm i\frac{\pi}{4}}}{\sqrt{2\omega_r(\sqrt{1-\zeta_r^2} - \zeta_r)}} \qquad (2\text{-}34)$$

From equation (2-34), it can be seen that in the case of proportional damping, the r^{th} 'A-normalised' modeshape for the case of viscous damping, $\{\psi\}_r$, is the corresponding mass-normalised modeshape $\{\phi\}_r$ for the case of hysteretic damping scaled by a factor of $\sqrt{2\omega_r(\sqrt{1-\zeta_r^2} - \zeta_r)}$ and a phase rotation of $\frac{\pi}{4}$. For the case of nonproportional damping, the relationship between these two damping models has been investigated in [35].

2.4.2 NUMERICAL EXAMPLE OF COMPLEX MODES

As discussed above, when the damping distribution of the structure is nonproportional, complex modes exist. However, the degree of complexity of a mode is mainly dependent on the closeness of the structure's natural frequencies. In order to illustrate these points, a numerical case study was carried out.

The system used in the numerical study is the 4DOF mass-spring system shown in Fig.2.11. The mass matrix [M], stiffness matrix [K] and hysteretic damping matrix [D] of the system are:

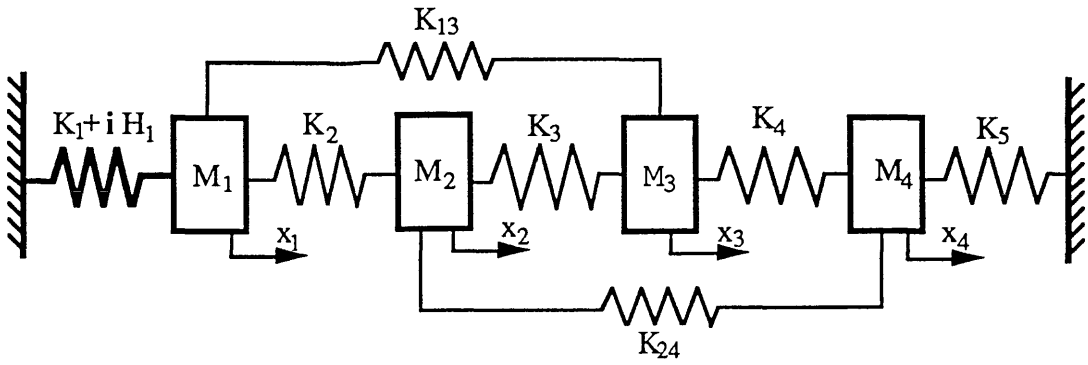


Fig. 2.11 - A 4DOF Mass-spring System

$$[M] = \begin{bmatrix} 0.800 & 0.000 & 0.000 & 0.000 \\ 0.000 & 1.005 & 0.000 & 0.000 \\ 0.000 & 0.000 & 1.000 & 0.000 \\ 0.000 & 0.000 & 0.000 & 0.800 \end{bmatrix} \text{(Kg)},$$

$$[K] = \begin{bmatrix} 3.00 & -1.00 & -1.00 & 0.00 \\ -1.00 & 3.25 & -1.25 & -1.00 \\ -1.00 & -1.25 & 3.25 & -1.00 \\ 0.00 & -1.00 & -1.00 & 3.00 \end{bmatrix} (\text{x}10^3 \text{ N/m})$$

$$[D] = \begin{bmatrix} 0.10 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 & 0.00 \end{bmatrix} (\text{x}10^3 \text{ N/m}).$$

The calculated eigenvalue matrix [$\dot{\lambda}$.] and eigenvector matrix [φ] are:

$$[\dot{\lambda}.] = \begin{bmatrix} 1066.2(1.0+\text{i}0.03) & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 3750.1(1.0+\text{i}0.00) & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 4517.3(1.0+\text{i}0.01) & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 4650.2(1.0+\text{i}0.026) \end{bmatrix}$$

$$[\varphi] = \begin{bmatrix} 0.495(-1.8°) & 0.006(175°) & 0.029(-37°) & 0.144(224°) \\ 0.676(0.5°) & 0.640(0.0°) & 0.712(180°) & 0.742(2.0°) \\ 0.387(-0.3°) & 0.640(180°) & 0.700(1.0°) & 0.668(176°) \\ 0.495(1.0°) & 0.474(0.0°) & 0.029(-46°) & 0.143(216°) \end{bmatrix}.$$

From the eigenvalue and eigenvector matrices, it can be seen that modes 1 and 2 are quite well separated and, as a result, their modeshapes are effectively real. While for modes 4 and 3, since they are very close in natural frequency, their modeshapes become quite complex when the damping is nonproportional, as it is in this case. Physically, complex modes can be explained as a kind of travelling wave which transfers energy from one part of the structure to another during vibration.

2.4.3 COMPLEX MODE FROM MEASUREMENTS

It is believed that for practical structures, most of the damping comes from joints [36]. Therefore, practical structures possess very nonproportional damping distribution and genuine complex modes exist. To demonstrate this, modal testing of a simple Beam/Absorber structure as shown in Fig.2.12 was carried out. The structure was found to be slightly nonlinear, as will be discussed again later on in this Chapter. However, during the test, the vibration amplitude of the structure was controlled to be constant at different sinusoidal excitation frequencies and, as a result, the FRF measured is exactly the FRF of a linear structure. One of the measured point FRFs (response and excitation at the same point) was analysed and, as shown in Fig.2.13, a mode complexity of about 15° is clearly demonstrated.

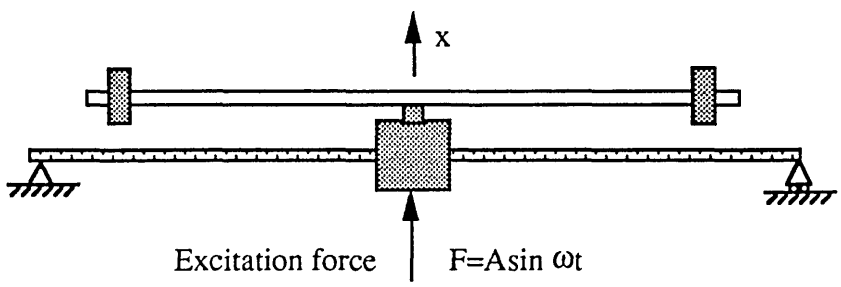


Fig.2.12 - The Beam/Absorber Structure*

* See Appendix III for more details

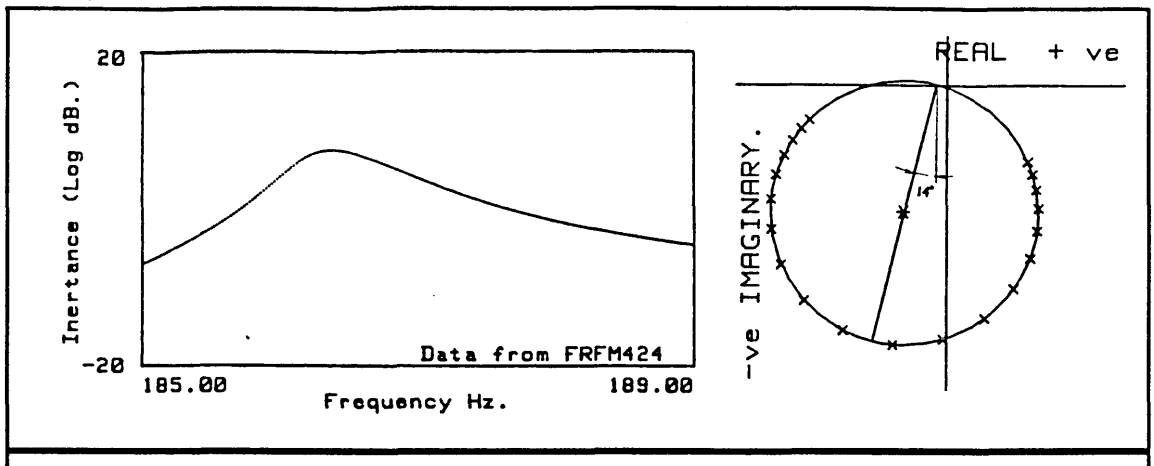


Fig.2.13 - Measured Point FRF with Mode Complexity

2.5 A NEW METHOD FOR NONLINEAR MODAL ANALYSIS OF COMPLEX MODES

So far, some of the most commonly-used nonlinear modal analysis methods have been reviewed and their limitations when applied to practically-measured data have been examined. In what follows, a new analysis method which avoids the aforementioned limitations will be proposed. The harmonic balance theory, on which the present new method is based, will be introduced together with its application conditions. In order to extend the method to MDOF systems, the residual effect (of other modes) on the analysis accuracy will be examined and the practical applicability of the method will be assessed by analysing data measured on practical nonlinear structures. Finally, the possibility of identifying physical characteristics of nonlinearity from analysed response amplitude-dependent modal data based on the new method, when an MDOF system is considered, will be discussed.

3.5.1 HARMONIC BALANCE THEORY

In the analysis of nonlinear systems, the harmonic balance method is frequently used where sustained oscillations exist. The theoretical basis of the harmonic balance analysis lies in the equivalent linearisation theory proposed by Krylov and Bogoliubov [28] for solving certain problems of nonlinear mechanics. To explain the concept of the harmonic balance method, an SDOF system with nonlinear restoring force $F(x, \dot{x})$ driven by a sinusoidal excitation is considered:

$$m\ddot{x} + F(x, \dot{x}) = F_0 \sin\omega t \qquad (2\text{-}35)$$

To solve the above problem by the harmonic balance approach, it is necessary to make a basic assumption that the variable x=x(t), appearing in the nonlinear function $F(x,\dot{x})$, is sufficiently close to a sinusoidal oscillation; that is,

$$x \approx A\sin(\omega t + \phi) \tag{2-36}$$

where the amplitude A, frequency ω and phase lag ϕ are constant. Therefore, the harmonic balance analysis belongs to those approximate methods of solving nonlinear differential equations which are based upon an assumed solution. As such, it requires that conditions for the assumed solution exist. Such an assumption is quite realistic since a nonlinear system may well exhibit periodic oscillations arbitrarily close to a pure sinusoid. If the variable x in the nonlinear function $F(x,\dot{x})$ has the sinusoidal form of (2-36), then the variable $y=F(x,\dot{x})$ is generally complex, but is also a periodic function of time. As such, it can be developed in a Fourier series:

$$y = N_0(\omega,A) + N_1(\omega,A)\,A\sin(\omega t + \phi) + \mathbf{i}\,N_2(\omega,A)\,A\sin(\omega t + \phi) + \text{harmonics} \tag{2-37}$$

When only the fundamental component is considered, the first three terms are

$$N_0 = \frac{1}{2\pi}\int_0^{2\pi} F(A\sin\psi,\omega A\cos\psi)\,d\psi \tag{2–38}$$

$$N_1 = \frac{1}{\pi A}\int_0^{2\pi} F(A\sin\psi,\omega A\cos\psi)\,\sin\psi d\psi \tag{2–39}$$

$$N_2 = \frac{1}{\pi A}\int_0^{2\pi} F(A\sin\psi,\omega A\cos\psi)\,\cos\psi d\psi \tag{2–40}$$

where $\psi = \omega t + \phi$. Coefficients N_0, N_1, N_2 are often referred as describing function coefficients.

If we consider the case where the nonlinear function $F(x,\dot{x})$ is symmetrical about the origin (although the analysis is equally applicable for the case of nonsymmetrical nonlinearities), the constant term N_0 in the Fourier series (2-37) is $N_0 \equiv 0$. The quantities defined in (2-39) and (2-40) are the coefficients of the describing function $N=N_1+iN_2$. To discuss the physical meaning of the describing function N defined above, suppose $F(x,\dot{x})=F(x)$ describes a backlash stiffness nonlinearity, as shown in Fig.2.14. Then, if

the input x (the response of the system) is a sinusoid, $x=\hat{x}\sin\omega t$, the output $y=F(x)$ will not be a pure sinusoid (assuming that $\hat{x}$ is big enough to exceed the system's linear regime). Expressing the output signal y in a Fourier series, the fundamental component y_1 will be

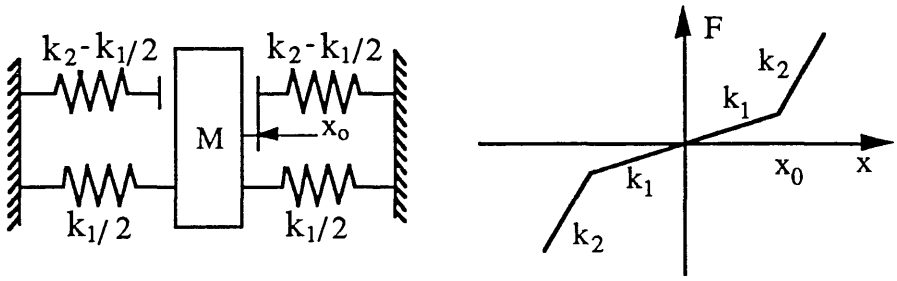


Fig.2.14 - Backlash Stiffness Nonlinearity

$$y_1 = \hat{y}_1 \sin\omega t \tag{2-41}$$

where $\hat{y}_1$ is the amplitude of the fundamental component which, according to the Fourier series theory, can be calculated as:

$$\hat{y}_1 = \frac{1}{\pi}\int_0^{2\pi} F(\hat{x}\sin\psi)\sin\psi d\psi \tag{2–42}$$

According to equation (2-39), N_1 has the form

$$N_1 = \frac{1}{\pi\hat{x}}\int_0^{2\pi} F(\hat{x}\sin\psi)\sin\psi d\psi \tag{2–43}$$

Therefore, the describing function $N=N_1$ is defined as the ratio between the amplitude $\hat{x}$ of the input signal x and the amplitude $\hat{y}_1$ of the fundamental component y_1 contained in the output; that is

$$N_1 = \frac{\hat{y}_1}{\hat{x}} \tag{2–44}$$

Compared with the definition of static stiffness, it can be seen that N_1 can be interpreted as the equivalent 'dynamic stiffness' of the nonlinear stiffness element corresponding to vibration amplitude $\hat{x}$. If the integral on the right hand side of equation (2-43) is

calculated with the explicit F(x) as shown in figure 2.14, we obtain the describing function coefficient $N_1=N_1(\hat{X})$ for the backlash stiffness case as:

$$N_1= k_1 - \frac{2}{\pi}(k_1 - k_2)[\arcsin(\frac{x_0}{\hat{X}}) + \frac{x_0}{\hat{X}}\sqrt{1 - \frac{x_0^2}{\hat{X}^2}}]\ ;\ \hat{X} \geq x_0 \qquad (2\text{-}45)$$

When the vibration amplitude $\hat{X}$ is given, the describing function coefficient $N_1(\hat{X})$ can be determined. In the case of nonlinear SDOF systems, since the mass property is usually linear, the vibration-amplitude dependent natural frequency $\omega_n(\hat{X})$ can be calculated from the describing function coefficient $N_1(\hat{X})$ as $\omega_n(\hat{X})=\sqrt{\frac{N_1(\hat{X})}{m}}$, as shown in Fig.2.15 for the above-mentioned system with backlash stiffness nonlinearity (k_1=5000N/m, k_2=10000N/m, m=1kg and x_0=0.001m). In fact, the describing function $N(\hat{X})=N_1(\hat{X})+iN_2(\hat{X})$ and the identified natural frequency $\omega_n(\hat{X})$ and damping coefficient $\eta(\hat{X})$ satisfy:

$$N_1(\hat{X}) = \omega_n^2(\hat{X})\, m \qquad (2\text{-}46)$$

$$N_2(\hat{X}) = \omega_n^2(\hat{X})\eta(\hat{X})\, m \qquad (2\text{-}47)$$

where m is the mass of the system which can be calculated from the identified modal constant A as m=1/A.

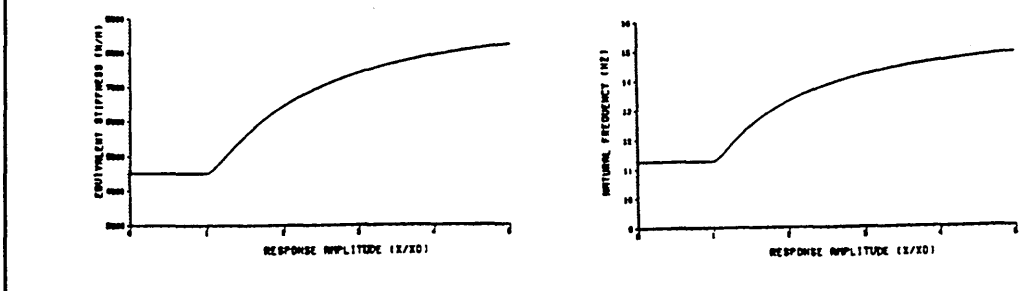


Fig.2.15 - Equivalent Stiffness and Natural Frequency of an SDOF Backlash System

To see how this harmonic balance theory can be applied to the measurement and analysis of nonlinear structures, the system shown in Fig.2.16 will be considered. When the system is excited by a sinusoidal force f(t)=Fsinωt, then after the transient dies away, the response of the system at any coordinate will be very close in its waveform to a sinusoid, as will be shown to be especially true when the excitation frequency is close to one of the resonance frequencies of the system. Therefore, corresponding to this specified excitation

condition, the equivalent stiffness which the nonlinear stiffness element exhibits can be calculated based on the harmonic balance theory. If, for example, the force amplitude F is kept constant as the excitation frequency varies, then corresponding to different frequencies, the vibration amplitude, and therefore the equivalent stiffness value of the nonlinear stiffness element, is different. The information on these different stiffness values is recorded in the measured first-order FRF data and by analysing these measured FRF data, the nonlinearity can be identified.

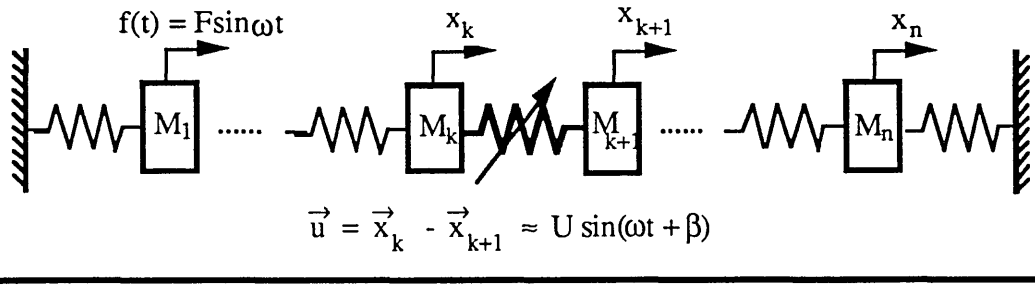


Fig.2.16 - A Nonlinear MDOF system

It should be noted that such a result is achieved only under the condition that it is not necessary to include the harmonics and combinational resonances in the response signal. For this condition to be valid, certain criteria should be satisfied by the linear part of a nonlinear system as well as the nature of the nonlinearity. These conditions are summarised here without mathematical proof, which can be found in [27]:

(i) the system to be analysed should have a narrow-band filter property so that the sub- and super-harmonic components will be heavily attenuated;
(ii) kω (k is an integer and ω is the excitation frequency) should not coincide with any of the natural frequencies of the system; and
(iii) the nonlinear function F(x,ẋ) should have finite partial derivatives with respect to x and ẋ.

Condition (i) can usually be satisfied because in the analysis of structural nonlinear systems, only the data points around the resonances are of interest and, therefore, the system acts as a very good narrow-band filter. As for condition (iii), most practically encountered nonlinearities, even for nonlinearities having relay (discontinuous) characteristics, have finite partial derivatives. However, condition (ii) is sometimes difficult to achieve because it depends on how the natural frequencies of the system are situated along the frequency axis and is thus the major source of analysis errors.

2.5.2 DESCRIPTION OF A NEW METHOD

Most practical nonlinearities (in both stiffness and damping elements) are response amplitude dependent and so if, in measurement, the response amplitudes at different frequencies are varied, then the effect of nonlinearities on the measured FRF data will be recorded. The main target of nonlinear modal analysis is to identify the nature of any nonlinearity by analysing thus-measured FRF data. As discussed, many different methods have been developed for detecting the existence of structural nonlinearities. Taking stiffness nonlinearity as an example, the nonlinearity can be exposed by observing the FRF data measured using different force or response control techniques, or by analysing the FRF data and examining the isometric damping plot, or by comparing the measured FRF data with their Hilbert transform pair [4], etc. With the more demanding objective of quantifying structural nonlinearity, the Inverse Receptance method was developed. However, the method was devised based on the assumption that the mode to be analysed should be real and the modal constant should be real and constant. As demonstrated, these assumptions are not usually valid when practical measured data are concerned. In order to remove these restrictions so that nonlinearities of practical structures can be analysed accurately, a new nonlinear modal analysis is introduced below.

According to harmonic balance theory, in the case of sinusoidal excitation, when a nonlinear structure vibrates at specific amplitude, there will be specific equivalent (linearised) stiffness and damping model as far as the first-order FRF is concerned. Therefore, measured FRF data generally contain information on a series of linear models. What the new method seeks to do is to calculate the modal parameters of these linear models together their corresponding response amplitudes so that the relationship between modal parameters and response amplitude can be established. Owing to the nature of resonance, it is always possible to find two frequency points in the measured FRF data - one on either side of the resonance - which have the same (or very similar) response amplitude. These two data points constitute a specific linear model corresponding to that specific response amplitude in the sense that all the modal parameters necessary to determine that linear(ised) model can be calculated just using these two receptance data points. The thus determined modal parameters are associated with that specific response level. Therefore, if there are many point pairs of different response amplitudes available around that resonance, a relationship between modal parameters of the mode and response amplitudes can be established.

Suppose $\alpha(\omega_1)$ and $\alpha(\omega_2)$ are known to correspond to a certain specific response level, one on either side of the resonance, then the following two mathematical equations can be established (assume the residual effect is negligible or has been removed at the moment and its influence on analysis accuracy will be discussed later)

$$\alpha(\omega_1) = \frac{A_r + iB_r}{\omega_r^2 - \omega_1^2 + i\eta_r\omega_r^2} \tag{2-48}$$

$$\alpha(\omega_2) = \frac{A_r + iB_r}{\omega_r^2 - \omega_2^2 + i\eta_r\omega_r^2} \tag{2-49}$$

Because equations (2-48) and (2-49) are complex algebraic equations, the four modal parameters ω_r, η_r, A_r and B_r can be determined. These parameters represent the linear model which corresponds to the chosen response amplitude. By examining different point pairs similar to $\alpha(\omega_1)$ and $\alpha(\omega_2)$, the characteristics $\omega_r(\hat{x})$, $\eta_r(\hat{x})$, $A_r(\hat{x})$ and $B_r(\hat{x})$ of the original nonlinear structure against vibration amplitude $\hat{x}$ can be revealed.

2.5.3 APPLICATION OF THE NEW METHOD TO THE ANALYSIS OF NONLINEAR SYSTEMS SIMULATED USING ANALOGUE CIRCUITS

The above method has been applied to several systems with various types of nonlinear stiffness or damping in order to assess fully the feasibility of the method. As an alternative means of solving nonlinear differential equations numerically, analogue circuits have been constructed to simulate nonlinear SDOF systems. Analogue computer FRF data with cubic stiffness and frictional damping as well as FRF data measured from the ETH/CIRP box (an electrical analogue computer device built to simulate nonlinearity) are employed and analysed.

For the case of a cubic stiffness nonlinearity, receptance FRF data with constant force input together with the calculated isometric damping plot are shown in Fig.2.17. From figure 2.17, the jump phenomenon which is typical of stiffness nonlinearity is clearly demonstrated and the isometric damping plot shows a well-defined surface distortion pattern. It can be seen that the nonlinearity is clearly indicated both from the receptance data plot and the damping plot. In practice, however, what is required is the quantification of the nonlinearity and not just its detection. By using this new method, the system's properties in terms of modal parameters versus response amplitude have been obtained as shown in Fig.2.18, from which it can be concluded that the damping is linear since the

damping coefficient does not change with response amplitude, while the stiffness is nonlinear and the degree of nonlinearity is quantified in the natural frequency versus response amplitude plot.

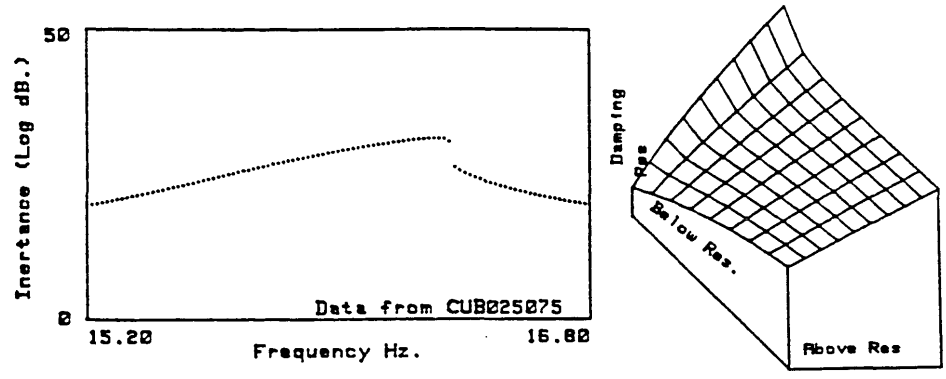


Fig.2.17 - FRF and Damping Plot of Analogue Computer Data

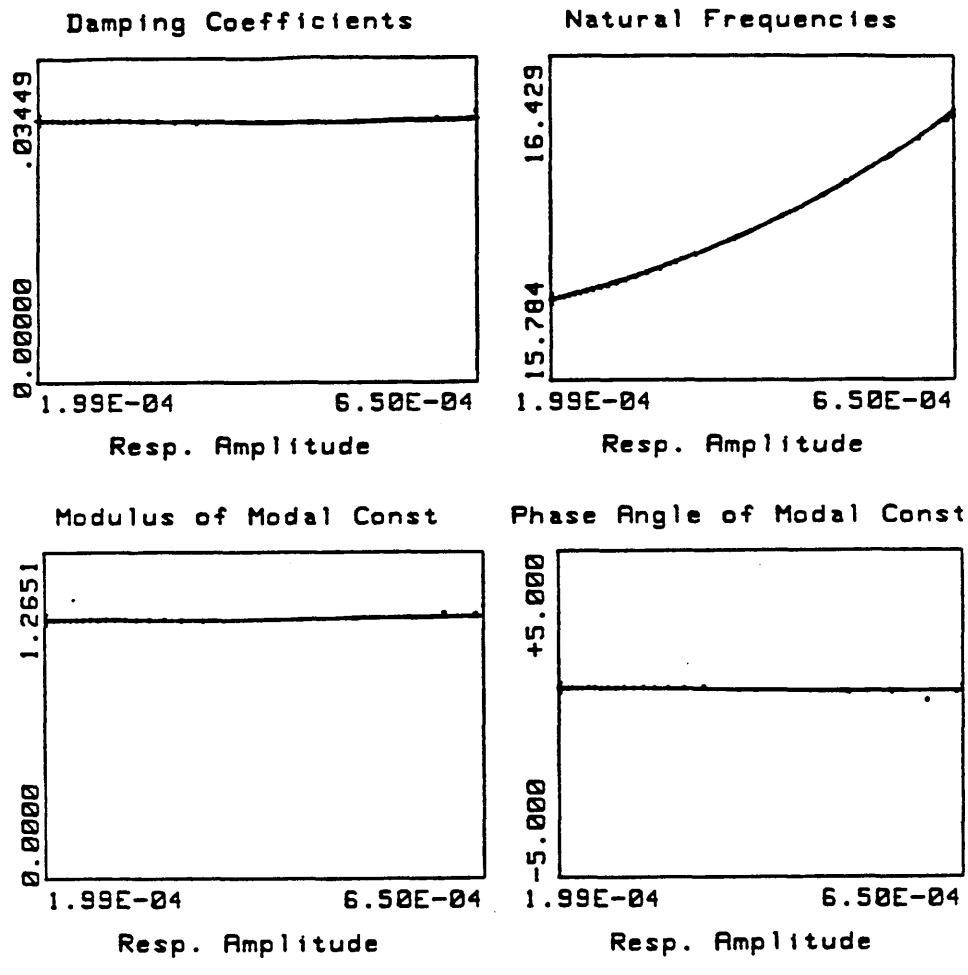


Fig.2.18 - Modal Parameters *versus* Vibration Amplitude

When the method is applied to the study of damping nonlinearity, the results are as encouraging as those for stiffness nonlinearity. As shown in Fig.2.19, the existence of dry friction damping can be suspected from the characteristic oval-shaped Nyquist plot and the distorted damping plot. As in the stiffness nonlinearity case, frequency response data measured from an analogue computer circuit with simulated frictional damping are analysed using this new method and the results shown in Fig.2.20. The damping coefficient decreases as response amplitude increases, which indicates damping of dry friction nature.

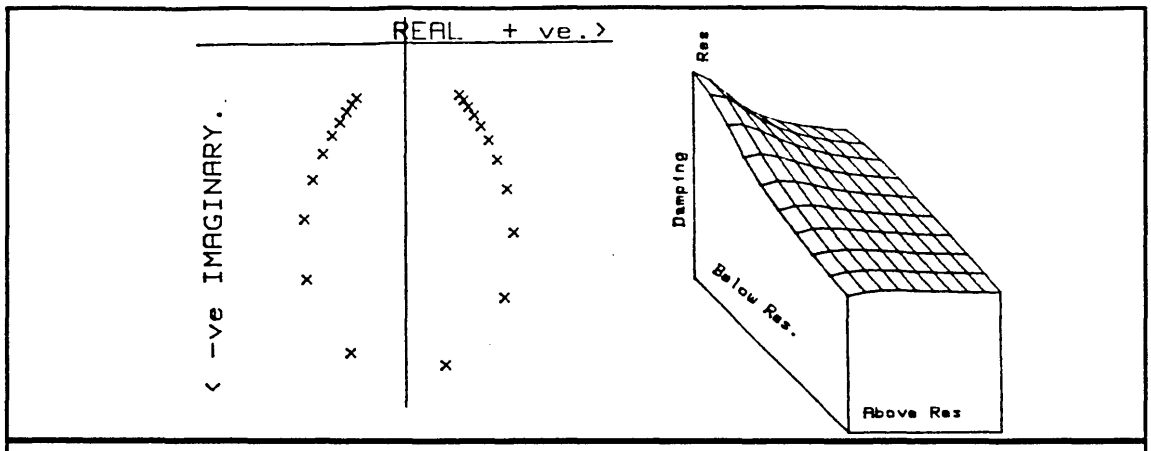


Fig.2.19 - Nyquist and Damping Plot of FRF Data with Frictional Damping

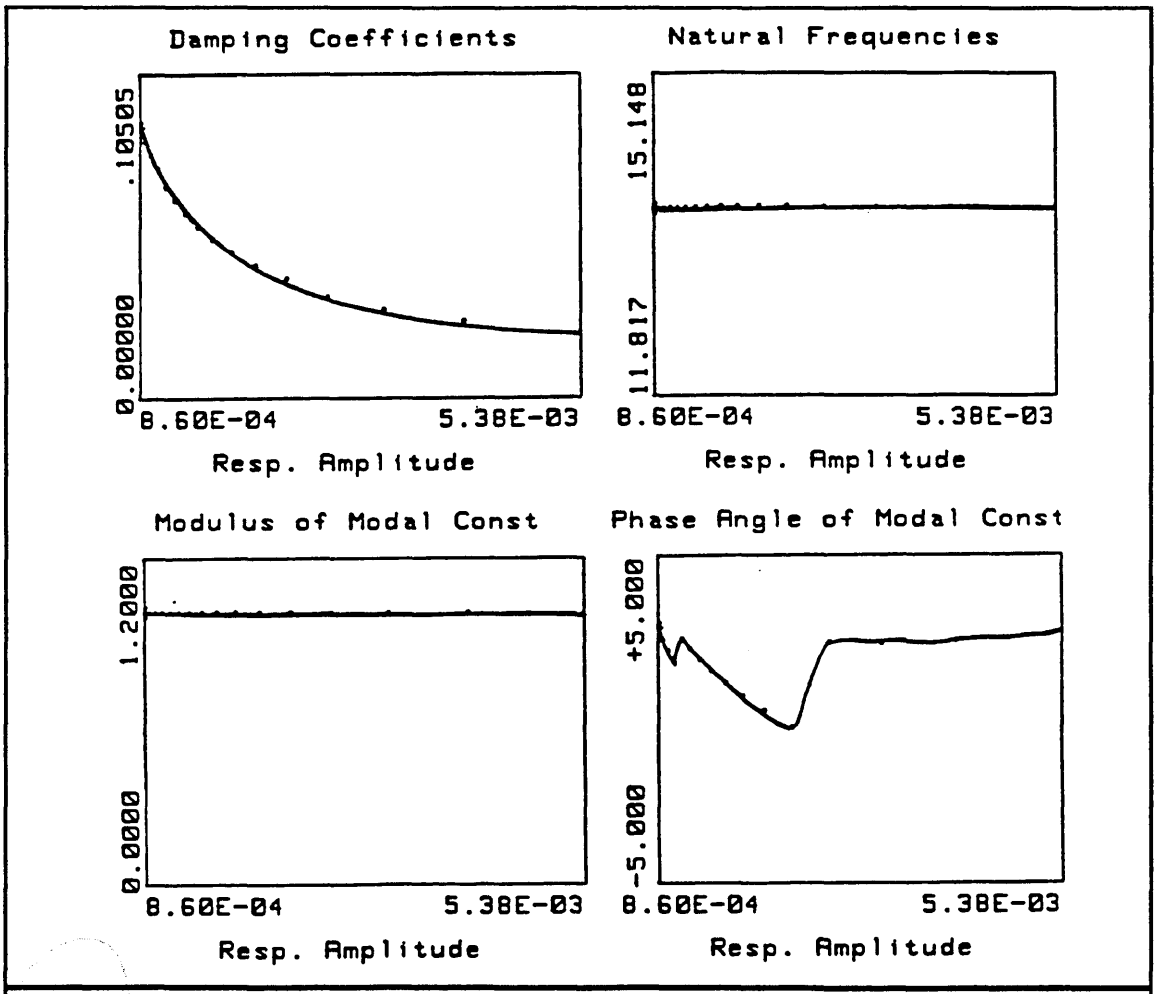


Fig.20 - Modal Parameters *versus* Vibration Amplitude (Simulated Friction Damping)

The ETH/CIRP box contains three circuits of unknown nonlinear SDOF systems and its purpose is to see whether these unknown nonlinear systems can be identified. The FRF data measured from one of the systems and the calculated damping plot are shown in

Fig.2.21. From these results, the existence of a stiffness nonlinearity is expected. By analysing the FRF data using the new method, the curves of natural frequency and damping coefficient *vs* response amplitude can be established as shown in Fig.2.22. As compared with the numerical calculation results of figure 2.15, it can be concluded that the system possesses backlash stiffness nonlinearity.

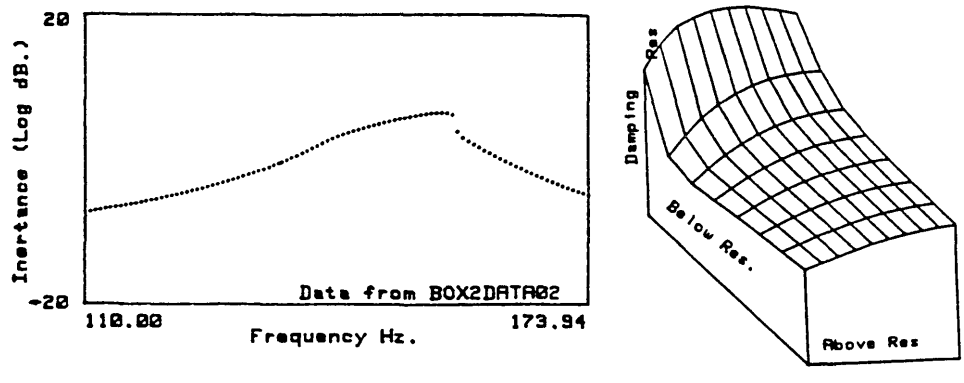


Fig.2.21 - FRF and Damping Plot of ETH/CIRP Box Data

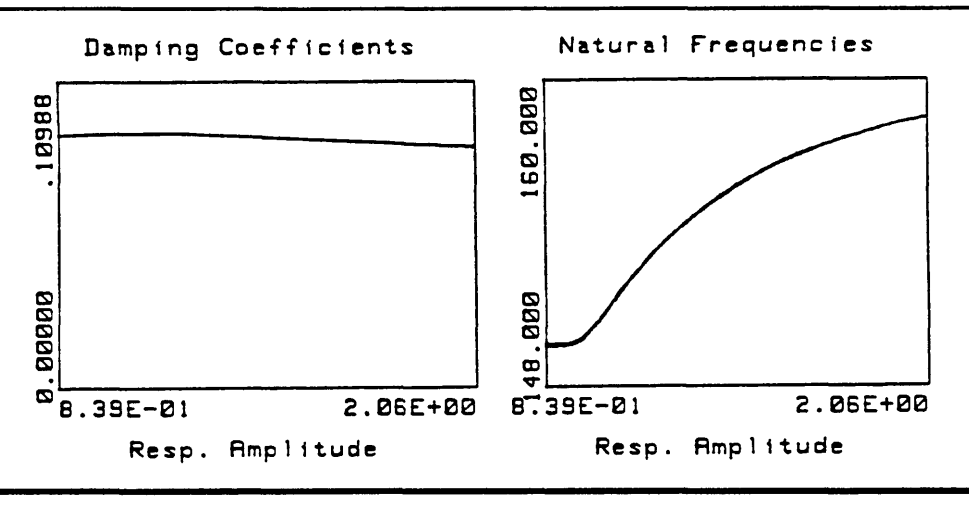


Fig.2.22 - Identified Natural Frequency and Damping Loss Factor (CIRP box)

As mentioned before, the new method is intended for cases in which measured mode is complex. The existence of mode complexity will not influence the analysis accuracy. To demonstrate this, 45° of mode complexity is added artificially to the FRF data shown in figure 2.17 and the data are analysed using the new method. The analysis results shown in Fig.2.23 are the same as those of figure 2.18 except that the complexity of the mode is 45° instead of 0°.

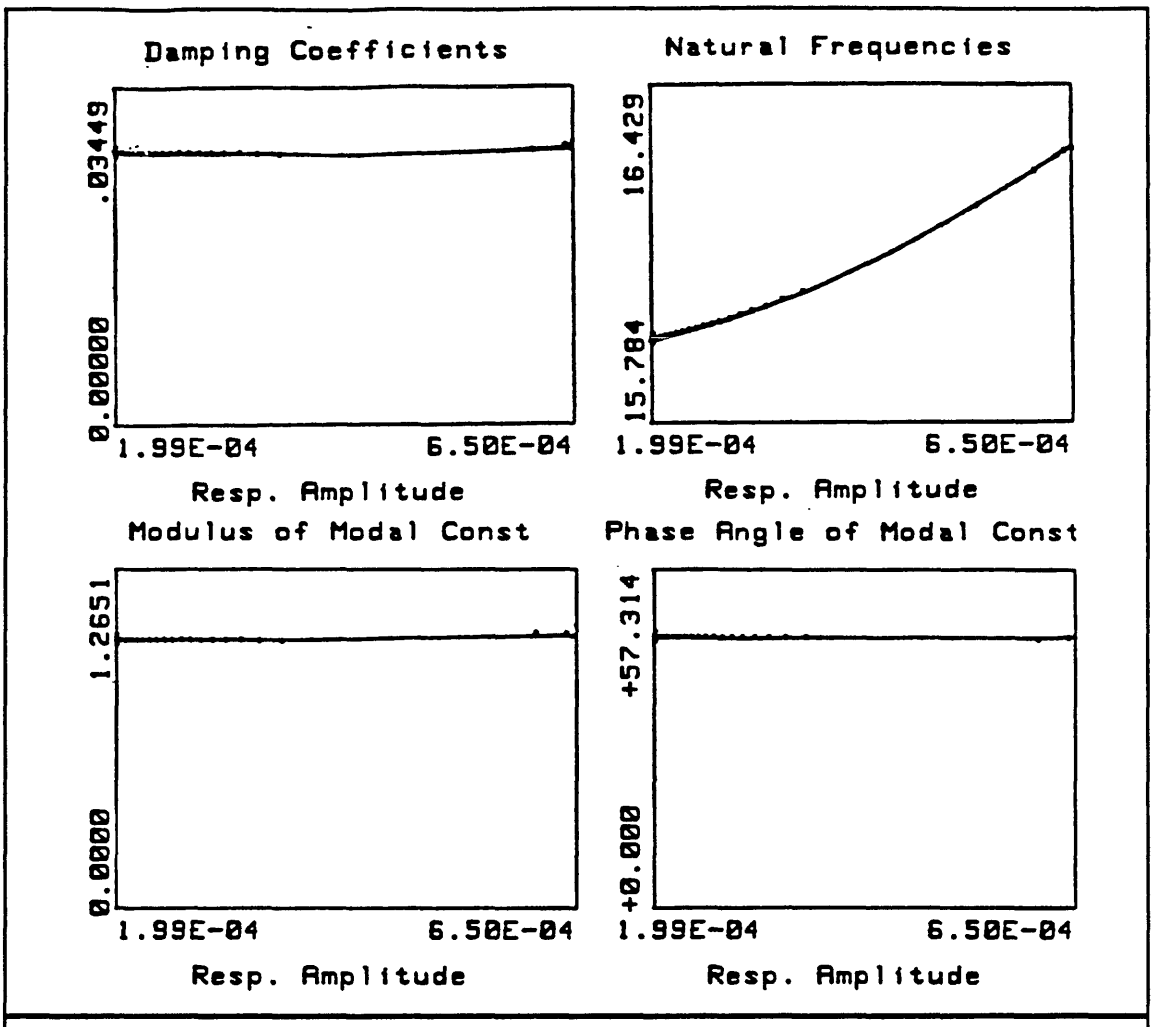


Fig.2.23 - Analysis Results of Data with Complexity

It also needs to be mentioned that although the phase angle of a mode is a measure of a linear system's complexity, linear modal analysis of the data from a nonlinear system can produce an erroneous phase angle which could be misinterpreted as complex mode. This is illustrated in Fig.2.24 where the estimated phase angle using the classical circle-fit is 29° while the true phase angle is 0°.

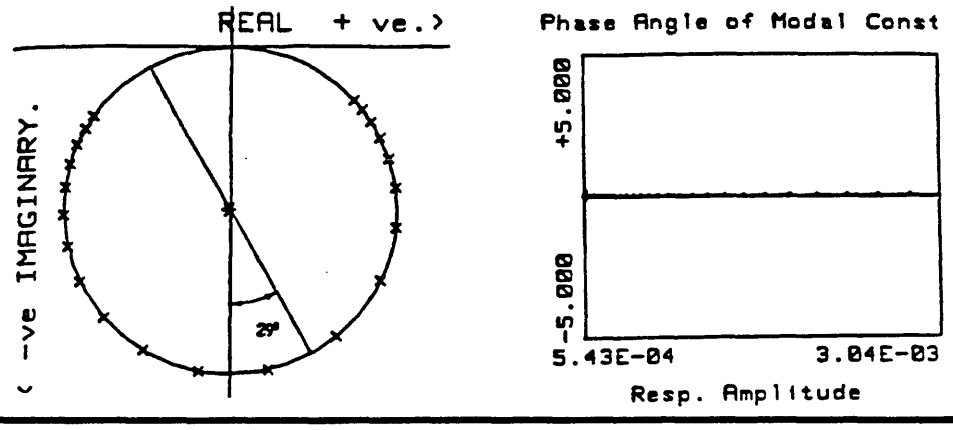


Fig.2.24 - Misestimation of Phase Angle Using Linear Modal Analysis Method

2.5.4 EXTENSION OF THE METHOD TO NONLINEAR MDOF SYSTEMS

In order to assess the applicability of the above method to MDOF systems, constant force FRFs of 2DOF systems with cubic stiffness nonlinearity have been analytically generated with (a) only one mode (the second mode is made linear by controlling the input force to be very low) and (b) both modes are nonlinear respectively. In this case, in order to analyse the mode accurately, the residual must be subtracted. The removal of the residual can be accomplished by the method called SIM [37] which analyses the neighbouring modes first and then subtracts the influence of these analysed modes from the one to be analysed. For the case of only one nonlinear mode, the residual can be removed almost completely as shown in Fig.2.25 in which the Nyquist circle passes through the origin (the Nyquist circle looks the same as those of SDOF systems). After the residual has been removed, the mode can be analysed accurately and the analysis results are shown in Fig.2.26. For the case of both modes being nonlinear, however, it becomes very difficult to remove the residual completely as shown in Fig.2.27 for the case of the first mode (the circle does not pass through the origin, the data points are not exactly on the circle and are not symmetrical with the imaginary axis) and therefore, the analysis results obtained could be in error. The main difficulty of removing the residual in this case lies in the wrong estimation of the phase angle of the neighbouring mode obtained by the linear modal analysis. This difficulty can be overcome in practical analysis by linearising the neighbouring modes in the measurement (by controlling the response amplitude) so that their modal parameters can be accurately estimated based on linear modal analysis method(s) and the residual can therefore be correctly subtracted.

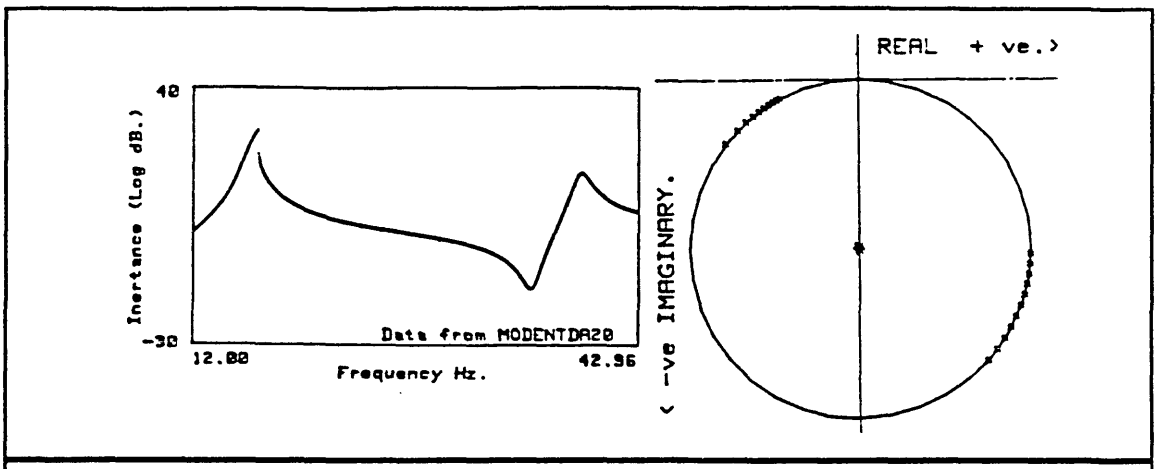


Fig.2.25 - FRF of 2DOF Nonlinear System

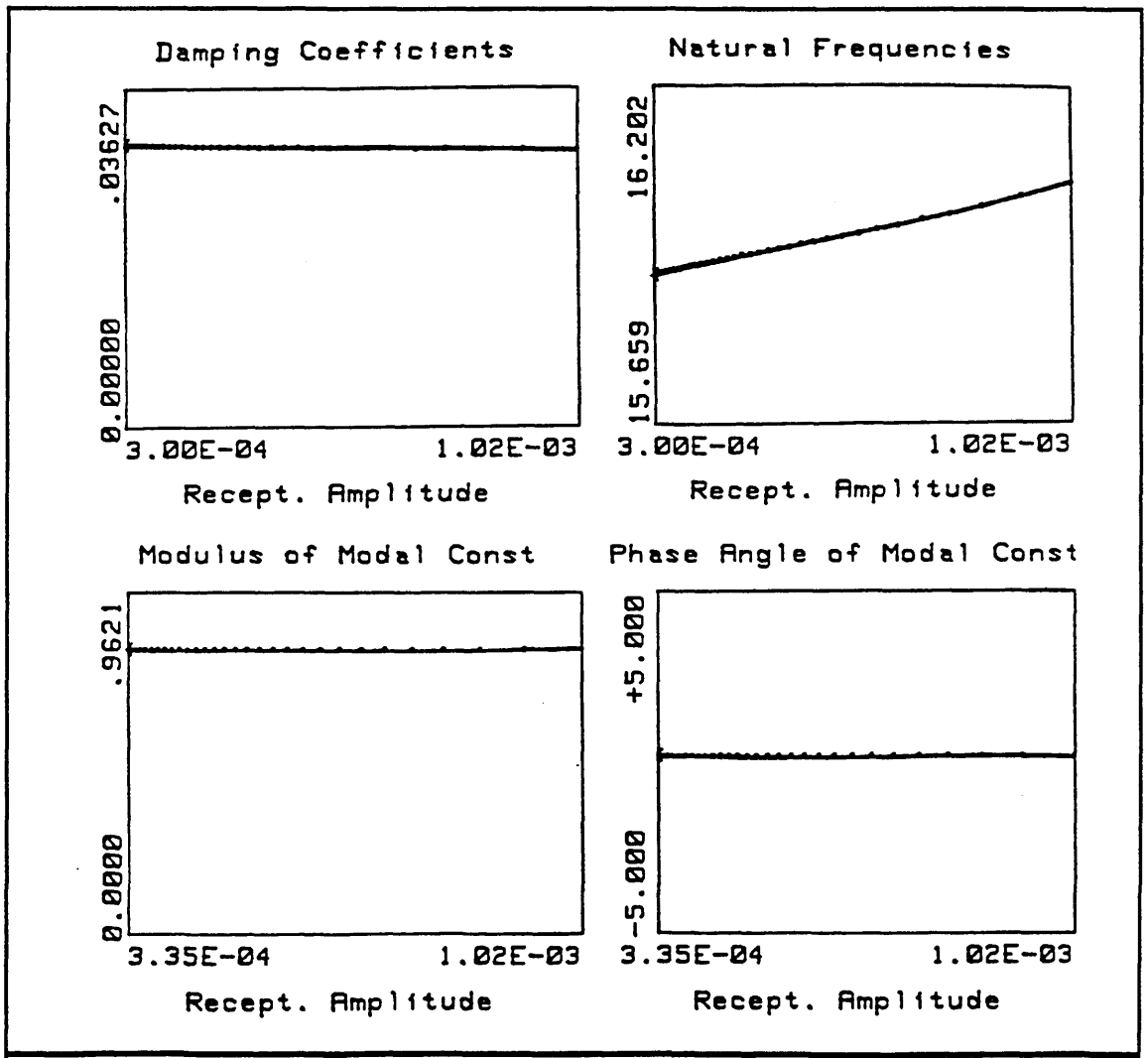


Fig.2.26 - Analysis Results of 2DOF Nonlinear System

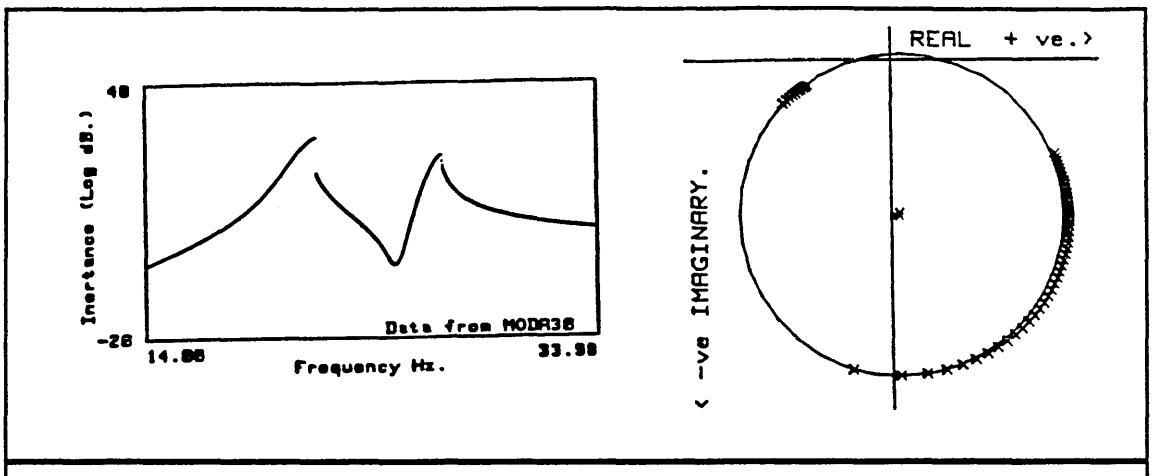


Fig. 2.27 - FRF of 2DOF System with both Modes Nonlinear

Although some measurement and analysis techniques can be employed to remove the residual effect, as mentioned above, when a structure is nonlinear, it is not possible for the data to be analysed to become completely residual free because in this case the residual is a function of response amplitude. Therefore, it is very important for a nonlinear modal analysis method based on the SDOF assumption to obtain satisfactory results even when a small amount of residual exists and it is necessary to undertake a residual analysis. For convenience, an assumption is made that the residual for the mode to be analysed is a complex constant (in fact, this is quite accurate for the case of separated modes). The mathematical expressions of the modal parameters obtained based on the proposed method are as follows:

$$\omega_r^2 = \frac{(R_2 - R_1)(R_2\omega_2^2 - R_1\omega_1^2) + (I_2 - I_1)(I_2\omega_2^2 - I_1\omega_1^2)}{(R_2 - R_1)^2 + (I_2 - I_1)^2} \tag{2-50}$$

$$\eta_r = -\frac{(I_2 - I_1)(R_2\omega_2^2 - R_1\omega_1^2) + (R_2 - R_1)(I_2\omega_2^2 - I_1\omega_1^2)}{\omega_r^2[(R_2 - R_1)^2 + (I_2 - I_1)^2]} \tag{2-51}$$

$$A_r = \frac{(\omega_2^2 - \omega_1^2)[(R_2 - R_1)(R_2R_1 - I_2I_1) + (I_2 - I_1)(R_1I_2 + R_2I_1)]}{(R_2 - R_1)^2 + (I_2 - I_1)^2} \tag{2-52}$$

$$B_r = \frac{(\omega_2^2 - \omega_1^2)[(R_2 - R_1)(I_2R_1 - R_2I_1) + (I_2 - I_1)(R_2R_1 - I_2I_1)]}{(R_2 - R_1)^2 + (I_2 - I_1)^2} \tag{2-53}$$

Where ω_r, η_r, A_r and B_r are the natural frequency, damping coefficient, real and imaginary parts of the modal constant respectively while ω_1, ω_2, R_1, R_2, I_1 and I_2 are the

frequencies, real parts of the receptances and imaginary parts of the receptances of the two points at either side of the resonance chosen.

Because of the similarity of these equations, equation (2-50) can be regarded as their representative for the residual analysis. For convenience, suppose that the receptances of the two points chosen satisfy: $R_1=-R_2=R>0$, $I_1=I_2=I<0$ (this is the case for a real mode with constant force input) and the complex constant for the residual is $\Delta R+i\Delta I$. Then the percentage error for the estimation of ω_r^2 based on equation (2-50) is:

$$\frac{\Delta\omega_r^2}{\omega_r^2} = \frac{(\omega_2^2 - \omega_1^2)[\Delta R(R_2 - R_1) + \Delta I(I_2 - I_1)]}{(R_2 - R_1)(R_2\omega_2^2 - R_1\omega_1^2) + (I_2 - I_1)(I_2\omega_2^2 - I_1\omega_1^2)} \tag{2-54}$$

If we denote $\frac{(\omega_2^2 + \omega_1^2)}{(\omega_2^2 - \omega_1^2)} = \beta^2$, then $\frac{\Delta\omega_r^2}{\omega_r^2}$ becomes:

$$\frac{\Delta\omega_r^2}{\omega_r^2} = \frac{\Delta R}{R\beta^2} \tag{2-55}$$

From equation (2-55), it can be seen that the percentage error for the estimation of ω_r^2 is proportional to the real residual ratio $\Delta R/R$ and the imaginary part of the residual has no effect on the estimate. Although this is true only for the case of natural frequency estimate, the percentage errors for estimation of the other parameters η_r, A_r and B_r are more or less at the same level of $|\Delta R/R| + |\Delta I/I|$. Also, from equation (2-55), it can be seen that the accuracy of the estimation of ω_r^2 can be improved if the frequency difference between the two selected points ($\Delta\omega=\omega_2-\omega_1$) is small so that β becomes large. Therefore, when the mode to be analysed is influenced by other modes, some measures can be taken at both measurement and analysis stages in order to obtain satisfactory results. At the measurement stage, (i) it is possible to linearise the neighbouring modes so that they can be analysed accurately using linear modal analysis method(s) and (ii) the response levels can be controlled so that it is possible to obtain enough points just around the resonance and so the values of R, I and β can be increased (in fact, it is possible to quantify structural nonlinearity by analysing FRF data measured at different response amplitudes at only two frequency points around resonance). At the analysis stage, on the other hand, the SIM method can be used to subtract the residual until it is at its minimum level.

To see how residual effects influence the analysis results, analogue computer FRF data representing dry friction damping nonlinearity and with a 1% artificially-added residual (here 1% residual means that the complex constant of the residual is

$(\Delta R+i\Delta I)=(A_r+iB_r)(\cos 45°+i\sin 45°)/100\eta_r\omega_r^2$ where (A_r+iB_r) is the modal constant of the mode to be analysed) were analysed using this proposed nonlinear modal analysis method and the results are shown in Fig.2.28. As compared with figure 2.20, the analysis results obtained are very similar indeed.

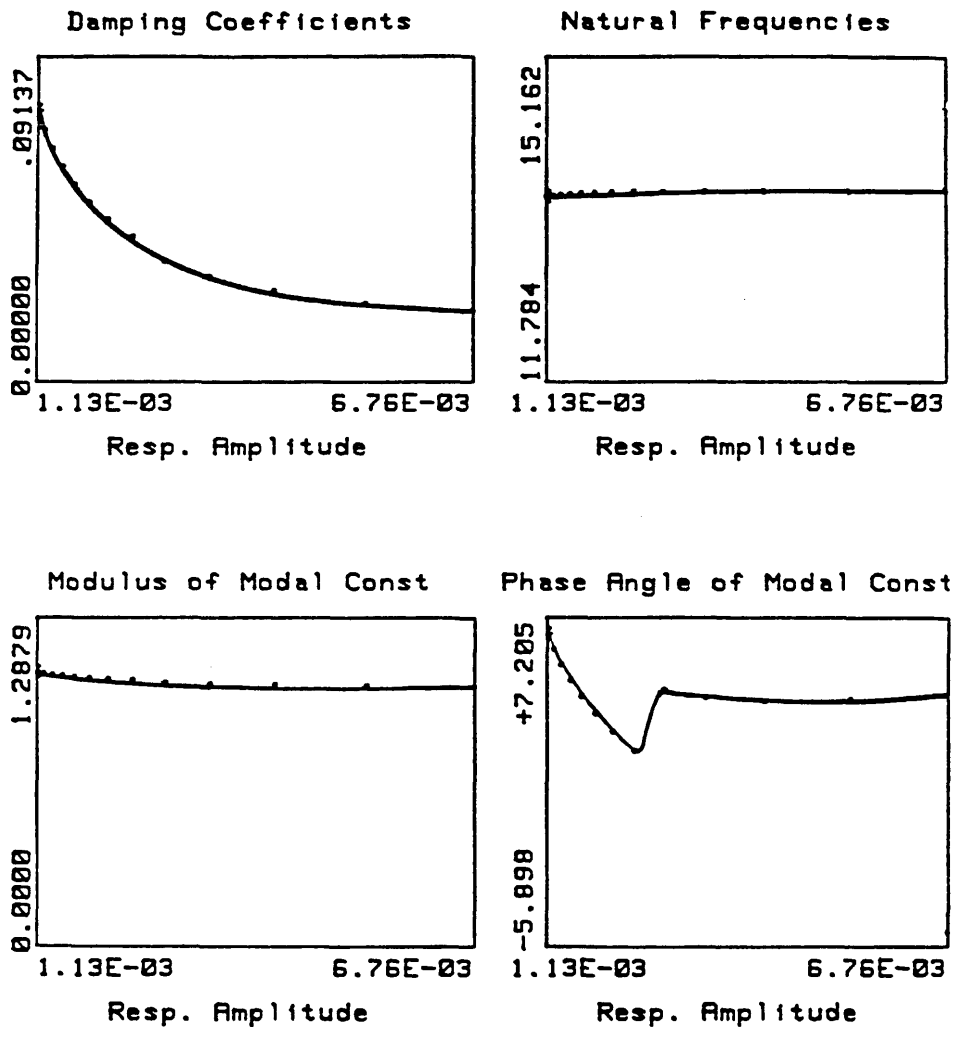


Fig.2.28 - Analysis Results of Frictional Damping with 1% Residual

2.5.5 APPLICATION OF THE METHOD TO PRACTICAL NONLINEAR STRUCTURES

The proposed new method has so far been successfully applied to the analysis of FRF data measured from nonlinear analogue circuits and analytically-generated FRF data for nonlinear MDOF systems. The assessment of residual effects on analysis accuracy has also been carried out. The method is now applied to the analysis of practical nonlinear

structures. First-order FRF data measured from three different practical structures (Beam/Absorber, Frame and NASTRAN Tower structures) are to be analysed.

The Beam/Absorber structure as shown in Fig.2.12 was designed for the experimental investigation of dynamic absorber systems. Some typical measured FRF data with constant force input are illustrated in Fig.2.29. In addition to the shift of resonance frequency which indicates the existence of a stiffness nonlinearity, the mode to be analysed is markedly complex (about 15° phase rotation of the Nyquist circle). One of the FRF curves shown in figure 2.29 was analysed using the proposed method and the results are shown in Fig.2.30. From the natural frequency *vs* response amplitude curve, it can be deduced that the structure probably possesses softening backlash stiffness nonlinearity which is physically due to the fact that, when the response amplitude increases, the supports at both ends of the beam (see fig.2.12) go from micro-slip (stiffness K_1) to slip (stiffness K_2, $K_1>K_2$) which reduces the system's stiffness and so the natural frequencies.

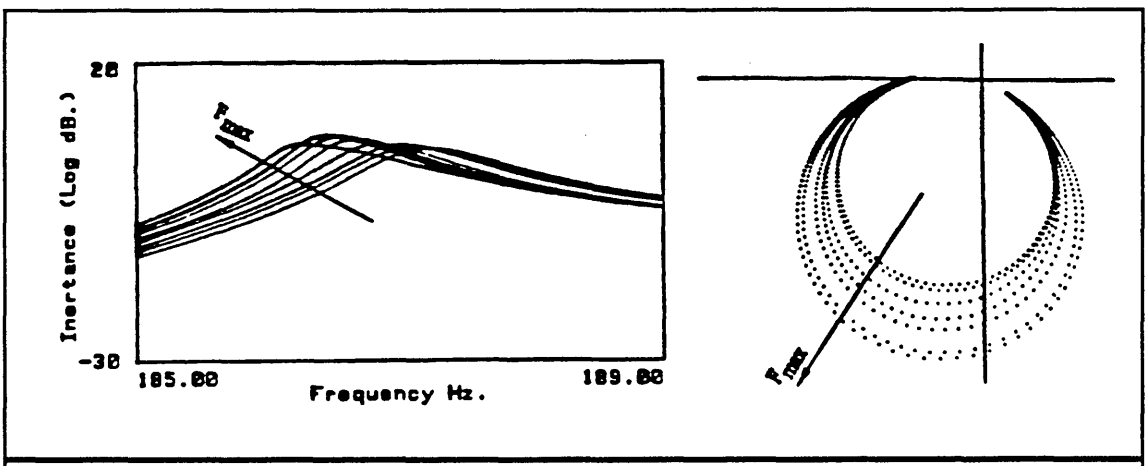


Fig.2.29 - FRFs of Beam/Absorber Structure

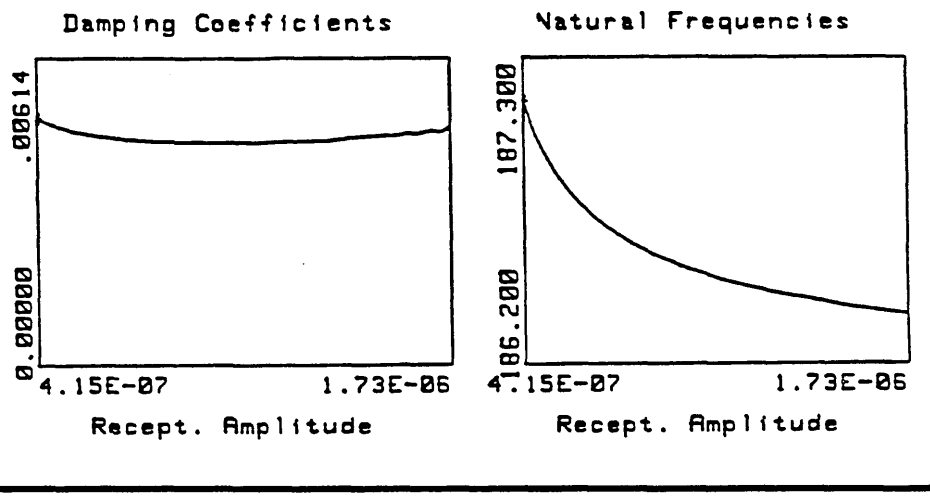


Fig.2.30 - Analysis Results of Beam/Absorber FRF Data

The Frame structure, as shown in figure 2.31, is an artificially-nonlinear structure which was designed for the purpose of the nonlinearity location study in Chapter 5. Typical measured FRF data with constant force input are shown in Fig.2.32, and from these, it can be determined that the structure exhibits both stiffness and damping types of nonlinearity. One of the FRF plots was analysed and the results are shown in Fig.2.33.

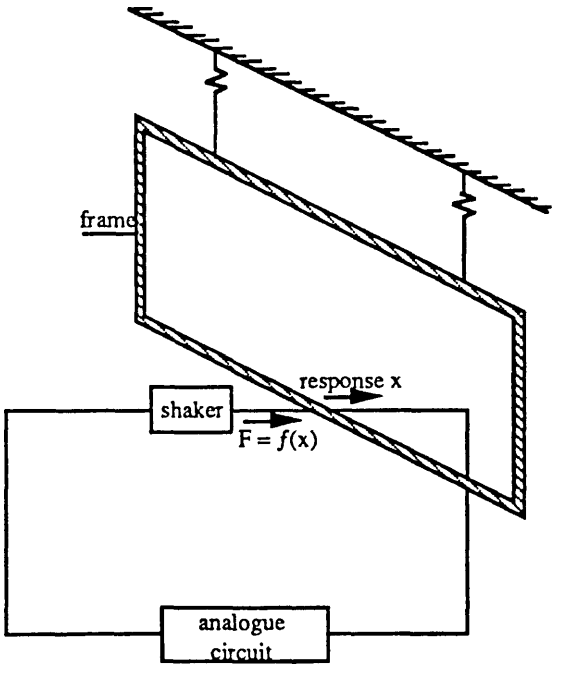


Fig.2.31 - Artificially-Nonlinear Frame Structure

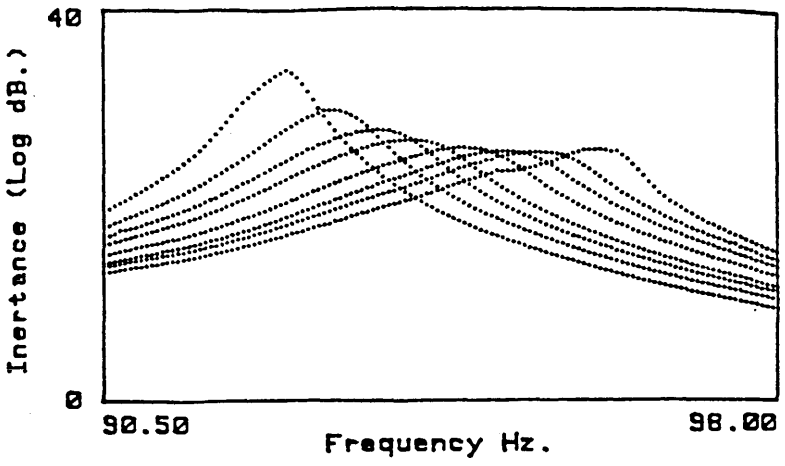


Fig.2.32 - Frequency Response Functions of Frame Structures

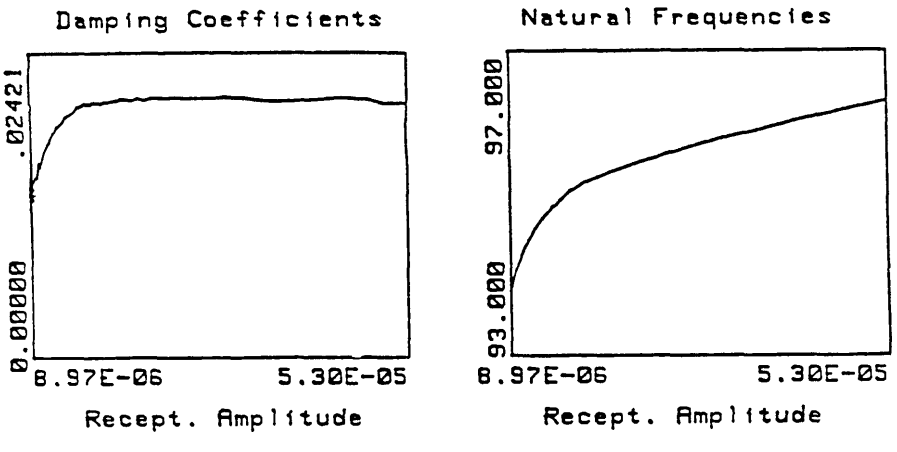


Fig.2.33 - Analysis of measured Frame FRF data

A practical structure known as the "NASTRAN Tower" was also investigated. The structure was known to possess certain type(s) of stiffness nonlinearity [38] although the exact nature of the nonlinearity is still unknown. In the measurement, neither force nor response control was used. The measured data and the analysis results are shown in Fig.2.34 from which, it can be seen that the structure possesses a softening stiffness nonlinearity, probably a softening type of backlash stiffness since the structure remains linear within certain response range and then loses stiffness as response level increases.

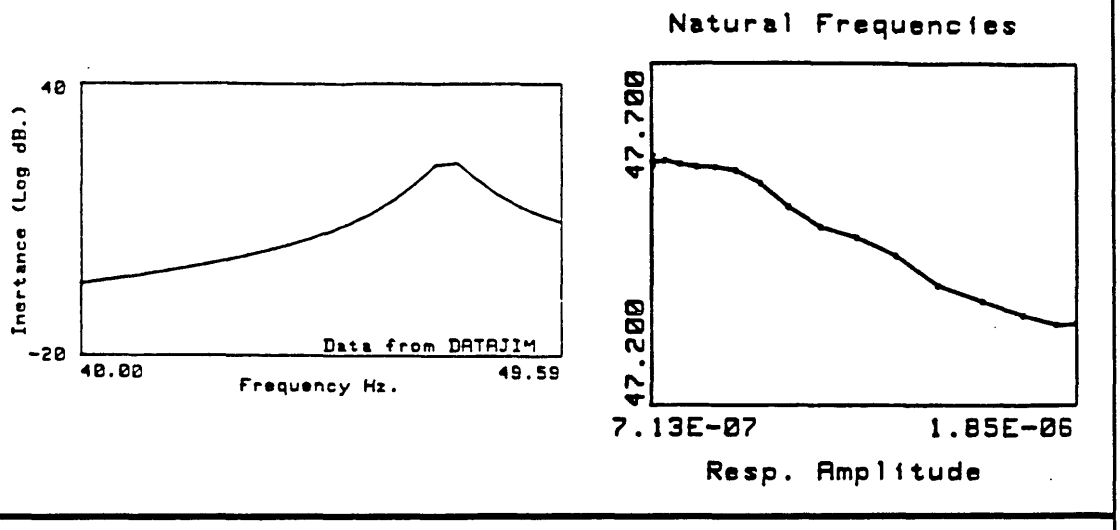


Fig. 2.34 - Analysis of Nonlinearity of NASTRAN Tower Structure

Although one can deduce the type of nonlinearity in some cases by examining the $\omega_r(\hat{X})$ *vs* $(\hat{X})$ and $\eta_r(\hat{X})$ *vs* $(\hat{X})$ relationships, as will be discussed next, the exact identification of the type of nonlinearity, will be difficult when most practical structures are considered.

2.5.6 IDENTIFICATION OF NONLINEAR PHYSICAL CHARACTERISTICS

Structural nonlinearities can now be analysed using the proposed method and relationships between modal parameters and response amplitudes can be established. The quantification of nonlinearity in modal space has thus been accomplished. When nonlinear SDOF systems are considered, according to harmonic balance theory, the describing function coefficients (linearised equivalent stiffness or damping) can be directly calculated from the identified modal data. For example, a nonlinear SDOF system's linearised equivalent stiffness (describing function coefficient) $N_1(\hat{X})$ can be calculated from the identified $\omega_n(\hat{X})$ as $N_1(\hat{X})=\omega_n^2(\hat{X})m$ (m is the mass of the system which can be calculated from the identified modal constant). Although there exists another step from $N_1(\hat{X})$ to the system's true stiffness K(x) (the physical characteristics of the nonlinearity), by comparing with known types of nonlinearity, K(x) can be conclusively identified in most cases from the calculated $N_1(\hat{X})$.

For nonlinear MDOF systems, however, the identification of $N_1(\hat{X})$ and thus of K(x) is not so straightforward. Considering an MDOF system with localised stiffness nonlinearity as shown in Fig.2.16, and supposing the r^{th} mode (which is sensitive to the thus introduced localised nonlinearity) is analysed and the relationship between the natural

frequency and response amplitude at certain reference coordinate is established, the describing function coefficient $N_1(\hat{X})$ of the nonlinear stiffness element cannot be calculated from these analysis results alone. Therefore, an identification of K(x) which is based on $N_1(\hat{X})$ will not be possible. If, on the other hand, the analysed modal data are interpreted as being from an SDOF system when identifying $N_1(\hat{X})$, then misleading results can be obtained because in this case the changes of measured modal parameters depend not only on the stiffness (or damping) changes due to nonlinearity, but also on the modification sensitivity where the nonlinear elements are located. Take the identified natural frequency as an example. The natural frequency change of a certain mode can be mathematically described by

$$\Delta\omega(\hat{X}) = \frac{\partial\omega(\hat{X})}{\partial N_1(\hat{X})}\,\Delta N_1(\hat{X}) = S(\hat{X})\,\Delta N_1(\hat{X}) \qquad (2\text{-}56)$$

Since $S(\hat{X})$ is unknown in the identification process and is a function of response amplitude ($\hat{X}$), except in the case of SDOF systems in which, $S(\hat{X})$ is known to be the identified modal constant 1/m, $\Delta N_1(\hat{X})$ cannot be calculated from the identified $\Delta\omega(\hat{X})$ and, as a result, the identification of K(x) is out of the question.

In fact, as will be discussed later on, in order to identify the describing function coefficients and thus the physical characteristics of nonlinear element(s) of a practical nonlinear structure, the nonlinearities have to be located first and then the linearised equivalent stiffness matrix $[K(\hat{X})]$ can be established by correlating the analytical model and measured dynamic testing data.

2.6 CONCLUSIONS

Once a structure is nonlinear, modal parameters obtained from the analysis of measured FRF data will, in general, be erroneous. In this case, a nonlinear modal analysis is required so that the structural nonlinearity can be taken into account. There are three main problems to be solved for a successful modal analysis of a nonlinear structure and they are: (i) detecting the existence; (ii) quantifying the extent and (iii) identifying the physical characteristics of the nonlinearity.

Commonly-used methods for the modal analysis of nonlinearity have been reviewed in this chapter. Bode plot and reciprocal receptance plot techniques detect nonlinearity by presenting measured FRF data in specific formats and then examining the systematic distortion(s) caused by the existence of structural nonlinearity. The isometric damping

plot method achieves its nonlinearity detection by calculating the damping matrix based on linear modal analysis theory and then examining the distortion of the damping plot due to the variation in response amplitude and so in the effective natural frequency differences of data points around Nyquist circle. These methods are convenient for the nonlinearity detection stage but not so applicable for nonlinearity quantification and identification. With a more ambitious objective of nonlinearity quantification, the Inverse Receptance method as discussed was developed by analysing stiffness and damping nonlinearity separately based on the real and imaginary parts of the inverse receptance data. However, as demonstrated in some detail in this Chapter, some assumptions have been made during the development of the method which are, in general, not valid for data measured on practical structures with nonlinearity and so the method is limited in terms of its practical applications.

The theoretical aspects of the existence of complex modes have been discussed. The necessary and sufficient condition for the existence of complex modes is that the damping distribution of the system is nonproportional. The effect of natural frequency spacing on the degree of complexity has been illustrated. Numerical as well as experimental examples are given.

The harmonic balance theory, which is the mathematical basis of the new nonlinear modal analysis method proposed in this Chapter, is presented together with its practical application conditions. The relationship between the analytical analysis of a nonlinear system based on harmonic balance theory and the experimental measurement of FRF data of a practical nonlinear structure has been discussed. Based on harmonic balance theory, a dynamic system having stiffness nonlinearity will take a different equivalent linearised stiffness values (describing function coefficient $N_1(\hat{X})$) for different response amplitudes, so that each FRF data point from a measurement with constant force actually relates to a specific FRF data curve measured with constant response, thereby containing all the information of the latter curves. Due to this specific characteristic of FRF data from constant force measurement, thus measured data can be analysed to quantify and identify the nonlinearity of the test structure if the force level is appropriately chosen.

With a theoretical basis of the harmonic balance analysis, a new method has been proposed to analyse nonlinearity from measured first-order FRF data. In addition to deriving an indication of the nonlinearity, the method aims at establishing the relationships between the modal parameters of interest and response amplitude from the FRF data measured using sinusoidal excitation. The final results of the analysis are the response-amplitude-dependent eigenvalues $\lambda(\hat{X})$ and eigenvectors $\{\phi(\hat{X})\}$ of nonlinear

systems. These identified modal data can be used subsequently to derive a linearised spatial model ([M], [K($\hat{x}$) and C($\hat{x}$) or D($\hat{x}$) matrices) of the structure. Also, it is necessary to mention that the condition of constant force is not necessary when measuring the FRF data for the subsequent modal analysis using this new method. In fact, satisfactory analysis can be carried out as long as the response amplitude varies sufficiently to expose the nonlinearity and embraces the range of displacements which must be described by the model.

The method has been extended to the analysis of nonlinear MDOF systems and the effect of residual on the analysis accuracy has been discussed. By linearising neighbouring modes at the measurement stage and applying the SIM method at the analysis stage, the residual of the mode to be analysed can be removed to its minimum level and so accurate analysis results can be obtained.

Although the quantification of nonlinearity in modal space has been achieved using the proposed method, identification of the describing function coefficient N($\hat{x}$) and so of the physical characteristics K(x) and/or C(x) from the identified modal parameters $\lambda(\hat{x})$ and $\{\phi(\hat{x})\}$, which will be discussed in later chapters, is by no means straightforward when MDOF systems are considered. Moreover, it has to be pointed out that, during sinusoidal excitation measurement, since the DC component of the response signal of a nonsymmetric nonlinear system has been filtered out, the nonsymmetric nonlinearity has been made symmetrical and the FRF data measured are the data from an equivalent symmetric nonlinear system. Due to this symmetrisation, for some nonsymmetric nonlinear systems, such as bilinear systems, the existing nonlinearity cannot be revealed from measured first-order FRF data which are effectively linear. Furthermore, except for the fundamental frequency component, the response of a nonlinear system usually contains super-, sub- and combinations of harmonics. However, in the first-order FRF analysis, all these harmonics, which are in some cases as important as the fundamental component in vibration analysis, are filtered out. Therefore, first-order FRF analysis is limited in the sense that it reveals the nature of the nonlinearity and in order to identify nonsymmetric nonlinear systems and take into account these super-, sub- and combinations of harmonics, higher-order FRF analysis becomes necessary and is to be introduced in next chapter.

CHAPTER 3

IDENTIFICATION OF NONLINEARITY USING HIGHER-ORDER FREQUENCY RESPONSE FUNCTIONS

3.1 INTRODUCTION

The identification of dynamic characteristics of linear structures from measured data is now well established. In order to characterise a linear system, what is required is the measurement of its impulse response functions (time-domain) or frequency response functions (frequency-domain). Unfortunately, as mentioned earlier, most practical engineering structures are nonlinear and the analysis of a nonlinear system is far more complicated than that of a linear system. As discussed in some detail in Chapter 2, a new analysis technique has been developed to identify nonlinear behaviour based on the analysis of measured classical first-order frequency response functions and has been found to be quite successful in cases where the effect of structural nonlinearities shows up in the measured data [39]. However, due to the symmetrisation effect and the approximate nature of the first-order FRF measurement, for some nonsymmetric nonlinear systems, the thus measured FRF data are the data from their equivalent

symmetric counterparts and the harmonic components which are usually present in the response signal of nonlinear systems are filtered out. This symmetrisation of nonsymmetrical nonlinearities and the elimination of harmonic components mean that the first-order frequency response function analysis is not very appropriate for the analysis of structures with nonsymmetrical nonlinearities. In fact, it will be shown that for some specific nonlinear systems, such as quadratic and bilinear systems, the analysis technique is incapable of analysing them at all. From the response prediction point of view, calculations made using the first-order frequency response functions only can be quite inaccurate in some cases as described in the application conditions of harmonic balance analysis in Chapter 2 because, mathematically, this means that only the linear term of the Taylor expression of a nonlinear function at certain point has been retained. These limitations of first-order frequency response analysis are illustrated next.

As mentioned, systems with nonsymmetrical nonlinearities such as quadratic and bilinear systems cannot be identified based on the first-order frequency response function analysis because these nonlinearities are such that the measured first-order frequency response functions based on sinusoidal excitation are effectively linear. Suppose the nonlinearities are of a stiffness type and their force-displacement relationships are shown in Fig.3.1, then the equivalent stiffness value corresponding to specific response amplitude can be calculated based on harmonic balance theory as discussed in Chapter 2. Assume the vibration to be sinusoidal as $x(t)=\hat{X}\sin\omega t$, then the describing function coefficients (equivalent stiffnesses) $N_q(\hat{X})$ for the case of quadratic stiffness and $N_b(\hat{X})$ for the case of bilinear stiffness can be calculated as:

$$N_q(\hat{X}) = \frac{1}{\pi\hat{X}} \int_0^{2\pi} (k\hat{X}\sin\omega t + k_1\hat{X}^2\sin^2\omega t)\sin\omega t \, d\omega t = k \qquad (3\text{-}1)$$

$$N_b(\hat{X}) = \frac{1}{\pi\hat{X}} \int_0^{\pi} k_1\hat{X}\sin\omega t \sin\omega t \, d\omega t + \frac{1}{\pi\hat{X}} \int_{\pi}^{2\pi} k_2\hat{X}\sin\omega t \sin\omega t \, d\omega t = \frac{k_1+k_2}{2} \qquad (3\text{-}2)$$

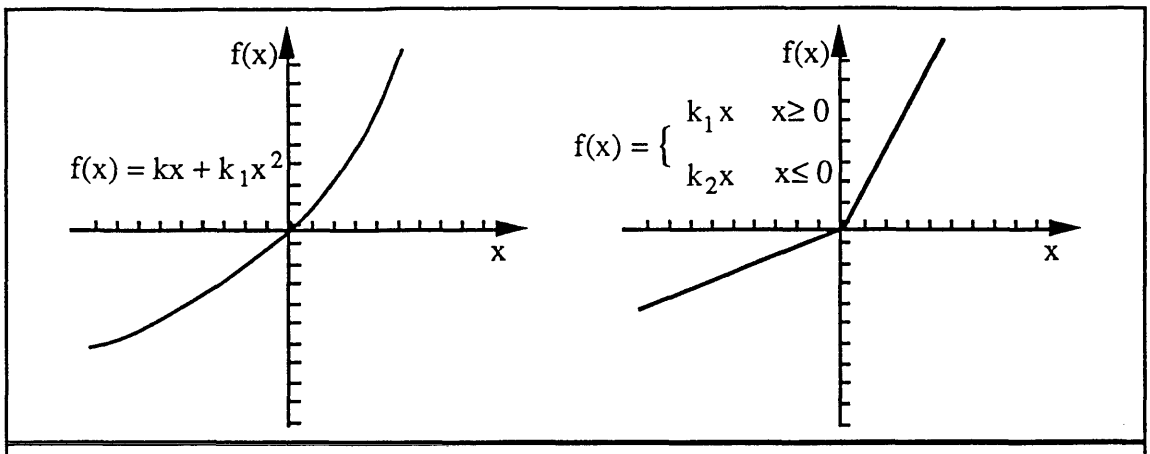


Fig. 3.1 - Force Displacement Relationships of Quadratic and Bilinear Stiffness

Since both $N_q(\hat{X})$ and $N_b(\hat{X})$, which are the equivalent stiffnesses, are constant (independent of the response amplitude), the measured first-order frequency response functions of these systems are linear with equivalent linear constant stiffnesses of k for the quadratic stiffness case and $(k_1+k_2)/2$ for the bilinear stiffness case.

On the other hand, the existence of nonlinear phenomena such as, sub-, super- and combinational resonances in nature is well known. Nayfeh [24] mentioned that Lefschetz described a commercial airplane in which propellors induced a subharmonic vibration in the wing which in turn induced subharmonic vibration in the rudder. The oscillations were violent enough to cause tragic consequences. Also, reports have been found in the literature that excessive vibrations were caused by superharmonic excitation and combinational resonances. In those cases, the analysis of the harmonic components becomes as important as that of the fundamental frequency component and the response predicted using first-order FRF data in such circumstances could be very inaccurate. To illustrate this point, superharmonic excitation is considered for the case of an SDOF system with cubic stiffness nonlinearity. When the external excitation frequency is far from one third of the natural frequency of the system (linear natural frequency, as if the cubic term were not introduced), the response prediction based on the first-order FRF is very accurate. When the excitation frequency is close to one third of the natural frequency, then the structural resonance will be excited by the third harmonic component generated by the cubic nonlinearity and as a result, the response prediction based on the first-order FRF in this case becomes very inaccurate. Comparisons of the true responses and the responses predicted using first-order frequency response function data when the excitation frequencies are of 4/3 and 1/3 the natural frequency are shown in Fig.3.2.

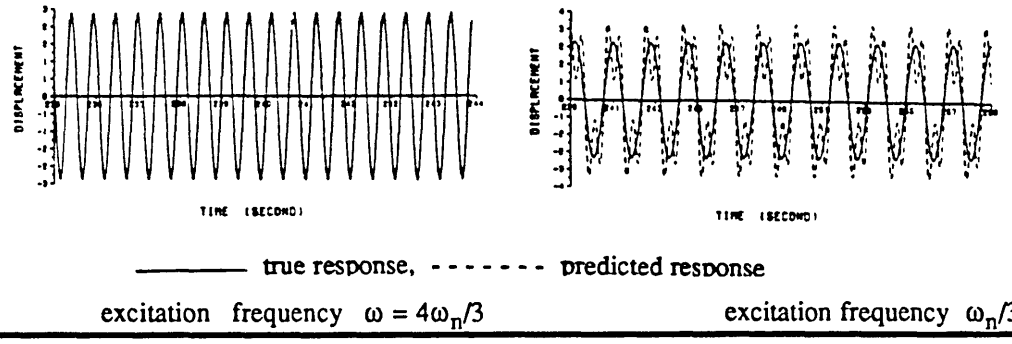


Fig.3.2 - Response Prediction Accuracy Using First Order FRF

All this means that the first-order frequency response function analysis is inadequate and even sometimes inappropriate for some nonlinear systems and more accurate representation of their dynamic characteristics becomes necessary. For this purpose, research work on the higher-order frequency response function analysis has been carried out and is described in this Chapter.

The mathematical basis of higher-order frequency response function analysis lies in the Volterra series theory which, as the functional series representation of nonlinear systems and with its rigorous mathematical base, has been found to be quite effective in the characterisation of general nonlinear systems. The theory was first introduced into nonlinear circuit analysis in 1942 by Wiener who later extended the theory [7] and applied it in a general way to a number of problems. Since Wiener's early work, many papers have been published dealing with this subject in system and communication engineering [40-42]. However, it was not until recently that the theory has been applied to the identification of nonlinear mechanical structures [8,43,44] and found to be quite useful. There is some literature available now on the identification of nonlinear mechanical structures based on the Volterra series theory, such as references [8,43,44,45]. However, most of the studies to date are still at the stage of numerical simulation of certain nonlinear systems and the difficulties in applying this theory to the identification of practical nonlinear mechanical structures have not been fully investigated although some experimental work based on specifically designed nonlinear structure has been carried out [44]. The research work presented in this Chapter introduces the basic theory of Volterra series and of their relation to the higher-order frequency response functions and how the higher-order frequency response functions generalise linear system theory to cover nonlinear systems. The harmonic probing method for the Volterra kernel measurement using multi-tone input [46] and correlation technique for the Wiener kernel measurement using random input [47] are investigated and the relationship between the Volterra and

Wiener kernels is studied. Possible ways of curve-fitting or surface-fitting the measured higher-order frequency response functions so that parametric or nonparametric model of the nonlinear structure can be established are discussed. Considerable attention is given to the practical assessment of the measurement of higher-order frequency response functions of realistic nonlinear mechanical systems, both in the case of sinusoidal and random inputs, by numerically simulating the measurement processes. The existing difficulties concerning the successful measurement of higher-order frequency response functions are discussed and possible ways of improving measurement results are suggested. The applications of higher-order frequency response function analysis in the identification of nonlinear mechanical systems are also discussed.

3.2 VOLTERRA SERIES REPRESENTATION OF NONLINEAR SYSTEMS

A nonlinear function $f(x)$ can in general be represented as a Taylor series at a certain point (e.g. $x=x_0$) and this series approaches $f(x)$ when the variable x is not far from that point. Similarly, a nonlinear system can in general be characterised by a Volterra series which converges when the nonlinearity of the system satisfies certain general conditions [6]. Before presenting the theory of Volterra series, it is necessary to examine some of the basic characteristics of nonlinear systems.

3.2.1 BASIC CHARACTERISTICS OF NONLINEAR SYSTEMS

Since a linear system must satisfy the principle of superposition (as discussed in Chapter 2), a sinusoid can be regarded as an eigenfunction of the system. For a sinusoid applied to a linear system, the system only changes its amplitude and phase angle without distorting its wave form. A nonlinear system however, is characterised by the transfer of energy between frequencies. For a sinusoidal input $f(t)=A\sin\omega t$ to the nonlinear system governed by equation

$$m\ddot{x} + c\dot{x} + kx + k_1x^2 + k_2x^3 = f(t) \qquad (3\text{-}3)$$

the system will generate harmonic frequency components response in addition to the fundamental frequency component, as shown in Fig.3.3 (the background curve is due to numerical inaccuracy). If a multi-tone input $x(t)=A\sin\omega_1 t+B\sin\omega_2 t$ is applied (the input signal has two or more frequency components where A, B can be complex numbers to accommodate the different phase shifts of these two waveforms), then in addition to the

fundamental frequencies (ω_1,ω_2) and their harmonics ($n\omega_1$,$n\omega_2$), there will also be combinational frequency components ($\omega_1+\omega_2$, $\omega_2-\omega_1$ etc.) as shown in Fig.3.4. In fact, for this specific system described by equation (3-3), there will be frequency components ($n_1\omega_1+n_2\omega_2$) present in the response x(t) for all integer values of n_1 and n_2. In order to establish an input/output model of a nonlinear system which can not only predict the fundamental frequency, but also the harmonics and combinational frequencies as well, the Volterra series theory of nonlinear systems was developed.

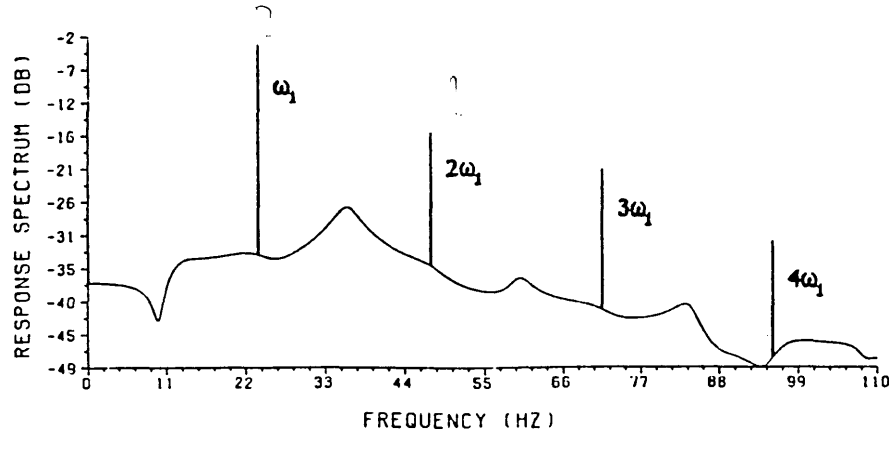


Fig.3.3 - Response Spectrum of an SDOF System with Input $f(t)=A\sin\omega_1 t$

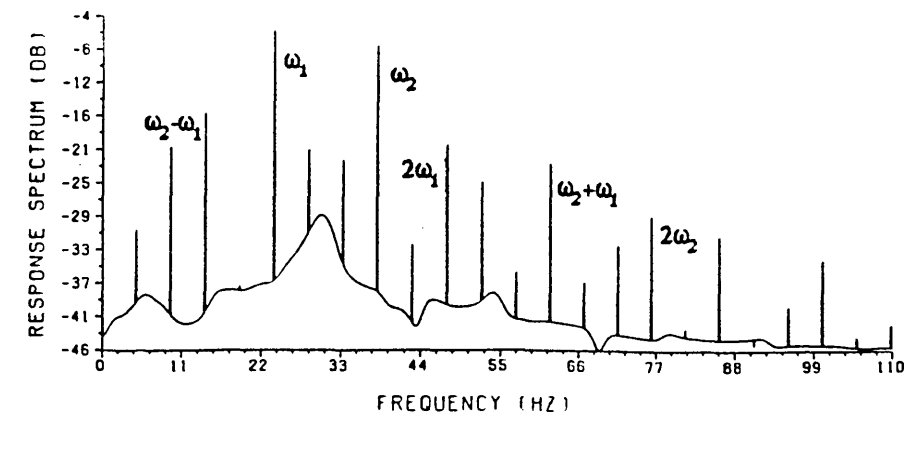


Fig.3.4 - Response Spectrum of an SDOF System with Input $f(t)=A_1\sin\omega_1 t + A_2\sin\omega_2 t$

3.2.2 THE VOLTERRA SERIES REPRESENTATION

Volterra series have been described as "power series with memory" which express the output of a nonlinear system in "powers" of the input. A wide class of nonlinear systems encountered in engineering can be represented as Volterra series. Given an input $f(t)$, the output x(t) of a time invariant system can, in general be expressed as

$$x(t) = \sum_{n=1}^{\infty} \int_{-\infty}^{+\infty} d\tau_1 \ldots \int_{-\infty}^{+\infty} d\tau_n \, h_n(\tau_1, \ldots, \tau_n) \prod_{r=1}^{n} f(t - \tau_r) \qquad (3\text{-}4)$$

where the kernels $h_n(\tau_1, \ldots, \tau_n)$ are the Volterra kernels which describe the system. It should be noted that the first-order kernel $h_1(\tau)$ is the impulse response due to the linear part of the nonlinear system and the higher-order kernels can thus be viewed as higher-order impluse responses which serve to characterise the various orders of nonlinearity. In the special case when the system is linear, all the higher-order kernels except $h_1(\tau)$ are zero. The Volterra series representation (3-4) of a nonlinear system is homogeneous. In order to illustrate this, rewrite equation (3-4) as $x(t) = x_1(t) + x_2(t) + \ldots + x_s(t) + \ldots$ where

$$x_s(t) = \int_{-\infty}^{+\infty} \ldots \int_{-\infty}^{+\infty} h_s(\tau_1, \ldots, \tau_s) \, f(t - \tau_1) f(t - \tau_2) \ldots f(t - \tau_s) \, d\tau_1 \, d\tau_2 \ldots d\tau_s \qquad (3\text{-}5)$$

From (3-5), it is easy to see that when the input changes from $f(t)$ to $\alpha f(t)$, then the s^{th} component of the output becomes $\alpha^s x_s(t)$ and the total output x(t) becomes $x(t) = \sum_{s=1}^{\infty} \alpha^s x_s(t)$. This homogeneous property of Volterra series representation has been applied to the measurement of Volterra kernels of electrical nonlinear circuits by repeating the measurements using different input levels of the same signal [48]. Since almost all physical systems, whether they are linear or nonlinear, are causal (a system is said to be causal if, for any input, the output at any instant of time does not depend upon the future input), all the kernels have to satisfy

$$h_n(\tau_1, \ldots, \tau_n) = 0 \quad \text{for any } \tau_s < 0 \quad ((s = 1, n), n = 1, \infty) \qquad (3\text{-}6)$$

Like a Taylor series representation of a nonlinear function, the Volterra series representation of a general nonlinear system is theoretically infinite and, as will be discussed later, the effort of computing the n^{th}-order kernel increases exponentially as n increases so that one has to be satisfied with the first few kernels only (usually, up to the third kernel). Fortunately, good approximations can be obtained for most engineering problems by just considering these first few kernels and this is why this theory has been widely applied to the characterisation of practical nonlinear systems.

3.3 HIGHER-ORDER FREQUENCY RESPONSE FUNCTIONS

3.3.1 THEORY OF HIGHER-ORDER FREQUENCY RESPONSE FUNCTIONS

The n^{th}-order Volterra kernel transform or n^{th}-order Volterra transfer function is simply defined as the n-dimensional Fourier transform of the n^{th}-order Volterra kernel

$$H_n(\omega_1, \omega_2, ..., \omega_n) = \int_{-\infty}^{+\infty} d\tau_1 \ ... \int_{-\infty}^{+\infty} d\tau_n \ h_n(\tau_1, ..., \tau_n) \ e^{-i(\omega_1\tau_1 + \omega_2\tau_2 + ... + \omega_n\tau_n)} \tag{3-7}$$

Since the n^{th}-order kernel $h_n(\tau_1,\tau_2,...,\tau_n)$ is real, symmetric (the value of $h_n(\tau_1,\tau_2,...,\tau_n)$ does not depend on the order of $\tau_1,\tau_2,...,\tau_n$, for example, in the case of second-order kernel, $h_2(\tau_1,\tau_2)= h_2(\tau_2,\tau_1)$) and causal, its Fourier transform $H_n(\omega_1,\omega_2,...,\omega_n)$ is symmetric and also possesses complex conjugate symmetry such that

$$H_n(\omega_1, \omega_2, ..., \omega_n) = H_n^*(-\omega_1, -\omega_2, ..., -\omega_n) \tag{3-8}$$

For a linear system, if the frequency response functions (only the first-order) have been determined, the output x(t) can be calculated for any form of inputs. The same argument holds if all the Volterra transfer functions $H_n(\omega_1,\omega_2,...,\omega_n)$ have been determined and since $H_n(\omega_1,\omega_2,...,\omega_n)$ are unique (independent of input and output of the system), the Volterra series representation is mathematically very attractive because under this representation, the identification of a nonlinear system reduces to the measurement of these unique Volterra kernel transforms. However, it will be shown that due to the interactions between kernels, these uniquely defined Volterra transfer functions cannot be uniquely measured in practice and all that can be measured are approximations which, in general, are input/output dependent.

The n^{th}-order frequency response function $H_n(\omega_1,\omega_2,...,\omega_n)$ is defined as the measured n^{th}-order Volterra kernel transform $H_n(\omega_1,\omega_2,...,\omega_n)$. The relationship between $H_n(\omega_1,\omega_2,...,\omega_n)$ and $H_n(\omega_1,\omega_2,...,\omega_n)$ is discussed below. The input-output relationship of a nonlinear system based on Volterra series representation has been discussed in detail [46] for different forms of input and only the sinusoidal form of input is considered here. To make the analysis convenient, it is necessary to introduce the partial differential operator given as:

$$D_\alpha^n \equiv \frac{\partial^n}{\partial\alpha_1 \ldots \partial\alpha_n} \Big| \alpha_1 = \alpha_2 = \ldots = \alpha_n = 0 \tag{3-9}$$

Under this operator, it can be seen that

$$\prod_{r=1}^{n} f(t - t_r) = D_\alpha^n \exp\Big[\sum_{s=1}^{n} \alpha_s f(t - t_s)\Big] = D_\alpha^n \frac{\Big[\sum_{s=1}^{n} \alpha_s f(t - t_s)\Big]^n}{n!} \tag{3-10}$$

Upon substitution, (3-4) becomes

$$x(t) = \sum_{n=1}^{\infty} \int_{-\infty}^{+\infty} d\tau_1 \ldots \int_{-\infty}^{+\infty} d\tau_n\, h_n(\tau_1, \ldots, \tau_n)\, D_\alpha^n \exp\Big[\sum_{s=1}^{n} \alpha_s f(t - t_s)\Big] \tag{3-11}$$

After some further mathematical manipulation, (3-11) can be written as:

$$x(t) = \sum_{n=1}^{\infty} \int_{-\infty}^{+\infty} d\tau_1 \ldots \int_{-\infty}^{+\infty} d\tau_n\, h_n(\tau_1, \ldots, \tau_n)\, D_\alpha^n \frac{\Big[\sum_{s=1}^{n} \alpha_s f(t - t_s)\Big]^n}{n!} \tag{3-12}$$

On the other hand, under the differential operator defined in (3-9), the n^{th}-order Volterra kernel transform can be rewritten as:

$$H_n(\omega_1, \omega_2, \ldots, \omega_n) = \frac{1}{n!} \int_{-\infty}^{+\infty} d\tau_1 \ldots \int_{-\infty}^{+\infty} d\tau_n\, h_n(\tau_1, \ldots, \tau_n)\, D_\alpha^n \prod_{r=1}^{n} A_n(\omega_r) \tag{3-13}$$

where

$$A_n(\omega) = \sum_{s=1}^{n} \alpha_s e^{-i\omega\tau_s} \tag{3-14}$$

To illustrate the validity of equation (3-13), consider the derivation of $H_2(\omega_1,\omega_2)$. In this case, n=2 and so $A_2(\omega)$ becomes

$$A_2(\omega) = \alpha_1 e^{-i\omega\tau_1} + \alpha_2 e^{-i\omega\tau_2} \tag{3-15}$$

and upon substitution,

$$D_\alpha^2 \prod_{r=1}^{2} A_2(\omega_r) = D_\alpha^2 [\,(\alpha_1 e^{-i\omega_1\tau_1} + \alpha_2 e^{-i\omega_1\tau_2})(\alpha_1 e^{-i\omega_2\tau_1} + \alpha_2 e^{-i\omega_2\tau_2})]$$

$$= e^{-i(\omega_1\tau_1 + \omega_2\tau_2)} + e^{-i(\omega_1\tau_2 + \omega_2\tau_1)} \tag{3-16}$$

Substituting (3-16) into (3-13) and considering the symmetry property of $H_2(\omega_1,\omega_2)$, equation (3-13) becomes

$$H_2(\omega_1, \omega_2) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} h_2(\tau_1, \tau_2)\, e^{-i(\omega_1\tau_1 + \omega_2\tau_2)}\, d\tau_1 d\tau_2 \tag{3-17}$$

Using this preliminary mathematics, it is now possible to establish an input-output relationship of a general nonlinear system when the input to the system is in the form of sinusoid.

When $f(t)=B\cos\omega t$, then

$$\sum_{s=1}^{n} \alpha_s f(t - \tau_s) = \frac{B}{2} \sum_{s=1}^{n} \alpha_s (e^{i\omega t - i\omega\tau_s} + e^{-i\omega t + i\omega\tau_s}) = \frac{B}{2} [e^{i\omega t} A_n(\omega) + e^{-i\omega t} A_n(-\omega)] \tag{3-18}$$

where $A_n(\omega)$ is given by (3.14). According to the binormial theorem,

$$\frac{[\sum_{s=1}^{n} \alpha_s f(t - t_s)]^n}{n!} = \frac{B^n}{2^n} \sum_{k=0}^{n} \frac{e^{i(2k-n)\omega t}}{k!\,(n-k)!} A_n^k(\omega)\, A_n^k(-\omega) \tag{3-19}$$

Substituting (3-19) into (3-12), and using (3-13) for $H_n(\omega_1,\omega_2,...,\omega_n)$, gives

$$x(t) = \sum_{n=1}^{\infty} \frac{B^n}{2^n} \sum_{k=0}^{n} \frac{n!}{k!\,(n-k)!} H_{k,n-k}(\omega)\, e^{i(2k-n)\omega t} \qquad (3\text{-}20)$$

where, $H_{k,n-k}(\omega)$ denotes the n^{th}-order Volterra kernel transform $H_n(\omega_1,\omega_2,...,\omega_n)$ with the first k of the ω_i values equal to $(+\omega)$ and the remaining (n-k) values equal to $(-\omega)$.

From (3-20), any frequency component which is present in x(t) due to input $f(t)=B\cos\omega t$ can be calculated. For example, the $e^{iN\omega t}$ component of x(t) is

$$e^{iN\omega t} \sum_{m=0}^{\infty} \frac{(N+2m)!\,(B/2)^{(N+2m)}}{m!\,(N+2m)!} H_{N+m,m}(\omega) \qquad (3\text{-}21)$$

If the input $f(t)=B\cos(\omega t+\phi)$, then in (3-21), ωt should be correspondingly replaced by $(\omega t+\phi)$.

The same type of argument shows that when $f(t)=X\cos\omega_1 t+Y\cos\omega_2 t$, the $e^{i(M\omega_1+N\omega_2)t}$ component of x(t) is (N, M ≥ 0)

$$e^{i(M\omega_1+N\omega_2)t} \sum_{l=0}^{\infty}\sum_{k=0}^{\infty} \frac{(M+2l+N+2k)!\,(X/2)^{(M+2l)}(Y/2)^{(N+2k)}}{(M+l)!\;l!\;(N+k)!\;k!} H_{M+l,l;N+k,k}(\omega_1,\omega_2) \qquad (3\text{-}22)$$

The four subscripts of $H_n(\omega_1,\omega_2,...,\omega_n)$ mean that n=M+2l+N+2k and the first (M+l) values of ω_i are equal to (ω_1), the next l values equal to $(-\omega_1)$, the next N+k values equal to (ω_2) and the last k values equal to $(-\omega_2)$. Similarly, when $f(t)=X\cos\omega_1 t+Y\cos\omega_2 t+Z\cos\omega_3 t$, the $e^{i(L\omega_1+M\omega_2+N\omega_3)t}$ frequency component in x(t) is (L, M &N ≥ 0)

$$\sum_{l=0}^{\infty}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty} \frac{(L+2l+M+2k+N+2j)!\,(X/2)^{(L+2l)}\,(Y/2)^{(M+2k)}\,(Z/2)^{(N+2j)}}{(L+l)!\;l!\;(M+k)!\;k!\;(N+j)!\;j!}$$

$$e^{i(L\omega_1+M\omega_2+N\omega_3)t} H_{L+l,l;M+k,k;N+j,j}(\omega_1,\omega_2,\omega_3) \qquad (2\text{-}23)$$

From (3-21), the leading terms for the frequency component ω are:

$$e^{i\omega t}\ [\frac{B}{2}H_1(\omega) + \frac{3B^3}{8}H_3(\omega,\omega,-\omega) + ...] + CC \qquad (3\text{-}24)$$

where CC means complex conjugate since the response component must be real.

The leading term for the frequency component $(\omega_1+\omega_2)$ in (3-22) if ω_1 and ω_2 are incommensurable (ω_1 and ω_2 are said to be incommensurable if ω_1/ω_2 cannot be expressed as n_1/n_2 where n_1 and n_2 are integers), is:

$$e^{i(\omega_1+\omega_2)t}\ \frac{XY}{2}H_2(\omega_1, \omega_2) + ... + CC \qquad (3\text{-}25)$$

Similarly, the leading term for the frequency component $(\omega_1+\omega_2+\omega_3)$ in (3-23) if ω_1, ω_2 and ω_3 are incommensurable, is:

$$e^{i(\omega_1+\omega_2+\omega_3)t}\ \frac{XYZ}{2}H_3(\omega_1, \omega_2, \omega_3) + ... + CC \qquad (3\text{-}26)$$

On the other hand, n^{th}-order frequency response function $H_n(\omega_1,\omega_2,...,\omega_n)$, which *is* experimentally measurable, is defined as the output component $X(\omega_1,\omega_2,...,\omega_n)$ of x(t) at frequency $\omega=\omega_1+\omega_2+...+\omega_n$ due to the input $x(t)=A_1\cos\omega_1 t+A_2\cos\omega_2 t+...+A_n\cos\omega_n t$ (here A_i can be complex to accommodate the different phase shifts) divided by the input spectra, that is

$$H_n(\omega_1, \omega_2, ..., \omega_n) = \frac{X(\omega_1, \omega_2, ..., \omega_n)}{\prod_{r=1}^{n} A_r} \qquad (3\text{-}27)$$

Comparing (3-24), (3-25) and (3-26) with the definition of the higher-order frequency response function of equation (3-27), it can be seen that the measured n^{th}-order frequency response function $H_n(\omega_1,\omega_2,...,\omega_n)$ is the first-order approximation of the n^{th}-order Volterra kernel transform $H_n(\omega_1,\omega_2,...,\omega_n)$. To illustrate this point, take the second-order frequency response function as an example. If only the leading term is considered in equation (3-25) and the contribution of other kernels (even-ordered kernels after the second) at frequency $(\omega_1+\omega_2)$ can be neglected, then it becomes clear that the measured second-order frequency response function $H_2(\omega_1,\omega_2)$ based on (3-27) will be the same as the second-order Volterra kernel transform $H_2(\omega_1,\omega_2)$. In general, however, there will be

some contribution from the higher even-ordered Volterra kernels and the estimated second-order frequency response function is an approximate of the uniquely defined second-order Volterra kernel transform. The same argument holds for other higher-order frequency response functions. Based on this observation, the Volterra kernel $h_n(\tau_1,\tau_2,...,\tau_n)$ and its transform $H_n(\omega_1,\omega_2,...,\omega_n)$ have direct physical meaning and interpretation.

It is worth pointing out here that the Volterra kernel transforms $H_n(\omega_1,\omega_2,...,\omega_n)$ are mathematically unique. However, the n^{th}-order frequency response functions $H_n(\omega_1,\omega_2,...,\omega_n)$ are usually input-output dependent like the classical first-order frequency response function $H_1(\omega)$ measured using a sine wave excitation. Since we are only able to deal with truncated series, these measured frequency response functions will, in some cases, give more accurate representation than the equivalent Volterra kernel transforms, which are by no means measurable.

3.3.2 ANALYTICAL CALCULATION OF FREQUENCY RESPONSE FUNCTIONS

So far, it has been shown how the output x(t) and input f(t) of a nonlinear system are related through the system's frequency response functions (or, more strictly, the Volterra kernel transforms), and it is appropriate here to investigate what forms and what characteristics the higher-order frequency response functions of typical nonlinear mechanical systems possess. There are some different methods for analytically calculating the frequency response functions of a known nonlinear system and what is discussed here is the harmonic probing method [46].

Suppose that the input f(t) is

$$f(t) = \sum_{r=1}^{n} A_r e^{i\omega_r t} \tag{3-28}$$

where the ω_r values are incommensurable and, for simplicity, let A_r=1(r=1,n) since the analytical n^{th}-order FRF, i.e. the n^{th}-order Volterra kernel transform (we define $H_n(\omega_1,\omega_2,...,\omega_n)$ as the analytical n^{th}-order FRF), is unique. Substituting into (3-4), then $H_n(\omega_1,\omega_2,...,\omega_n)$ is given [46] by

$H_n(\omega_1,\omega_2,...,\omega_n)$ = { coefficient of $e^{i(\omega_1 + \omega_2 + ... + \omega_n)t}$ term in the expression of x(t) } (3-29)

Based on (3-29), it is possible to compute $H_1(\omega_1)$, $H_2(\omega_1,\omega_2)$,... of a nonlinear mechanical system successively. To illustrate this, first consider an SDOF system given by

$$m\ddot{x} + c\dot{x} + kx + k_1x^2 + k_2x^3 = f(t) \tag{3-30}$$

Let $f(t) = e^{i\omega t}$ and substitute into (3-4), then

$$x(t) = \sum_{n=1}^{\infty} \frac{1}{2^n} \sum_{k=0}^{n} \frac{n!}{k!\,(n-k)!} H_{k,n-k}(\omega)\, e^{i(2k-n)\omega t} \tag{3-31}$$

$$\dot{x}(t) = \sum_{n=1}^{\infty} \frac{1}{2^n} \sum_{k=0}^{n} \frac{n!}{k!\,(n-k)!} H_{k,n-k}(\omega)\, i(2k-n)\,\omega\, e^{i(2k-n)\omega t} \tag{3-32}$$

$$\ddot{x}(t) = \sum_{n=1}^{\infty} \frac{1}{2^n} \sum_{k=0}^{n} -\frac{n!}{k!\,(n-k)!} H_{k,n-k}(\omega)\, (2k-n)^2\,\omega^2\, e^{i(2k-n)\omega t} \tag{3-33}$$

The first few terms of $x^2(t)$ and $x^3(t)$ are

$$x^2(t) = H_1^2(\omega)\, e^{i2\omega t} + 2\,H_1(\omega)\,H_2(\omega,\omega)\, e^{i3\omega t} + H_2^2(\omega,\omega)\, e^{i4\omega t} \tag{3-34}$$

$$x^3(t) = H_1^3(\omega)\, e^{i3\omega t} + 3\,H_1^2(\omega)\,H_2(\omega,\omega) e^{i4\omega t} + 3\,H_1^2(\omega)\,H_2^2(\omega,\omega)\, e^{i5\omega t} \tag{3-35}$$

Substitute (3-31)-(3-35) into (3-30) and set the coefficients of $e^{i\omega t}$ at both sides to be equal:

$$H_1(\omega) = \frac{1}{k - m\omega^2 + ic\omega} \tag{3-36}$$

The first-order Volterra kernel transform is independent of the nonlinear terms present in the equation of motion and represents the dynamic characteristics of the linear part of the nonlinear system.

Now, let $f(t)=e^{i\omega_1 t}+e^{i\omega_2 t}$ and substitute into (3-4), then

$$x(t) = \sum_{N=0}^{\infty} \sum_{M=0}^{\infty} \sum_{l=0}^{\infty} \sum_{k=0}^{\infty} e^{i(M\omega_1 + N\omega_2)t}\, 2^{-(M+N+2l+2k)} \frac{(M+2l+N+2k)!}{(M+l)!\, l!\, (N+k)!\, k!}$$

$$\cdot H_{M+l,l;N+k,k}(\omega_1, \omega_2) \qquad (3\text{-}37)$$

Differentiate x(t) to get $\dot{x}(t)$ and $\ddot{x}(t)$ and then substitute into (3-30) in similar way as for the $H_1(\omega)$ calculation, and let the coefficients of $e^{i(\omega_1+\omega_2)t}$ on both sides be equal:

$$H_2(\omega_1, \omega_2) = - k_1 H_1(\omega_1) H_1(\omega_2) H_1(\omega_1+\omega_2) \qquad (3\text{-}38)$$

From (3-38) it can be seen that $H_2(\omega_1,\omega_2)$ has all the poles which $H_1(\omega)$ has and is proportional to the coefficient of the quadratic nonlinearity term k_1.

Similarly, if we let $f(t)=e^{i\omega_1 t}+e^{i\omega_2 t}+e^{i\omega_3 t}$, then

$$x(t) = \sum_{N=0}^{\infty} \sum_{M=0}^{\infty} \sum_{L=0}^{\infty} \sum_{l=0}^{\infty} \sum_{k=0}^{\infty} \sum_{j=0}^{\infty} H_{L+l,l;M+k,k;N+j,j}(\omega_1,\omega_2,\omega_3)$$

$$e^{i(M\omega_1 + N\omega_2 + L\omega_3)t}\, 2^{-(N+M+L+2l+2k+2j)} \frac{(L+2l+M+2k+N+2j)!}{(L+l)!\, l!\, (M+k)!\, k!\, (N+j)!\, j!} \qquad (3\text{-}39)$$

Upon substituting into (3-30) and letting the coefficients of $e^{i(\omega_1+\omega_2+\omega_3)t}$ on both sides be equal:

$$H_3(\omega_1, \omega_2, \omega_3) = - \frac{2k_1}{3} H_1(\omega_1 + \omega_2 + \omega_3) [\, H_1(\omega_3) H_1(\omega_1, \omega_2) + H_1(\omega_2) H_2(\omega_1, \omega_3) +$$

$$H_1(\omega_1) H_1(\omega_2, \omega_3) \,] - k_2 H_1(\omega_1 + \omega_2 + \omega_3) H_1(\omega_1) H_1(\omega_2) H_1(\omega_3) \qquad (3\text{-}40)$$

In fact, it has been established in [46] that for a physically realisable system specified by the nonlinear differential equation as

$$F(d/dt)\, y + \sum_{m=2}^{\infty} a_m y^m = f(t) \qquad (3\text{-}41)$$

where F(x) is a function of the differential operator d/dt, the n^{th}-order Volterra kernel transform is given by

$$H_n(\omega_1, \omega_2, ..., \omega_n) = \frac{1}{n!} \frac{\sum_{m=2}^{n} a_m H_n^{(m)}(\omega_1, \omega_2, ..., \omega_n)}{F(i\omega_1 + i\omega_2 + ... + i\omega_n)} \tag{3-42}$$

where $H_n^{(m)}(\omega_1, \omega_2, ..., \omega_n)$ is defined in such a way that

$$H_2^{(2)}(\omega_1, \omega_2) = 2! H_1(\omega_1) H_1(\omega_2)$$

$$H_3^{(2)}(\omega_1,\omega_2,\omega_3) = (2!)^2 [H_1(\omega_1) H_2(\omega_2,\omega_3) + H_1(\omega_2)H_2(\omega_1,\omega_3) + H_1(\omega_3)H_2(\omega_1,\omega_2)]$$

$$H_4^{(2)}(\omega_1,\omega_2,\omega_3,\omega_3) = 3! 2! H_1(\omega_1) H_3(\omega_2,\omega_3,\omega_4) + ... \tag{3-43}$$

It should be noted here that higher nonlinear terms $a_m(m>n)$ have no influence on the lower-order Volterra kernel transforms $H_s(\omega_1,\omega_2,...,\omega_s)$ $(s\leq n)$ while the lower nonlinear terms $(m\leq r)$ *do* have an influence on all the higher-order Volterra kernel transforms, $(s\geq r)$. Therefore, a system with cubic stiffness nonlinearity does not possess second-order Volterra kernels while a system with quadratic stiffness nonlinearity has, in general, all the Volterra kernels. The second- and third-order Volterra kernel transforms of the SDOF nonlinear system described by equation (3-30) are calculated and are shown in Figs.3.5-3.7.

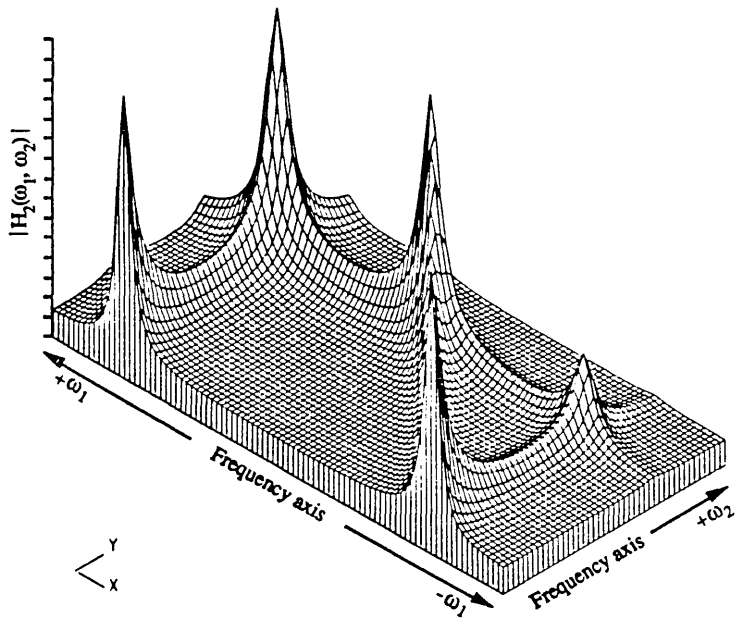


Fig.3.5 - Analytical Second Order Frequency Response Function of an SDOF Nonlinear System (Modulus Linear Scale, x axis ω_1: -275 — 275, y axis ω_2: 0 — 275 rad./s)

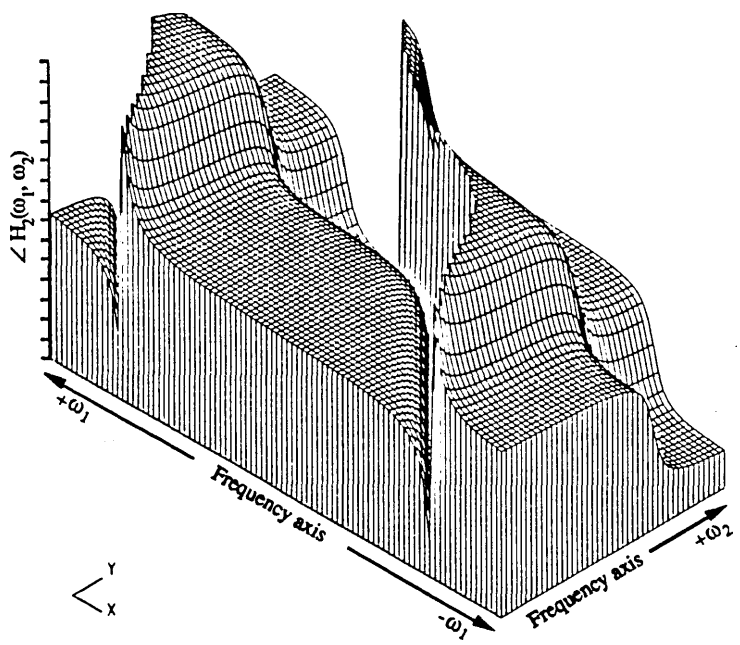


Fig.3.6 - Analytical Second Order Frequency Response Function of an SDOF Nonlinear System (Phase Linear Scale, x axis ω_1: -275 — 275, y axis ω_2: 0 — 275 rad./s)

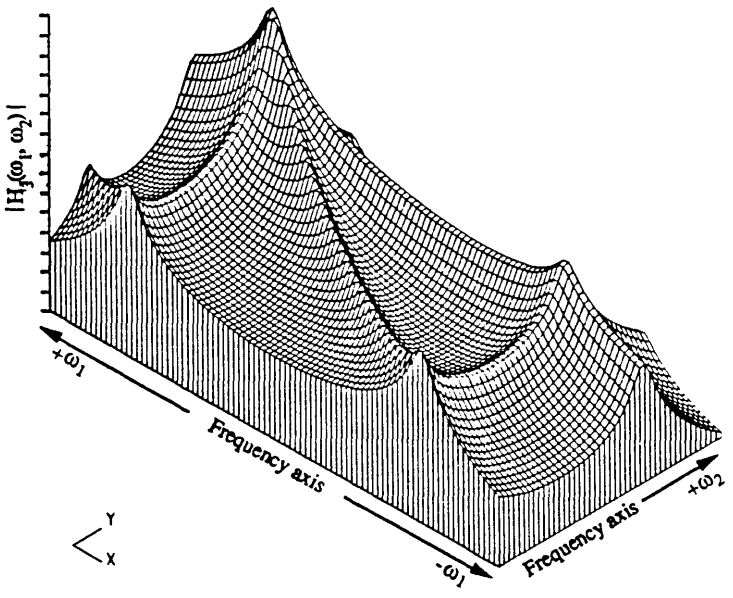


Fig.3.7 - Analytical Third Order Frequency Response Function of an SDOF Nonlinear System (Modulus Log. Scale, x axis ω_1: -275 — 275, y axis ω_2: 0 — 275 rad./s, $\omega_3=\omega_2$)

To illustrate the physical interpretation of higher-order frequency response functions, the second-order frequency response function shown in figure 3.5 is discussed. As shown in Fig.3.8, the components of the second-order frequency response function near both frequency axes represent the fundamental frequency components of the response. While components along both diagonals are the static components ($\omega_1=-\omega_2$) and second harmonic components ($\omega_1=\omega_2$) of the response respectively. All the other components defined on the ω_1 *vs* ω_2 plane are the combinational frequency components of the response which, as will be discussed later, are important in cases of nonlinear MDOF systems because these combinational frequency components can excite the system into its resonances when they coincide with some of the natural frequencies of the system.

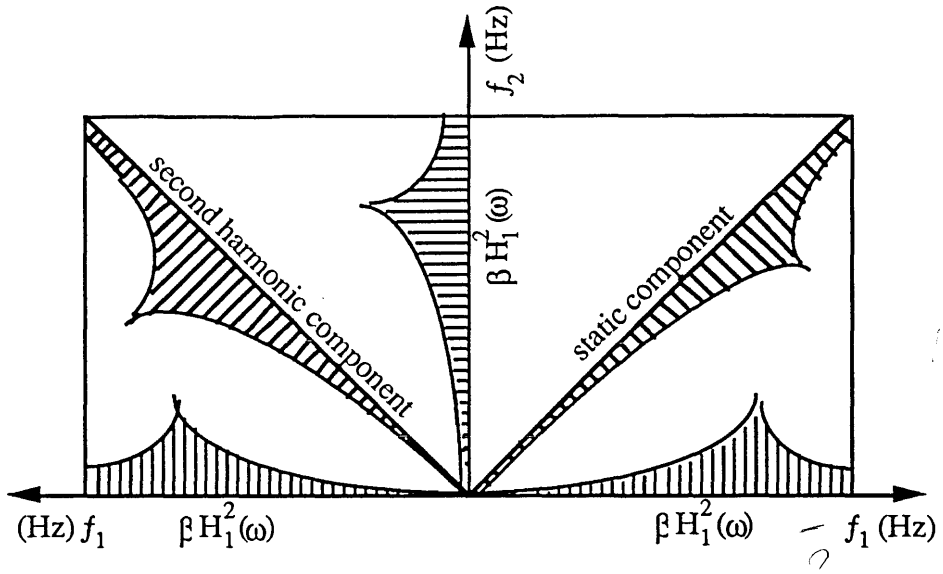


Fig.3.8 - Physical Interpretation of Second-Order FRF

Also, one special case which has been treated in the literature [49] and is discussed here is the square-law system given as:

$$m\ddot{z} + c\dot{z} + kz = f(t) \tag{3-44}$$

$$x(t) = z^2(t) \tag{3-45}$$

This is an ideal Volterra system for which all the kernels except the second one are zero and it is easy to prove that its second-order Volterra kernel transform is

$$H_2(\omega_1, \omega_2) = H_1(\omega_1)\, H_1(\omega_2) \tag{3-46}$$

where $H_1(\omega)$ is given by (3-36). This second-order frequency response function of the square-law system is shown in Fig.3.9. Because of its purely quadratic nature, the system's response is dominated by the second harmonic and static components.

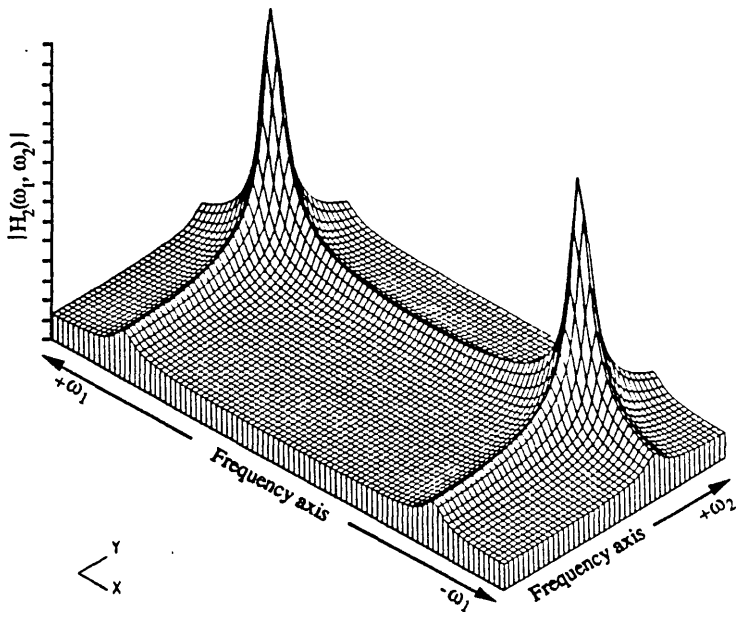


Fig.3.9 - Analytical Second Order Frequency Response Function of the Square-law System (Modulus Linear Scale, x axis ω_1: -275 — 275, y axis ω_2: 0 — 275 rad./s)

The same argument holds for MDOF nonlinear systems, although the analytical calculation becomes a bit complicated. To see this, the 3DOF nonlinear system as shown in Fig.3.10 is considered. The governing differential equations of the motion are given as

$$\ddot{x}_1 + 2\alpha\dot{x}_1 + 2kx_1 - \alpha\dot{x}_2 - kx_2 + \beta x_1^2 = f(t) \tag{3-47}$$

$$\ddot{x}_2 + 3\alpha\dot{x}_2 + 3kx_2 - \alpha\dot{x}_1 - \alpha\dot{x}_3 - kx_1 - kx_3 + \beta x_2^2 = 0 \tag{3-48}$$

$$\ddot{x}_3 + 2\alpha\dot{x}_3 + 2kx_3 - \alpha\dot{x}_2 - kx_2 + \beta x_3^2 = 0 \tag{3–49}$$

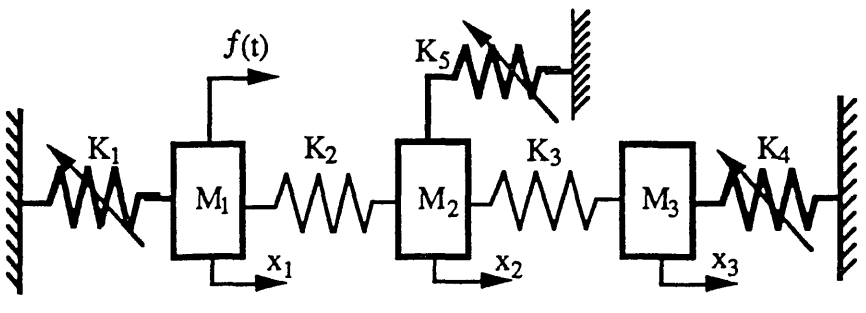


Fig.3.10 - A Three Degrees-of-freedom Nonlinear System

The first-order Volterra kernel transforms of the system are the frequency response functions of the linear system (β=0) and here only the second-order Volterra kernel transforms are going to be calculated. Similarly to the SDOF case, let $f(t)=e^{i\omega_1 t}+e^{i\omega_2 t}$, then

$$x_r(t) = \sum_{N=0}^{\infty} \sum_{M=0}^{\infty} \sum_{l=0}^{\infty} \sum_{k=0}^{\infty} \frac{(M+2l+N+2k)!\ (X/2)^{M+2l}\ (Y/2)^{N+2k}}{(M+l)!\ l!\ (N+k)!\ k!}$$

$$e^{i(M\omega_1 + N\omega_2)t}\ H^{rl}_{M+l,l;N+k,k}(\omega_1, \omega_2) \qquad (3\text{-}37)$$

where $H^{rl}_{M+l,l;N+k,k}(\omega_1, \omega_2)$ (r=1,2,3) are the transfer Volterra kernel transforms between coordinates x_1 and x_r. Substitute $x_r(t)$ and its derivatives $\dot{x}_r(t)$ and $\ddot{x}_r(t)$ into (3-47)-(3-49) and let the coefficients of $e^{i(\omega_1 t+\omega_2 t)}$ on both sides of the equations be equal respectively, so that the following algebraic equations are obtained:

$$[\ (\omega_1+\omega_2)^2 - i2\alpha(\omega_1+\omega_2) - 2k]\ H^{11}_2(\omega_1, \omega_2) + [\ i\alpha(\omega_1+\omega_2) + k]\ H^{21}_2(\omega_1, \omega_2) =$$

$$\beta\ H^{11}_1(\omega_1)\ H^{11}_1(\omega_2) \qquad (3\text{-}51)$$

$$[\ (\omega_1+\omega_2)^2 - i3\alpha(\omega_1+\omega_2) - 3k]\ H^{21}_2(\omega_1, \omega_2) + [\ i\alpha(\omega_1+\omega_2) + k]\ H^{11}_2(\omega_1, \omega_2) +$$

$$[\ i\alpha(\omega_1+\omega_2) + k]\ H^{21}_2(\omega_1, \omega_2) = \beta\ H^{21}_1(\omega_1)\ H^{21}_1(\omega_2) \qquad (3\text{-}52)$$

$$[\ (\omega_1+\omega_2)^2 - i2\alpha(\omega_1+\omega_2) - 2k]\ H^{31}_2(\omega_1, \omega_2) + [\ i\alpha(\omega_1+\omega_2) + k]\ H^{21}_2(\omega_1, \omega_2) =$$

$$\beta\ H^{31}_1(\omega_1)\ H^{31}_1(\omega_2) \qquad (3\text{-}53)$$

From (3-51)-(3-53), the three unknowns $H^{r1}_2(\omega_1,\omega_2)$ (r=1,2,3) can be calculated. The analytical second-order point frequency response function $H^{11}_2(\omega_1,\omega_2)$ (excitation and response are at the same coordinate) is shown in Fig.3.11. In this case, the appearance of combinational resonances (peaks which do not lie on the two diagonals) is clearly demonstrated.

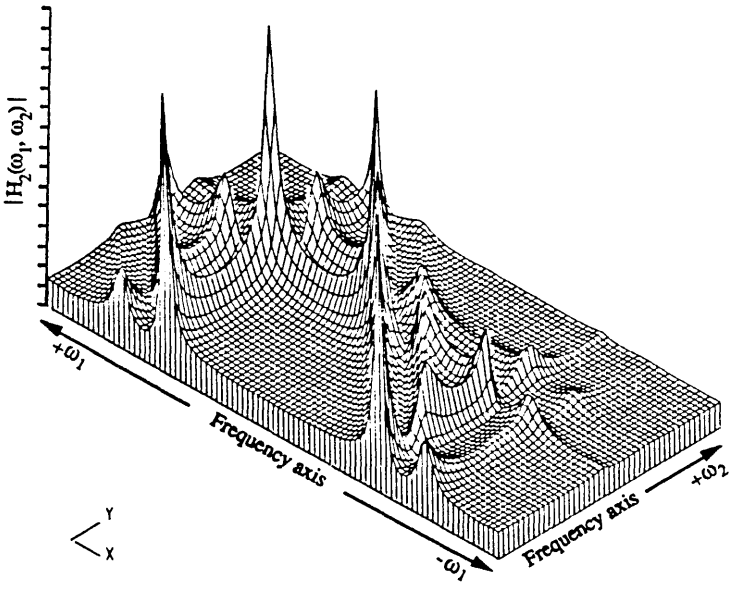


Fig.3.11 - Analytical Second Order Frequency Response Function of a 3DOF Nonlinear System (Modulus Linear Scale, x-axis ω_1: -275 — 275, y-axis ω_2: -275 — 275 rad./s)

Since the above calculations are based on the definition of (3-29), the calculated $H_n(\omega_1,\omega_2,...,\omega_n)$ are of receptance-like frequency response functions. If x(t) is changed into $\dot{x}(t)$ in (3-29), then the calculated Volterra kernel transforms are the mobility-like frequency response functions and it can be easily seen that the receptance-like and mobility-like Volterra kernel transforms are related by

$$H_n^R(\omega_1, \omega_2, ..., \omega_n) = \mathbf{i}(\omega_1 + \omega_2 + ...+ \omega_n)\, H_n^M(\omega_1, \omega_2, ..., \omega_n) \qquad (3\text{-}54)$$

3.3.3 MEASUREMENT OF HIGHER-ORDER FREQUENCY RESPONSE FUNCTIONS USING HARMONIC PROBING METHOD

As in the case of the measurement of frequency response functions of a linear system, different measurement techniques can be employed to measure the frequency response functions of a nonlinear system, such as (i) multi-impulse technique [30], (ii) harmonic probing method [48], (iii) correlation analysis using random input [47] and (iv) NARMAX time series modelling technique [18] etc.. Among these different methods, the

harmonic probing method and correlation analysis using random input are the most commonly referred methods and have been given much discussion in the literature. The theoretical basis for the measurement of higher-order frequency response functions by the harmonic probing method has already been presented and what is to be discussed next is the practical applicability of the method through numerical case study. The theoretical basis of the correlation technique and the numerically simulated case studies of higher-order frequency response function (Wiener kernel transform) measurement will be given later in this Chapter.

As explained earlier, if the input to a nonlinear system is $f(t)=A_1\cos\omega_1 t+...+A_n\cos\omega_n t$, where the frequency components ω_i (i=1,2,..,n) are incommensurable, then the n^{th}-order frequency response function $H_n(\omega_1,\omega_2,...,\omega_n)$, which is the first-order approximation of the n^{th}-order Volterra kernel transform $H_n(\omega_1,\omega_2,...,\omega_n)$, can be estimated using (3-27). On the thus- estimated $H_n(\omega_1,\omega_2,...,\omega_n)$, with the exception of $H_n(\omega_1,\omega_2,...,\omega_n)$ which is the leading term, all the higher kernel transforms $H_{n+2l}(\omega_1,\omega_2,...,\omega_{n+2l})$ (l=1,2,...) may have a contribution. This, as already pointed out, may make the measured $H_n(\omega_1,\omega_2,...,\omega_n)$ a better representation of the system than the corresponding Volterra kernel transforms $H_n(\omega_1,\omega_2,...,\omega_n)$ since we are only able to deal with the truncated series. Measurement of the first-order frequency response functions of nonlinear systems has already been well established [39] and here the measurement of higher-order (mainly second-order) frequency response functions is discussed. Although the technique can theoretically be readily extended to the the measurement of higher- (than the second) order terms, because of the time and effort involved, it is hardly practical to measure beyond the third and for most practical nonlinear systems, only the lower few terms are in general required in order to provide accurate representation of the system.

Based on the above-mentioned theory, numerical simulation of the measurement of the second-order frequency response functions for a square-law system described by (3-34)-(3-35) and an SDOF nonlinear system given by (3-3) was carried out. For computational convenience, the input was set to be $f(t)=A\sin\omega_1 t+B\sin\omega_2 t$ where A is real and equal to B (in real practical measurements, A and B can be set independently and can be complex to accommodate the relative phase difference of these two sinusoids). The response of the system x(t) was calculated using a numerical integration technique and, after the transient dies away, the signal was sampled and Fourier transformed to find the $2\omega_1$, $2\omega_2$, $(\omega_1+\omega_2)$ and $(\omega_1-\omega_2)$ frequency components of x(t). Suppose the frequency components of x(t) at $2\omega_1$, $2\omega_2$, $(\omega_1+\omega_2)$ and $(\omega_1-\omega_2)$ are Y_1, Y_2, Y_3 and Y_4 respectively, then based on (3-27), the following four points of $H_2(\omega_1,\omega_2)$ on ω_1-ω_2 plane are given by:

$$H_2(\omega_1, \omega_1) = \frac{2Y_1}{A^2} \tag{3-55}$$

$$H_2(\omega_2, \omega_2) = \frac{2Y_2}{B^2} \tag{3-56}$$

$$H_2(\omega_1, \omega_2) = \frac{Y_3}{AB} \tag{3-57}$$

$$H_2(\omega_1, -\omega_2) = \frac{Y_4}{AB} \tag{3-58}$$

The measured $H_2(\omega_1,\omega_1)$ and $H_2(\omega_2,\omega_2)$ lie on the diagonal of $\omega_1=\omega_2$ of ω_1 *vs* ω_2 plane. When ω_1 and ω_2 are varied, the value of $H_2(\omega_1,\omega_2)$ in any desired region on ω_1 *vs* ω_2 plane can be obtained. Considering the mathematical symmetry of $H_2(\omega_1,\omega_2)$, if the frequency range of interest is ω_s-ω_n, then what needs to be measured is the triangular region where ω_1:ω_s-ω_n and ω_2:ω_s-ω_n as shown in Fig.3-12.

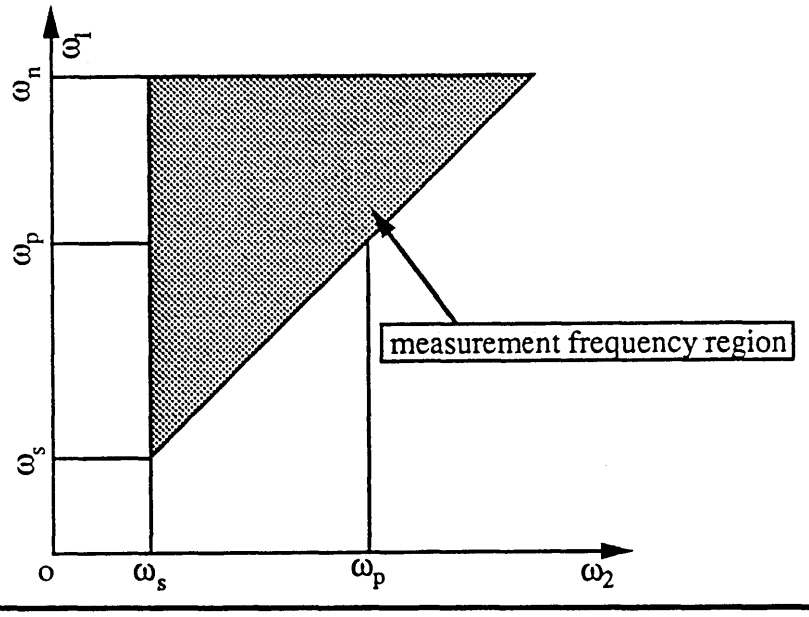


Fig.3.12 - Illustration of Measurement Frequency Region Required

In this way, the second-order frequency response functions of the square-law system and the SDOF nonlinear system are measured and they are shown in Figs.3.13 and 3.14. When comparing figure 3.13 with its analytical counterpart figure 3.9, it can be seen that except for some spurious spikes at both diagonals of $\omega_1=\omega_2$ and $\omega_1=-\omega_2$, the measurement results are quite acceptable. The spurious spikes appear because in the numerical simulation, ω_1 and ω_2 were chosen to be integers and so the condition that ω_1 and ω_2 should be incommensurable was violated and for this specific system, it can be shown that such violation only causes errors when $\omega_1=\omega_2$ and then, $H_2(\omega_1,\omega_2)$ is overestimated by 100%. This problem can be removed by measuring the diagonal elements of $H_2(\omega_1,\omega_2)$ at $\omega_1=\omega_2$ and $\omega_1=-\omega_2$ using one single sinusoid input and then

measuring the second harmonic component (on the diagonal $\omega_1=\omega_2$) and the DC component (on the diagonal $\omega_1=-\omega_2$) of the output. On the other hand, in the case of a nonlinear SDOF system, the situation becomes somewhat complicated. When comparing figure 3.14 with its analytical counterpart figure 3.5, we see that the measured results are not very bad except for some small spurious spikes appearing on the plane, again showing the effects of the violation of the condition that ω_1 and ω_2 must be incommensurable.

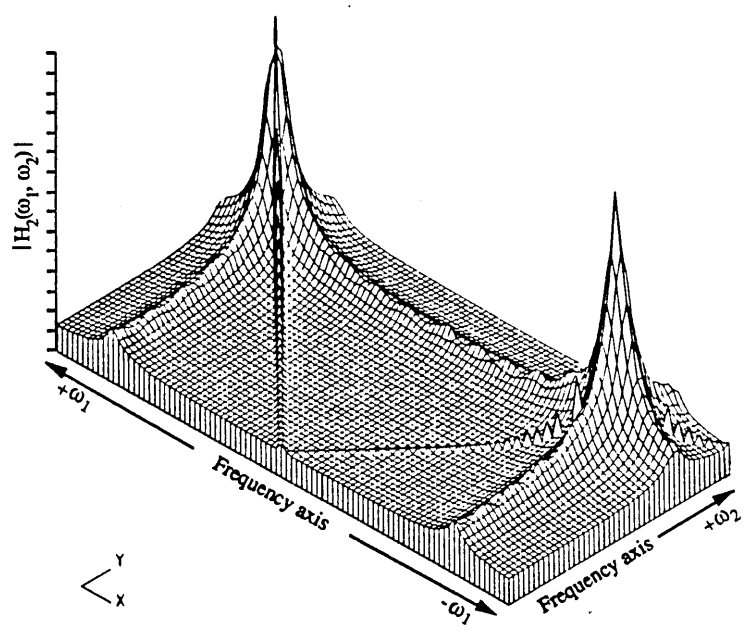


Fig.3.13 - Measured Second Order FRF of the Square-law System Using Harmonic Probing Method (Modulus Linear Scale, x-axis ω_1:-260 — 260, y-axis ω_2: -260 — 260 rad./s)

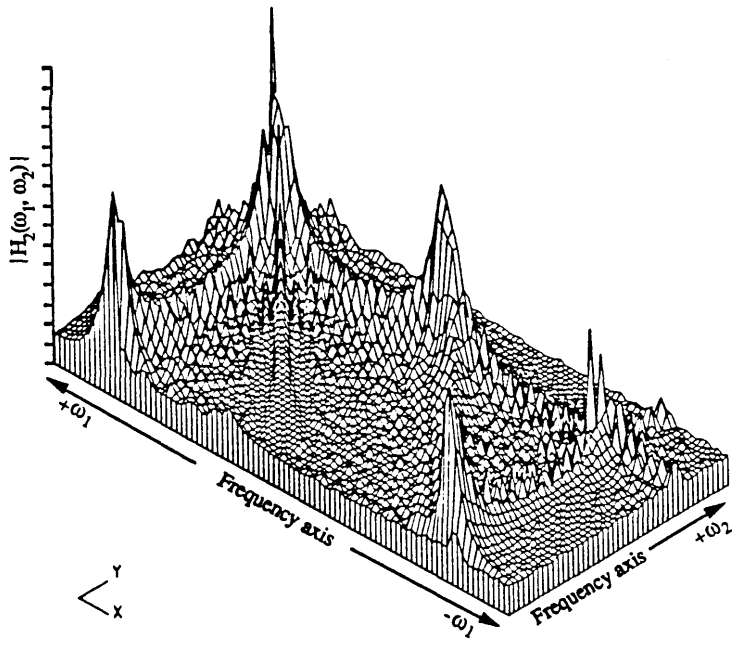


Fig.3.14 - Measured Second Order FRF of the SDOF Nonlinear System Using Harmonic Probing Method (Modulus Linear Scale, x-axis ω_1:-260 — 260, y-axis ω_2: -260 — 260 rad./s)

Of course, it is possible to set ω_1 and ω_2 to be incommensurable in the measurement, but then there may be a leakage problem in the DFT of response x(t) because, in this case, it is not possible to make all these major frequency components ω_1, ω_2, $2\omega_1$, $2\omega_2$, $\omega_1+\omega_2$, $\omega_1-\omega_2$,... contained in the x(t) coincide with frequency lines. Some further research is needed to investigate how cleanly and consistently the second-order frequency response functions of a nonlinear system can be measured using the harmonic probing method based on DFT algorithms. However, one possible way of getting around the leakage problem, which the author suggests here, is to use the correlation technique. If, say, the frequency component Y_3 of x(t) at $(\omega_1+\omega_2)$ is of interest, then instead of obtaining Y_3 using the DFT (which can sometimes cause serious leakage errors), the correlation technique can be used by multiplying $\sin(\omega_1+\omega_2)t + i\cos(\omega_1+\omega_2)t$ to x(t) and then integrating the product with time as:

$$Y_3 = \lim_{T\to\infty} \frac{2}{(\omega_1 + \omega_2)\, T} \int_0^T x(t) \left[\sin(\omega_1 + \omega_2)t + i\cos(\omega_1 + \omega_2)t \right] dt \qquad (3\text{-}59)$$

Based on this correlation technique, together with the diagonal elements of $H_2(\omega_1,\omega_2)$ to be measured using single sinusoid input, clean and consistent measurement results could be obtained.

3.4 CORRELATION ANALYSIS USING RANDOM INPUT

The correlation method for the measurement of frequency response functions using random input has been widely used in structural modal testing because of its convenience. In the study of nonlinear structures, as discussed in Chapter 2, first-order frequency response functions (first-order Wiener kernel transforms) can be measured using random excitation. Corresponding to different excitation levels (or input power spectra), the thus measured first-order frequency response functions of a nonlinear system are in general different and, therefore, the existence of nonlinearity can be detected by comparing FRF data measured at different excitation levels. The theoretical aspects of this first-order frequency response function analysis based on random input are given in refs. [22,50]. Nevertheless, anything beyond the detection in the identification of nonlinear systems by application of the thus measured first-order frequency response functions is difficult. However, in addition to the auto- and cross-correlation analysis which is used in the calculation of first-order FRF, if we can do higher-order correlation analysis, then as in the case of the Volterra kernel measurement, there is a systematic way of characterising a nonlinear system by measuring its higher-order Wiener kernels using random input. The theory behind this practice is the Wiener series of nonlinear systems.

3.4.1 THE WIENER SERIES

In the Wiener theory of nonlinear systems, if the input $f(t)$ is a white Gaussian time series with autocorrelation function $\phi_{ff}(\tau)=A\delta(\tau)$, then the output x(t) of a nonlinear system can be expressed by the orthogonal expression:

$$x(t) = \sum_{n=1}^{\infty} G_n [\, k_n; f(t) \,] \tag{3-60}$$

in which $\{k_n(\tau_1,\tau_2,...,\tau_n)\}$ is the set of Wiener kernels of the nonlinear system which, like the set of Volterra kernels $\{h_n(\tau_1,\tau_2,...,\tau_n)\}$, serve to describe the system and $\{G_n\}$ is a complete set of orthogonal functionals. For a linear system, all the higher-order kernels except k_0 and $k_1(\tau)$ are zero. Unlike the n^{th} Volterra functional, which is homogeneous and defined as

$$T_n [h_n; f(t)] = \int_{-\infty}^{+\infty} d\tau_1 \ ... \int_{-\infty}^{+\infty} d\tau_n \, h_n (\tau_1, ..., \tau_n) \prod_{r=1}^{n} f(t - \tau_r) \qquad (3\text{-}61)$$

the Wiener G-functionals are a set of nonhomogeneous Volterra functionals defined as:

$$G_n [k_n; f(t)] = k_{0(n)} + \sum_{r=1}^{n} \int_{-\infty}^{+\infty} d\tau_1 \ ... \int_{-\infty}^{+\infty} d\tau_r \, k_{r(n)} (\tau_1, ..., \tau_r) \prod_{s=1}^{r} f(t - \tau_s) \qquad (3\text{-}62)$$

where $k_{n(n)} \equiv k_n$ is known as the n^{th}-order Wiener kernel and $k_{n-1(n)}$, ..., $k_{0(n)}$ are known as the derived Wiener kernels of the Wiener G-functional. $G_n [k_n; f(t)]$ satisfy

$$\overline{T_m [h_m; f(t)] G_n [k_n; f(t)]} = 0 \qquad \text{for } m < n \qquad (3\text{-}63)$$

where the over bar means taking the average of the process. Theoretically, all the derived Wiener kernels $k_{n-1(n)}$, ..., $k_{0(n)}$ of n^{th}-order can be determined uniquely by the leading n^{th}-order Wiener kernel k_n when (3-63) is satisfied for all integer values of m<n and, therefore, in the notation $G_n[k_n;f(t)]$, only the leading term k_n is specified as in the case of Volterra functional $T_n[h_n; f(t)]$. The first few $G_n[k_n; f(t)]$ of a general nonlinear system, are given as [6]

$$G_0 [k_0; f(t)] = k_0 \qquad (k_0 \text{ is a constant}) \qquad (3\text{-}64)$$

$$G_1 [k_1; f(t)] = \int_{-\infty}^{+\infty} k_1(\tau_1) f(t - \tau_1) \, d\tau_1 \qquad (3\text{-}65)$$

$$G_2 [k_2; f(t)] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} k_2(\tau_1, \tau_2) f(t - \tau_1) f(t - \tau_2) \, d\tau_1 \, d\tau_2 - A \int_{-\infty}^{+\infty} k_2(\tau_1, \tau_1) \, d\tau_1 \qquad (3\text{-}66)$$

$$G_3 [k_3; f(t)] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} k_3(\tau_1, \tau_2, \tau_3) f(t - \tau_1) f(t - \tau_2) f(t - \tau_3) \, d\tau_1 \, d\tau_2 \, d\tau_3$$

$$- 3A \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} k_3(\tau_1, \tau_2, \tau_2)\, f(t - \tau_1)\, d\tau_1\, d\tau_2 \qquad (3\text{-}67)$$

The relationship between the Volterra kernel $h_n(\tau_1,\tau_2,...,\tau_n)$ and the Wiener kernel $k_n(\tau_1,\tau_2,...,\tau_n)$ is that the system's n^{th}-order Volterra kernel is equal to the system's n^{th}-order Wiener kernel plus the sum of all the (even or odd order) derived Wiener kernels that are of the n^{th}-order, that is

$$h_n(\tau_1,\tau_2,...,\tau_n) = k_n(\tau_1,\tau_2,...,\tau_n) + k_{n-2(n)}(\tau_1,\tau_2,...,\tau_{n-2}) + ... \qquad (3\text{-}68)$$

From (3-68), it can be seen that since the derived Wiener kernels are determined uniquely by their leading Wiener kernel, a given system's Volterra kernels can be obtained uniquely from the system's Wiener kernels (leading Wiener kernels). Also, it should be noted from equations (3-66) and (3-67) that as the input level $A \rightarrow 0$, the derived kernels approach zero and the leading Wiener kernels approach the Volterra kernels. On the other hand, it should be pointed out that unlike the Volterra kernels, which are mathematically unique, the Wiener kernels are input-output dependent and since the Volterra kernels $h_n(\tau_1,\tau_2,...,\tau_n)$ which uniquely determine the system are uniquely determined by the Wiener kernels $k_n(\tau_1,\tau_2,...,\tau_n)$, the measured Wiener kernels $k_n(\tau_1,\tau_2,...,\tau_n)$ also uniquely determine the system.

3.4.2 DETERMINATION OF WIENER KERNELS BY CROSS-CORRELATION

As in the case of Volterra series representation, under Wiener series representation, the problem of identifying a nonlinear system becomes one of determining all the Wiener kernels which describe the system. The orthogonality property of G_n and the statistical properties of Gaussian noise enable the Wiener kernels to be determined using a cross-correlation technique. The first four kernels are given [47] as

$$k_0 = \overline{x(t)} \qquad (3\text{-}69)$$

$$k_1(\tau) = \frac{1}{A}\, \overline{x(t)\, f(t - \tau)} \qquad (3\text{-}70)$$

$$k_2(\tau_1, \tau_2) = \frac{1}{2A^2}\, \overline{[x(t) - k_0]\, f(t - \tau_1)\, f(t - \tau_2)} \qquad (3\text{-}71)$$

$$k_3(\tau_1, \tau_2, \tau_3) = \frac{1}{3!\,A^3} \overline{[x(t) - G_1[k_1; f(t)]]\, f(t-\tau_1)\, f(t-\tau_2)\, f(t-\tau_3)} \tag{3-72}$$

Equations (3-69)-(3-72) serve as a basis for the measurement of Wiener kernels. To illustrate the derivation of these equations, consider the calculation of the second-order kernel $k_2(\tau_1,\tau_2)$. From equation (3-60), $\overline{x(t)f(t-\tau_1)f(t-\tau_2)}$ becomes

$$\overline{x(t)f(t-\tau_1)f(t-\tau_2)} = \overline{\{\sum_{n=0}^{\infty} G_n[k_n; f(t)]\}\, f(t-\tau_1)\, f(t-\tau_2)} \tag{3-73}$$

and since, from the orthogonality relationship of (3-63), the functionals G_n for n>2 are orthogonal to $f(t-\tau_1)f(t-\tau_2)$, which is a homogeneous functional of second degree, therefore equation (3-73) can be rewritten as

$$\overline{x(t)f(t-\tau_1)f(t-\tau_2)} = \overline{\{\sum_{n=0}^{2} G_n[k_n; f(t)]\}\, f(t-\tau_1)\, f(t-\tau_2)} \tag{3-74}$$

For n = 0, the average involving G_0 is:

$$\overline{G_0[k_0; f(t)]\, f(t-\tau_1)\, f(t-\tau_2)} = \overline{k_0\, f(t-\tau_1)\, f(t-\tau_2)} = k_0\, \delta(\tau_1 - \tau_2) \tag{3-75}$$

The average for n=1 is:

$$\overline{G_1[k_1; f(t)]\, f(t-\tau_1)\, f(t-\tau_2)} =$$

$$\overline{[\int_{-\infty}^{+\infty} k_1(\sigma_1)\, f(t-\sigma_1)\, d\sigma_1]\, f(t-\tau_1)\, f(t-\tau_2)} =$$

$$\int_{-\infty}^{+\infty} k_1(\sigma_1)\, \overline{f(t-\sigma_1)\, f(t-\tau_1)f(t-\tau_2)}\, d\sigma_1 = 0 \tag{3-76}$$

since the average of the product of an odd number of zero-mean Gaussian variables is zero. Finally the average for n = 2 is:

$$\overline{G_2[k_2; f(t)]\, f(t-\tau_1)\, f(t-\tau_2)} =$$

$$\overline{\left[\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} k_2(\tau_1,\tau_2) f(t-\tau_1) f(t-\tau_2)\, d\tau_1 d\tau_2 - A\int_{-\infty}^{+\infty} k_2(\tau_1,\tau_1)\, d\tau_1\right] f(t-\sigma_1) f(t-\sigma_2)}$$

$$= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} k_2(\sigma_1,\sigma_2)\, \overline{f(t-\sigma_1) f(t-\sigma_2) f(t-\tau_1) f(t-\tau_2)}\, d\sigma_1 d\sigma_2 -$$

$$A^2 \delta(\tau_1-\tau_2) \int_{-\infty}^{+\infty} k_2(\sigma_2,\sigma_2)\, d\sigma_2$$

$$= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} A^2 [\delta(\sigma_1-\sigma_2)\delta(\tau_1-\tau_2) + \delta(\sigma_1-\tau_1)\delta(\sigma_2-\tau_2) + \delta(\sigma_1-\tau_2)\delta(\sigma_2-\tau_1)]$$

$$k_2(\sigma_1,\sigma_2)\, d\sigma_1 d\sigma_2 - A^2 \delta(\tau_1-\tau_2) \int_{-\infty}^{+\infty} k_2(\sigma_2,\sigma_2)\, d\sigma_2 =$$

$$A^2 [\delta(\tau_1-\tau_2) \int_{-\infty}^{+\infty} k_2(\sigma_1,\sigma_1)\, d\sigma_1 + k_2(\tau_1,\tau_2) + k_2(\tau_2,\tau_1) - \delta(\tau_1-\tau_2) \int_{-\infty}^{+\infty} k_2(\sigma_2,\sigma_2)\, d\sigma_2]$$

$$= 2A^2 k_2(\tau_1,\tau_2) \qquad (3\text{-}77)$$

Combining equations (3-73)-(3-77), equation (3-71) is obtained.

From the process of the above derivation, it can be seen that care must be taken in applying the Wiener theory to practical problems because it is strictly valid only when the averages of infinite time series are considered and Gaussian white noise input is assumed. What can be obtained in practical calculation is an estimate of the true Wiener kernel and the accuracy of the estimation depends on the length of the averaging time, the

characteristics of the system to be investigated and the closeness of the input signal to the white Gaussian noise process. These points will be discussed further later on.

If the n^{th}-order Wiener kernel $k_n(\tau_1,\tau_2,...,\tau_n)$ has been measured, then its corresponding n^{th}-order frequency response function $H_n(\omega_1,\omega_2,...,\omega_n)$ is defined as

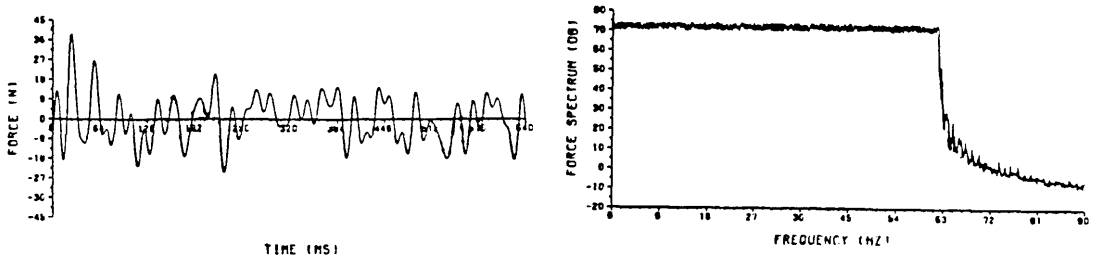
$$H_n(\omega_1,\omega_2,...,\omega_n) \equiv K_n(\omega_1,\omega_2,...,\omega_n) \tag{3-78}$$

where $K_n(\omega_1,\omega_2,...,\omega_n)$ is the n^{th}-order Wiener kernel transform of $k_n(\tau_1,\tau_2,...,\tau_n)$. As discussed, when the input is low (the power spectrum of input $A\rightarrow 0$), the measured Wiener kernels approach their corresponding Volterra kernels and, therefore, the measured $H_n(\omega_1,\omega_2,...,\omega_n)$ based on (3-78) approaches the Volterra kernel transform $H_n(\omega_1,\omega_2,...,\omega_n)$.

3.5 MEASUREMENT OF WIENER KERNEL TRANSFORMS

3.5.1 MEASUREMENT OF WIENER KERNELS USING CORRELATION ANALYSIS WITH RANDOM INPUT

So far, the theoretical basis for Wiener kernel measurement has been introduced and the possibility of measuring these kernels from practical nonlinear structures now needs to be assessed. Such an assessment can be carried out by simulating the measurement of Wiener kernels of realistic nonlinear mechanical systems. The input random signal is a band-limited Gaussian noise (the effect of non-white Gaussian noise input on the estimation of Wiener kernels is discussed in [47]) since according to the sampling theorem [51], the maximum valid frequency ω_n (the Nyquist frequency) is limited by the sampling rate $1/\Delta t$ and a true white Gaussian noise signal is therefore impossible to achieve. The process of generating band-limited Gaussian noise is done by passing the sampled standard white Gaussian noise data with sampling frequency at $1/\Delta t$ through a band-limited filter to remove the higher frequency components. The numerical realisation of this process is to interpolate the standard Gaussian noise data by a smoothing function $h(t)=\sin t/t$ which, in the frequency domain, is an ideal low-pass filter [51]. The time history and its power spectrum of one of the input band-limited noise signals with sampling frequency at 62.5Hz are shown in Fig.3.15.



Time History Power Spectrum

Fig.3.15 - Time History and Power Spectrum of a Band-limited Input Signal

Again, as in the case of second-order Volterra kernel transform measurement, the second-order Wiener kernel transforms of the square-law system, the SDOF nonlinear system, the 3DOF nonlinear system and a bilinear system are calculated and are shown in Figs.3.16-3.19. Also, as a typical example, the second Wiener kernel (time domain) of the SDOF system is illustrated in Fig.3.20. When comparing the measured frequency response functions (Wiener kernel transforms) with their corresponding analytical Volterra kernel transforms (figs.3.5, 3.9 & 3.11), it can be seen that the results are quite good. For the bilinear system, the measured second-order frequency response function looks very much like the second-order frequency response functions of the square-law system. Therefore, for a bilinear system, the quadratic component is quite substantial in the response x(t). However, it should be pointed out that although the linear contribution can theoretically be averaged out by including sufficient data points (increasing the averaging time), this is often difficult to do in practice since the computation resources required are considerable. During calculation of the second-order Wiener kernels of the SDOF and the 3DOF nonlinear systems, the linear contributions (the response component due to the linear part of the nonlinear network) are removed first before the correlation process takes place because the convergence seems to be very slow in these cases. In order to calculate second-order Wiener kernels of a nonlinear system efficiently, removal of the linear contribution becomes necessary. A possible way of removing linear contribution and therefore increasing the computational efficiency is proposed and discussed next.

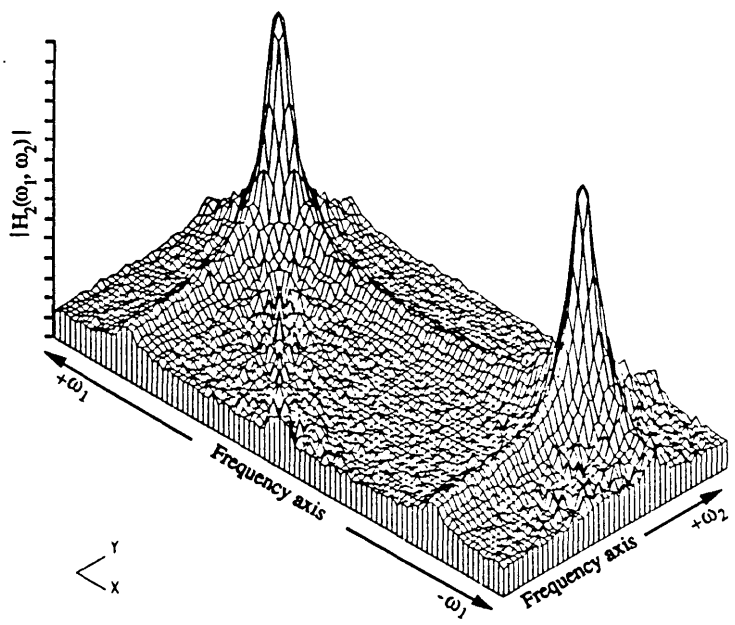


Fig.3.16 - Measured Second Order FRF of the Square-law System Using Correlation Analysis (Modulus Linear Scale, x-axis ω_1: -275 — 275, y-axis ω_2: -275 — 275 rad./s)

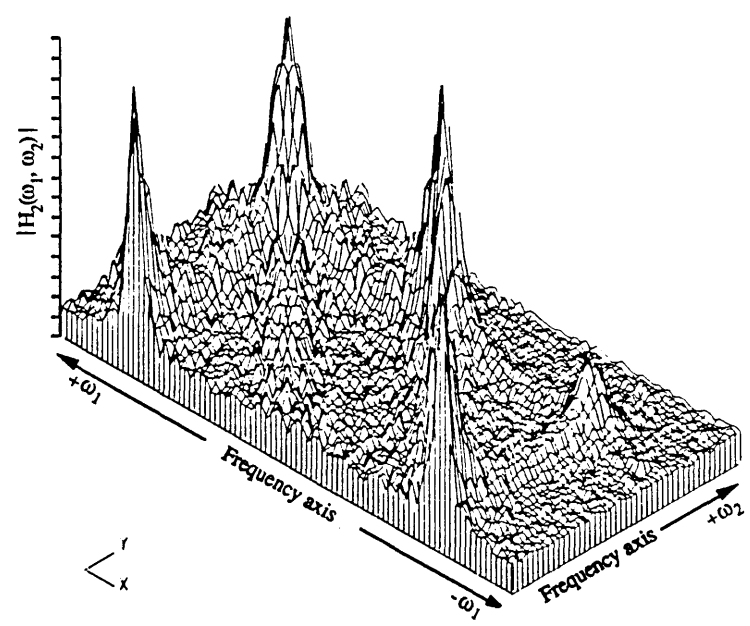


Fig.3.17 - Measured Second Order FRF of the SDOF Nonlinear System Using Correlation Analysis (Modulus Linear Scale, x-axis ω_1: -275 — 275, y-axis ω_2: -275 — 275 rad./s)

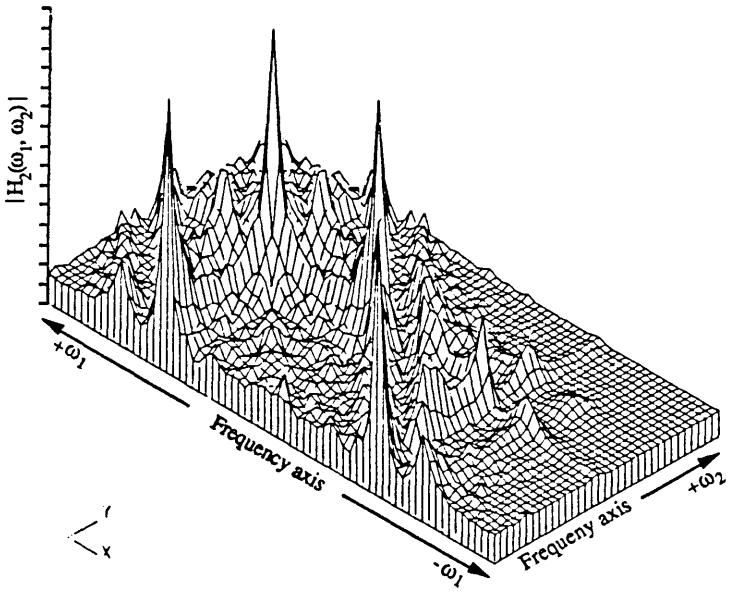


Fig.3.18 - Measured Second Order FRF of the 3DOF Nonlinear System Using Correlation Analysis (Modulus Linear Scale, x-axis ω_1: -275 — 275, y-axis ω_2: -275 — 275 rad./s)

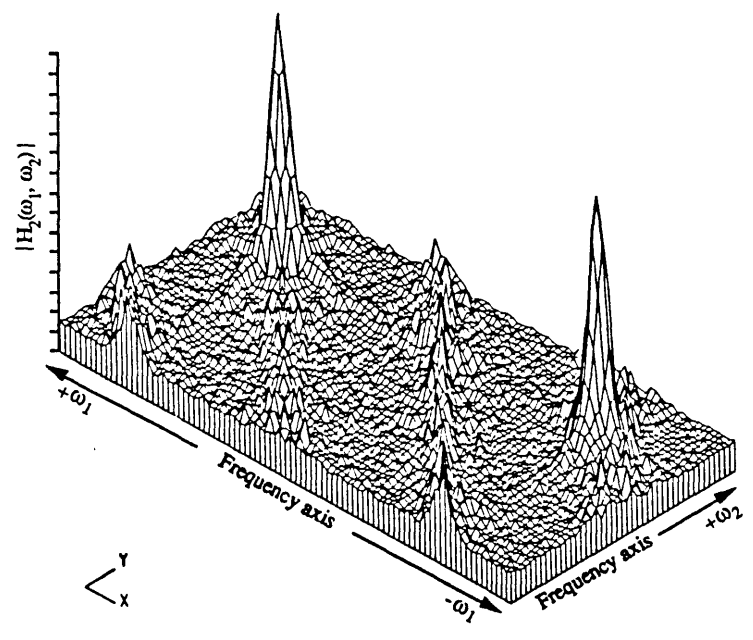


Fig.3.19 - Measured Second Order FRF of a Bilinear System Using Correlation Analysis (Modulus Linear Scale, x-axis ω_1: -275 — 275, y-axis ω_2: -275 — 275 rad./s)

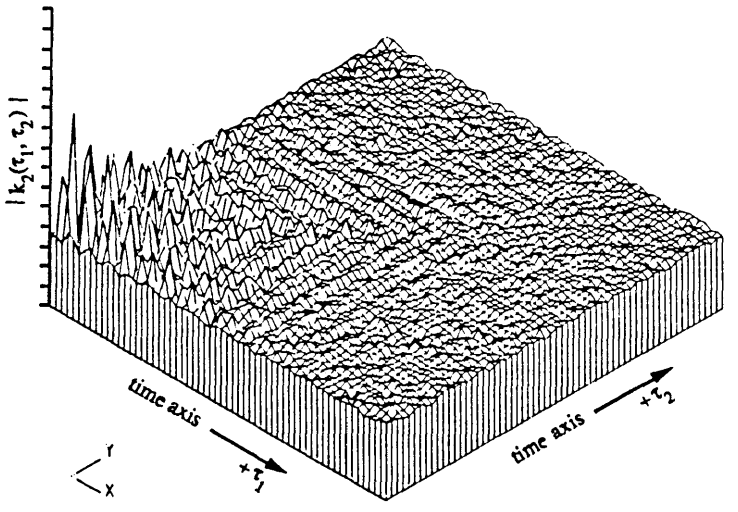


Fig.3.20 - Measured Second Order Wiener kernel of the SDOF Nonlinear System (Modulus Linear Scale, x-axis τ_1: 0 — 1, y-axis τ_2: 0 — 1s)

3.5.2 REMOVAL OF LINEAR CONTRIBUTION

It has been demonstrated in the numerical case studies that in order to calculate higher-order (second-order) Wiener kernel transforms of a nonlinear system efficiently, removal of the linear contribution in the response signal x(t) becomes necessary. This is due to the fact that for nonlinear systems like the SDOF and 3DOF systems in the numerical case studies the nonlinearities are such that the response component due to the nonlinear terms (mainly quadratic term in the estimation of second-order Wiener kernels) is usually of second order when compared with that of the linear contribution in the response x(t). As a result, although the term expressed in equation (3-76) (due to the linear contribution) for the estimation of second-order Wiener kernel should mathematically go to zero as averaging time increases, the time required for this to become valid could be very long indeed and so to improve the calculation efficiency, it is necessary to remove the linear contribution first, before the correlation process takes place.

It is suggested here that this removal of the linear contribution from the system response x(t) can be achieved by performing the averaging process in the frequency domain rather than in the time domain. The whole procedure is discussed next. Suppose the response component due to the quadratic and higher even terms of nonlinearity $x_2(t)$ be expressed

as $x_2(t)=x(t)-x_1(t)-x_0$ (where x_0 is the DC component which is supposed to be removed and x_1 is the linear contribution), then, by replacing $x(t)-k_0$ with $x_2(t)=x(t)-x_1(t)$, equation (3-71) can be rewritten as

$$k_2(\tau_1, \tau_2) = \frac{1}{2A^2} \overline{[x(t) - x_1(t)] f(t - \tau_1) f(t - \tau_2)} \tag{3-79}$$

Instead of taking the time domain average, let the averaging be done in the frequency domain, Fourier transform both sides of (3-79), then

$$K_2(\omega_1, \omega_2) = \frac{1}{A^2} [X(\omega_1+ \omega_2) - X_1(\omega_1+ \omega_2)] F^*(\omega_1) F^*(\omega_2)$$

$$= \frac{1}{2A^2} [X(\omega_1+ \omega_2) - K_1(\omega_1+ \omega_2) X_1(\omega_1+ \omega_2)] F^*(\omega_1) F^*(\omega_2) \tag{3-80}$$

With $K_1(\omega)$, which is the calculated first-order frequency response function based on (3-70), to be available beforehand, $K_2(\omega_1, \omega_2)$ can be derived based on equation (3-80) and in this way, the computational efficiency can be improved.

3.6 IDENTIFICATION OF NONLINEARITY USING HIGHER-ORDER FREQUENCY RESPONSE FUNCTIONS

So far, the theoretical basis of and measurement techniques for higher-order frequency response functions have been discussed in some detail and the remaining question which needs to be answered is: "what information about the nature of nonlinearity of a system can be derived from measured higher-order frequency response functions?" First, the existence of second-order frequency response functions indicates the existence of nonsymmetric nonlinearity of a system - a task that, for some systems such as qudratic and bilinear systems as mentioned before, cannot be achieved based on the analysis of classical first-order frequency response functions. Secondly, as is discussed in some detail next, parameters of a nonlinear system can also be identified based on the analysis of higher-order frequency response functions together with the first-order ones .

In the following discussion, only the analysis of second-order frequency response functions is presented. Depending on whether the physical parameters or the modal parameters of the system are of interest, a 'state-space analysis' method [44] or a 'modal-space analysis' method [45] can be developed.

In the 'state-space analysis', a-priori information about the total number of degrees of freedom and the physical connectivity of the system to be analysed should be given. Since, mathematically, measured first- and second-order FRF data are functions of all the physical parameters (mass, nonlinear stiffness and damping elements), given known measured first- and second-order FRFs, these parameters can, in theory, be calculated provided enough data have been measured. The physical parameter identification problem of a nonlinear system can therefore be formulated mathematically as the solution of the following linear algebraic equation:

$$[A]\{p\} = \{b\} \tag{3-81}$$

where $\{p\}$ is the unknown vector of physical parameters, and $[A]$ and $\{b\}$ are the coefficient matrix and vector formed using the measured first- and second-order frequency response functions. To illustrate this process, take the nonlinear SDOF system described by (3-3) as an example. The system mass m, linear stiffness k and damping c can be calculated based on the familiar analysis of the measured first-order FRF $H_1(\omega)$, while the coefficient of the second-order nonlinear term k_1 can be obtained from equation (3-38) using the measured second-order FRF $H_2(\omega_1,\omega_2)$. In fact, one data point on the ω_1 *vs* ω_2 plane is enough to determine k_1, although more data points are recommended in practice in order to have a reliable averaged estimation.

The 'modal-space analysis' is based on the mathematical observation that, in general, the second-order frequency response function (second-order Volterra kernel transform) can be decomposed as [45]

$$H_2(\omega_1, \omega_2) = \sum_{n=1}^{2N} \frac{A_n}{i\omega_1 + \omega_n} \sum_{m=1}^{2N} \frac{A_m}{i\omega_2 + \omega_m} \sum_{l=1}^{2K} \frac{C_l}{i\omega_1 + i\omega_2 + \omega_l} \tag{3-82}$$

where N is the number of degrees of freedom of the system and 2K represents the number of poles corresponding to "nonlinear coupling modes" which are the combinational resonances of the system [45]. When one of the variables (ω_1 or ω_2) is fixed, then equation (3-82) reduces to the following polynormial form:

$$H_2(\omega_1, \omega_2) = \frac{P(\omega_1, \omega_n, \omega_m, \omega_l)}{Q(\omega_1, \omega_n, \omega_m, \omega_l)}; \; n = 1, 2N; \; m = 1; 2N; \; l = 1, 2K \tag{3-83}$$

Curve-fitting of this polynomial function can be made using the well-developed polynomial curve-fitting algorithms [52] used in linear modal analysis, then all the ω_n and ω_m which are the natural frequencies of the system can be obtained and the analytical model of $H_2(\omega_1,\omega_2)$ in its polynormial expression can be established. Such analytical models can be used for further applications such as response prediction, as shown in Fig.3.21 for the system described by (3-3) with $k_2=0$ and with input being a pure sinusoid of frequency half of the natural frequency of the system. The improvement of prediction accuracy by including second-order FRF is clearly demonstrated.

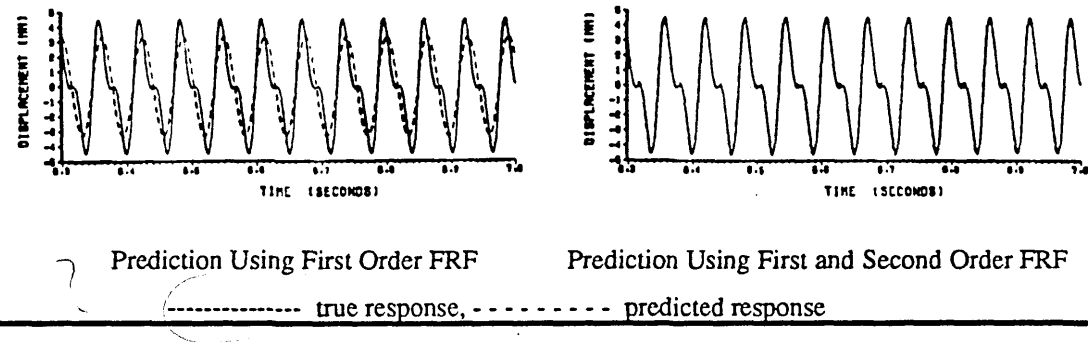


Fig.3.21 - Response Prediction Using Second Order Frequency Response Function

3.7 CONCLUSIONS

In this Chapter, the basic theory of Volterra and Wiener series of nonlinear systems has been introduced and the measurement of higher-order frequency response functions has been discussed. The relationships between the Volterra and Wiener kernels and their corresponding measured frequency response functions have been demonstrated. By extending the classical first-order frequency response function analysis to higher-order frequency response function analysis, it can be seen that the linear system theory is extended in a logical way to cover nonlinear systems.

From the system identification point of view, the measured higher-order frequency response functions provide considerable information about the nature of the nonlinearity of the system which the classical first-order frequency response functions cannot provide. Among them are the following:

(a) since for certain nonsymmetric nonlinear systems, such as the quadratic and bilinear systems mentioned in this Chapter, the measured first-order frequency response functions are effectively linear and therefore, cannot be used to detect existence of nonlinearity, the

measured higher- (second-) order frequency response functions give an indication of the nonlinearity in the system;

(b) the different characteristics of higher-order frequency functions may give categorization and so identification of common mechanical nonlinear systems by comparing the measured higher-order frequency response functions with those analytical ones of known nonlinear systems;

(c) from a system response prediction point of view, the higher-order frequency response functions together with the first-order frequency response functions give more accurate response prediction to any input than just using the first-order frequency response functions and;

(d) the measured higher-order frequency response functions can be analysed in a similar way to the case of first-order frequency response functions in order to identify either the physical parameters or the modal parameters of a nonlinear system so that its mathematical model can be established.

The existing numerical difficulties concerning the successful measurement of higher-order frequency response functions have been discussed and possible ways of overcoming these difficulties suggested, both in the case of measurement using harmonic probing technique and correlation analysis with random input. In the harmonic probing method, the main problem involved is leakage in the DFT of the response signal. This leakage problem can be overcome using the correlation technique as suggested. For the correlation analysis with random input, the main problem involved is the removal of the linear contribution from the total response so that the computational efficiency can be improved. For this purpose, averaging in the frequency domain instead of the time domain as discussed is recommended.

CHAPTER 4

IDENTIFICATION OF CHAOTIC VIBRATIONAL SYSTEMS

4.1 PRELIMINARIES

So far, techniques for identifying structural nonlinearity based on the measured first- and higher- order frequency response functions have been developed and discussed in some detail in Chapter 2&3. If the first-order and higher-order (usually second-order) frequency response functions of a nonlinear structure are measured in the ways as have been discussed in previous Chapters then, in most cases, the structural nonlinearity can be detected, quantified and identified by analysing these measured FRFs. However, for some nonlinear systems (chaotic systems, which are discussed in this Chapter), the dynamic behaviour is so complex that the first- and higher-order FRF analyses, which are largely based on the assumption of periodic input periodic output, become inadequate. In order to analyse these systems, the development of yet more analysis techniques becomes necessary.

In the case of a single sinusoidal input, the possible behaviour of the response spectrum of a nonlinear dynamic system can be illustrated in Fig.4.1. The first-order frequency response function analysis only considers the fundamental frequency components, while the higher-order frequency response function analysis takes the harmonic (both

subharmonic and superharmonic) and combinational frequency components into account. However, except for these fundamental and harmonic frequency components, the subharmonics of some nonlinear systems bifurcate in such a way that the response spectrum due to a single sinusoid input changes from discrete (periodic) to continuous (nonperiodic) distribution. Such newly-discovered strange behaviour of nonlinear system - a deterministic system exhibiting apparently random behaviour - is called *chaos* and is one of the most exciting research topics in nonlinear systems research.

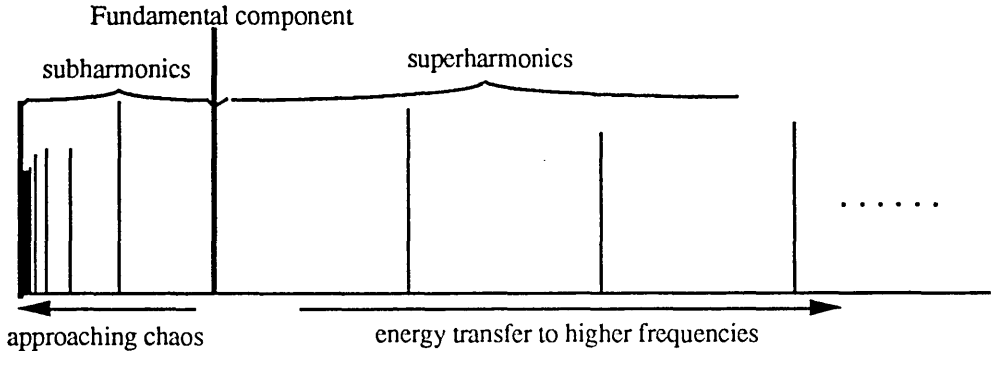


Fig.4.1 - Response Frequency Components of a Nonlinear System

In the last fifteen years, clues to the emergence of randomlike motion in deterministic dynamic systems have been uncovered by new topological methods in mathematics. At the same time, experimental measurements and numerical simulations have provided supporting evidence to the mathematical analysis which shows that many physical systems may exhibit chaotic behaviour without random inputs. Research on chaos has become an interdisciplinary subject and applications of the study have been found in almost all engineering subjects. In mechanical engineering, it has been well known that Duffing's system [53] with negative stiffness, such as that which represents mechanical structure of pre-stressed buckled beams, and, some impact mechanical oscillators [54], exhibit chaotic behaviour under certain excitation and initial conditions. These systems represent very special types of mechanical nonlinear structures which are not commonly encountered in practice. In order to investigate the possible chaotic behaviour of practical mechanical structures with more commonly-encountered nonlinearities, a mechanical system with backlash stiffness nonlinearity is considered in this Chapter. Extensive numerical as well as experimental research work has been carried out and apparently, it is the first time in literature that the chaotic nature of mechanical system with backlash stiffness nonlinearity under realistic system parameters has been revealed. Such a nonlinear mechanism as backlash stiffness represents an important and extensive group of mechanical structures with manufacturing clearances such as gearing systems. Based on

the mechanical backlash system, the basic theory of chaotic vibration is introduced and qualitative as well as quantitative ways of identifying chaotic behaviour of a nonlinear system are presented. Possible engineering applications of the study presented in this Chapter are suggested.

4.2 INTRODUCTION TO CHAOTIC VIBRATION THEORY

Chaos is an nonlinear phenomenon which permeates all fields of science. Although identified as an important research area only recently, chaos has existed from time immemorial. It is now known that chaos can readily occur not only in man-made systems, but in all natural and living systems where nonlinearity is present.

Roughly speaking, chaos is an exotic *steady-state response*. The steady-state response of a system is what remains after the transient has decayed to zero and from what has been taught in linear system theory, it can either be an equilibrium point, periodic or quasi-periodic solution. This basic principle of linear systems has been so deeply-rooted in the mind of most engineers that they may subconsciously extrapolate it for nonlinear systems as well. We know that for some nonlinear systems, there exist a wide range of parameters for which the steady-state responses are bounded, but are *not* periodic. In fact, the response waveform becomes erratic with a broad continuous frequency spectrum (rather than discrete, as in the periodic case). Moreover, the response is so sensitive to initial conditions that unless a computer with infinite word length is used in the simulation, no long-term prediction of the precise waveform is possible.

For mechanical systems, the study of chaotic vibration is important for several reasons. First, in the design of mechanical control systems, it is essential to avoid the occurrence of chaotic oscillation at design stage because chaos means unpredictability and so uncontrollability. Secondly, the random nature of the response to a deterministic (usually periodic) excitation of a mechanical structure makes life prediction difficult and statistical stress/fatigue analysis becomes necessary. Finally, from a machine monitoring point of view, that a broad continuous response spectrum can be due a single sinusoidal input makes the reliable diagnosis in most cases difficult and suggests that the development of new techniques is required.

In the following section, the basic theory of chaotic vibration is introduced based the well-known Duffing's and van der Pol's systems. Ingredients which are essential for understanding chaos are presented.

4.2.1 DYNAMICAL SYSTEMS THEORY

In general, dynamic systems can be divided into three different categories: *autonomous dynamic systems*, *nonautonomous dynamic systems* and *discrete-time dynamic systems*. Both autonomous and nonautonomous dynamic systems are described by differential equations (ordinary or partial differential equations) while discrete-time dynamic systems are expressed in terms of iterative maps. All three types of systems are defined and discussed and some of the useful facts from the theory of differential equations are presented.

An n^{th}-order autonomous dynamical system is defined by the state equation as

$$\dot{\mathbf{x}} = f(\mathbf{x}) \qquad \mathbf{x}(t_0) = \mathbf{x}_0 \tag{4-1}$$

where $\dot{\mathbf{x}} \equiv \frac{d}{dt}\mathbf{x}$, $\mathbf{x}(t) \in \mathbb{R}^n$ is the state at time instant t and f: $\mathbb{R}^n \rightarrow \mathbb{R}^n$ is called the vector field. Since the vector field does not depend on time, the initial time may always be taken as $t_0=0$. The solution of equation (4-1) with initial condition $\mathbf{x}=\mathbf{x}_0$ at time t=0 is called a trajectory (in n-dimensional space) and is denoted by $\phi_t(\mathbf{x}_0)$. The mapping (which is continuous as compared with the discrete mapping of discrete-time dynamical systems) ϕ_t: $\mathbb{R}^n \rightarrow \mathbb{R}^n$ is called the flow of the system since ϕ_t is a continuous trajectory starts at $\mathbf{x}_0$ and is like fluid flowing in the state-space. The dynamic system described in (4-1) is linear if $f(\mathbf{x})$ is a linear function of state variable $\mathbf{x}$. Free vibrational mechanical systems belong to this category.

An n^{th}-order nonautonomous dynamical system, on the other hand, is defined by the time-varying state equation as

$$\dot{\mathbf{x}} = f(\mathbf{x},t) \qquad \mathbf{x}(t_0) = \mathbf{x}_0 \tag{4-2}$$

The vector field depends on time and, unlike the autonomous case, the initial time cannot be arbitrarily set to 0. The solution of (4-2) passing through the point $\mathbf{x}_0$ at $t=t_0$ is denoted as $\phi_t(\mathbf{x}_0,t_0)$. Again, the system is linear if $f(\mathbf{x})$ is linear with respect to $\mathbf{x}$.

If for a nonautonomous system, there exists a constant $T > 0$ such that $f(\mathbf{x},t) = f(\mathbf{x},t+T)$ for all $\mathbf{x}$ and all t, then the system is said to be *time periodic* with period T. The smallest such T is called the *minimal period*. In this Chapter, all nonautonomous systems are assumed to be time periodic e.g., time invariant systems with periodic input force.

An n^{th}-order time periodic nonautonomous system can always be converted to an $(n+1)^{th}$-order autonomous system by appending an extra state $\theta=2\pi t/T$. Therefore, the corresponding autonomous system is given by

$$\dot{\mathbf{x}} = f(\mathbf{x},\theta T/2\pi) \qquad \mathbf{x}(0) = \mathbf{x}_0 \tag{4-3}$$

$$\dot{\theta} = 2\pi/T \qquad \theta(0) = 2\pi t_0/T \tag{4-4}$$

Since f is time periodic with period T, the new system described by (4-3) and (4-4) is periodic in θ with period 2π. Therefore, the state-space is transformed from Euclidean space $\mathbb{R}^{n+1}$ to cylindrical space $\mathbb{R}^n \times S$ where $S \equiv [0, 2\pi)$ is a circle. The solution in the new state-space is

$$\begin{Bmatrix} x(t) \\ \theta(t) \end{Bmatrix} = \begin{Bmatrix} \phi_t(x_0,t_0) \\ \frac{2\pi t}{T} \bmod 2\pi \end{Bmatrix} \tag{4-5}$$

where the modulo function (x mod y gives the remainder of x divided by y, e.g., 3 mod 2 =1) restricts θ to be within the semi-closed interval $[0,2\pi)$. Using this transformation, results for autonomous systems can be applied to the time periodic nonautonomous case.

As for discrete-time dynamic systems, any map $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ defines a discrete-time dynamic system by the state equation

$$\mathbf{x}_{k+1} = f(\mathbf{x}_k), \quad k = 0, 1, 2, \cdots\cdots \tag{4-6}$$

where $\mathbf{x}_k$ is called the state, and f maps state $\mathbf{x}_k$ to the next state $\mathbf{x}_{k+1}$. Starting with an initial condition $\mathbf{x}_0$, repeated applications of the map f gives rise to a sequence of points $\{\mathbf{x}_k\}_{k=0}^{\infty}$ called an orbit of the discrete-time system. Examples of discrete-time dynamic systems are given below.

Although the research presented in this Chapter focuses on continuous time vibrational systems, discrete-time systems will be discussed for two reasons. First, the Poincaré mapping technique, which replaces the analysis of flow of continuous-time system with the analysis of a discrete-time system, is an extremely useful tool for studying dynamical systems. Second, due to this correspondence between flows (of continuous-time dynamic systems) and maps (of discrete-time dynamic systems), maps will be used to illustrate important concepts without getting into details of solving differential equations.

The simplest one-dimensional discrete-time system - the population growth model - which has been found to be chaotic is described by the logistic equation

$$x_{k+1} = \lambda x_k (1 - x_k) \tag{4-7}$$

For some values of λ, after certain iterations until the transient component dies, the x_k will settle to one specific value (period one solution). While for other values of λ, x_k oscillates between 2 values (period 2 solution), 4 values (period 4 solution) and so on. However, there are some parameter regions in which x_k never repeats its value as iteration continues, as shown in Fig.4.2 and such phenomenon is the earliest observation of what we call *chaos* today.

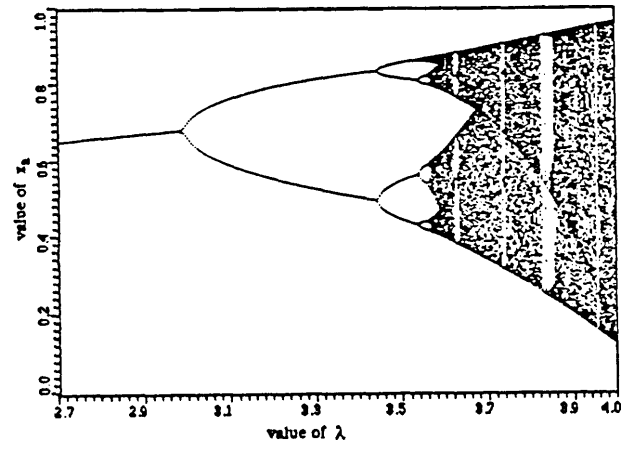


Fig.4.2 - Bifurcation Digram of Logistic Map for $2.7 \leq \lambda \leq 4.0$

Another discrete-time system which exhibits chaotic behaviour is the quadratic map studied by Hénon [55]

$$x_{k+1} = 1 - ax_k^2 + y_k \tag{4-8}$$

$$y_{k+1} = bx_k \tag{4-9}$$

In the case when a=1.4 and b=0.3, for initial condition (x_0,y_0), the sequence of points generated by the mapping $\{{x \atop y}\}_{k=0}^{\infty}$ is shown in Fig.4.3. Although the sequence of points never repeats, they settle to restricted areas on the x-y plane and exhibit a very well-constructed pattern (as will be discussed, the pattern is very finely defined as it is fractal).

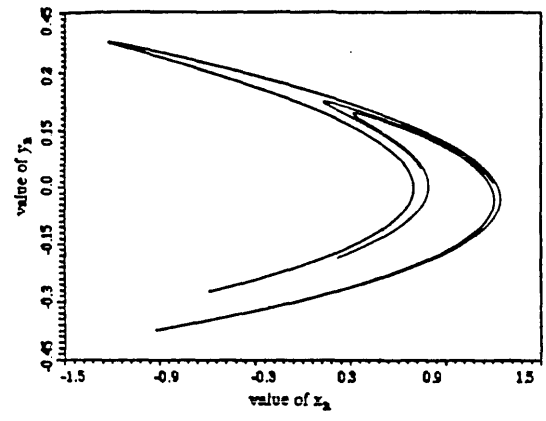


Fig. 4.3 - The Hénon Attractor

The characteristics of logistic map and Hénon map are briefly discussed here because they will be referred to in later discussions.

4.2.2 STEADY-STATE BEHAVIOUR AND LIMIT SETS OF DYNAMIC SYSTEMS

Dynamic systems are classified in terms of their steady-state solutions and limit sets. Steady state refers to the asymptotic behaviour of the solution of a dynamic system as time $t \to \infty$. The difference between the solution and its steady state is called the transient.

A point **y** is defined as the *limit point* of **x** if, for every neighbourhood U of **x**, flow $\phi_t(\mathbf{x})$ repeatedly enters U as $t \to \infty$, e.g., the equilibrium point of a dynamic system.

The set of all limit points is called the *limit set* $L(\mathbf{x})$ of **x**. Limit sets are closed and invariant under the flow ϕ_t (a set L is invariant under ϕ_t if, for all $\mathbf{x} \in L$ and all t, $\phi_t(\mathbf{x}) \in L$), e.g., the limit cycle of a nonlinear system.

A limit set L is attracting if there exists an open neighbourhood U of L such that $L(\mathbf{x})=L$ for all $\mathbf{x} \in U$. The limit cycle of a nonlinear system is an attracting limit set.

The *basin of attraction* B(L) of an attracting set L is defined as the union of all such neighbourhoods U. Every trajectory starting in B(L) tends towards L as $t \to \infty$. These definitions are briefly illustrated in Fig.4.4.

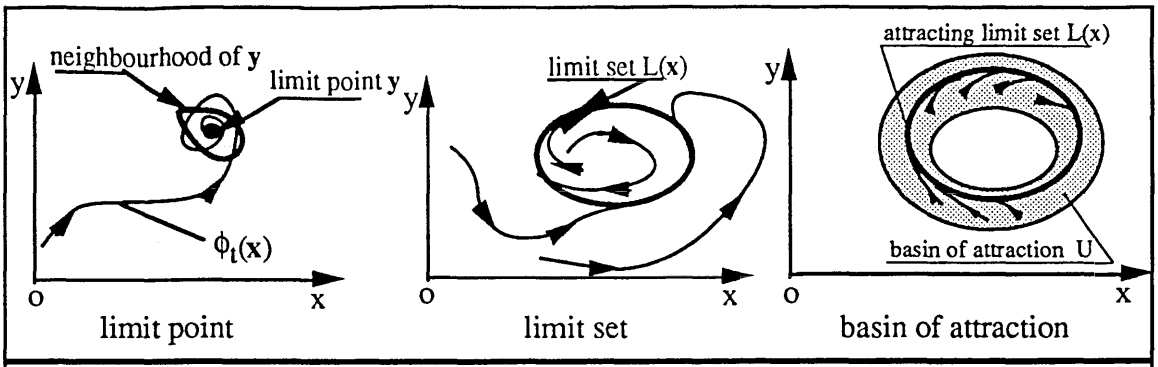


Fig .4.4 - Illustration of Limit Point, Limit Set and Basin of Attraction

In a stable linear system, there is only one limit set corresponding to specific input and therefore the steady-state behaviour is independent of initial conditions. In a typical nonlinear system however, there can be several attracting limit sets, each with a different basin of attraction. In this case, the initial condition determines in which limit set the system eventually settles.

The concept of limit sets is very useful in understanding different classical types of steady-state behaviour such as equilibrium points, limit cycles and quasi-periodic solutions. However, as will be shown, it is far too simple to describe the complex steady-state behaviour found in chaotic systems and some new mathematical concepts such as fractal dimension and Lyapunov exponent need to be introduced when steady-state chaotic behaviour (strange attractor) is considered. In what follows, different types of steady-state behaviour are discussed based on the well-known Duffing's and van der Pol's systems. Each state will be described from three different points of view: in the time domain, in the frequency domain and as a limit set in state-space domain.

An equilibrium point $\mathbf{x}^e$ is related with an autonomous system (a nonautonomous system does not have equilibrium points because the vector field $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ varies with time) and is the constant solution of equation (4-1), $\phi_t(\mathbf{x}^e)=\mathbf{x}^e$ for all time t. In general, $f(\mathbf{x})=\mathbf{0}$ implies that $\mathbf{x}$ is an equilibrium point of the system. A simple example is the damped free vibration system given by

$$m\ddot{x} + c\dot{x} + kx = 0 \qquad (4\text{-}10)$$

It is well known that the system possesses an equilibrium point which is $(x,\dot{x})=(0,0)$. This equilibrium point can be obtained by solving $f(x)=0$ as follows:
Rewrite (4-10) into its state-space form as

$$\dot{x} = y \tag{4-11}$$

$$\dot{y} = -\frac{c}{m}y - \frac{k}{m}x \tag{4-12}$$

Therefore, $f(x)=0$ means that y=0 and $-\frac{c}{m}y - \frac{k}{m}x = 0 \Rightarrow (x,\dot{x})=(0,0)$.

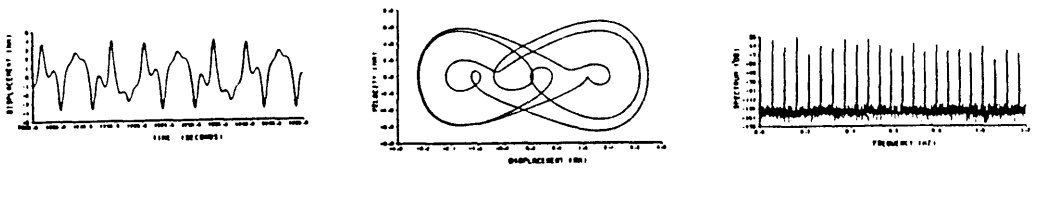
Both autonomous and nonautonomous systems can have periodic solutions under certain initial and excitation conditions. A solution $\phi_t(x^*,t_0)$ is a periodic solution if

$$\phi_t(x^*,t_0) = \phi_{t+T_m}(x^*,t_0) \tag{4-13}$$

for all time t and some minimal period $T_m > 0$. In general, a periodic solution of a dynamic system has a Fourier transform consisting of a fundamental frequency component at $f=1/T_m$ and evenly spaced harmonics at k/T_m, k=2,3,.... The amplitudes of some of these spectral components may be zero. For a nonautonomous system, T_m is typically some multiple of forcing period $T_m=kT$ and the periodic solution is usually referred to as a k^{th} subharmonic. To illustrate this point, a periodic (periodic 3) solution of the well-known Duffing's equation with $\gamma=0.1$, B=9.8 and $\omega=1$ (all units appear in this chapter are supposed to be normalised except where physical units are given).

$$\ddot{x} + \gamma\dot{x} + x^3 = B\cos\omega t \tag{4-14}$$

was calculated and is shown in Fig.4.5.



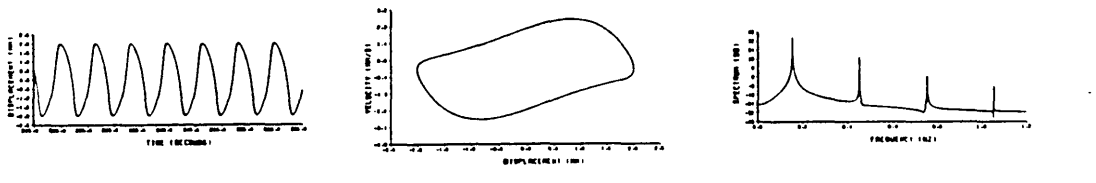
time domain state-space domain frequency domain

Fig. 4.5 - Period 3 Solution of Duffing's System

Also, periodic solutions exist in autonomous systems and in this case, the periodic solution is called a *limit cycle*. A limit cycle is a self-sustained oscillation and cannot occur in a linear system. One classical example of limit cycle is found in van der Pol's equation

$$\ddot{x} + (x^2 - 1)\,\dot{x} + x = 0 \qquad (4\text{-}15)$$

The existence of a stable van der Pol limit cycle is shown in Fig.4.6 and can be physically explained in terms of the damping mechanism of the system. When $|x| \leq 1$, the damping of the system is negative and therefore, the solution is expanding. While on the other hand, when the solution becomes $|x| \geq 1$, it is contracting. As a result, the solution will eventually settle down a limit cycle.



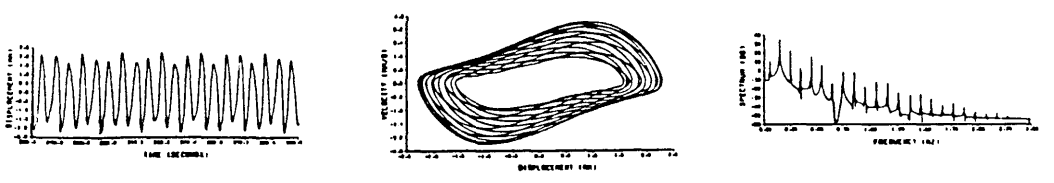
time domain state-space domain frequency domain

Fig. 4.6 - Limit Cycle of the van der Pol's System

Another type of steady-state solution which exists in some nonlinear systems is the *quasi-periodic* solution (often referred to as *almost periodic* solution) which is the sum of periodic functions with their fundamental frequencies (the reciprocals of their minimal periods) to be incommensurable. To see how quasi-periodic solutions arise in dynamic systems, again consider the van der Pol's equation with external forcing as

$$\ddot{x} + (x^2 - 1)\,\dot{x} + x = B\cos\omega t \qquad (4\text{-}16)$$

The system has a limit cycle oscillation with fundamental frequency ω_L. If the forcing frequency ω is incommensurable with ω_L, then a quasi-periodic solution occurs. The quasi-periodic solution of equation (4-16) with B=1.0 and $\omega=\pi/2$ is shown in Fig.4.7.



time domain state-space domain frequency domain

Fig. 4.7 - Quasi-periodic Solution of van der Pol's Equation

Mathematically, a quasi-periodic trajectory which contains n different incommensurable fundamental frequencies lies on an n-dimensional torus. Taking the two-periodic trajectory (contains two incommensurable fundamental frequencies) as an example, the trajectory lies on a two dimensional torus SxS as shown in Fig.4.8 with each S representing one of the base frequencies. Since a trajectory is a curve while SxS is a surface, not every point on the torus lies on the trajectory. However, it can be shown that the trajectory repeatedly passes arbitrarily close to any point on the torus and, therefore, the torus is the limit set of the quasi-periodic behaviour.

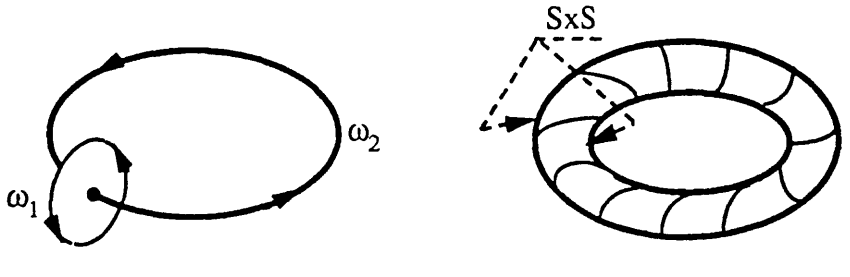


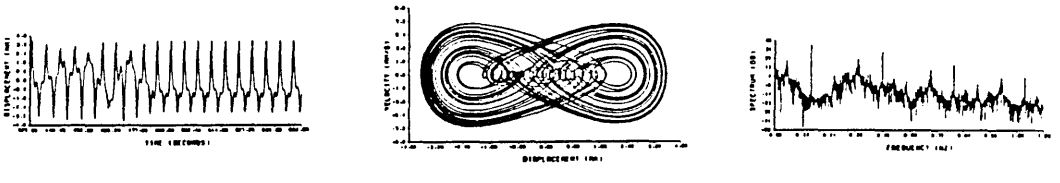
Fig. 4.8 - Two Periodic Behaviour Lies on Two Dimensional Torus SxS

4.2.3 CHAOTIC ATTRACTOR

There is no generally-accepted definition of a chaotic attractor. From a practical point of view, chaotic solution can be defined as none of the above mentioned steady-state solutions; that is, as bounded steady-state behaviour which is neither an equilibrium point, nor periodic, and not a quasi-periodic limit set either. For this reason, chaotic attractors are often referred to as "strange attractors". Since the solution is nonperiodic (which means that the solution contains some random components) while the system is deterministic (there are no random parameters involved in describing the system), chaotic

systems are very often described as "deterministic systems that exhibit random behaviour".

The chaotic behaviour of discrete-time dynamic systems such as the one-dimensional logistic equation and the two-dimensional Hénon map has been briefly discussed and the bifurcation diagram for the logistic equation and the strange attractor for the Hénon map are shown in figures 4.2&4.3. Now, a chaotic solution of Duffing's equation (4-14) with c=0.1, B=10 and ω=1 is calculated and is shown in Fig.4.9. It is evident from this that the trajectory is indeed bounded and nonperiodic. However, one should be careful to note that boundedness and nonperiodicity do not necessarily mean that the solution is chaotic because a quasi-periodic solution is bounded and nonperiodic as well. In order to distinguish chaotic solutions from quasi-periodic ones, the frequency spectrum of the signal needs to be calculated. For a quasi-periodic signal, the spectrum only contains discrete frequency components while a chaotic solution has a spectrum with a continuous, broad-band nature, as shown in figure 4.9. This noise-like spectrum is characteristic of chaotic systems.



time domain state-space domain frequency domain

Fig.4.9 - Chaotic Solution of Duffing's Equation

Unlike the classical types of attractor that are associated with classical geometric objects such as an equilibrium state with a point, the periodic motion or limit cycle with a closed curve and a quasi-periodic motion with a surface in multi-dimensional space, the limit set of chaotic behaviour is related to a new geometric object called a fractal set [56], which will be discussed later on.

Another property of chaotic systems is *sensitive dependence on initial conditions*: given two initial conditions arbitrarily close to one another, the trajectories emanating from these initial conditions diverge at an exponential rate (which is some kind of characteristic value of the system) until for all practical purposes, they are uncorrelated. This sensitive dependence on initial condition for Duffing's equation is illustrated in Fig.4.10 (the two initial conditions differ by only 1%). In practice, the initial state of the system can never

be specified exactly, but only within some tolerance ε and therefore, if two initial conditions x_0 and x_0^* lie within ε of one another, they cannot be distinguished. However, after a finite amount of time, flows $\phi_t(x_0)$ and $\phi_t(x_0^*)$ will diverge and become uncorrelated. As a result, no matter how precisely the initial condition is known, the long-term behaviour of a chaotic system can never be predicted (of course, the more accurate the initial conditions are, the longer the prediction can be, but since the divergence is exponential with time, unless the initial condition could be specified to infinite precision, accurate long-term prediction becomes impossible).

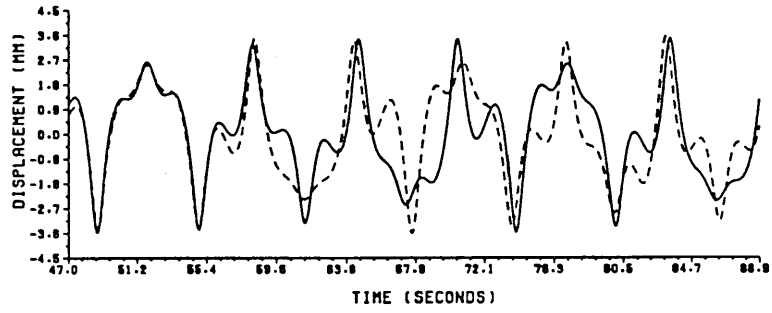


Fig.4.10 - Illustration of Sensitive Dependence on Initial Conditions

4.2.4 THE POINCARÉ MAPS

A very useful classical technique for analysing dynamical systems was developed by Poincaré. The technique replaces the flow of a continuous-time dynamical system with a discrete map called Poincaré map. For autonomous and nonautonomous systems, the definitions of the Poincaré map are slightly different and two cases are treated separately.

Consider an n^{th}-order autonomous system with a limit cycle Γ as shown in Fig.4.11. Let x^* be a point on the limit cycle Γ and let T be the minimal period of the limit cycle. Take an (n-1)-dimensional hyper-plane Σ (a plane has a dimension more than two) transverse to Γ at x^*. The trajectory emanating from x^* will hit Σ at x^* in T seconds. Due to the continuity of ϕ_t with respect to the initial conditions, trajectories starting on Σ in a sufficiently small neighbourhood of x^* will, in approximately T seconds, intersect Σ in the vicinity of x^*. Therefore, vector field f and hyper-plane Σ define a mapping P of some neighbourhood $U \in \Sigma$ of x^* onto another neighbourhood $V \in \Sigma$ of x^*. The thus-defined P is called the Poincaré map of an autonomous system.

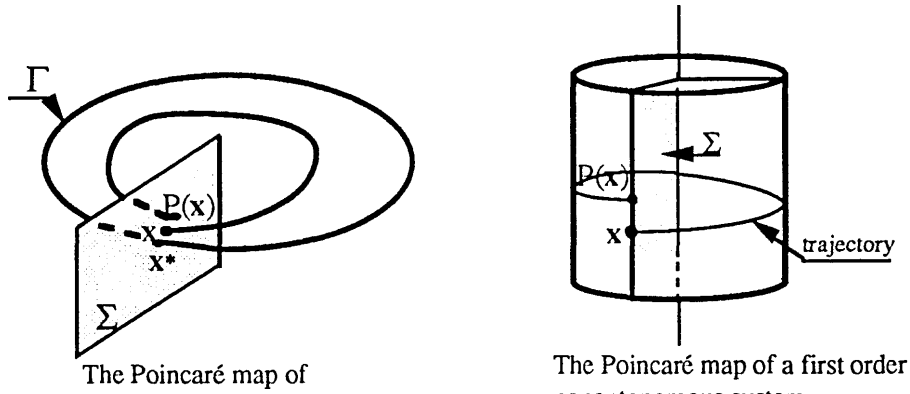
For nonautonomous systems, as shown in section 4.2.1, an n^{th}-order nonautonomous system in $\mathbb{R}^n$ Euclidean space with period T may be transformed into an (n+1)-dimensional autonomous system in cylindrical state-space $\mathbb{R}^n$xS. Consider the n-dimensional hyper-plane Σ in $\mathbb{R}^n$xS defined by

$$\Sigma \equiv \{(\mathbf{x}, \theta)\ \mathbb{R}^n\text{xS} \mid \theta = \theta_0\}. \tag{4-17}$$

Every T seconds, the trajectory intersects Σ as shown in Fig.4.11. Thus a map P: $\Sigma \rightarrow \Sigma$ is defined by $P(\mathbf{x})=\phi_T(\mathbf{x},t_0)$ where P is called the Poincaré map. Such a Poincaré map can be thought of in following two ways:

(i) P(**x**) indicates where the flow takes **x** after a T seconds and this is called T advance mapping; or

(ii) the orbit (sequence of points) $\{P^k(\mathbf{x})\}_{k=1}^{\infty}$ is a sampling of a single trajectory every T seconds; that is $P^k(\mathbf{x}_0) = \phi_{kT}(\mathbf{x}_0,t_0)$ for k = 1, 2, ...



The Poincaré map of autonomous system

The Poincaré map of a first order nonautonomous system

Fig.4.11 - Poincaré Maps of Autonomous and Nonautonomous Systems

The usefulness of the Poincaré map derives from the fact that there is one-to-one correspondence between the different types of steady-state behaviour of the underlying continuous-time dynamic system and the steady-state behaviour of mapping P. Therefore, from the steady-state behaviour of mapping P, the steady-state behaviour of the continuous dynamic system can be deduced.

As is clear from the definition of P, a period one solution of the underlying flow corresponds to a fixed point of the Poincaré map. For nonautonomous systems, a period K solution (contains K^{th} subharmonic) corresponds to K different points on the Poincaré map.

The Poincaré map can also be used to detect quasi-periodic solutions. As mentioned earlier, two-periodic solution (meaning two incommensurable minimal periods and is different from period two solution) lies on a two dimensional torus SxS as shown in figure 4.8. Using coordinate (θ_1,θ_2) on the torus, a two-periodic trajectory may be written as

$$\begin{Bmatrix}\theta_1(t)\\ \theta_2(t)\end{Bmatrix} = \begin{Bmatrix}\omega_1 t \bmod 2\pi\\ \omega_2 t \bmod 2\pi\end{Bmatrix} \tag{4-18}$$

where ω_1 and ω_2 are incommensurable. In the nonautonomous case, one of the frequencies, say, ω_1 is the forcing frequency of the system. An orbit of the Poincaré map corresponds to sampling (4-18) every $2\pi/\omega_1$ seconds

$$\begin{Bmatrix}\theta_1(2\pi k/\omega_1)\\ \theta_2(2\pi k/\omega_1)\end{Bmatrix} = \begin{Bmatrix}2\pi k \bmod 2\pi\\ 2\pi k\omega_2/\omega_1 \bmod 2\pi\end{Bmatrix} = \begin{Bmatrix}0\\ 2\pi k\omega_2/\omega_1 \bmod 2\pi\end{Bmatrix}, \quad k = 1, 2,... \tag{4-19}$$

Since ω_1 and ω_2 are incommensurable, $\{\theta_2(2\pi k/\omega_1)\}_{k=1}^{\infty}$ is not periodic and repeatedly comes arbitrarily close to every point in $[0, 2\pi)$. Therefore, in the (θ_1,θ_2) coordinates, the limit set of the Poincaré map is the circle S. In the original Euclidean coordinates, the limit set is a closed curve. To illustrate this point, the Poincaré map of quasi-periodic solution shown in figure 4.7 with sampling frequency equal to the the forcing frequency is shown in Fig.4.12.

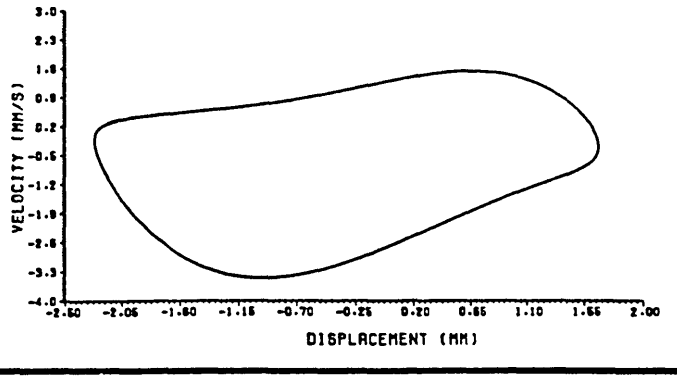


Fig. 4.12 - The Poincaré Map of a Quasi-periodic Solution of van der Pol's System

For chaotic trajectories, the steady-state Poincaré maps are distinctive and often quite beautiful. In order to illustrate this, the chaotic attractor (Poincaré map) of Duffing's equation with c=0.1, B=10 and ω=1 is calculated and is shown in Fig.4.13. Looking at these orbits, it becomes immediately clear that the steady-state orbits do not lie on a

simple geometrical form as is the case with periodic and quasi-periodic behaviour. The attractor has fine structure which is *fractal* as will be discussed later on. Such fine structure of Poincaré map is typical of chaotic systems.

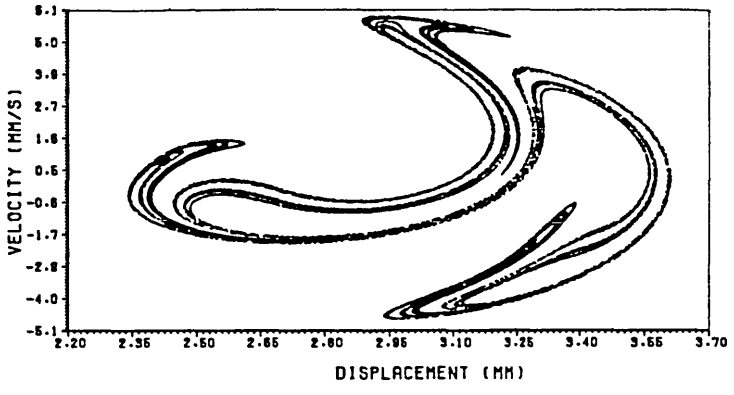


Fig.4.13 - Strange Attractor of Duffing's System

4.2.5 STABILITY OF LIMIT SETS AND LYAPUNOV EXPONENTS

The study of the stability of limit sets is important because only the attracting limit sets (structurally stable) can be physically observed. In this section, the conditions for a limit set to be stable will be discussed both in the case of equilibrium points and periodic solutions. The Lyapunov exponents which can be used to determine the stability of any type of steady-state behaviour, including quasi-periodic and chaotic solutions, will be introduced.

Consider an equilibrium point $\mathbf{x}^e$ of equation (4-1). It is well-known that the local behaviour (for small perturbations) of a nonlinear system near the equilibrium point is determined by linearising f at $\mathbf{x}^e$ as

$$\delta\dot{\mathbf{x}} = Df(\mathbf{x}^e)\,\delta\mathbf{x} \tag{4-20}$$

where $Df(\mathbf{x}^e)$ is the Jacobian matrix at point $\mathbf{x}^e$ and $\delta\mathbf{x} = (\mathbf{x} - \mathbf{x}^e)$. The thus derived linear vector field (4-20) governs the time evolution of perturbations near $\mathbf{x}^e$ of the original nonlinear system. In particular, the stability of the flow near $\mathbf{x}^e$ can be determined based on the linear vector field by examining the real parts of the eigenvalues of the Jacobian matrix $Df(\mathbf{x}^e)$.

Suppose [λ.] and [ϕ] are the eigenvalue and eigenvector matrices of the Jacobian matrix, then mathematically, the trajectory with initial condition $\mathbf{x}^e+\delta\mathbf{x}^e$ is, to the first order,

$$\phi_t(\mathbf{x}^e+\delta\mathbf{x}^e) = \phi_t(\mathbf{x}^e) + \delta\mathbf{x}(t) = \mathbf{x}^e + e^{Df(\mathbf{x}^e)t}\delta\mathbf{x}^e = \mathbf{x}^e + \sum_{r=1}^{n} c_r\{\phi\}_r e^{\lambda_r t} \quad (4\text{-}21)$$

where $\{c_r\}_{r=1}^n$ contains scalar constants chosen to achieve the correct initial conditions. From equation (4-21), it can be seen that the real part of λ_i gives the rate of expansion (if $Re(\lambda_i) > 0$) or contraction (if $Re(\lambda_i) < 0$) in the neighbourhood of the equilibrium point along the direction of $\{\phi\}_i$.

If $Re(\lambda_i) < 0$ for all λ_i, then all sufficiently small perturbations will die out as $t \to \infty$ and $\mathbf{x}^e$ is *asymptotically stable*. If $Re(\lambda_i) > 0$ for some λ_i, then $\mathbf{x}^e$ is not stable. If one of the eigenvalues has zero real part, then the stability cannot be determined from the linearised vector field and higher terms need to be included in the expression of (4-20).

The stability of a periodic solution is determined by its *characteristic multiplier*. Characteristic multipliers are a generalisation of the eigenvalues at an equilibrium point. Since a periodic solution corresponds to a fixed point of the Poincaré map P, the stability of the periodic solution is determined by the stability of the fixed point of the Poincaré map. By analogy with the equilibrium point, the stability of fixed point $\mathbf{x}^*$ of P (corresponding to initial condition $\mathbf{x}_0$) is determined by linearising P at $\mathbf{x}^*$. The linear discrete-time system

$$\delta\mathbf{x}_{k+1} = DP(\mathbf{x}^*)\,\delta\mathbf{x}_k \quad (4\text{-}22)$$

governs the local behaviour of map P near $\mathbf{x}^*$. Again as in the case of equilibrium point, the orbit of P for an initial condition $\mathbf{x}_0+\delta\mathbf{x}_0$ is, to the first order,

$$\mathbf{x}_k = \mathbf{x}^* + \delta\mathbf{x}_k = \mathbf{x}^* + DP(\mathbf{x}^*)\,\delta\mathbf{x}_k = \mathbf{x}^* + \sum_{r=1}^{n} c_r \mu_r^k \{\phi\}_r \quad (4\text{-}23)$$

where [μ.] and [ϕ] are the eigenvalue and eigenvector matrices of $DP(\mathbf{x}^*)$ and $\{c_r\}_{r=1}^n$ are scalar constants chosen to achieve the correct initial conditions. The eigenvalues [μ.] are the characteristic multipliers of the fixed point and determine the amount of contraction (if $|\mu_i| < 1$) or expansion (if $|\mu_i| > 1$) near $\mathbf{x}^*$ in the direction of $\{\phi\}_i$ for one iteration of the map P.

The characteristic multipliers determine the stability of the periodic solution. If all the μ_i lie within the unit circle (μ_i is in general, complex), then the periodic solution is asymptotically stable. If some of the μ_i lie outside the unit circle, then the periodic solution is not stable. If some of the characteristic multipliers lie on the unit circle, then the stability of the periodic solution cannot be determined by the multipliers alone.

Lyapunov exponents are a generalisation of the eigenvalues at an equilibrium point and of characteristic multipliers. They are used to determine the stability of any type of steady-state behaviour, including quasi-periodic and chaotic solutions. The definition of the Lyapunov exponent is as follows. Let [m_i] be the eigenvalues of a $\Phi_t(\mathbf{x}_0)$ (a square matrix which is a function of time), then the Lyapunov exponents are defined by

$$\lambda_i \equiv \lim_{t\to\infty} \frac{\ln |m_i(t)|}{t}, \quad i = 1, 2, ..., n \tag{4-24}$$

if the limit exists.

To explain the physical meaning, the Lyapunov exponents of an equilibrium point are calculated. Let [μ_i] be the eigenvalues of $D\boldsymbol{f}(\mathbf{x}^e)$, then for flow $\Phi_t(\mathbf{x}^e) = e^{D\boldsymbol{f}(\mathbf{x}^e)t}$ which is a linearised vector field, $m_i(t) = e^{\mu_i t}$ and

$$\lambda_i \equiv \lim_{t\to\infty} \frac{\ln |m_i(t)|}{t} = \lim_{t\to\infty} \frac{\ln |e^{\mu_i t}|}{t} = \mathrm{Re}(\mu_i) \tag{4-25}$$

Therefore, in this special case, the Lyapunov exponents are equal to the real parts of the eigenvalues of $D\boldsymbol{f}(\mathbf{x})$ at the equilibrium point and indicate the rate of contraction ($\lambda_i < 0$) or expansion ($\lambda_i > 0$) near the equilibrium point.

Lypunov exponents are convenient for categorising steady-state behaviour. For an attractor (including a chaotic attractor), contraction must outweigh expansion and therefore $\sum_{r=1}^{n} \lambda_i < 0$. Attractors are classified in terms of Lyapunov exponents as follows. For a stable equilibrium point, $\lambda_i < 0$ for all i. For a stable limit cycle, $\lambda_1 = 0$ and $\lambda_i < 0$ for i = 2, 3, ..., n. For a torus, $\lambda_1 = \lambda_2 = 0$ and $\lambda_i < 0$ for i = 3, 4, ..., n.

One feature of chaos, as mentioned earlier, is its sensitive dependence on initial conditions. Sensitive dependence occurs in an expanding flow, as is illustrated below.

Consider a nonautonomous system with a contracting flow ϕ_t as shown in Fig.4.14(a). Suppose that the state of the system can be measured to within an accuracy of ε, then it is clear that it is more accurate to predict the state at time t_2 using the measured state at time t_1 than to measure the state at t_2. The larger the elapse time (t_2-t_1), the greater the accuracy of the prediction. Thus for a contracting system, the predictive value of the initial condition increases with time. On the other hand, consider the opposite case of an expanding flow as shown in Fig.4.14(b). It is more accurate to measure the state at t_2 than to predict it using the measured state at t_1 and the predictive value of the initial condition deteriorates with time. This means that expanding systems exhibit sensitive dependence on initial conditions, but a purely expanding flow also implies unbounded behaviour. By definition, a chaotic trajectory is bounded, and therefore it follows that a chaotic system must contract in some directions and expand in others with the contraction outweighing the expansion (here we only consider the dissipative/damped dynamical systems). Hence, for a chaotic/strange attractor, at least one of the Lyapunov exponents must be positive and this existence of positive Lyapunov exponents distinguishes a strange attractor from other types of attractor and is one of the main criteria for detecting chaos.

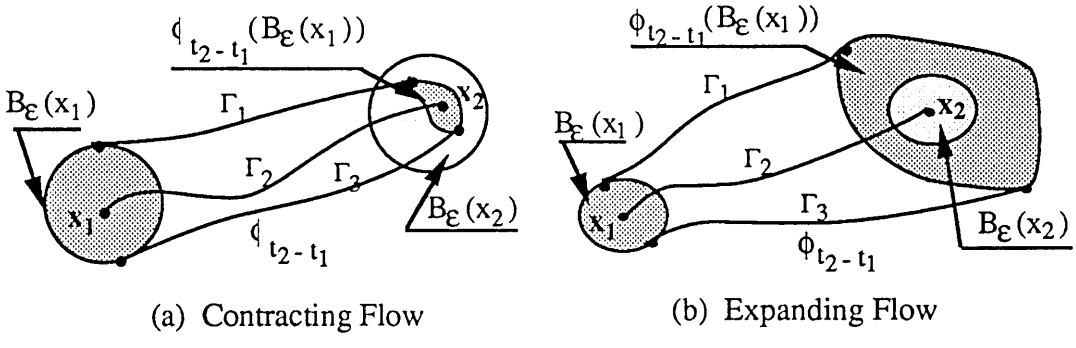


Fig.4.14 - Illustration of Contracting Flow and Expanding Flows

4.2.6 THE DIMENSION OF AN ATTRACTOR

As discussed above, Lyapunov exponents can be used to categorise different types of limit set and here in this section another important concept - the concept of dimension of an attractor which serves to quantify the complexity of a given attractor - is introduced. An attractor could be defined to be n-dimensional if, in the neighbourhood of every point on the attractor, it looks like an open subset of $\mathbb{R}^n$. This is how the dimension of a manifold is defined in differential topology. For example, a limit cycle is one-dimensional

because it looks locally like an interval. A two-dimensional torus has a dimension of 2 because for every point, locally, it resembles an open set of $\mathbb{R}^2$. An equilibrium point is considered to have zero dimension. However, as shown in figure 4.13, the neighbourhood of any point of a strange attractor has a very finely defined structure and does not resemble any Euclidean space. Therefore, strange attractors are not manifolds and do not have integer dimension. There are several ways to generalise the dimension to the general fractional case and in this section, only the capacity dimension is presented.

The simplest dimension is the capacity dimension. To illustrate how the capacity dimension can be calculated, let us consider a long time trajectory in phase space as shown in Fig.4.15. First, time sample the trajectory so that a large number of points on the trajectory are obtained. Then place a sphere (or cube) of radius (or length) ε at some point of the orbit and count the number of points within the sphere $N(\varepsilon)$. The probability of finding a point in this sphere is then defined as

$$P(\varepsilon) = \frac{N(\varepsilon)}{N_0} \tag{4-26}$$

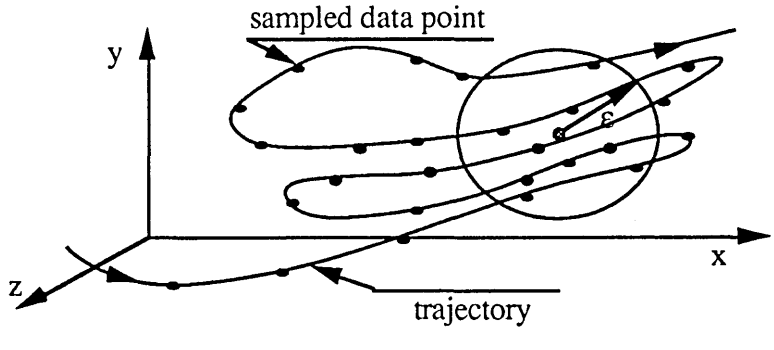


Fig.4.15 - Time Sampled Data points of a Trajectory

where N_0 is the total number of sampled time data points. For a one-dimensional orbit, such as a closed periodic orbit, $P(\varepsilon)$ will be linear in ε as $\varepsilon \to 0$ and $N_0 \to \infty$; $P(\varepsilon) \approx \beta\varepsilon$ (where β is a constant). If the orbit is quasi-periodic, two-periodic for example, then the probability $P(\varepsilon)$, as $\varepsilon \to 0$ and $N_0 \to \infty$ will be $P(\varepsilon) \approx \gamma\varepsilon^2$ (where γ is a constant). These observations lead one to define the capacity dimension of an orbit at point $\mathbf{x}_i$ by measuring the relative percentage of time that the orbit spends in the small sphere; that is,

$$D_c^i = \lim_{\varepsilon \to 0} \frac{\ln P(\varepsilon, \mathbf{x}_i)}{\ln \varepsilon} \tag{4-27}$$

In general, thus calculated D_c^i will be dependent on x_i and, therefore, averaging is required in order to calculate the capacity dimension of the orbit

$$D_c = \frac{1}{M} \sum_{i=1}^{M} D_c^i \tag{4-28}$$

where M is the number of points which have been averaged. In this way, the capacity dimension of the strange attractor shown in Fig.4.3 was calculated to be $D_c = 1.26$. Since D_c is not an integer, the attractor is indeed chaotic.

For nonautonomous time periodic systems, the capacity dimension of the Poincaré map of an attractor is often used to detect the existence of chaos and to quantify the complexity of the motion. If the calculated capacity dimension D_c of the Poincaré map is independent of the phase of the Poincaré map (the phase angle $0 \leq \theta < 2\pi$) and satisfies $0 < D_c < 2$, then the dimension of the complete attractor is just $d = 1 + D_c$.

4.3 CHAOTIC VIBRATION OF NONLINEAR MECHANICAL SYSTEM WITH BACKLASH

So far, the basic theories which are required in order to understand chaotic vibration of dynamic systems have been reviewed and summarised. In the following sections, the research carried out on the chaotic vibration of mechanical systems with a backlash stiffness nonlinearity is presented. Apparently, it is the first time in literature that the chaotic behaviour of such a general, yet so simple a nonlinear mechanical system has been revealed. Based on such mechanical backlash systems, qualitative as well as quantitative ways of analysing chaotic behaviour are presented. Possible practical applications of the research presented are discussed and suggested.

4.3.1 INTRODUCTION

In recent years, the study of chaotic dynamic behaviour in nonlinear deterministic systems has become a major research topic in nonlinear dynamic system analysis and new discoveries of chaos have been reported in several engineering applications such as nonlinear circuit design in electrical engineering [57], turbulence modelling in fluid dynamics [58] and chemical reaction process modelling in chemical engineering [59]. In mechanical engineering, systems modelled by Duffing's equation, such as pre-stressed buckled beams, have been studied extensively and it has been found that under certain excitation and initial conditions, chaotic vibrations can occur [53,61,62,63]. In particular,

the chaotic behaviour of mechanical impact oscillators (oscillators with rigid motion constraints), both single and double oscillators, have been studied by Shaw and Holmes [54,64]. However, it should be noted that these systems represent very special types of nonlinear mechanical system. In the case of Duffing's system, although chaotic vibration has been observed experimentally when the linear stiffness of the system is negative, such as for pre-stressed buckled beams, when the linear stiffness becomes positive - which is the more realistic case of some practical nonlinear structures with a stiffness nonlinearity - on the other hand, only when the vibration amplitude becomes excessively high that chaotic vibration occurs. Under practical service conditions therefore, chaotic vibration cannot in general occur for Duffing's system with positive linear stiffness. For impact oscillators, practical nonlinear structures rarely possess infinite stiffness and hence the impact oscillator model is, in general, not realistic of mechanical structures. Therefore, the possible existence of chaotic vibration in a general and practically realistic nonlinear mechanical structure has not been investigated to date. The research work presented below seeks to demonstrate that it is possible for chaotic vibration to occur in a general mechanical system with backlash stiffness nonlinearity which represents a group of mechanical systems with manufacturing clearances.

The classical analysis of the vibration behaviour of mechanical system with backlash stiffness nonlinearity is treated in standard texts on nonlinear oscillations, such as that of Minorsky [65], and an investigation of harmonic and superharmonic resonances of this type of systems was carried out by Maezawa [66]. The present analysis concentrates on the chaotic behaviour of the mechanical backlash system with realistic system parameters under sinusoidal excitation and presents both numerical and experimental results of the research. The fourth-order Runge-Kutta method with precision control was used in the numerical simulations. It was found that both periodic and chaotic vibrations exist under different forcing conditions.

4.3.2 THE GENERAL SYSTEM

The system studied is the simple nonlinear mechanical system shown in Fig.4.16. When the vibration amplitude $|x|$ is less than a certain value, x_0, the system is linear. However, when the vibration amplitude $|x| \geq x_0$, the system becomes nonlinear. The equation of motion of the system excited by a sinusoidal force is written as:

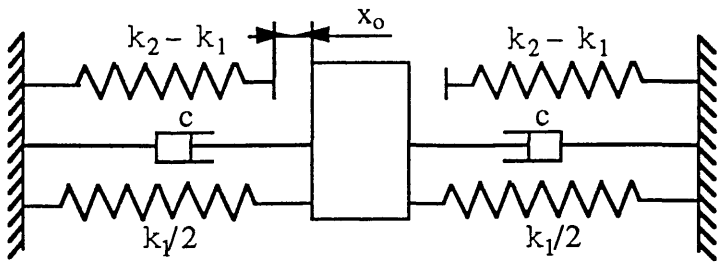


Fig.4.16 - Nonlinear Mechanical System with Backlash Stiffness Nonlinearity

$$m\ddot{x} + 2c\dot{x} + F(x) = A\cos\omega t \tag{4-29}$$

where F(x) is given as

$$F(x) = \begin{cases} k_2(x + x_0) & (x \geq x_0) \\ k_1 x & (|x| \leq x_0) \\ k_2(x - x_0) & (x < -x_0) \end{cases} \tag{4-30}$$

and is shown in Fig.4.17(a). Since the transient solution of (4-29) will decay due to the existence of damping (as illustrated in Fig.4.17(b), for which m=1kg, k_1=5000N/m, k_2=10000N/m, c=4N.s/m and x_0=0.005m), only the steady-state solution of (4-29) is of interest. When the forcing amplitude A and frequency ω satisfy the following relationship

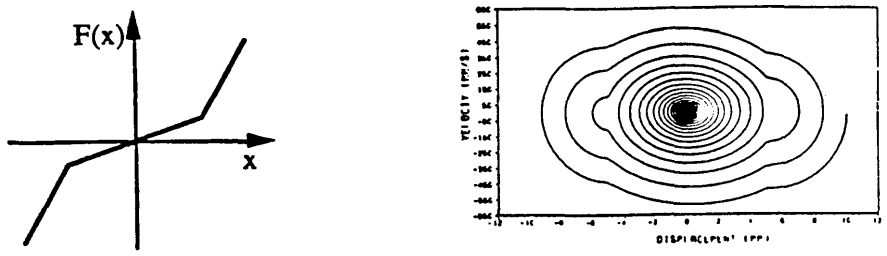
$$\frac{A}{\sqrt{(k_1 - m\omega^2)^2 + 4c^2\omega^2}} \leq x_0 \tag{4-31}$$

the system will behave exactly like a linear system for which the steady-state solution is given by

$$x(t) = X\cos(\omega t + \phi) \tag{4-32}$$

$$\text{where } X = \frac{A}{\sqrt{(k_1 - m\omega^2)^2 + 4c^2\omega^2}} \quad \text{and } \phi = \text{arctg}\left(\frac{2c\omega}{k_1 - m\omega^2}\right)$$

However, when (4-31) is not satisfied, the system becomes nonlinear and an analytical solution of (4-29) becomes mathematically impossible because an explicit analytical expression for the returning times $\tau(x)\,|_{x=\pm x_0}$ does not exist and numerical methods have to be employed.



(a) force displacement relationship (b) phase plane trajectory of free vibration

Fig.4.17 - Force Displacement Relationship and Phase Plane Trajectory

As discussed in some detail in Chapter 2, measured first-order frequency response functions can be analysed to detect and to quantify structural nonlinearity. Here, they are used to give a rough indication as to whether and when chaotic vibration will possibly occur in a mechanical backlash system. The first-order FRFs corresponding to the above-mentioned parameter settings are calculated for various excitation amplitudes and are shown in Fig.4.18. From figure 4.18, it can be seen that when the forcing amplitude is either large or small, the system becomes effectively linear and this gives the indication that if chaos is to exist in such a system, the forcing amplitudes should be of intermediate values.

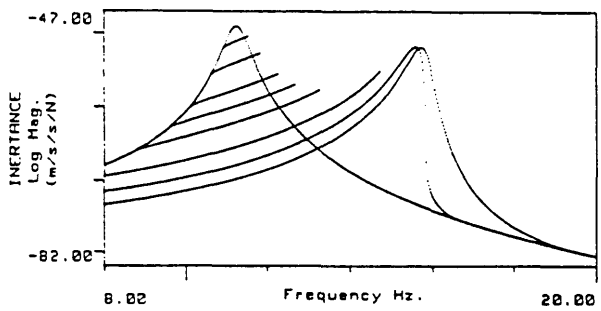
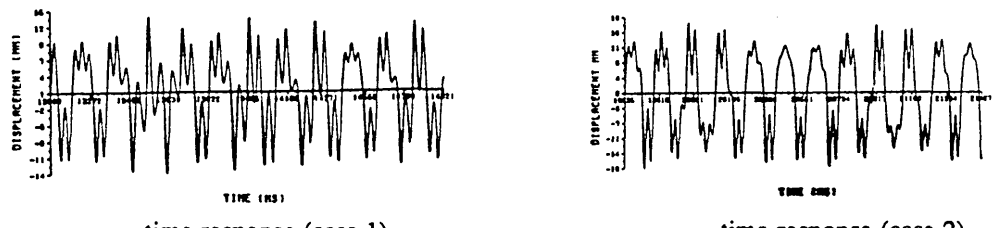


Fig.4.18 - First Order Frequency Response Functions of Mechanical Backlash System

4.3.3 CHAOTIC MOTION AND STRANGE ATTRACTING SETS

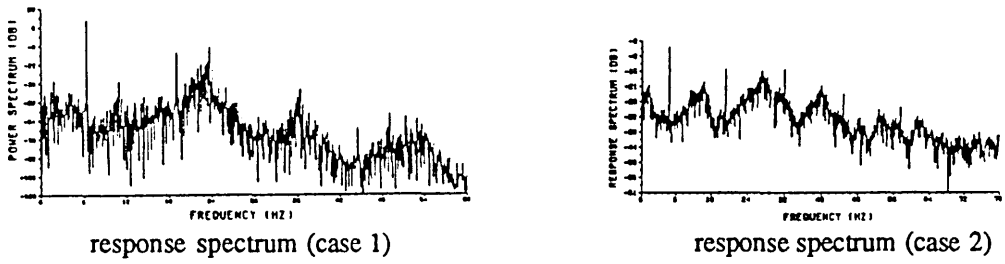
Such a simple system as described by equation (4-29) is found to be chaotic under certain excitation conditions. Here, the chaotic behaviour of the system with two different sets of system parameters is studied (case 1: m=1kg, k_1=0N/m, k_2=40000N/m, c=4N.s/m and x_0=0.005m and case 2: m=1kg, k_1=10000N/m, k_2=40000N/m, c=4N.s/m and x_0=0.005m) and a number of typical results are presented in the time, frequency and state-space domains. It has been found that there exist large forcing parameter (A,ω)

regions in which chaotic (bounded, nonperiodic) solutions exist and, from these, the chaotic solutions for case 1 with A=100N; ω=40rad./s and for case 2 with A=240N; ω=40rad./s (the excitations are pure sinusoids) are presented and shown in Figs.4.19-4.22. The Poincaré maps shown in figure 4.22 are plots of discrete state-space trajectories with sampling frequency equal to that of the excitation. These figures give a visual impression of what a chaotic motion looks like. From the time domain plots (figure 4.19) and the continuous state-space trajectory plots (figure 4.20) of the solutions, it can be seen that the motions contain some form of random components (nonperiodic) and this is confirmed by the broad-band frequency components appearing in the response spectra (figure 4.21). The well-defined patterns of Poincaré maps (figure 4.22) give rigorous confirmation that the solutions *are* indeed nonperiodic and, hence, chaotic.



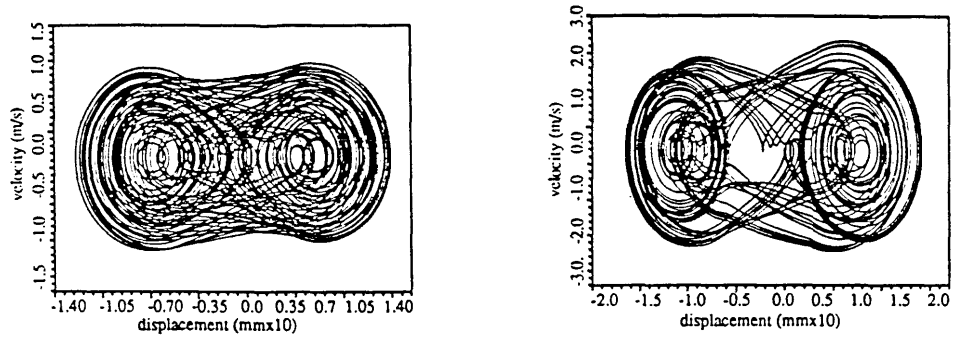
time response (case 1) time response (case 2)

Fig.4.19 - Time Response of Chaotic Backlash System



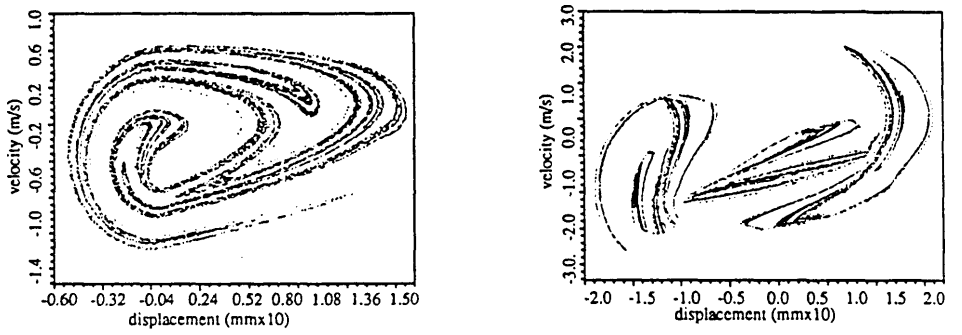
response spectrum (case 1) response spectrum (case 2)

Fig.4.20 - Response Spectrum of Chaotic Backlash System



state-space trajectory (case 1) state-space trajectory (case 2)

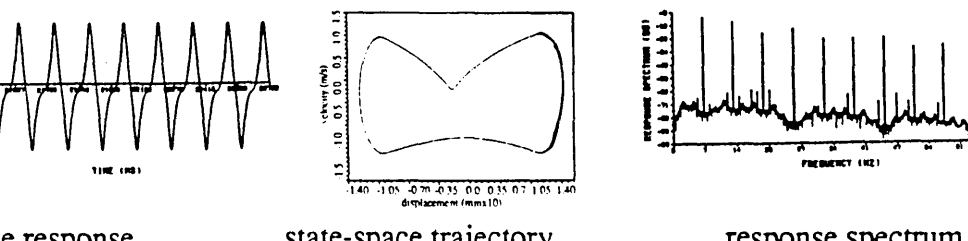
Fig.4.21 - State-space Trajectory of Chaotic Response



Poicaré map (case 1) Poicaré map (case 2)

Fig.4.22 - Poincaré map of Chaotic Response

Also, for comparison with these chaotic solutions, the period 1 (the period of the response is the same as that of the force) solution for case 1 with A=10N; ω=40rad./s is shown Fig.4.23. The time-domain plot and state-space trajectory show clearly the periodicity of the resulting motion, and the effective absence of any broad-band component (only harmonic components are present) in the response spectrum demonstrates the clear difference from the response spectrum of a chaotic response.



time response state-space trajectory response spectrum

Fig.4.23 - Period one solution of Backlash System

During the numerical simulation, it was found that before the onset of chaos, as the forcing parameters change gradually, a series of periodic doublings (bifurcations) occurred, as is the normal route to chaos. This was shown clearly in the logistic map of figure 4.2 in which, as the parameter λ changes, the period 1 solution bifurcates into period 2 and then to period 4 and then to period 8 and so on, until chaos sets in. However, in the present case, since the nonlinearity is symmetric, odd periodic solutions (e.g. period 3) also exist and therefore it is difficult to say in this case that the route to chaos is via period doubling.

Typical chaotic behaviour of a nonlinear backlash system has now been presented in the time, frequency and state-space domains. The existence of chaotic behaviour of a nonlinear system can, in general, be detected, as shown above, either using the response spectrum or more rigorously, using the the Poincaré map of the motion. The quantitative analysis of chaotic behaviour is to be discussed next.

4.3.4 FRACTAL DIMENSION OF STRANGE ATTRACTORS

As discussed in the introductory section, an attractor is defined in system dynamics as a well-defined structure in the state-space plot after the decay of transients due to the existence of damping. There are three classical types of dynamic motion and they are: (i) equilibrium, (ii) periodic motion/limit cycle and (iii) quasi-periodic motion. These states are called attractors since, after the transient decays, the system is attracted to one of the above states. Classical types of attractor are all associated with classical geometric forms in state-space; the equilibrium state with a point, the periodic motion/limit cycle with a closed curve and a quasi-periodic motion with a hyper-surface (a surface has a dimension of more than 3). However, a chaotic motion rides on a chaotic or strange attractor which is a stable structure of a long-term trajectory in a bounded region of state-space, which folds the bundle of trajectories back onto itself, resulting in a mixing and divergence of nearby states [67]. The strange attractor is associated with a new geometric form called a fractal set which has a dimension of noninteger value known as the 'fractal dimension'. For each chaotic motion, based on its Poincaré map, the fractal dimension can be calculated and the value of this fractal dimension gives quantitative measure of the complexity (or chaos) of the motion. As mentioned before, there are some different measures of the dimension of a set of points in space and the most intuitive one is the capacity dimension. The detailed procedure of calculating the fractal (capacity) dimension of a given chaotic attractor (Poincaré map) was presented in section §4.2.6.

Based on equations (4-27) and (4-28), the fractal dimensions have been calculated for the Poincaré maps of figure 4.22 and found to be $D_c = 1.206$ for case 1 and $D_c = 1.165$ for case 2, respectively. It should be mentioned here that although ε should theoretically be as small as possible, according to (4-27), different values of ε must be tried due to numerical and/or experimental inaccuracies until the calculated dimension becomes independent of ε as shown in Fig.4.24. The noninteger values of these dimensions show that the attractors of figure 4.22 indeed have fractal/self-similar structures and that the motions riding on them *are* chaotic. Furthermore, it is worth mentioning that, in addition to the quantification of the complexity of the chaotic motion, the calculation of fractal dimension is very important for the modelling of chaotic systems because it is from this value that the number of degrees of freedom necessary to model a practical chaotic system can be determined so that all the topological nature of the attractor can be preserved.

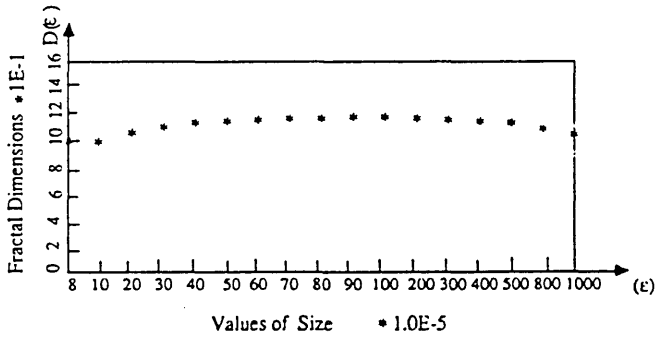


Fig.4.24 - Capacity Dimension Versus the Size of ε

4.3.5 SENSITIVITY TO INITIAL CONDITIONS AND LYAPUNOV EXPONENTS

As discussed in section §4.2.5, chaos in dynamics implies a sensitivity in the outcome of a dynamic process to small changes in the initial conditions. When a system becomes chaotic, the accurate prediction of long-term response becomes impossible because, in this case, a small initial condition uncertainty will be magnified exponentially as time goes on and, as a result, two originally indistinguishable initial conditions will lead to completely different long-term solutions. This sensitivity to initial conditions for case 1 with A=100N; ω=40rad./s is illustrated in Fig.4.25 (the time interval between two successive points for these two trajectories is a forcing period).

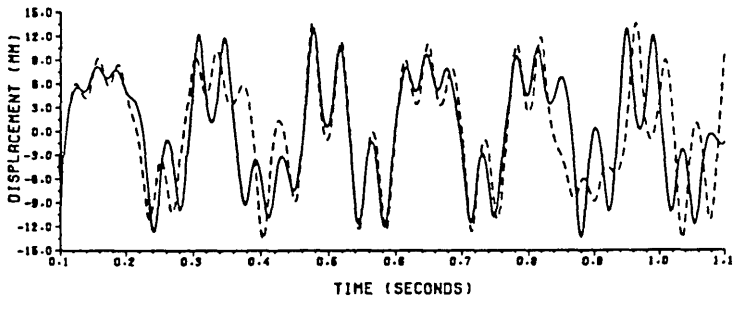


Fig.4.25 - Sensitivity on Initial Conditions of Chaotic Solutions

In order to quantify this sensitivity to the initial conditions, the Lyapunov exponent of the motion needs to be calculated. Imagine a set of initial conditions within a sphere of radius ε in phase space, then the chaotic motion trajectories originating in the sphere will map the sphere into an ellipsoid whose major axis grows as $d=\varepsilon e^{\lambda t}$, where λ is known as a Lyapunov exponent. As mentioned in §4.2.5, for regular motions, $\lambda \leq 0$, while for chaotic motions, $\lambda > 0$. Thus, the sign of λ is a criterion of chaos. The numerical method for calculating the Lyapunov exponent was well explained in [68]. Suppose we have two chaotic trajectories, Γ_1 and Γ_2, starting with very close initial conditions as shown in Fig.4.26, then λ can be calculated as:

$$\lambda = \lim_{N\to\infty} \frac{1}{t_N - t_0} \sum_{k=1}^{N} \ln \frac{d_k}{d_{k-1}} \tag{4-33}$$

where d_k is the separation of these two trajectories at time t_k.

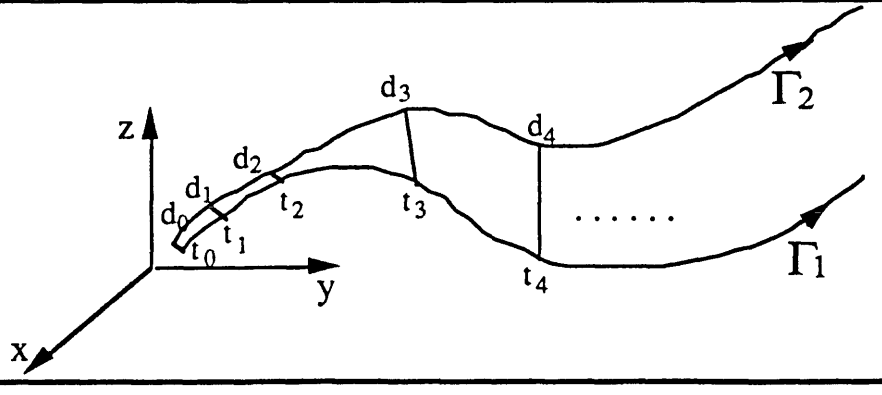


Fig.4.26 - Trajectories Starting with Close Initial Conditions

Based on (4-33), the Lyapunov exponents for case 1 with A=100N; ω=40rad./s and case 2 with A=240N; ω=40rad./s are calculated to be 0.532 and 0.624. During the calculation,

an integration time of 200 cycles was used and, as shown in Fig.4.27, the calculated value can be considered to be reliable because by that time, λ virtually does not change. These positive Lyapunov exponents give quantitative measure that the trajectories diverge, on average, at an exponential rate of $\lambda = 0.532$ for case 1 and $\lambda=0.624$ for case 2.

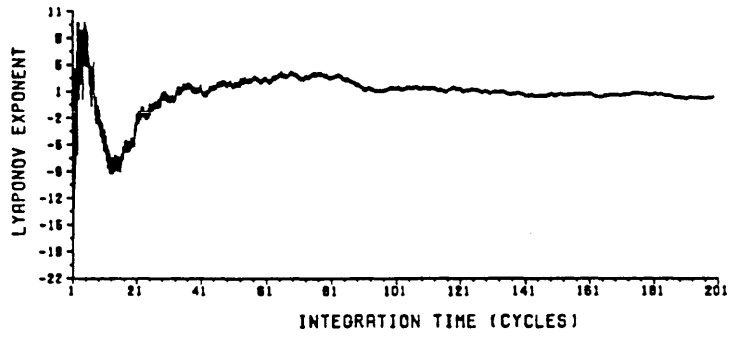


Fig.4.27 - Calculated Lyapunov Exponents Versus Integration Time

4.3.6 EFFECT OF FORCING PARAMETER AND DAMPING ON CHAOS

Once the chaotic nature of a nonlinear system has been established, what is then of interest is to know under what forcing conditions chaotic vibrations will occur because if the necessary conditions for chaos have been determined, then it is possible to avoid them or to employ them if there are some advantages of doing so. At present, the determination of the forcing parameter field of a nonlinear system in which chaotic vibrations occur is generally achieved by experiment although analytical predictions for some specific chaotic systems such as Duffing's system have been undertaken. In the present study, the forcing parameter field for the existence of chaotic vibration of the system described by equation (4-29) (case 1) was determined by numerical experiment, results of which are shown in Fig.4.28. It has been found - as expected - that chaotic vibrations occur when the forcing amplitudes are of moderate values for all the excitation frequencies tried. For the higher excitation frequencies, although it cannot be proven because of limited calculation capacity, it was found that no chaotic vibrations occurred when ω was greater than 85 rad./s.

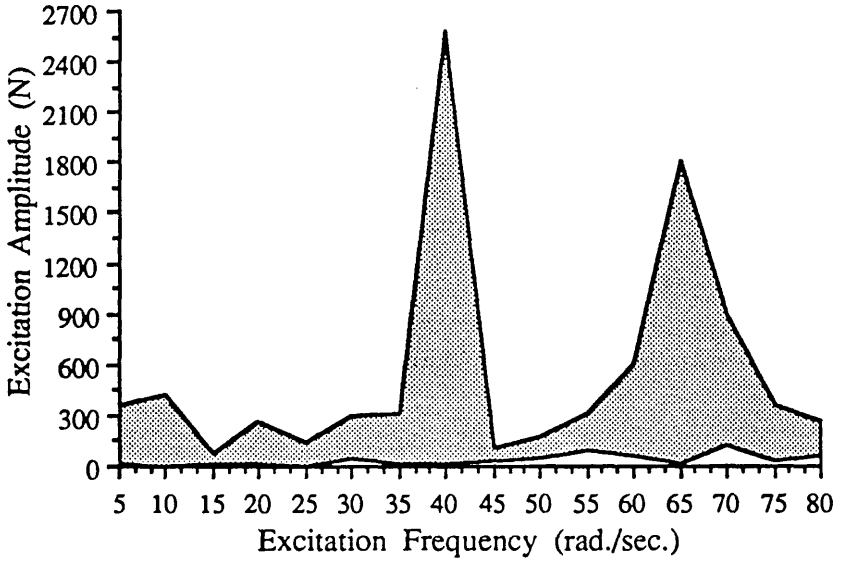
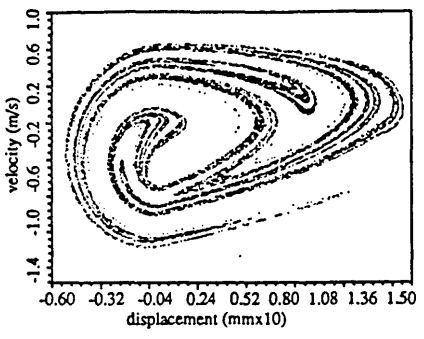
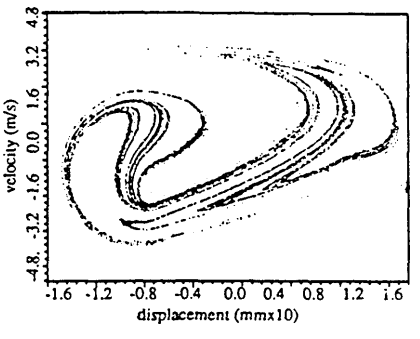


Fig.4.28 - Forcing Parameter Field for the Existence of Chaotic Vibration

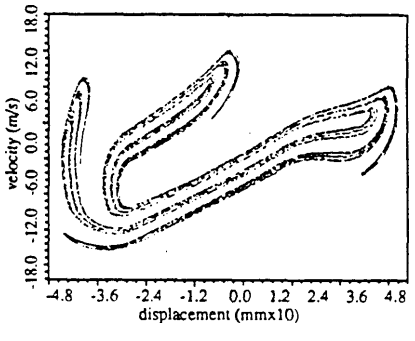
To see how chaotic motion changes when the forcing amplitude increases, the Poincaré maps of different forcing amplitudes at an excitation frequency ω=40rad./s are calculated and are shown in Fig.4.29. The calculated fractal dimensions show that although all the motions are chaotic, they become more 'regular' as forcing amplitude increases.



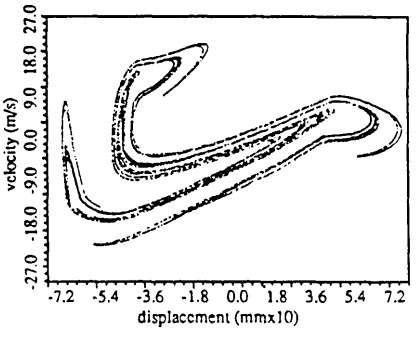
A = 100, D_c=1.206



A=200, D_c=1.187



A = 1000, D_c=1.156



A=2000, D_c=1.143

Fig.4.29 - The Poincaré Maps of Different Forcing Amplitudes

To assess the effects of damping on chaotic vibration, different damping levels were introduced for case 1 with A=100N; ω=40rad./s. As expected, an increase in damping was found to reduce the 'randomness' in the chaotic motion and the Poincaré map of the motion becomes more compact as the damping increases, as shown in Fig.4.30. The fractal dimensions of the Poincaré maps for different values of damping c were also calculated and the results shown in Fig.4.31. Clearly, the introduction of damping is an effective way of avoiding unsatisfactory motions of chaotic systems.

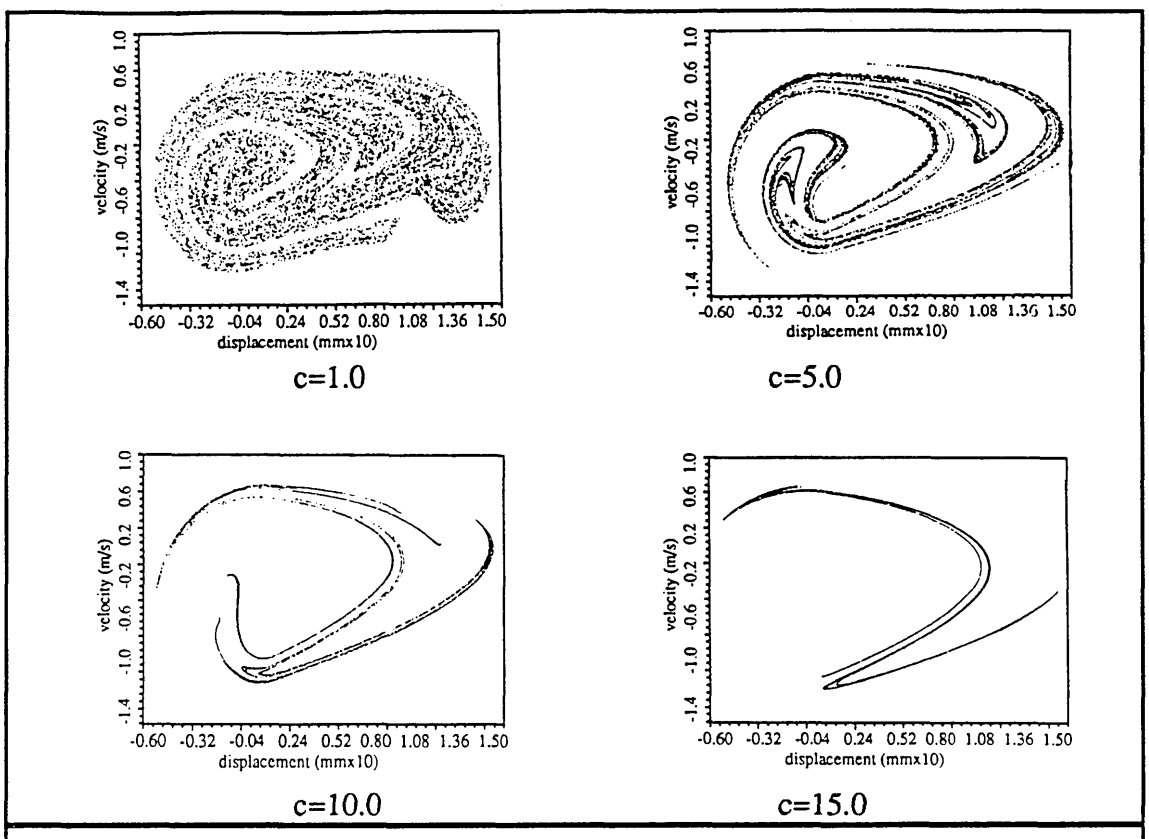


Fig.4.30 - The Poincaré Maps at Different Damping Levels

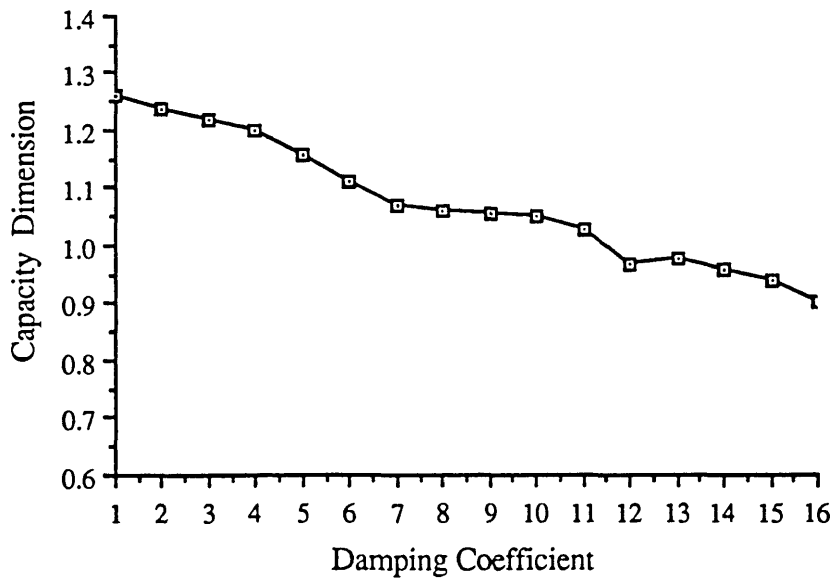


Fig.4.31 - Fractal Dimension D_c Versus Damping Level c

4.3.7 TWO-DEGREE-OF-FREEDOM BACKLASH SYSTEM

So far, the chaotic behaviour of an SDOF mechanical system with backlash stiffness nonlinearity has been investigated in some detail. In order to study chaotic vibrations of a nonlinear MDOF system, the 2DOF system with backlash stiffness nonlinearity as shown in Fig.4.32 is considered. It is generally believed that if chaotic vibration occurs in a MDOF system, the motion will become yet more complex than that in an SDOF system with same type of nonlinearity because, in this case, the interactions between all the degrees of freedom act as chaotic excitations and these chaotic excitations make the resulting motion more complex. As in the case of the SDOF backlash system, it was found that for very low and very high excitation levels, the motions are periodic. However, there exists a wide range of forcing parameters in which chaotic vibrations occur.

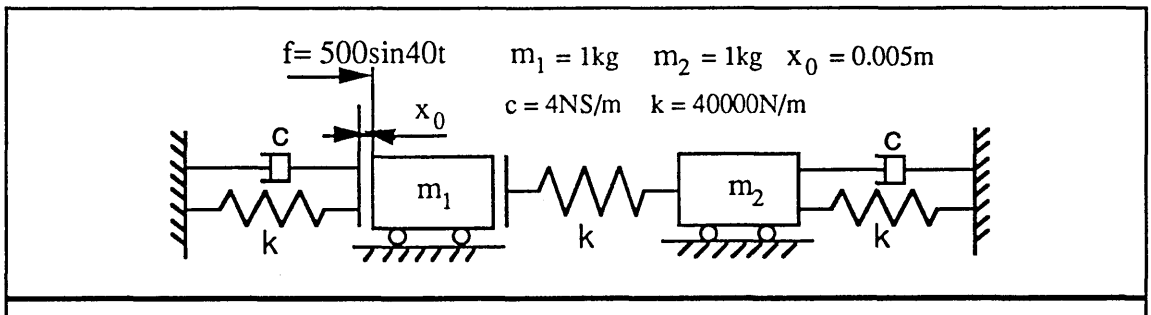


Fig.4.32 - 2DOF Mechanical System with Backlash Stiffness Nonlinearity

The system parameters are as shown in figure 4.32 and it was found that when the forcing amplitude A=500N and forcing frequency ω=40rad./s, the motions become chaotic. As before, the results are presented in the time-, frequency- and state-space domains as shown in Fig.4.33. However, as the 2DOF system is a four-dimensional system since there are four state variables $\{x_1,\dot{x}_1,x_2,\dot{x}_2\}^T$, the projection of the four-dimensional Poincaré map onto the x_1 *vs* $\dot{x}_1$ plane (2-dimensional) disguises the fractal/self-similar properties of the true Poincaré map.

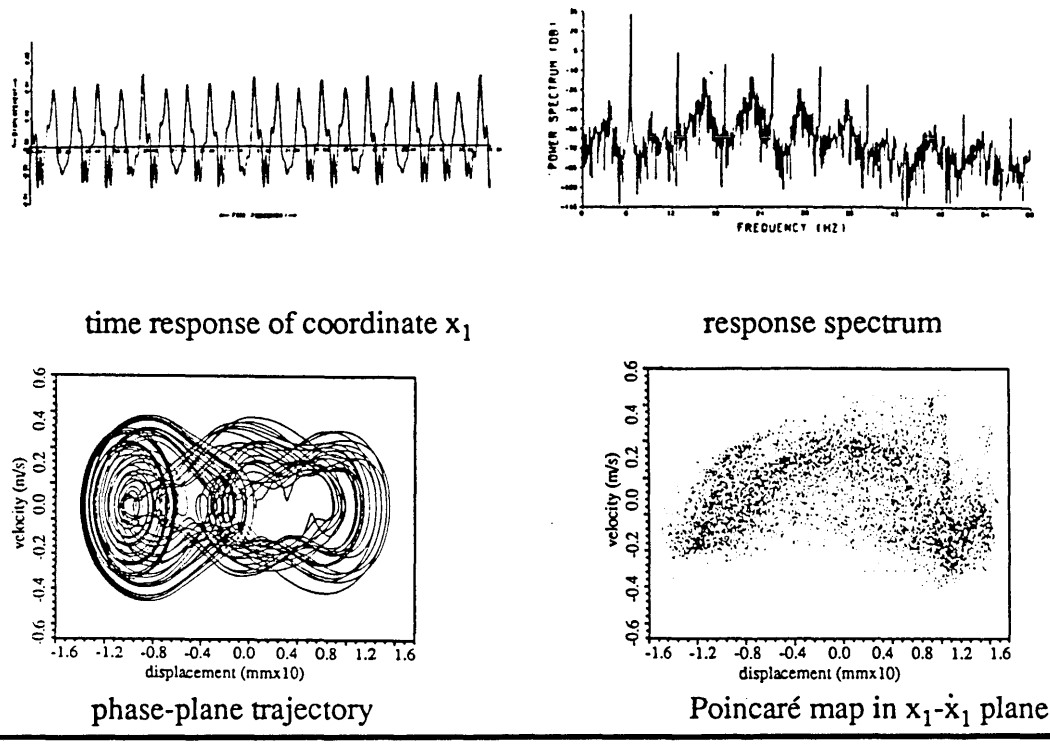


Fig.4.33 - Chaotic Response of 2DOF Backlash System

4.3.8 EXPERIMENTAL INVESTIGATION

Detailed numerical studies on the chaotic vibration of backlash system have been carried out and reported alone. However, the question remains: "do chaotic vibration exists in real practical backlash systems?" To answer this question, an experiment was designed based on a test structure comprising a simply-supported beam with a mass at its midpoint to simulate the SDOF system, as shown in Fig.4.34. The first natural frequency of the structure was designed to be around 20Hz with the second mode much higher so that when the excitation is around 20Hz, the structure behaves effectively like an SDOF system. The backlash stiffness nonlinearity was introduced by providing motion constraints on both sides of the mass (figure 4.34) so that the stiffness characteristics of the equivalent SDOF system can be represented by that shown in figure 4.17(a). The response was detected by a strain gauge attached on the beam near the mass such that the measured strain is proportional to the displacement of the mass. A sinusoidal excitation force is produced by an electro-magnetic shaker acting on the mass.

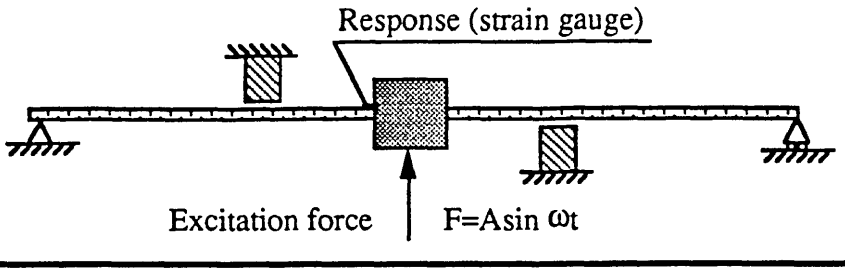


Fig.4.34 - Experimental Model of Backlash System

As in the numerical studies, periodic as well as chaotic responses were found to exist at different forcing amplitudes and frequencies. The time response and auto-spectrum of a chaotic response at excitation frequency f=25Hz are presented in Fig.4.35. A pseudo-Poincaré map (x(n+1) *vs* x(n), where x(n) is the sampled response signal with sampling frequency equal to the excitation frequency which is 25Hz), is shown in Fig.4.36 because in the experimental case, usually only one signal (displacement or velocity) is available (the simultaneous measurement of displacement and velocity is practically difficult in some cases and it has been mathematically established [58] that the same amount of information about the motion of the system can be obtained from the pseudo-Poincré map instead of the true Poincaré map). From these results, the chaotic nature of the response is clearly demonstrated. However, it is difficult to compare these results with those from the numerical calculations because the necessary system parameters of the experimental rig are difficult to determine.

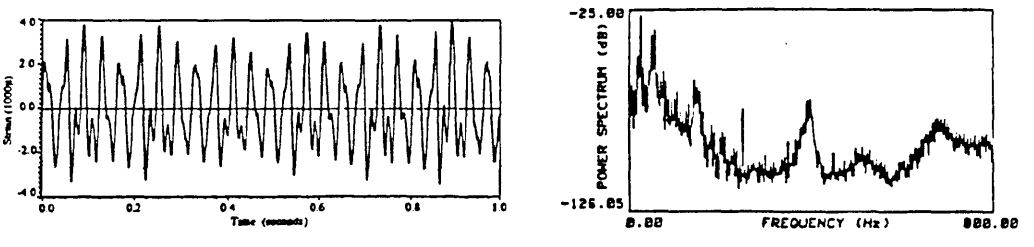


Fig.4.35 - Time Response and Auto-spectrum of Measured Chaotic Response

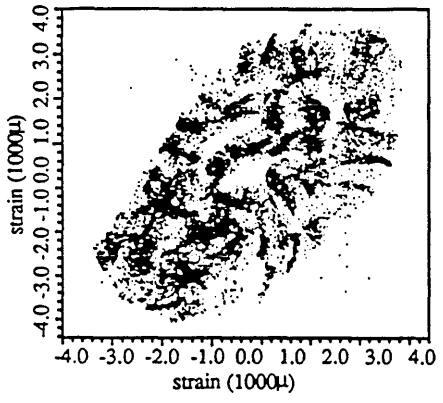


Fig.4.36 - Pseudo-Poincaré Map of Experimental Chaotic Response

4.4 CONCLUSIONS

In this chapter, the basic theory of chaotic vibration has been summarised and ingredients which are required in order to understand chaotic behaviour of dynamic systems have been illustrated. Together with Chapters 2 & 3, a complete picture of all the probable nonlinear phenomena in structural dynamics and the analysis techniques for identifying them have been presented.

For the first time, the hidden chaotic behaviour of nonlinear mechanical systems with backlash stiffness nonlinearity has been studied in some detail both numerically and experimentally. Particular attention has been paid to the identification of chaotic vibration in such nonlinear systems. Indeed, as shown in the numerical simulations, there exist wide parameter regions, both in the system parameters and the external forcing conditions, for which chaotic vibrations occur. Qualitative as well as quantitative ways of identifying chaotic vibration in nonautonomous nonlinear systems are presented.

The chaotic behaviour is explained in time-, frequency- and state-space domains. For detecting the existence of chaotic vibrations, the response spectrum or, more rigorously, the Poincaré map of the motion, is employed. The fractal dimensions of strange attractors are calculated and serve the purpose of quantifying the complexity of the motion. The sensitivity of chaotic motions to initial conditions is examined and the Lyapunov exponents are calculated, giving further indication of the existence of chaotic vibration.

The system studied is a singularly simple system whose equation of motion is very easy to understand physically. Also, as shown in this paper, an experimental model can readily be constructed to demonstrate the predicted behaviour. Such a system is likely to become a paradigm for further research into chaos in nonlinear dynamical systems. In mechanical structures, such nonlinear mechanisms represent the intermittent contact between components due to manufacturing clearances, and therefore it is expected that many mechanical systems might exhibit chaotic behaviour under appropriate operating conditions. Since one of the major consequences of chaos is unpredictability of the response, it is therefore recommended that statistical methods should be applied to stress/fatigue analysis when such conditions are anticipated. Furthermore, from a condition monitoring view point, if a broad-band response can be caused by a purely sinusoidal excitation (e.g., due to the eccentricity of rotational components), this makes reliable diagnosis in most cases difficult and creates the necessity of a new understanding of such nonlinear systems and the development of new techniques so that reliable diagnosis can be achieved. Further, in the design of mechanical control systems such as robots, where such backlash stiffness nonlinearity is very likely to exist, care must clearly be taken at the design stage so that under normal service conditions, undesirable or unpredictable chaotic motion will not occur.

CHAPTER 5

LOCATION OF STRUCTURAL NONLINEARITIES

5.1 PRELIMINARIES

So far, different types of dynamic phenomenon observed in nonlinear structures have been discussed and techniques for analysing them have been presented. For practical structures whose nonlinearities are such that the measured first-order FRFs using sinusoidal excitation (when the amplitude of the excitation force is kept constant) display the nonlinear behaviour (most practical nonlinearities are of this nature), a new identification method has been developed in Chapter 2 which can not only quantify the extent of the nonlinearity, but also identify its type in some cases. On the other hand, for some nonsymmetric nonlinearities such as the quadratic and bilinear stiffness nonlinearities, as discussed in Chapter 3, the measured first-order FRFs are effectively linear and analysis of the higher-order FRFs becomes necessary so that such nonlinearities can be identified in practice. However, both first-order and higher-order FRF analyses are largely based on the assumption of periodic-input periodic-output and for some nonlinear structural systems (chaotic systems), this assumption is no longer valid. Based on a system with backlash stiffness nonlinearity, qualitative as well as quantitative ways of identifying chaotic vibrational systems have been presented in Chapter 4. Some important engineering applications of the techniques developed have been discussed.

In this and later Chapters, we shall confine ourselves to the analysis of first-order FRFs only although we accept that such analysis could, in some cases, be very approximate. It will be shown how measured first-order FRF data (or their derived modal data), together with an analytical model of the structure (usually an FE model), can be used to locate the structure's localised nonlinearity (Chapter 5) and later how an accurate mathematical model of a dynamic structure can be established by correlating an analytical model and measured dynamic test data (Chapters 6 & 7).

5.2 NECESSITIES AND REQUIREMENTS FOR NONLINEARITY LOCATION

It is usually believed that, if they exist, structural nonlinearities are localised in terms of spatial coordinates as a result of the nonlinear dynamic characteristics of structural joints, nonlinear boundary conditions and nonlinear material properties such as plasticity. The ability to locate a structure's localised nonlinearity thus has some important engineering applications. First, the information about where the structural nonlinearity is may offer opportunities to separate the structure into linear and nonlinear subsystems so that these can be analysed separately based on nonlinear substructuring analysis [69]. Second, since nonlinearity is often caused by the improper connection of structural joints, its location may give an indication of a malfunction or of poor assembly of the system. Third, from a materials property point of view, the stress at certain parts of the structure during vibration can become so high that the deformation of that part becomes plastic and the dynamic behaviour becomes nonlinear. In this case, location of the nonlinearity may offer the possibility of failure detection. Finally, as will be discussed in detail in Chapters 6&7, location information is essential if a nonlinear mathematical model of the structure is to be established.

In practical measurements, the data measured are usually quite limited (both measured modes and coordinates are incomplete) and this is especially true when a nonlinear structure is considered, as will be discussed in some detail in Chapter 6. It is therefore believed that the task of locating a structure's localised nonlinearity can only become possible by correlating an analytical model, which may contain modelling errors but can represent the structure to some accuracy, and the results from dynamic test of the structure. To illustrate the above argument, consider a typical nonlinear structure (two linear components connected by a nonlinear joint) as shown in Fig.5.1. Mathematically, the structure possesses a mass matrix [M], which is constant, and a stiffness matrix $[K(\hat{X})]$, which is a function of of response amplitude, if stiffness nonlinearity is

considered (for the convenience of discussion, the structure is assumed to be undamped). Clearly, if the impedance matrix $[Z(\omega)]$ of the structure can be measured, then the nonlinearity location becomes straightforward as shown in Fig.5.2. However, what can be measured in practice is the the receptance matrix $[\alpha(\omega)]$, which is the inverse of $[Z(\omega)]$, and in this inverse format the localised stiffness change due to the nonlinearity at different response amplitudes is spread over the whole matrix, as shown schematically in Fig.5.3. Of course, one may obtain $[Z(\omega)]$ by inverting the measured $[\alpha(\omega)]$ but, unfortunately, such a process is found to be extremely sensitive to measurement noise and, therefore, is often not implementable in practice. Hence, it becomes clear that in order to locate the nonlinearity based on measured data only, *all* the coordinates and *all* the modes of the structure should be measured so that the mass and stiffness matrices of the structure at different response amplitudes, which are necessary to the location of nonlinearity, can be reconstructed. This demonstrates the difficulties of using measured data only to do the location task.

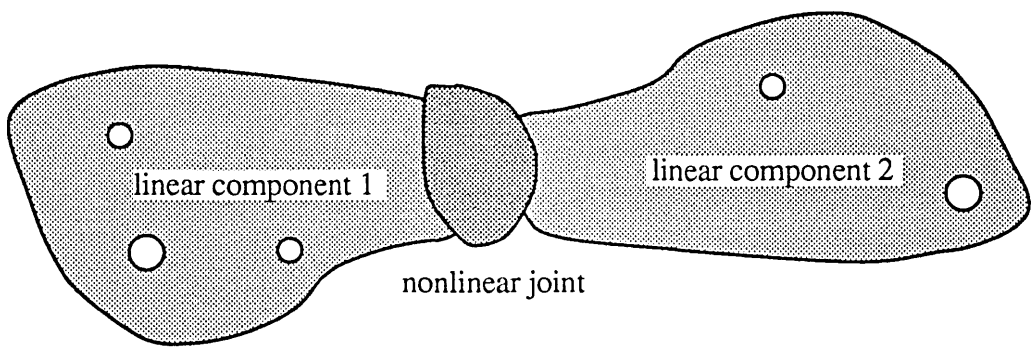


Fig.5.1 - A typical Nonlinear Structure

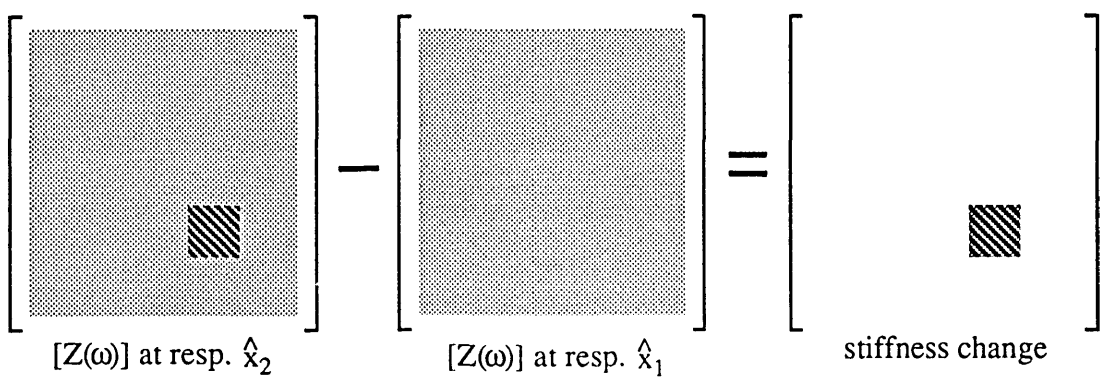


Fig.5.2 - Schematical Illustration of Location Process

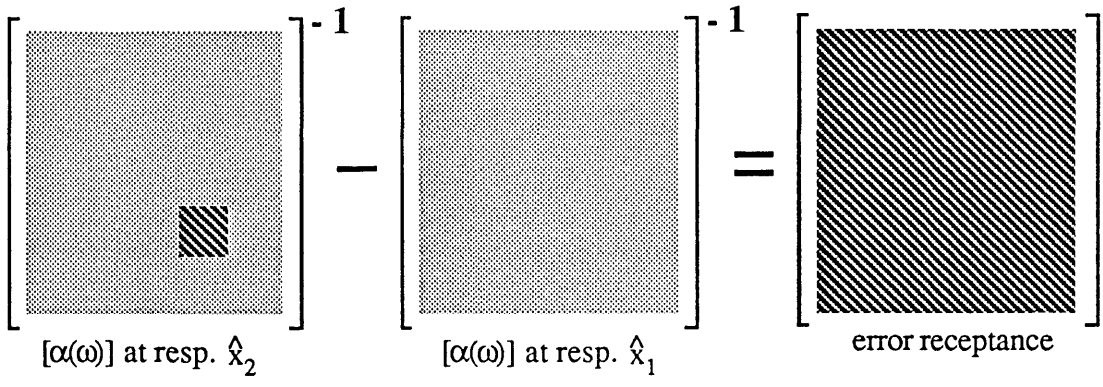


Fig.5.3 - Schematical Illustration of Inverse Process

Fortunately, due to the development of analytical modelling techniques, an analytical model of a structure can be employed. Although it may contain modelling errors, these are usually of second order when compared with the analytical model itself in the sense of the Euclidean norm. With such an analytical model available, it will be shown in this Chapter that by correlating the analytical model and the measured dynamic test data, location of nonlinearity can be achieved.

5.3 TECHNIQUE FOR THE LOCATION OF STRUCTURAL NONLINEARITY

5.3.1 LOCATION USING MEASURED MODAL DATA

The location method developed in this Chapter is based on the correlation between an analytical model which contains modelling errors and dynamic test data which are measured at different response levels. A nonlinearity location method based on the use of measured modal data is discussed first. This method is then extended to the case of using measured FRF data.

Before discussion, it is necessary to mention that the 'modes' of a nonlinear structure are difficult to define (if indeed they exist at all) in an exact mathematical sense [70-73] because of the existence of harmonic response components, and so the term 'modes of a nonlinear structure' is used in this Chapter to mean the natural frequencies and modeshapes which are derived from the analysis of measured first-order FRFs in which only the fundamental frequency component of the response is of interest. For most nonlinear mechanical structures, the thus-obtained natural frequencies and modeshapes

are response level dependent. As far as stiffness nonlinearity is concerned, the stiffness matrix of the structure corresponding to different response levels will be different and, therefore, if this difference in stiffness matrix can be calculated in some way, the problem of nonlinearity location can be resolved.

Suppose that the eigenvalues and eigenvectors of the r^{th} mode (which is sensitive to the localised nonlinearity) corresponding to a lower response level, $\hat{x}_1$, are $\lambda_1;\{\phi_1\}$ and those corresponding to a higher response level, $\hat{x}_2$, are $\lambda_2;\{\phi_2\}$ and that these have been obtained from the analysis of measured first-order FRFs (either based on the new nonlinear modal analysis method discussed in Chapter 2 or based on standard linear modal analysis methods by linearising the structure using response control). Suppose also that the analytical model which contains second-order modelling errors (corresponding to lower response level) is available. Then, from the eigendynamic equations, the following relationship can be established:

$$(- ([M_a] + [\Delta M]) \lambda_1 + [K_a] + [\Delta K]) \{\phi_1\} = \{0\} \tag{5-1}$$

$$(- ([M_a] + [\Delta M]) \lambda_2 + [K_a] + [\Delta K] + [\Delta K_n]) \{\phi_2\} = \{0\} \tag{5-2}$$

Post-multiply (5-2) by $\{\phi_1\}^T$, then

$$(- \lambda_2 [\Delta M] + [\Delta K] + [\Delta K_n]) \{\phi_2\}\{\phi_1\}^T = - (- \lambda_2 [M_a] + [K_a]) \{\phi_2\}\{\phi_1\}^T \tag{5-3}$$

Post-multiply (5-1) by $\{\phi_2\}^T$, we have

$$(- \lambda_1 [\Delta M] + [\Delta K]) \{\phi_1\}\{\phi_2\}^T = - (- \lambda_1 [M_a] + [K_a]) \{\phi_1\}\{\phi_2\}^T \tag{5-4}$$

Subtract (5-4) from (5-3) and rearrange, then

$$[\Delta M] (- \lambda_2 \{\phi_2\}\{\phi_1\}^T + \lambda_1 \{\phi_1\}\{\phi_2\}^T) + [\Delta K] (\{\phi_2\}\{\phi_1\}^T - \{\phi_1\}\{\phi_2\}^T) +$$

$$[\Delta K_n] (\{\phi_2\}\{\phi_1\}^T) =$$

$$(- \lambda_1 [M_a] + [K_a]) \{\phi_1 \} \{\phi_2\}^T - (- \lambda_2 [M_a] + [K_a]) \{\phi_2\} \{\phi_1\}^T \tag{5-5}$$

Since $\{\phi_2\}$ is a perturbed modeshape of $\{\phi_1\}$, due to the stiffness change of nonlinearity, $\{\phi_2\}\{\phi_1\}^T - \{\phi_1\}\{\phi_2\}^T$ is of second order compared with $\{\phi_2\}\{\phi_1\}^T$ in the sense of the Euclidean norm (one can notice that all the diagonal elements of $\{\phi_2\}\{\phi_1\}^T - \{\phi_1\}\{\phi_2\}^T$

are zero), as a result, if the modelling errors $[\Delta K]$, $\lambda_1[\Delta M]$ and stiffness change due to nonlinearity $[\Delta K_n]$ are of the same order of magnitude (also in the sense of the Euclidean norm), then to the first order approximation, (5-5) becomes

$$[\Delta K_n]\{\phi_2\}\{\phi_1\}^T \approx (-\lambda_1[M_a]+[K_a])\{\phi_1\}\{\phi_2\}^T - (-\lambda_2[M_a]+[K_a])\{\phi_2\}\{\phi_1\}^T \quad (5\text{-}6)$$

As a special case in which $[\Delta K]=[O]$ and $[\Delta M]=[O]$ (no modelling errors), then (5-6) becomes an exact statement for $[\Delta K_n]$. The principle of the nonlinearity location process based on equation (5-6) is illustrated in Fig.5.4. If the nonlinearity is localised, then $[\Delta K_n]$ will be a very sparse matrix (only those elements where the structural nonlinearity is located are nonzero) and, as shown in figure 5.4, the dominant nonzero elements of the resultant matrix after the matrix multiplication will indicate the location of localised nonlinearity. Also, it should be noticed that during the location process, only one measured mode is required and it is recommended that the mode which is the most sensitive to nonlinearity in the measured frequency range should be used. Extra modes can be used to check the consistency and reliability of the location results.

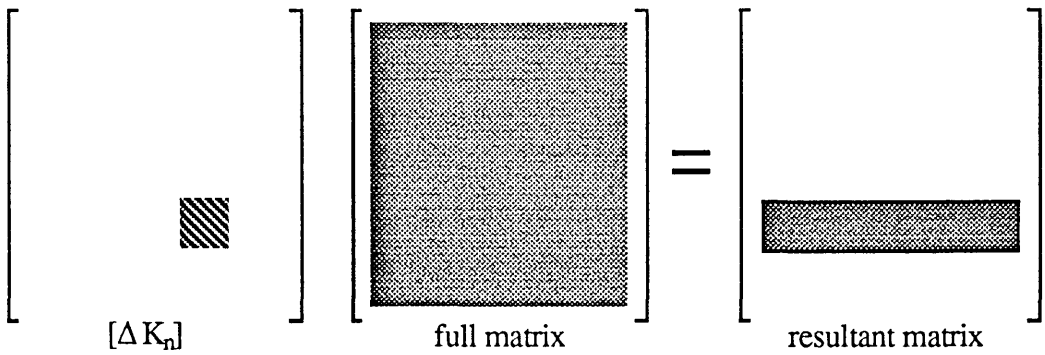


Fig.5.4 - Illustration of Nonlinearity Location Process

5.3.2 EXPANSION OF UNMEASURED COORDINATES

In the theoretical development of the location method, it is assumed that the measured coordinates are complete. In practice, however, this is very difficult to achieve because certain coordinates are physically inaccessible, such as internal DOFs, and the rotational coordinates are very difficult to measure and so the unmeasured coordinates have to be interpolated first before the location process can be carried out. This interpolation of

unmeasured coordinates can be achieved by using the analytical model itself based on Kidder's expansion method [74].

Although the analytical model contains modelling errors, in order to interpolate the unmeasured coordinates, it is assumed that the following relationship between the analytical model and the r^{th} measured and unmeasured sub-modes holds:

$$-\omega_r^2 \begin{bmatrix} [M_{mm}] & [M_{ms}] \\ [M_{sm}] & [M_{ss}] \end{bmatrix} \begin{Bmatrix} \{\phi_m\} \\ \{\phi_s\} \end{Bmatrix} + \begin{bmatrix} [K_{mm}] & [K_{ms}] \\ [K_{sm}] & [K_{ss}] \end{bmatrix} \begin{Bmatrix} \{\phi_m\} \\ \{\phi_s\} \end{Bmatrix} = \begin{Bmatrix} \{0\} \\ \{0\} \end{Bmatrix} \qquad (5\text{-}7)$$

where $\{\phi_m\}$ and $\{\phi_s\}$ are the r^{th} measured and unmeasured sub-modes. Upon multiplying out (5-7), the following two equations are established

$$(-\omega_r^2 [M_{mm}] + [K_{mm}]) \{\phi_m\} + (-\omega_r^2 [M_{ms}] + [K_{ms}]) \{\phi_s\} = \{0\} \qquad (5\text{-}8)$$

$$(-\omega_r^2 [M_{sm}] + [K_{sm}]) \{\phi_m\} + (-\omega_r^2 [M_{ss}] + [K_{ss}]) \{\phi_s\} = \{0\} \qquad (5\text{-}9)$$

Theoretically, $\{\phi_s\}$ can be calculated from either (5-8) or (5-9). However, when the number of measured coordinates is less than that of the unmeasured ones, which is quite usual in practice, (5-8) becomes underdetermined in terms of the solution of $\{\phi_s\}$ (the coefficient matrix is rank deficient), and it is therefore recommended that (5-9) should be used to interpolate $\{\phi_s\}$ as follows:

$$\{\phi_s\} = (-\omega_r^2 [M_{ss}] + [K_{ss}])^{-1} (\omega_r^2 [M_{sm}] - [K_{sm}]) \{\phi_m\} \qquad (5\text{-}10)$$

It has been found that the interpolation of unmeasured coordinates based on (5-10) is quite accurate for the lower modes of vibration (this will be further discussed in Chapter 7) and from the nonlinearity location point of view, if some coordinates have been measured where the structural nonlinearity is located, the thus-interpolated modeshapes can be used to achieve a successful nonlinearity location. Also, it can be shown mathematically that the located errors in the above mentioned nonlinearity location process will only occur in the measured coordinates if the unmeasured coordinates are interpolated based on (5-10). This is briefly illustrated below.

Since $(\lambda_1[M_a] + [K_a])\{\phi_1\}$ (and similarly $(\lambda_2[M_a] + [K_a])\{\phi_2\}$) on the RHS of (5-6) can be re-written as:

$$(\lambda_1 [M_a] + [K_a]) \{\phi_1\} = - \hat{\omega}_r^2 \begin{bmatrix} [M_{mm}] & [M_{ms}] \\ [M_{sm}] & [M_{ss}] \end{bmatrix} \begin{Bmatrix} \{\hat{\phi}_m\} \\ \{\hat{\phi}_s\} \end{Bmatrix} + \begin{bmatrix} [K_{mm}] & [K_{ms}] \\ [K_{sm}] & [K_{ss}] \end{bmatrix} \begin{Bmatrix} \{\hat{\phi}_m\} \\ \{\hat{\phi}_s\} \end{Bmatrix}$$

$$= \begin{Bmatrix} (- \hat{\omega}_r^2 [M_{mm}] + [K_{mm}]) \{\hat{\phi}_m\} + (- \hat{\omega}_r^2 [M_{ms}] + [K_{ms}]) \{\hat{\phi}_s\} \\ (- \hat{\omega}_r^2 [M_{sm}] + [K_{sm}]) \{\hat{\phi}_m\} + (- \hat{\omega}_r^2 [M_{ss}] + [K_{ss}]) \{\hat{\phi}_s\} \end{Bmatrix} \quad (5\text{-}11)$$

where parameters with ^ are the modal parameters corresponding to lower response level. When $\{\hat{\phi}_s\}$ is interpolated based on (5-10), then it is easy to see that in (5-11), the elements corresponding to the unmeasured coordinates (the lower part) are zero and (5-11) becomes

$$\begin{Bmatrix} (- \hat{\omega}_r^2 [M_{mm}] + [K_{mm}]) \{\hat{\phi}_m\} + (- \hat{\omega}_r^2 [M_{ms}] + [K_{ms}]) \{\hat{\phi}_s\} \\ (- \hat{\omega}_r^2 [M_{sm}] + [K_{sm}]) \{\hat{\phi}_m\} + (- \hat{\omega}_r^2 [M_{ss}] + [K_{ss}]) \{\hat{\phi}_s\} \end{Bmatrix} = \begin{Bmatrix} \{R_1\} \\ \{0\} \end{Bmatrix} \quad (5\text{-}12)$$

Upon substitution of (5-12), (5-6) becomes

$$[\Delta K_n] \{\phi_2\} \{\phi_1\}^T = \begin{Bmatrix} \{R_1\} \\ \{0\} \end{Bmatrix} \{\phi_2\}^T - \begin{Bmatrix} \{R_2\} \\ \{0\} \end{Bmatrix} \{\phi_1\}^T = \begin{bmatrix} [R] \\ [0] \end{bmatrix} \quad (5\text{-}13)$$

5.3.3 SENSITIVITY OF MODAL PROPERTIES TO LOCALISED NONLINEARITY

In order to make the location more reliable, it is recommended that a mode which is sensitive to the localised nonlinearity should be used in the location process. In order to determine which mode is the most sensitive one in the measurement frequency range (corresponding to specific excitation point), first-order constant-force FRFs can be measured and, as discussed in Chapter 2, the degree of distortion of these measured FRF data around each mode can be used to give an indication of which is the most sensitive to the nonlinearity. Accordingly, the mode sensitivity to localised nonlinearity can be established theoretically. Suppose that a stiffness nonlinearity is introduced between coordinates x_i and x_j and a unit (harmonic) force is applied at x_k, then to a first-order approximation, the maximum relative displacement between x_i and x_j for the r^{th} mode, d_r, can be expressed as

$$d_r = \frac{r\phi_k}{\eta_r \omega_r^2} \left| r\phi_i - r\phi_j \right| \qquad (5\text{-}14)$$

Since nonlinearities of practical structures are usually displacement dependent, d_r can be used to quantify the sensitivity of different modes to localised structural nonlinearity. From (5-14), it can be seen that if the structure is nonlinear, some of the lower modes will appear to be nonlinear due to their particular modeshapes while the higher modes are likely to appear linear.

5.3.4 NUMERICAL CASE STUDIES

A 30DOF mass-spring 'chain' model shown in Fig.5.5 is used in the numerical case study. Cubic stiffness nonlinearity is introduced between x_{15} and x_{16} and 40% analytical stiffness modelling errors are introduced between coordinates x_1-x_3 and x_{28}-x_{30}. When the excitation coordinate is chosen to be x_1, mode 2 is found to be the most sensitive mode (d_1=1.93E-8, d_2=3.37E-5, d_3=4.83E-7 and d_4=3.33E-5 based on (5-14) with 1% (η_r=0.01) proportional damping) to the thus-introduced nonlinearity and therefore is used in the location process. The point receptances of coordinate x_1 with a 25% stiffness change in the nonlinear stiffness element are shown in Fig.5.6, demonstrating the sensitivity of mode 2 to the localised stiffness nonlinearity (clear shift of the natural frequency of mode 2).

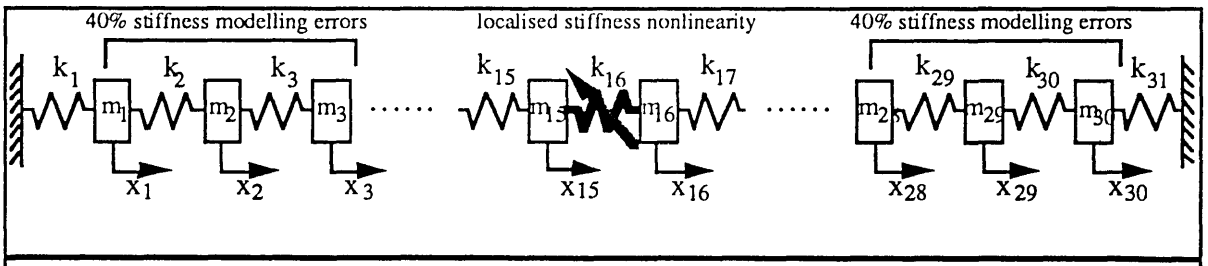


Fig.5.5 - A Mass-spring Model Used in Numerical Case Studies

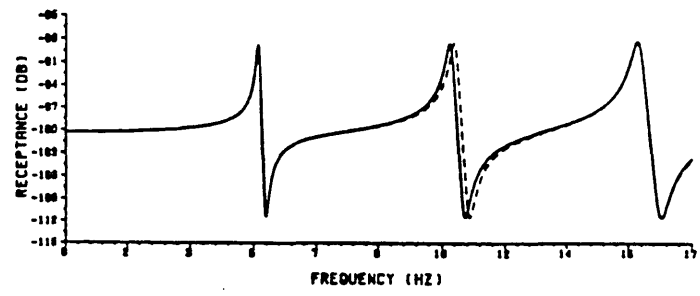


Fig.5.6 - FRFs of $\alpha_{11}(\omega)$ with 25% Stiffness Change of k_{16}

When the measured coordinates are complete, the location results are as shown in Figs.5.7-5.9 (3-dimensional plots showing the absolute values of the elements of a matrix against its two dimensions) using the modal data of the second mode 'measured' at two different response levels (the stiffness change of the nonlinear stiffness element corresponding to the lower and higher response levels is 25% of its original value). Fig.5.7 shows the stiffness modelling errors $[\Delta K]\{\phi_1\}\{\phi_2\}^T$ calculated based on (5-4) and Fig.5.8 shows the stiffness modelling errors $[\Delta K]$ and stiffness change $[\Delta K_n]$ due to nonlinearity $([\Delta K]+[\Delta K_n])\{\phi_2\}\{\phi_1\}^T$ calculated based on (5-3). By taking the difference of these two error matrices, shown in figure 5.7 and Figure 5.8 based on (5-5), the location of the nonlinearity becomes clear, as shown in Fig.5.9.

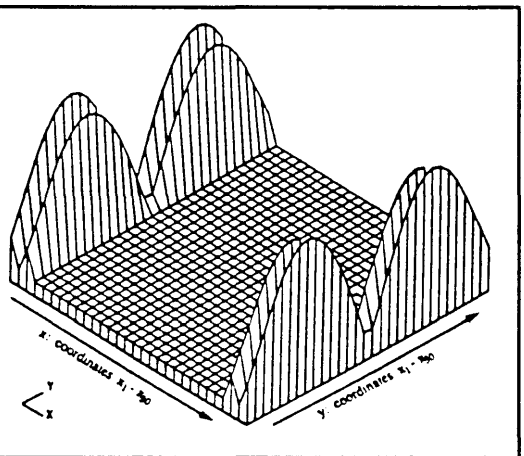


Fig.5.7 - Stiffness Error $[\Delta K]\{\phi_1\}\{\phi_2\}^T$

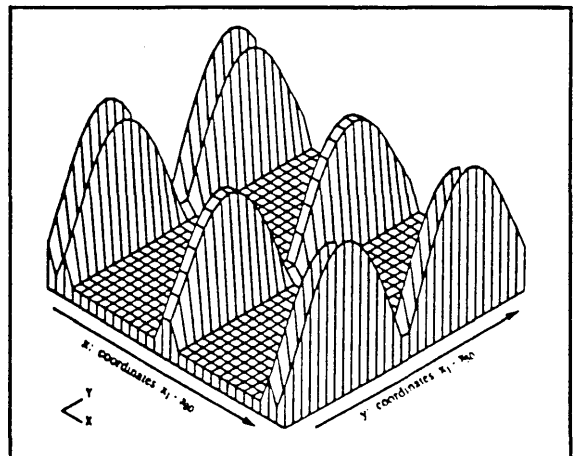


Fig.5.8 - Stiffness Error $([\Delta K]+[\Delta K_n])\{\phi_2\})\{\phi_1\}^T$

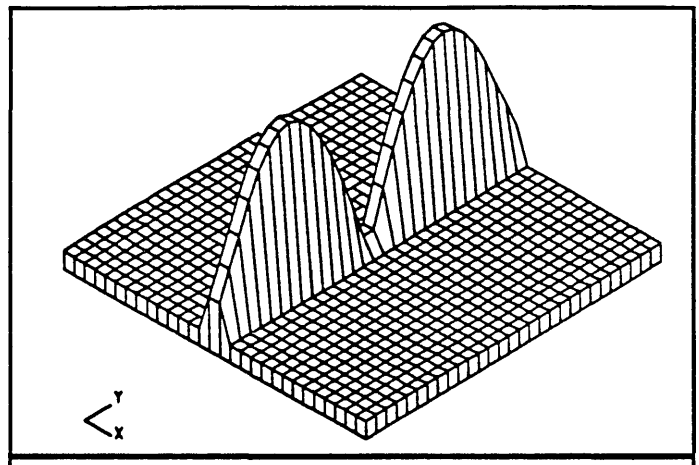


Fig.5.9 - Stiffness Error Due to Nonlinearity $[\Delta K_n]\{\phi_2\}\{\phi_1\}^T$

However, as mentioned, it is unlikely that all the coordinates (which are specified in the analytical model) will be measured in practice. To simulate coordinate incompleteness of practical measurements, only the odd numbered coordinates are included as 'measured' and the unmeasured (even-numbered) coordinates are interpolated based on (5-10) using the analytical model itself. The location results for this case are shown in Figs.5.10-5.12. In this case, since coordinate x_{16} is not measured, the located nonlinearity error is shifted to coordinates x_{15} and x_{17} instead of x_{15} and x_{16}, as shown in Fig.5.12, although the location task is effectively successfully completed.

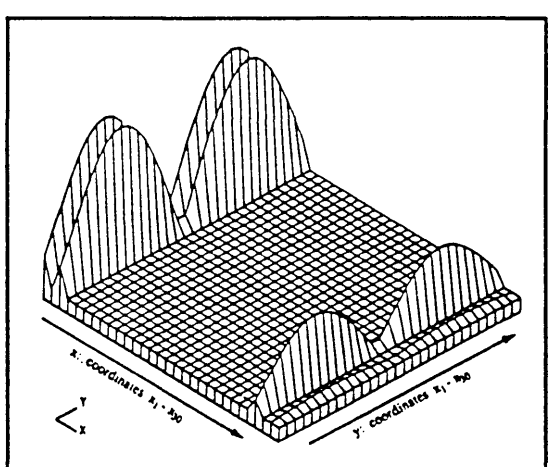


Fig.5.10 - Stiffness Error $[\Delta K]\{\phi_1\}\{\phi_2\}^T$

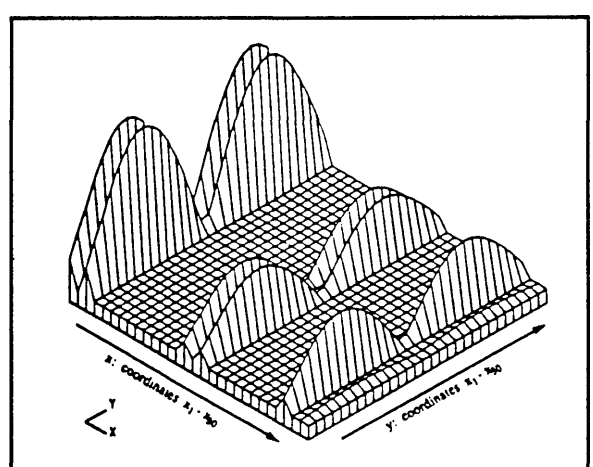


Fig.5.11 - Stiffness Error $([\Delta K]+[\Delta K_n])\{\phi_2\}\{\phi_1\}^T$

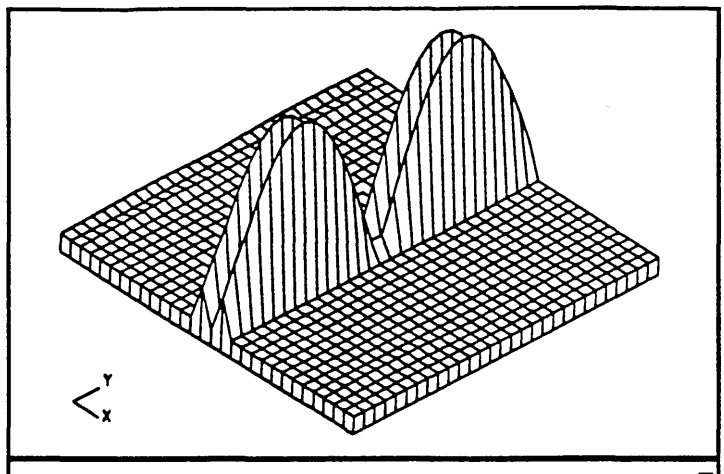


Fig.5.12 - Stiffness Error Due to Nonlinearity $[\Delta K_n]\{\phi_2\}\{\phi_1\}^T$

5.4 EXPERIMENTAL INVESTIGATION

5.4.1 SIMULATION OF STRUCTURAL SYSTEM WITH LOCALISED NONLINEARITY

To demonstrate the practical applicability of this proposed nonlinearity location method, an experimental investigation was carried out. The experimental system is an essentially linear frame structure made of mild steel coupled to an SDOF system with nonlinear stiffness. This nonlinear SDOF system is simulated using an electro-dynamic shaker by feeding the displacement signal of its moving table through a nonlinear analogue circuit and then back to the shaker to produce a force which satisfies a prescribed nonlinear function $F=f(x)$. Fig.5.13 illustrates the setup of the simulated nonlinear structure.

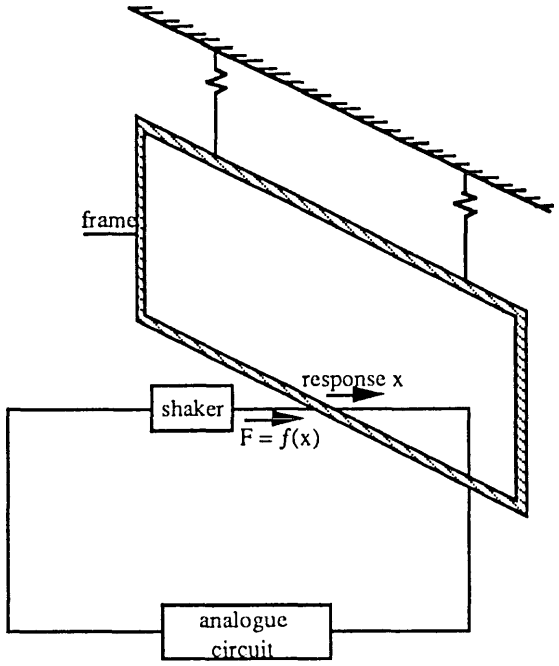


Fig.5.13 - Simulated Nonlinear Structure with Localised Stiffness Nonlinearity

FRAME GEOMETRY

The frame structure is made of mild steel with Young's modulus assumed to be $E=2.1x10^9N/m^2$ and density $\rho=7800kg/m^3$. The geometry of the frame is illustrated in Fig.5.14.

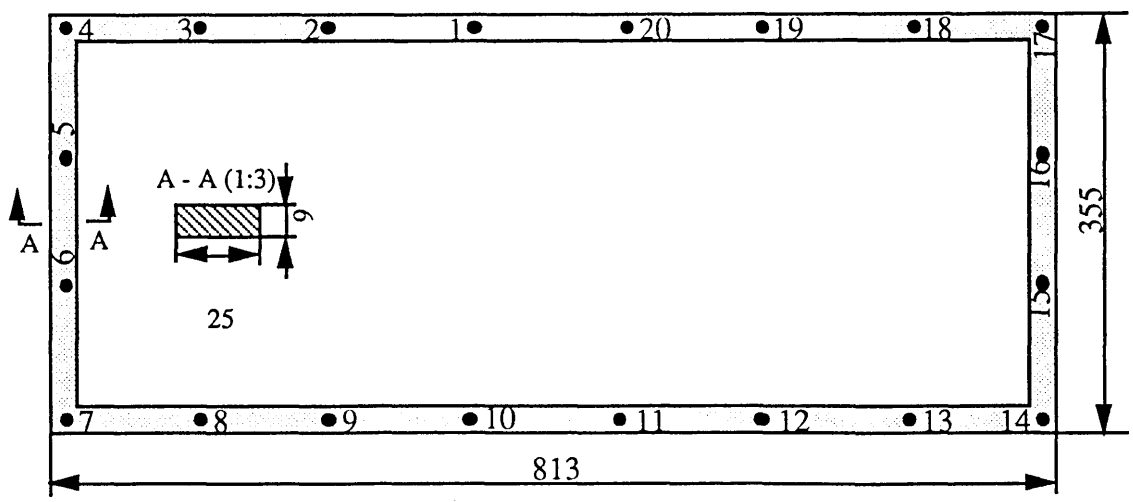


Fig.5.14 - Geometry of the Frame Structure

SHAKER PROPERTIES

A Ling Dynamics shaker (type No.403) was used to simulate the nonlinear SDOF system attached to the frame. The effective mass of the moving system of the shaker is m_s=0.2kg and the spring/mass resonance is f_n=36Hz. The effective stiffness is therefore k_s=10230N/m ($k_s=4\pi^2 f_n^2 m_s$). The SDOF system model of the shaker is shown in Fig.5.15.

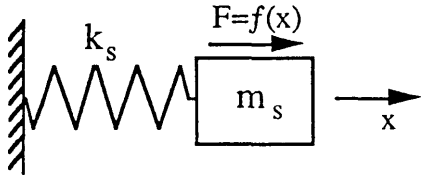


Fig.5.15 - Shaker SDOF Model

According to electro-magnetic dynamics, the force F produced on the coil (figure 5.15) is proportional to the current I in the coil: F=kI. However, due to the motion of the coil in the magnetic field, a back emf is produced which is proportional to the velocity of the coil: $\varepsilon=k\dot{x}$. As a result, the relationship between the force F produced on the coil and the voltage V applied on the coil becomes $F= k(V- \varepsilon)/(R+i\omega L)$ (R, L and ω are coil resistance, coil inductance and excitation frequency, respectively). Clearly, if F=f(x) (a prescribed nonlinear function) is to be satisfied, the effect of this back emf must be compensated so that the force produced is proportional to the voltage applied to the power amplifier. This compensation can be achieved by using a current power amplifier which produces a current output (and so the force F) proportional to voltage input regardless of the loading impedance (which changes dramatically around structural resonances). The current amplifier and shaker can be looked on as a single unit, as shown in Fig.5.16. The output/input (I/V) characteristic of this unit when the frame is attached is measured as shown in Fig.5.17. From figure 5.17, it can be seen that although the effective impedance of the shaker changes dramatically around the resonance, due to the back emf effect, the force produced on the coil is always proportional to the voltage applied to the current power amplifier.

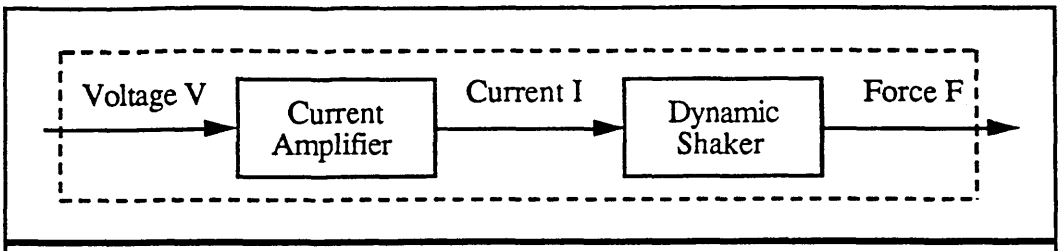


Fig.5.16 - Amplifier Shaker Unit

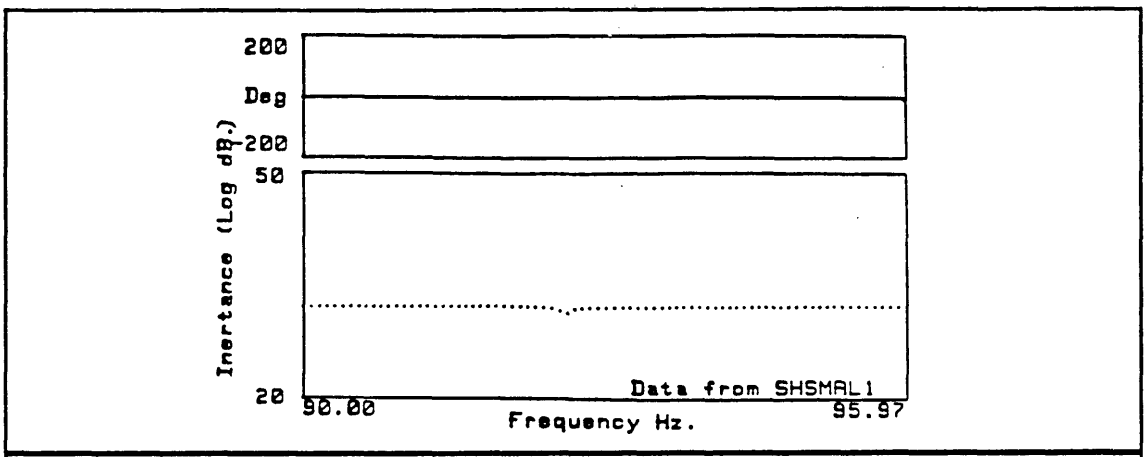


Fig.5.17 - Characteristics of Amplifier-shaker Unit around Structural Resonance

ANALOGUE CIRCUIT

The simulated stiffness nonlinearity is required to be a cubic stiffness described by $f(x)=\beta x^3$ and the acceleration signal is proportional to the displacement signal when sinusoidal input is considered. Therefore, the analogue circuit is required simply to multiply the input signal twice to obtain a βx^3 output as shown in Fig.5.18.

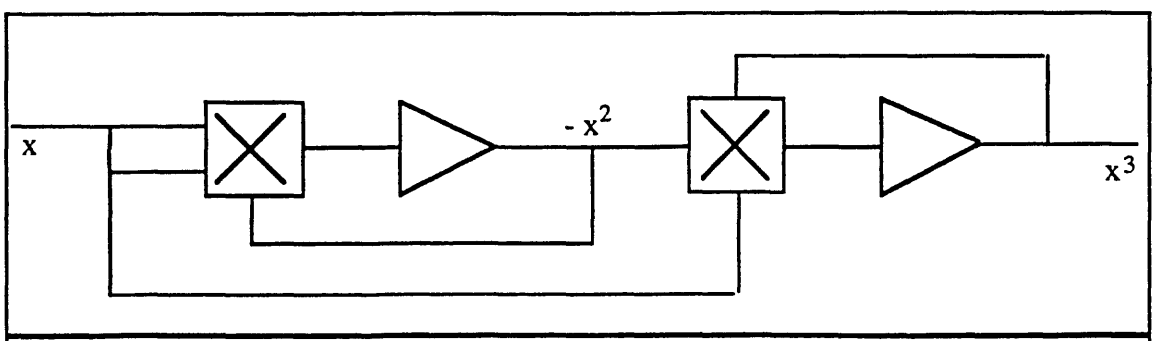


Fig.5.18 - Analogue Computer Circuit

MEASUREMENT SETUP

The measurement setup is illustrated in Fig.5.19. A Solartron 1254 Frequency Analyser was used to obtain the frequency response functions of the structure and an HP 9816 computer was used to store and analyse the measured data.

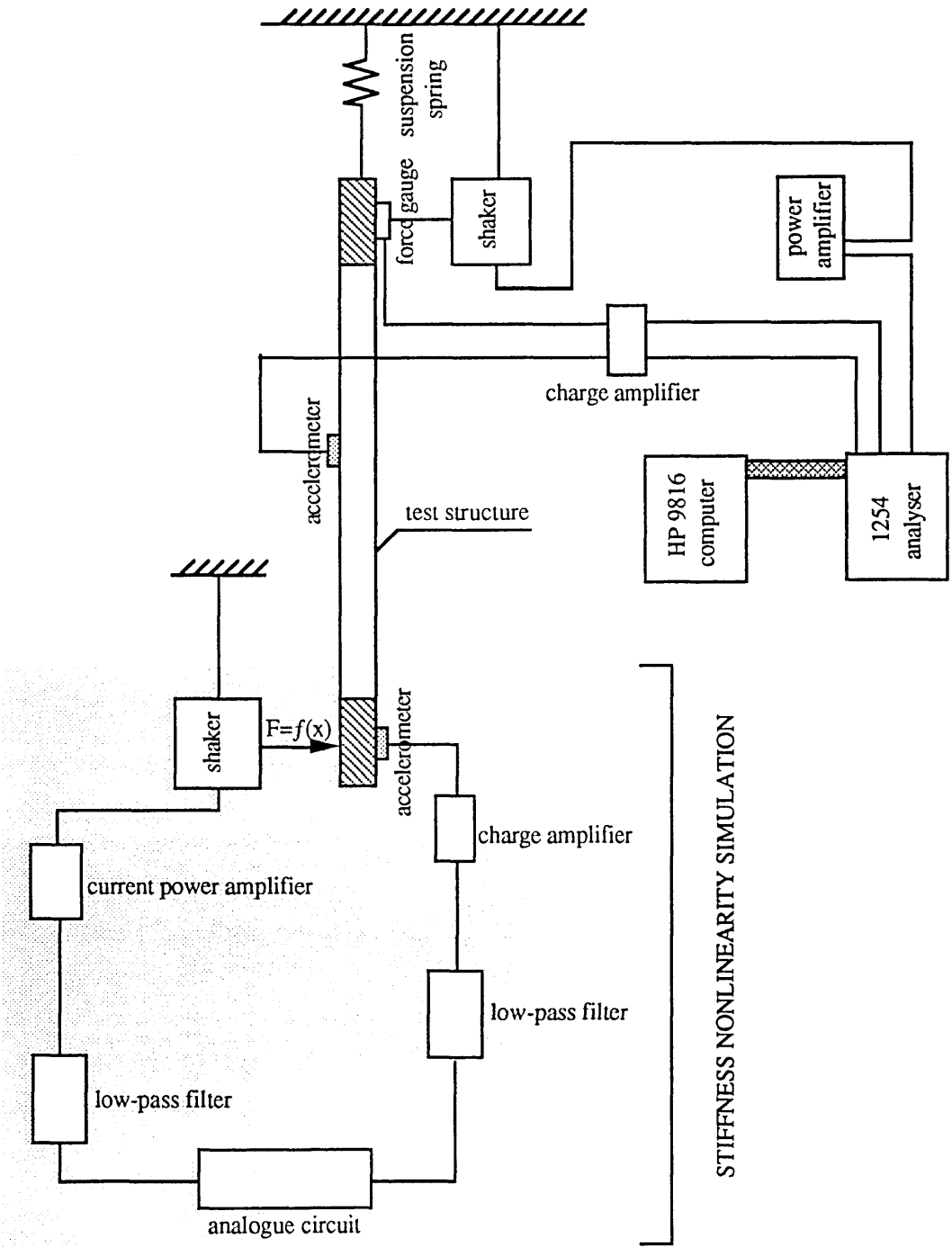


Fig.5.18 - Measurement Setup

5.4.2 MEASUREMENT RESULTS

As a preliminary measurement, the frame was tested without the nonlinear SDOF system (shaker) attached. All 20 translational coordinates around the frame were measured using sinusoidal excitation in the frequency range of 30-300Hz with excitation at coordinate x_1 (figure 5.14). The calibration results of the measurement system and typical frequency response functions measured are shown in Figs.5.19-5.20. In the measurement frequency range of 30-300Hz, 6 modes were clearly identified (the first resonance is in fact two close modes as shown in Fig.5.21). All these modes were analysed and the modal parameters are tabulated in Table 5.1 (in modal analysis, the first two close modes were treated as single mode because the frequency resolution of the measured FRF data is not enough for them to be accurately identified and it is not our purpose to do so).

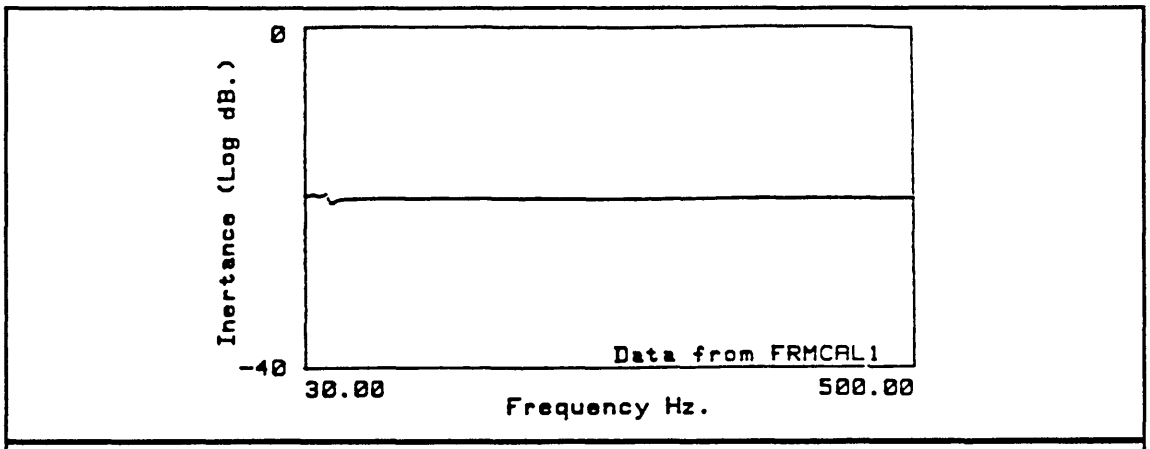


Fig.5.19 - Calibration Results of Measurement

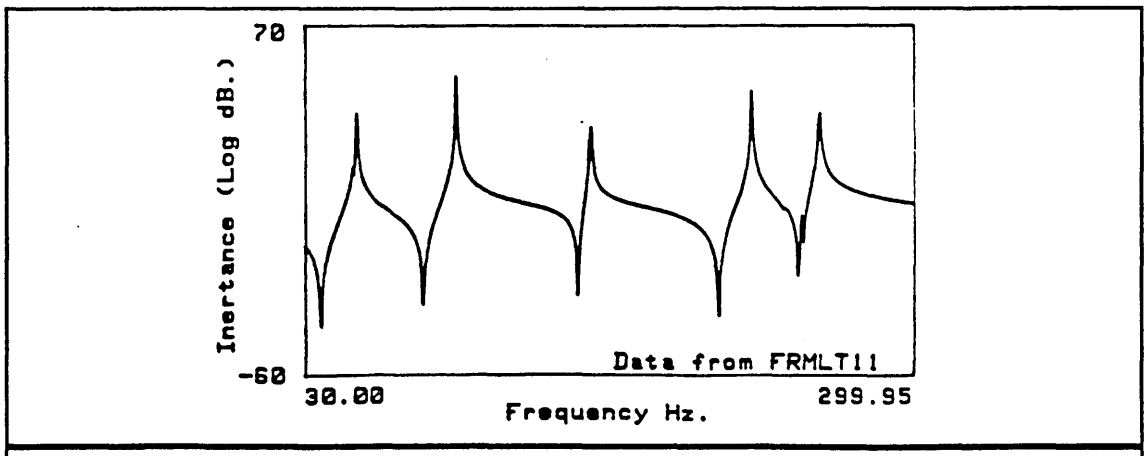


Fig.5.20 - Typical Measured FRF Data

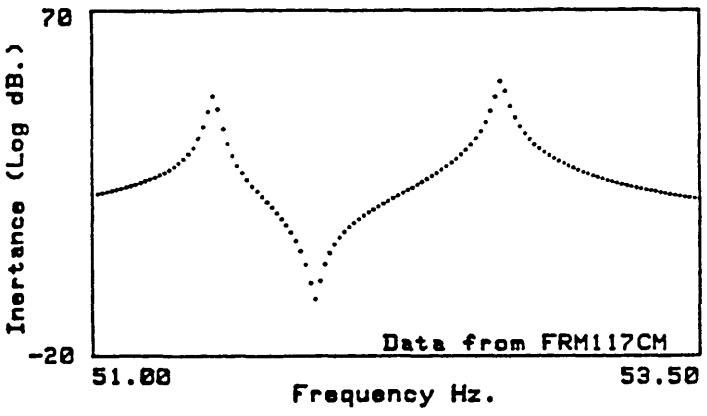


Fig.5.21 - The First 2 Close Modes of the Frame

From figure 5.20, it can be seen that mode 2 ($f_2 \cong 97$ Hz) is the strongest mode in the measurement frequency range and is the most sensitive one to the stiffness nonlinearity introduced between x_{10} and x_{11} because of its particular modeshape (see table 5.1). Accordingly, mode 2 was chosen to study the nonlinear effects of the simulated nonlinear structure. Frequency response functions in the vicinity of mode 2 corresponding to different response levels (linearised) are measured (in fact, FRFs at different constant force levels are also measured and analysed as shown in §2.5.5) at all 20 translational coordinates around the structure and some of these are shown in Fig.5.22. It can be seen that a natural frequency shift of approximately 4Hz is caused by the stiffness change of the system when it vibrates at different response amplitudes. The measured modal data of mode 2 corresponding to two different response levels (the response signals were set to be 0.1v for low vibration level and 2.0v for high vibration level) are analysed and tabulated in Table 5.2. From Table 5.2, it can be seen that in addition to the 4Hz shift in the the natural frequency, approximately 10% changes in the modeshapes are observed due to stiffness change of the system.

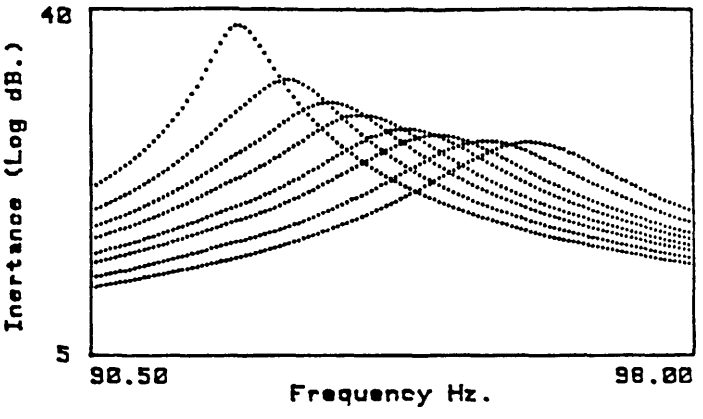


Fig.5.22 - Measured FRFs at Different Response Amplitudes

5.4.3 LOCATION OF NONLINEARITY

A Finite Element analysis of the structure was performed using the PAFEC package and the modal parameters of the first 6 modes of the FE model are shown in Table 5.3. A correlation between the measured modal data and those from the FE model of the frame is carried out and the MAC value matrix is shown in Table.5.4. From Table 5.4, it can be seen that good correlation has been obtained for modes 2, 3 and 4. However, it is surprising that good correlation has also been obtained for the first mode (in fact, 2 close modes) even when it is treated as a single mode in the experimental modal analysis.

MAC matrix	FE mode 1	FE mode 2	FE mode 3	FE mode 4	FE mode 5	FE mode 6
EX mode 1	0.923	0.135	0.000	0.000	0.001	0.048
EX mode 2	0.000	0.004	0.948	0.000	0.001	0.005
EX mode 3	0.000	0.002	0.003	0.985	0.000	0.023
EX mode 4	0.000	0.000	0.006	0.001	0.959	0.005
EX mode 5	0.006	0.001	0.000	0.000	0.004	0.775

Table 5.4 - MAC Value Matrix

The mass and stiffness matrices of the frame were generated using PAFEC. However, before these matrices are used to correlate with the measured modal data to locate the nonlinearity, the shaker characteristics have to be compensated in the FE model. Since what is of interest is mode 2, the compensation of the shaker properties for this mode can be illustrated in Fig.5.23. An effective mass m_e and stiffness k_e, can be calculated based

on equivalent kinetic and potential energy as $m_e=0.605m_s$ and $k_e=0.605k_s$ (m_s and k_s are the table mass and suspension stiffness of the shaker and they are $m_s=0.2kg$ and $k_s=10230N/m$).

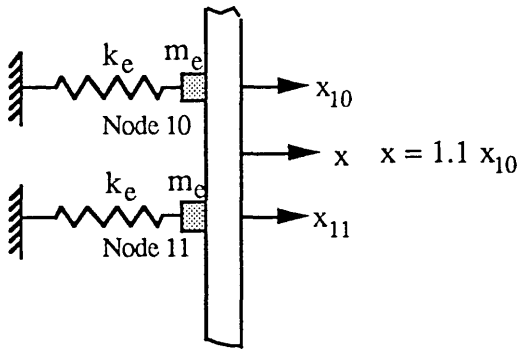


Fig.5.23 - Illustration of Shaker Property Compensation

On the other hand, since in the FE analysis, 3 DOFs (one translational and two rotational - bending and torsion) are considered at each point while in the measurement, only the translational degree of freedom is measured, the measured modeshapes have to be interpolated first using Kidder's method as mentioned earlier before they can be used to correlate with the FE model to locate the structural nonlinearity. The calculated location results are shown in Fig.5.23. Theoretically, the errors should be contained between two translational coordinates (x_{10}-x_{11}) right in the middle of the plot. However, the errors are distributed so that almost a third of the coordinates have been contaminated. One reason for this is that since the exact coordinate where the localised stiffness nonlinearity is introduced (between x_{10} and x_{11}) has neither been measured nor included in the FE model, this missing coordinate is expected to cause spatial leakage. Another reason is that the measured modal data contain measurement errors and these errors may cause this spread of location results. To check this later possibility, 2% random errors were added to the measured modal data and the location results (Fig.5.24) show that although the location results have not been much affected, measurement errors do have the effect of spreading the location results.

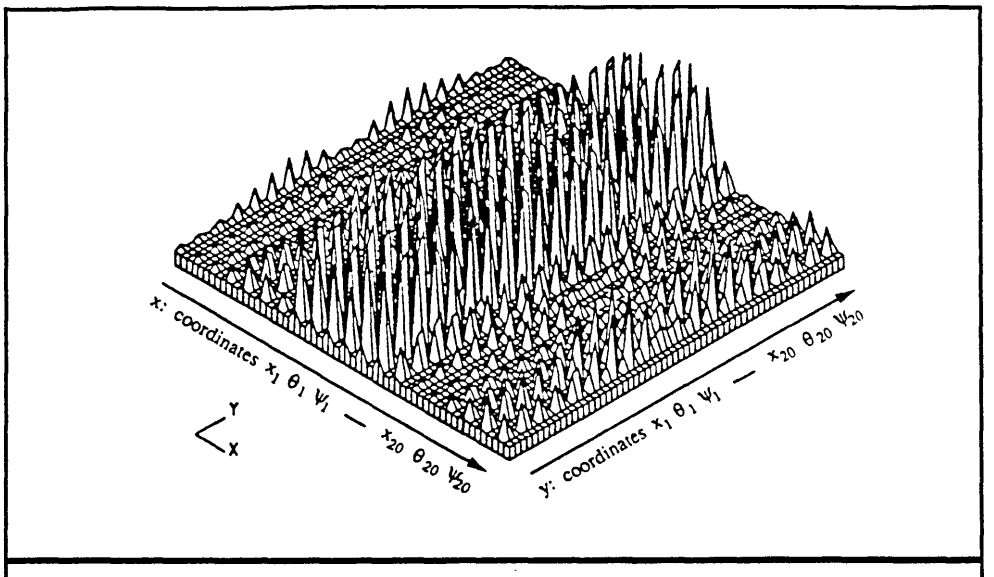


Fig.5.24 - Located Stiffness Nonlinearity

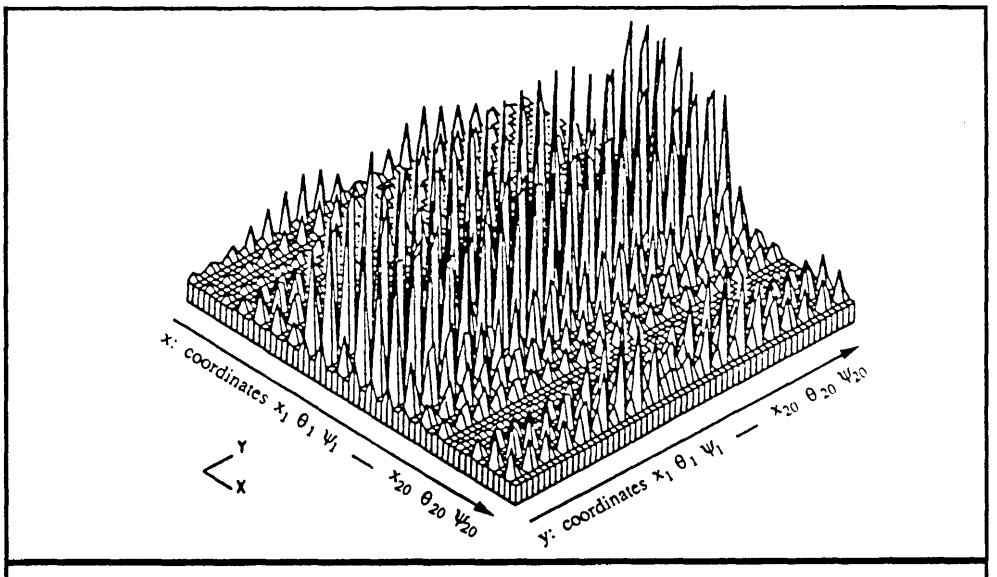


Fig.5.25 - Located Stiffness Nonlinearity (2% Noise on Measured Data)

5.5 EXTENSION OF THE TECHNIQUE TO MEASURED FRF DATA

The above developed location technique can be generalised when measured FRF data are used. Suppose that the i^{th} column $\{\alpha_i(\omega)\}_1$ of the receptance matrix (corresponding to lower response level $\hat{x}_1$) and $\{\alpha_i(\omega)\}_2$ (corresponding to higher response level $\hat{x}_2$) around the r^{th} mode (which is sensitive to nonlinearity) have been measured and, again, the analytical model which contains second order modelling errors (corresponding to

lower response level $\hat{x}_1$) is available. Since the impedance and receptance matrices of a system satisfy

$$[Z(\omega)]\,[\alpha(\omega)] = [\,I\,] \tag{5-15}$$

by taking the i^{th} column of both sides of (5-15), following equations can be established

$$(-\omega_j^2([M_a]+[\Delta M]) + [K_a] + [\Delta K])\{\alpha_i(\omega_j)\}_1 = \{e_i\} \tag{5-16}$$

$$(-\omega_k^2([M_a]+[\Delta M]) + [K_a] + [\Delta K] + [\Delta K_n])\{\alpha_i(\omega_k)\}_2 = \{e_i\} \tag{5-17}$$

where $\{e_i\}$ is a vector with its i^{th} element equal to unity and all the others zero and ω_j and ω_k are the measured FRF data points chosen. Post-multiply (5-17) by $\{\alpha_i(\omega_j)\}_1^T$, then

$$(-\omega_k^2[\Delta M] + [\Delta K] + [\Delta K_n])\{\alpha_i(\omega_k)\}_2\{\alpha_i(\omega_j)\}_1^T =$$

$$-(-\omega_k^2[M_a] + [K_a])\{\alpha_i(\omega_k)\}_2\{\alpha_i(\omega_j)\}_1^T + \{e_i\}\{\alpha_i(\omega_j)\}_1^T \tag{5-18}$$

Post-multiply (5-16) by $\{\alpha_i(\omega_k)\}_2^T$, and we have

$$(-\omega_j^2[\Delta M] + [\Delta K])\{\alpha_i(\omega_j)\}_1\{\alpha_i(\omega_k)\}_2^T =$$

$$-(-\omega_j^2[M_a] + [K_a])\{\alpha_i(\omega_j)\}_1\{\alpha_i(\omega_k)\}_2^T + \{e_i\}\{\alpha_i(\omega_k)\}_2^T \tag{5-19}$$

Subtract (5-18) by (5-19) and rearrange, then

$$[\Delta M]\,(\omega_j^2\{\alpha_i(\omega_j)\}_1\{\alpha_i(\omega_k)\}_2^T - \omega_k^2\{\alpha_i(\omega_k)\}_2\{\alpha_i(\omega_j)\}_1^T) +$$

$$[\Delta K]\,(\{\alpha_i(\omega_k)\}_2\{\alpha_i(\omega_j)\}_1^T - \{\alpha_i(\omega_j)\}_1\{\alpha_i(\omega_k)\}_2^T) + [\Delta K_n]\{\alpha_i(\omega_k)\}_2\{\alpha_i(\omega_j)\}_1^T$$

$$= [M_a]\,(\omega_k^2\{\alpha_i(\omega_k)\}_2\{\alpha_i(\omega_j)\}_1^T - \omega_j^2\{\alpha_i(\omega_j)\}_1\{\alpha_i(\omega_k)\}_2^T) +$$

$$[K_a]\,(\{\alpha_i(\omega_j)\}_1\{\alpha_i(\omega_k)\}_2^T - \{\alpha_i(\omega_k)\}_2\{\alpha_i(\omega_j)\}_1^T) + \{e_i\}(\{\alpha_i(\omega_j)\}_1^T - \{\alpha_i(\omega_k)\}_2^T) \tag{5-20}$$

Unlike the case of location using modal data, here we have the chance to choose ω_j and ω_k properly in the measured frequency range as shown in Fig.5.26 so that following function is minimised

$$\min_{\substack{\omega_j \in \Delta\omega_j \\ \omega_k \in \Delta\omega_k}} \frac{\| \{\alpha_i(\omega_k)\}_2 \{\alpha_i(\omega_j)\}_1^T - \{\alpha_i(\omega_j)\}_1 \{\alpha_i(\omega_k)\}_2^T \|}{\| \{\alpha_i(\omega_k)\}_2 \{\alpha_i(\omega_j)\}_1^T \|} \tag{5-21}$$

If the modelling errors $[\Delta K]$, $\omega_j^2[\Delta M]$ and stiffness change due to nonlinearity $[\Delta K_n]$ are of the same order of magnitude in the sense of the Euclidean norm, then to a first-order approximation, (5-20) becomes

$$[\Delta K_n] \{\alpha_i(\omega_k)\}_2 \{\alpha_i(\omega_j)\}_1^T \approx [M_a] (\omega_k^2 \{\alpha_i(\omega_k)\}_2 \{\alpha_i(\omega_j)\}_1^T - \omega_j^2 \{\alpha_i(\omega_j)\}_1 \{\alpha_i(\omega_k)\}_2^T)$$

$$+ [K_a] (\{\alpha_i(\omega_j)\}_1 \{\alpha_i(\omega_k)\}_2^T - \{\alpha_i(\omega_k)\}_2 \{\alpha_i(\omega_j)\}_1^T) + \{e_i\}(\{\alpha_i(\omega_j)\}_1^T - \{\alpha_i(\omega_k)\}_2^T) \tag{5-22}$$

We state here that when $\omega_j=\omega_{r1}$(the r^{th} natural frequency of the structure corresponding to the lower response level) and $\omega_k=\omega_{r2}$ (the r^{th} natural frequency corresponding to the higher response level), (5-22) will degenerate to (5-6) but we shall leave the mathematical proof of this relationship between (5-18) and (5-6) to Chapter 6. The principle of the nonlinearity location process based on equation (5-22) is the same as that of (5-6).

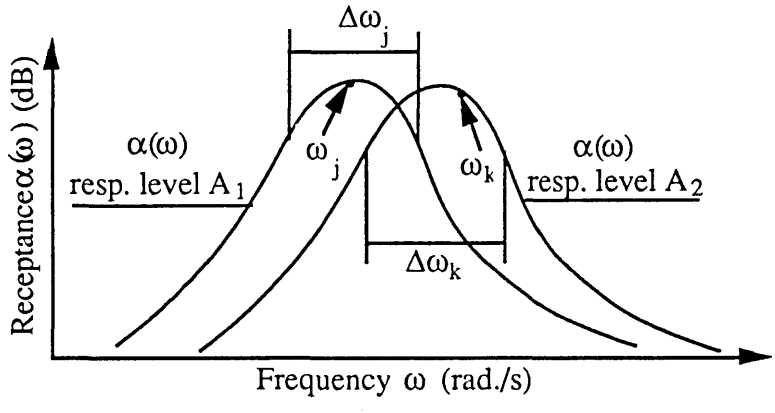


Fig.5.26 - Illustration of the Choice of Frequency Points

When the measured coordinates are incomplete, the receptance terms of the unmeasured coordinates can be interpolated based on following equation which is derived from (5-15) (the i^{th} column of (5-15)):

$$-\omega^2 \begin{bmatrix} [M_{mm}] & [M_{ms}] \\ [M_{sm}] & [M_{ss}] \end{bmatrix} \begin{Bmatrix} \{\alpha_m(\omega)\} \\ \{\alpha_s(\omega)\} \end{Bmatrix} + \begin{bmatrix} [K_{mm}] & [K_{ms}] \\ [K_{sm}] & [K_{ss}] \end{bmatrix} \begin{Bmatrix} \{\alpha_m(\omega)\} \\ \{\alpha_s(\omega)\} \end{Bmatrix} = \begin{Bmatrix} \{e_i\} \\ \{0\} \end{Bmatrix} \tag{5-23}$$

where $\{\alpha_m(\omega)\}$ and $\{\alpha_s(\omega)\}$ are the measured and unmeasured receptances of the i^{th} column of the receptance matrix. Upon multiplying out (5-23), $\{\alpha_s(\omega)\}$ can be calculated as

$$\{\alpha_s(\omega)\} = (-\omega^2 [M_{ss}] + [K_{ss}])^{-1} (\omega^2 [M_{sm}] - [K_{sm}]) \{\alpha_m(\omega)\} \quad (5\text{-}24)$$

5.6 LOCATION OF NONLINEARITY AND MODELLING ERRORS IN ANALYTICAL MODEL IMPROVEMENT

So far, the nonlinearity location technique has been developed and has been verified numerically as well as experimentally and it will be advantageous here to examine the relationship between the two applications of (i) nonlinearity location and (ii) modelling error location of a linear structure in the analytical model updating practice. The discussion is made in this chapter for the location of modelling errors and will be referred to from time to time in later chapters because of the mathematical similarity of nonlinearity location and modelling error location. In modelling error location, $[\Delta K_n] = [O]$ and what needs to be located are the dominant modelling errors in $[\Delta M]$ and $[\Delta K]$ (assume the modelling errors are localised as they usually are in practice).

Following the same reasoning as that used for nonlinearity location, it can be seen that the localised modelling errors in $[\Delta M]$ and $[\Delta K]$ can be located based on (5-4). Since in the modelling error location case, $[\Delta K_n]=[O]$, therefore $\{\phi_2\}=\{\phi_1\}= \{\phi\}_r$ and (5-4) becomes

$$(\lambda_r [\Delta M] + [\Delta K]) \{\phi\}_r \{\phi\}_r^T = - (\lambda_r [M_a] + [K_a]) \{\phi\}_r \{\phi\}_r^T \quad (5\text{-}25)$$

The modelling error location technique based on (5-25) has been reported in [74] and we will show here that it can be generalised when measured FRF data are used.

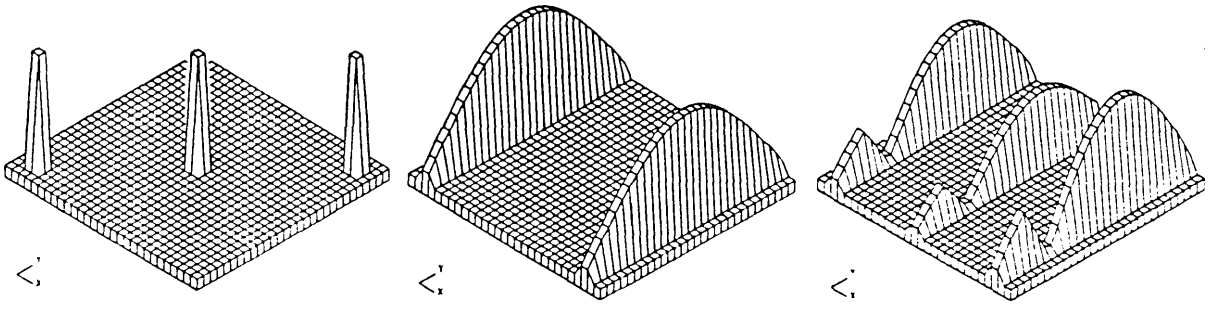
Using measured FRF data, let $\{\alpha_i(\omega_j)\}_1=\{\alpha_i(\omega_k)\}_2=\{\alpha_i(\omega)\}$, so that (5-19) becomes

$$(-\omega^2 [\Delta M] + [\Delta K]) \{\alpha_i(\omega)\} \{\alpha_i(\omega)\}^T =$$

$$- (-\omega^2 [M_a] + [K_a]) \{\alpha_i(\omega)\} \{\alpha_i(\omega)\}^T + \{e_i\} \{\alpha_i(\omega)\}^T \quad (5\text{-}26)$$

Again, it can be shown that when $\omega \rightarrow \omega_r$, (5-26) will degenerate to (5-25). The location technique given in (5-26) not only generalises the technique developed in [74], but because every measured FRF data point theoretically contains a contribution of all the

modes of the structure, the location of modelling errors becomes more consistent. This is illustrated in Fig.5.27 based on the same mass-spring system used in the numerical case studies of nonlinearity location. Fig.5.27(a) shows the exact stiffness modelling error $[\Delta K]$, Fig.5.27(b) shows the error location results $[\Delta K]\{\phi\}_1\{\phi\}_1^T$ based on (5-25) and Fig.5.27(c) shows the error location results $[\Delta K]\{\alpha_i(\omega)\}\{\alpha_i(\omega)\}^T$ based on (5-26). Due to the specific modeshape of the first mode, the stiffness error introduced between coordinates x_{15}-x_{16} has been totally missed out as shown in Fig.5.27(b) while it is located as shown in Fig.5.27(c) when FRF data are used.



exact stiffness errors (a) location using mode 1 (modal data) (b) location using FRF data round mode 1 (c)

Fig.5.27 - Comparison of Modelling Error Location Using Modal and FRF Data

5.7 CONCLUSIONS

Most mechanical structures are nonlinear to some extent and the nonlinearites are usually localised. The ability to locate a structure's localised nonlinearity has some important engineering applications. In this chapter, a nonlinearity location technique has been developed based on the correlation between an analytical model of the structure (which contains modelling errors) and modal test data which are measured at different response levels. In the practical case where the measured coordinates are incomplete, an interpolation technique to estimate the unmeasured coordinates based on the analytical model has been discussed. The sensitivity of certain modes to localised structural nonlinearity has been established. It is recommended that a sensitive mode should always be used in the location process so that reliable location can be obtained. Numerical case studies have been undertaken to verify the technique developed.

To assess the practical feasibility of the location technique, an experiment was carried out. A localised stiffness nonlinearity was simulated using a electro-dynamic shaker and

analogue computer circuit based on feed-back system control theory. The experimental results demonstrate the practical applicability of the proposed method.

It has been shown that the location technique can be generalised when measured FRF data are employed. Also, the relationship between nonlinearity location and modelling error location in analytical model improvement has been examined and it can be shown that the modelling error location method given in [74] can be generalised when FRF data are used.

The information concerning the location of structural nonlinearities on practical structures can be used subsequently in nonlinear substructuring analysis, system failure and malfunctioning detection and mathematical modelling of nonlinear structures.

Mode No.		1	2	3	4	5
Nat. Freq. (Hz.)		52.80	97.10	156.59	228.58	258.55
mode shapes	x_1	0.657	0.674	0.345	0.408	0.415
	x_2	0.490	0.291	0.692	0.682	-0.162
	x_3	0.152	-0.274	0.547	0.150	-0.605
	x_4	-0.336	-0.703	-0.134	-0.386	-0.312
	x_5	-0.502	-0.245	-0.467	-0.229	0.650
	x_6	-0.667	0.241	-0.579	0.232	0.569
	x_7	-0.776	0.760	-0.277	0.553	-0.500
	x_8	-0.172	0.274	0.448	-0.190	-0.595
	x_9	0.280	-0.303	0.679	-0.779	-0.134
	x_{10}	0.604	-0.693	0.398	-0.345	0.465
	x_{11}	0.703	-0.712	-0.303	0.436	0.408
	x_{12}	0.543	-0.303	-0.570	0.829	-0.127
	x_{13}	0.150	0.279	-0.426	0.196	-0.513
	x_{14}	-0.329	0.708	0.271	-0.700	-0.385
	x_{15}	-0.512	0.264	0.523	-0.276	0.696
	x_{16}	-0.691	-0.238	0.474	0.235	0.586
	x_{17}	-0.763	-0.742	0.278	0.490	-0.565
	x_{18}	-0.190	-0.291	-0.436	-0.141	-0.479
	x_{19}	0.249	0.277	-0.796	-0.882	-0.199
	x_{20}	0.606	0.682	-0.340	-0.476	0.357

Table 5.1 Measured Modal Parameters of the Frame

Resp. Level		Low resp. level (0.1v)	High resp. level (2.0v)
Nat. Freq. (Hz.)		92.05	96.03
mode shapes	x_1	0.715	0.630
	x_2	0.325	0.276
	x_3	-0.266	-0.249
	x_4	-0.776	-0.688
	x_5	-0.283	-0.229
	x_6	0.211	0.227
	x_7	0.725	0.691
	x_8	0.345	0.271
	x_9	-0.169	-0.251
	x_{10}	-0.523	-0.605
	x_{11}	-0.520	-0.602
	x_{12}	-0.167	-0.247
	x_{13}	0.338	0.265
	x_{14}	0.739	0.705
	x_{15}	0.211	0.226
	x_{16}	-0.285	-0.230
	x_{17}	-0.772	-0.684
	x_{18}	-0.269	-0.253
	x_{19}	0.303	0.258
	x_{20}	0.721	0.635

Table 5.2 Measured Modal Parameters of Mode 2 at Different Response Amplitudes

Mode No.		1	2	3	4	5	6
Nat. Freq. (Hz)		50.36	51.41	94.85	152.40	222.88	255.55
mode shapes	x_1	0.1624	0.6826	0.7189	0.3553	0.4250	0.4528
	x_2	0.4543	0.4306	0.3168	0.7605	0.7680	-0.1704
	x_3	0.6611	-0.0144	-0.2832	0.4958	0.1755	-0.6282
	x_4	0.7697	-0.5566	-0.7878	-0.2796	-0.7259	-0.4528
	x_5	0.2756	-0.6030	-0.2690	-0.5225	-0.2913	0.7256
	x_6	-0.2756	-0.6030	0.2690	-0.5225	0.2913	0.7256
	x_7	-0.7697	-0.5566	0.7878	-0.2796	0.7259	-0.4528
	x_8	-0.6611	-0.0144	0.2832	0.4958	-0.1755	-0.6282
	x_9	-0.4543	0.4306	-0.3168	0.7605	-0.7680	-0.1704
	x_{10}	-0.1624	0.6826	-0.7189	0.3553	-0.4250	0.4528
	x_{11}	0.1624	0.6826	-0.7189	-0.3553	0.4250	0.4528
	x_{12}	0.4543	0.4306	-0.3168	-0.7605	0.7680	-0.1704
	x_{13}	0.6611	-0.0144	0.2832	-0.4958	0.1755	-0.6282
	x_{14}	0.7697	-0.5566	0.7878	0.2796	-0.7259	-0.4528
	x_{15}	0.2756	-0.6030	0.2690	0.5225	-0.2913	0.7256
	x_{16}	-0.2756	-0.6030	-0.2690	0.5225	0.2913	0.7256
	x_{17}	-0.7697	-0.5566	-0.7878	0.2796	0.7259	-0.4528
	x_{18}	-0.6611	-0.0144	-0.2832	-0.4958	-0.1755	-0.6282
	x_{19}	0.4543	0.4306	0.3168	-0.7605	-0.7680	-0.1704
	x_{20}	-0.1624	0.6826	0.7189	-0.3553	-0.4250	0.4528

Table 5.3 First 6 Modes of the FE Model

CHAPTER 6

IDENTIFICATION OF MATHEMATICAL MODELS OF DYNAMIC STRUCTURES

6.1 PRELIMINARIES

As with the identification of the dynamic characteristics of a linear structure, the ultimate target involved in the identification of a nonlinear structure is obviously to establish its nonlinear mathematical model which is a function of vibration amplitude ([M], [K($\hat{x}$)] and [C($\hat{x}$)] or [D($\hat{x}$)]). As will be shown, the establishment of such a nonlinear mathematical model becomes possible only when, on the one hand, an accurate linear mathematical model (corresponding to very low vibration amplitude and therefore, can be regarded as linear model of the nonlinear structure) is available and, on the other, the location information of the localised nonlinearity is given. Such requirements in the modelling of a nonlinear structure, which are different from those for the modelling of a linear structure, are due to the fact that the mathematical model of a nonlinear structure has to be established on a mode by mode basis as will be explained later in this chapter. Since an accurate linear model is an essential pre-requisite for the modelling of a nonlinear structure, most of the effort will be devoted in this chapter to discussions of how an accurate linear model can be be established.

As discussed in Chapter 5, techniques for locating the nonlinearities (which are usually localised) of practical structures have been developed and these techniques are further employed in the modelling of nonlinear structures in this Chapter.

On the other hand, for the modelling of linear structures (in fact, as will be shown, the problem of modelling a nonlinear structure is essentially the same as that for a linear structure except that in the former case, a series of linearised models need to be established), a new method is presented in this chapter which tackles the modelling problem by using the measured frequency response data directly. The method is then extended to the modelling of nonlinear structures by combining the updated linear model (corresponding to very low vibration amplitude) and the nonlinearity location results.

6.2 MODELLING OF LINEAR AND NONLINEAR STRUCTURES

Mathematical models of practical continuous structures, both linear and nonlinear, play an important role in dynamic analysis. They are frequently used in response and load prediction, modification/sensitivity and stability analysis, structural coupling etc.. Due to the development of mathematical and physical sciences, the closed-form modelling of some basic mechanical components such as uniform beams and plates has now become possible. However, for complicated practical structures, there do not in general exist analytical solutions (or if they do, it is extremely difficult to find such analytical solutions) and therefore some discrete approximate models which can well represent the structure under given conditions have to be sought so that the dynamic characteristics of the structure can be analysed mathematically. According to Berman [75], such a discrete model can be considered as a good model if it will not only predict responses over the frequency range of interest, but will also be representative of the physical characteristics of the structure. Thus it must have the capability to predict the effects of changes in physical parameters and to represent correctly the structure when it is treated as a component of a large system. It is the establishment of such a physical model of the structure which is discussed in this chapter.

Basically, there are two ways of establishing a discrete mathematical model of a practical structure and they are (i) experimental modelling and (ii) theoretical modelling (Finite Element Modelling). Due to the advances made in measurement instrumentation and techniques, nowadays it is usually agreed that measured data should be considered as the true representation of the structure while the FE model - because of the idealisation involved, lack of knowledge about the structure and difficulty of modelling of boundary conditions - is usually considered to be inaccurate and therefore should be updated using measured data if possible. Based on the assumption that the measured data are correct, two approaches to modelling exist: (i) using experimental data only to establish a mathematical model in terms of measured coordinates of interest and (ii) correlating the

FE model and dynamic testing data to update the FE model. Both modelling activities will be reviewed and further possible development will be discussed.

As for the mathematical modelling of nonlinear structures, correlation between analytical model and dynamic testing data becomes essential because, in general, structural nonlinearities cannot be foreseen and, therefore, cannot be analytically modelled but can be measured. It is believed that the modelling of a nonlinear structure becomes possible only when an accurate linear mathematical model of the nonlinear structure is available. In addition to the availability of an accurate linear model, location information on the structural nonlinearity is also essential in most cases so that the number of unknowns involved in the modelling process can be reduced because, unlike the modelling of a linear structure, in which the model to be sought is unique and so all the data measured are consistent and can be used at the same time, the mathematical model of a nonlinear structure has to be established based on a mode by mode basis so that only one measured mode is available each time as illustrated below.

Consider the system shown in Fig.6.1. In measurement, in order to linearise the structure, the response amplitude of a chosen point (x_1, for example) is kept constant. Then, when the excitation frequency $\omega \approx \omega_i$ (around the i^{th} mode), the displacements of all coordinates are determined by the i^{th} modeshape and therefore, the relative vibration amplitude of the nonlinear stiffness, which determines the value of equivalent linearised stiffness, can be considered to be proportional to the vibration amplitude of x_1. This is especially true when the mode to be analysed is well separated from its neighbours. However, when the excitation frequency is around the j^{th} mode, $\omega \approx \omega_j$, for the same vibration amplitude of x_1 (assuming this to be possible), the relative vibration amplitude of the nonlinear stiffness will also be proportional to the amplitude of x_1, but will be different in magnitude from that of the i^{th} mode because these two modeshapes are different. Therefore, even when the response amplitude of a certain coordinate (x_1 in this case) is constant, the data measured around different modes could be the data from different linearised systems and this means that only one measured mode can be used each time in the modelling of a nonlinear structure due to the inconsistency of measured data.

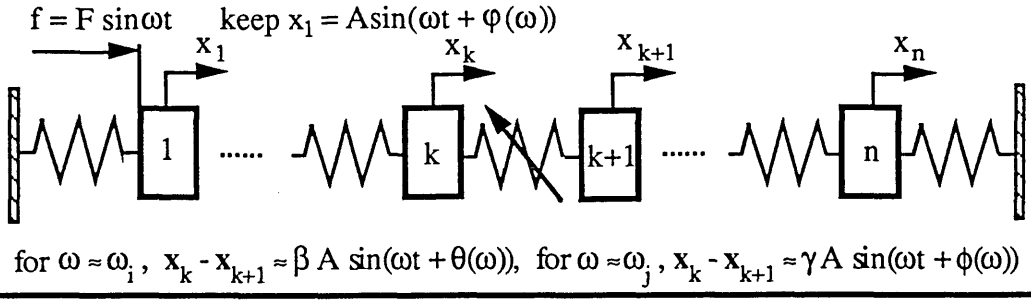


Fig.6.1 - Illustration of Inconsistency of Measured Data of a Nonlinear Structure

Although the inevitable limitations of measured data could present a problem, the greatest difficulty involved in the modelling of a nonlinear structure, as will be shown, is the establishment of an accurate linear model. When such a linear model is available, by using the measured frequency response data together with location results of the nonlinearity, the modelling problem can be resolved in most cases. For this reason, most of the space in this Chapter will be devoted to a discussion of how to obtain an accurate linear model of a structure by correlating the analytical model and measured dynamic test data.

When the identification of structural dynamic characteristics is undertaken by both theoretical analysis (normally FE analysis) and experimental modal testing, discrepancies often exist between the vibration characteristics predicted by the theoretical model and those identified experimentally. In such cases, the analytical model needs to be modified, if practically necessary, so that it represents more accurately the dynamic characteristics of the actual structure. When both the analytical model of a structure and experimental modal testing data are available, analytical model improvement can be mathematically formulated as described below.

Given the analytical mass matrix $[M_a]$ and stiffness matrix $[K_a]$ (containing modelling errors $[\Delta M]$ and $[\Delta K]$) and dynamic testing data, which can be either modal data $[\cdot\omega_r^2.]$ and $[\phi]$ or frequency response data $[\alpha_x(\omega)]$ (but in both cases incomplete in terms of measured modes and/or coordinates), then by correlating the analytical model and measured data, the former can be updated so that it better represents the dynamic characteristics of the actual structure.

6.3 REVIEW OF ANALYTICAL MODEL IMPROVEMENT METHODS

As mentioned earlier, there exist two different branches of modelling activity, these being modelling using experimental data only and modelling based on the correlation between an analytical model and experimental data. For mathematical modelling via direct use of

measured modal data, there have been attempts in recent years to use results of dynamic testing to identify the parameters in the equations of motion directly, but it is generally believed that unless the number of measured modes is greater than or equal to the number of coordinates of interest, a mathematically unique model cannot be obtained [76-79]. A survey of the work on the mathematical modelling using measured modal data only was conducted in 1969 by Young and On [76]. In 1971, Berman [77], assumed a simple form for the mass matrix and then used the limited measured modes to construct a mass matrix by invoking the orthogonality equations. This mass matrix is then used together with the measured modes to construct the so-called 'incomplete' stiffness matrix: 'incomplete' because the contribution of all the unmeasured higher modes (which is usually the major part) to the stiffness matrix has not been included. Link [78] used a multi-point excitation technique (more than one excitation forces) to establish the mass, incomplete stiffness and incomplete damping matrices using measured force and response data. Luk [79] employed the incomplete modal data set and used the pseudoinverse to calculate system matrices which are minimum 2-norm least-squares solution in a mathematical sense. The possibility of establishing a complete mathematical model when the measured data contain less modes than the number of coordinates remain to be investigated.

Alongside these studies, a number of methods have been published in the literature to deal with analytical model improvement by correlating FE models and measured data. The philosophy behind this practice is that the analytical model, while containing modelling errors, is assumed to represent the structure with some accuracy so that the limited measured test data available will offer the possibility of updating it (otherwise the modelling problem will become the same as that of using experimental data only). Based on an inverse first-order sensitivity analysis, and considering the random nature of measurement errors, Collins et al [80] employed an iterative procedure to adjust their analytical model so that the difference between the measured and analytical modal data is minimised in terms of the Euclidean norm. Later, Chen/Garba [81] modified this procedure by introducing matrix perturbation concepts to avoid the need for an eigensolution at every iteration (which is required in the formulation in [80]). Lallement [82] extended the method in [81] to pinpoint first where the significant errors are located and then to reduce the number of unknowns to improve the solution condition. On the other hand, based on the assumption that the mass matrix is correct, Baruch/Itzhack [83-84] introduced a kind of objective function together with an orthogonality property so that the analytical modes are optimised in such a way that they are closest to the measured ones in a weighted Euclidean sense. These optimised analytical modes are then used to derive the updated stiffness and flexibility matrices. Berman later extended this theory to the case of mass matrix updating [85]. Having recognised the mathematical difficulty of

whole system matrix updating, simple eigendynamic equations are used in [86] to locate the major modelling errors first and then to employ the limited measured modes to turn the updating problem into an overdetermined one. All these above-mentioned activities are based on the correlation between an analytical model and measured modal data and the completeness of measured coordinates is, in most cases, critical.

Recently, there have also been publications on the identification of mass, stiffness and damping matrices in terms of measured coordinates of a system from measured FRF data. Fritzen used an Instrumental Variable method to identify system matrices based on the measured force and response data [87]. Mottershead/Stanway extended the time domain invariant imbedding filter to the frequency domain to estimate system parameters [88]. Foster/Mottershead later extended the method in [88] to allow correction of reduced-order finite element model (Guyan-reduced and so fully-populated) by minimising the difference between the analytical and identified models [89].

In the following section, a new advantageous model updating method is developed which tackles the problem by using the measured frequency response function data directly. The new method is then extended to the case where the structure to be modelled is nonlinear. The advantages of using FRF data over modal data to update an analytical model are demonstrated. It is shown that model updating methods based on modal data are, in a broad sense, discrete versions of the present generalised method where only FRF data at resonance frequencies are employed. Based on this method, the uniqueness of the updating problem is discussed in some mathematical rigour. Special attention is given to the application of the method to the case where both measured modes and coordinates are incomplete. The practical applicability of the method is assessed based on the GARTEUR exercise which is intended to represent practical problems in terms of the incompleteness of both measured modes and coordinates.

6.4 MODEL UPDATING USING FRF DATA - A NEW GENERALISED METHOD

6.4.1 DESCRIPTION OF THE METHOD

The development of the model updating method described here is based on the following mathematical identity:

$$[[A] + [B]]^{-1} = [A]^{-1} - [[A] + [B]]^{-1} [B] [A]^{-1} \qquad (6\text{-}1)$$

where [A] and [B] are two complex matrices satisfying the condition that both [A] and ([A]+[B]) are nonsingular. To check the validity of equation (6-1), premultiply both sides of (6-1) by ([A]+[B]), since it is nonsingular, so that (6-1) becomes:

$$[\,I\,] = [\,I\,] + [B]\,[A]^{-1} - [B]\,[A]^{-1} = [\,I\,] \tag{6-2}$$

Now, if we assume that [A] is the impedance matrix of the analytical system $[Z_a(\omega)]$ and that ([A]+[B]) is the impedance matrix of the experimental system $[Z_x(\omega)]$, then equation (6-1) becomes:

$$[Z_x(\omega)]^{-1} = [Z_a(\omega)]^{-1} - [Z_x(\omega)]^{-1}\,([Z_x(\omega)] - [Z_a(\omega)])\,[Z_a(\omega)]^{-1} \tag{6-3}$$

Rewriting (6-3) in its more familiar receptance and impedance error form, we have

$$[\Delta\alpha(\omega)] = [\alpha_x(\omega)] - [\alpha_a(\omega)] = -\,[\alpha_x(\omega)]\,[\Delta Z(\omega)]\,[\alpha_a(\omega)] \tag{6-4}$$

where $[\Delta Z(\omega)]$ is the impedance error matrix defined as $[\Delta Z(\omega)]=[Z_x(\omega)]-[Z_a(\omega)]$. In what follows, it will be shown how equation (6-4) can be used to solve the updating problem uniquely when one complete column of the receptance matrix is measured and how it can be extended to cases where the measured coordinates are incomplete.

In the case where one complete column (the i^{th}) has been measured, then (6-4) can be rewritten in terms of measured and analytical receptance terms as

$$\{\alpha_a(\omega)\}_i^T - \{\alpha_x(\omega)\}_i^T = \{\alpha_x(\omega)\}_i^T\,[\Delta Z(\omega)]\,[\alpha_a(\omega)] \tag{6-5}$$

this formula being obtained by simply taking the i^{th} row of both sides of equation (6-4). In analytical model updating, the physical connectivity of the analytical model should usually be respected and therefore the updated model should have the same physical connectivity as that of the analytical model, i.e. the modelling errors can only occur where the elements of mass or stiffness matrix are nonzero. Further discussions on the necessity and mathematical validity of preserving the physical connectivity of the analytical model will be given in Chapter 7. Also, in general, the physical connectivity of the damping matrix can be assumed to be the same as that of the stiffness. Upon substitution of $[\Delta Z(\omega)]$ in (6-5) in terms of [ΔM], [ΔK] and [D] and transpose, equation (6-5) becomes:

$$\{\alpha_a(\omega)\}_i - \{\alpha_x(\omega)\}_i = [\alpha_a(\omega)]\,(-\omega^2[\Delta M] + [\Delta K] + \mathbf{i}[D])\,\{\alpha_x(\omega)\}_i \tag{6-6}$$

Consider the physical connectivity of the analytical model and let the design variable changes in the mass matrix be Δm_r (r=1, L_1 where L_1 is the total number of independent design variables in mass matrix which could be individual nonzero elements, as in the case of mass spring systems, or coefficients of element mass matrices as in finite element models), in the stiffness matrix be Δk_s (s=1, L_2) and in the damping matrix to be d_s (s=1, L_2). Then, it can be shown that every element $b_j(\omega)$ (j=1,N where N is the number of total degrees of freedom specified in the analytical model) of the RHS of (6-6), which is a vector, can be expressed as a linear combination of the changes of all design variables as:

$$b_j(\omega) = \sum_{r=1}^{L_1} c_m^r(\omega)\,\Delta m_r + \sum_{s=1}^{L_2} c_k^s(\omega)\,\Delta k_s + \sum_{s=1}^{L_2} c_h^s(\omega)\,d_s = \{C_j(\omega)\}^T \{P\} \quad (6\text{-}7)$$

Here $\{C_j(\omega)\}$ is the known coefficient vector formed using $\{\alpha_x(\omega)\}$ and $[\alpha_a(\omega)]$ and $\{P\}$ is an unknown vector of design variable changes. Also, for a specific frequency value ω, $b_j=b_j(\omega)$ (j=1,N) are known as the difference between analytical and experimental receptance of coordinate x_j (with excitation being applied at x_i). Therefore, from (6-6), the following linear algebraic equations in terms of unknown design variable vector $\{P\}$ can be established

$$[C(\omega)]\,\{P\} = \{B(\omega)\} \quad (6\text{-}8a)$$

where $[C(\omega)] = \begin{bmatrix} \{C_1(\omega)\}^T \\ \{C_2(\omega)\}^T \\ \vdots \\ \{C_N(\omega)\}^T \end{bmatrix}$, $\{P\} = \begin{Bmatrix} \Delta m_1 \\ \vdots \\ \Delta m_{L_1} \\ \Delta k_1 \\ \vdots \\ \Delta k_{L_2} \\ d_1 \\ \vdots \\ d_{L_2} \end{Bmatrix}$ and $\{B(\omega)\} = \begin{Bmatrix} b_1(\omega) \\ b_2(\omega) \\ \vdots \\ b_N(\omega) \end{Bmatrix}$.

In the presence of damping, $[C(\omega)]$ and $\{B(\omega)\}$ become complex, while $\{P\}$ is known to be real, and equation (6-8a) can be reformulated to solve for $\{P\}$ as:

$$\begin{bmatrix} \mathrm{Re}([C(\omega)]) \\ \mathrm{Im}([C(\omega)]) \end{bmatrix} \{P\} = \begin{Bmatrix} \mathrm{Re}(\{B(\omega)\}) \\ \mathrm{Im}(\{B(\omega)\}) \end{Bmatrix} \quad (6\text{-}8b)$$

To illustrate how the coefficient matrix [C(ω)] can be obtained, consider a simple two-mass and three-spring 2DOF system with its analytical mass and stiffness matrices as follows:

$$[M_a] = \begin{bmatrix} m_1 & 0 \\ 0 & m_2 \end{bmatrix} \text{ and } [K_a] = \begin{bmatrix} k_1+k_2 & -k_2 \\ -k_2 & k_2+k_3 \end{bmatrix} \tag{6-9}$$

Suppose the first column of the receptance matrix of the 'experimental' system is measured, then (6-6) becomes:

$$\begin{Bmatrix} {}_a\alpha_{11}(\omega)-{}_x\alpha_{11}(\omega) \\ {}_a\alpha_{21}(\omega)-{}_x\alpha_{21}(\omega) \end{Bmatrix} = \begin{bmatrix} {}_a\alpha_{11}(\omega) & {}_a\alpha_{12}(\omega) \\ {}_a\alpha_{21}(\omega) & {}_a\alpha_{22}(\omega) \end{bmatrix} \begin{bmatrix} -\omega^2\Delta m_1+\Delta k_1+\Delta k_2 & -\Delta k_2 \\ -\Delta k_2 & -\omega^2\Delta m_2+\Delta k_2+\Delta k_3 \end{bmatrix} \begin{Bmatrix} {}_x\alpha_{11}(\omega) \\ {}_x\alpha_{21}(\omega) \end{Bmatrix}$$

$$= \begin{bmatrix} -\omega^2{}_a\alpha_{11}\,{}_x\alpha_{11} & -\omega^2{}_a\alpha_{21}\,{}_x\alpha_{21} & {}_a\alpha_{11}\,{}_x\alpha_{11} & ({}_a\alpha_{11}-{}_a\alpha_{21})({}_x\alpha_{11}-{}_x\alpha_{21}) & {}_a\alpha_{21}\,{}_x\alpha_{21} \\ -\omega^2{}_a\alpha_{21}\,{}_x\alpha_{21} & -\omega^2{}_a\alpha_{22}\,{}_x\alpha_{21} & {}_a\alpha_{21}\,{}_x\alpha_{21} & ({}_a\alpha_{21}-{}_a\alpha_{22})({}_x\alpha_{11}-{}_x\alpha_{21}) & {}_a\alpha_{22}\,{}_x\alpha_{21} \end{bmatrix} \begin{Bmatrix} \Delta m_1 \\ \Delta m_2 \\ \Delta k_1 \\ \Delta k_2 \\ \Delta k_3 \end{Bmatrix} \tag{6-10}$$

where the frequency term has been dropped from the RHS of (6-10) to save space. Comparing (6-10) with (6-8a), the coefficient matrix [C(ω)] and vector {B(ω)} are obtained.

Equation (6-8a) (or (6-8b) in the case where damping exists) is obtained using the analytical and measured receptance data at one frequency point, but when j frequency points are used, then the total number of linear algebraic equations become j times as many as that of using one frequency point and (6-8a) (or (6-8b)) becomes a set of overdetermined algebraic equations. In order to solve for {P} in this case, the best technique available is the Singular Value Decomposition (SVD) which is described in Appendix I. Since no approximation has been made during the formulation of the problem, {P} can be solved directly. After {P} is calculated, and together with the analytical model itself, the updated system matrices can be determined uniquely. Also, it should be noted that in this case, no assumption has been made concerning the magnitudes of the error matrices [ΔM] and [ΔK].

The assumption that measurements are made in all the coordinates which are specified in the analytical model is, in most cases, unrealistic because in many cases: certain coordinates are physically inaccessible, such as the internal DOFs, and the rotation coordinates are notoriously difficult to measure. When the measured coordinates are

incomplete, as will be discussed in Chapter 7, direct solution of the updating problem is, in general, impossible and some sort of iteration scheme has to be introduced. Suppose one *incomplete* column $\{\alpha_x(\omega)\}_i$ has been measured, then the multiplication of the RHS of (6-6), which requires the complete vector $\{\alpha_x(\omega)\}_i$, can no longer be carried out exactly and some approximation has to be introduced.

Filling the unmeasured coordinates of $\{\alpha_x(\omega)\}_i$ in (6-6) with their analytical counterparts and then carrying out the multiplication of the RHS of (6-6) in the same way as for the complete coordinate case leads to the following linear algebraic equations which are the first order approximation:

$$[U(\omega)]\{P\} = \{V(\omega)\} \tag{6-11}$$

where $[U(\omega)]$ and $\{V(\omega)\}$ are obtained in a similar way to $[C(\omega)]$ and $\{B(\omega)\}$ in the case where measured coordinates are complete. However, in addition to the the approximate nature of (6-11), the total number of linear algebraic equations involved in (6-11) is n when data one frequency point are considered while in (6-8), the number is N ($n \leq N$ where n is the number of measured coordinates and N is total number of degrees of freedom specified in the analytical model). Again when more frequency points are used, (6-11) can be turned into an overdetermined set and a least-squares method (SVD) can be used to solve for $\{P\}$. Of course, the thus obtained $\{P\}$ is only the first order approximation and an iteration scheme has to be introduced in order to obtain the exact solution.

Also, it should be mentioned that in the case where the measured coordinates are incomplete, the updating problem formulated in (6-11) is, mathematically, based on a particular form of matrix perturbation analysis. Therefore, in order to guarantee convergence of the iteration process, some restriction may have to be made on the difference between the analytical and experimental models. This restriction is that the Frobenius norms of the error matrices should be of second order when compared with those of analytical mass and stiffness matrices (it is found that the convergence is largely determined by the difference of these two models in the modal domain rather than the spatial domain and the Frobenius norm is more closely related to the modal characteristics of the system than any other norms), that is:

$$\frac{\|\Delta M\|_F}{\|M_a\|_F} \leq \varepsilon \text{ and } \frac{\|\Delta K\|_F}{\|K_a\|_F} \leq \varepsilon \text{ where } \|A\|_F \equiv \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}^2} \tag{6-12}$$

Although ε varies from system to system, computational experience shows that for dynamical systems, the maximum value of ε can reach 0.3. Since the norm is taken in its Frobenius form and the modelling errors are usually localised, the relative amplitudes of modification for individual design variables can be more than 100% as shown in the numerical case studies below.

6.4.2 NUMERICAL CASE STUDIES

The first system studied here is an 8DOF mass-spring system as shown in Fig.6.2. This model has been used previously by several authors in model updating analysis exercises [86] and [90]: one of its noticeable features is that translational as well as rotational motions are permitted, and that the mass matrix is not diagonal.

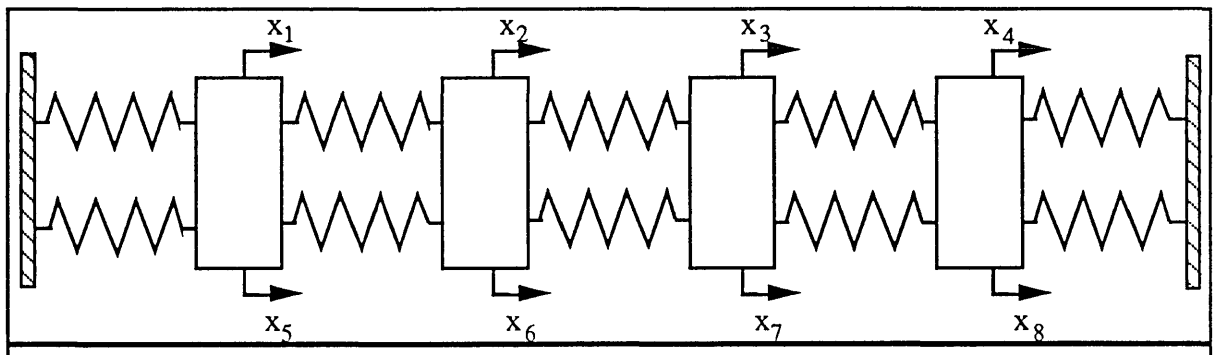


Fig.6.2 - An 8DOF System Used in Numrical Case Studies

When one complete column of the receptance matrix (column 1 in this numerical study) has been measured, the solution of the updating problem is unique and the procedure involved is direct since no approximation has been introduced in formulating the problem as discussed in §6.4.1. The analytical model is an undamped system with mass and stiffness matrices shown in Tables 6.1&6.2. The 'experimental' model consists, in general, of mass, stiffness and damping matrices which are

$$[M_x] = \frac{1}{1.25}[M_a]\,;\,[K_x] = 1.25\,[K_a] \text{ and } [D] = 0.02\,[K_a] \qquad (6\text{-}13)$$

1.0833	0	0	0	0.9166	0	0	0
0	1.0833	0	0	0	0.9166	0	0
0	0	1.0833	0	0	0	0.9166	0
0	0	0	1.0833	0	0	0	0.9166
0.9166	0	0	0	1.0833	0	0	0
0	0.9166	0	0	0	1.0833	0	0
0	0	0.9166	0	0	0	1.0833	0
0	0	0	0.9166	0	0	0	1.0833

Table 6.1 - Analytical Mass Matrix (kg)

4.0E+5	-2.0E+5	0	0	0	0	0	0
-2.0E+5	4.0E+5	-2.0E+5	0	0	0	0	0
0	-2.0E+5	4.0E+5	-2.0E+5	0	0	0	0
0	0	-2.0E+5	4.0E+5	0	0	0	0
0	0	0	0	4.0E+5	-2.0E+5	0	0
0	0	0	0	-2.0E+5	4.0E+5	-2.0E+5	0
0	0	0	0	0	-2.0E+5	4.0E+5	-2.0E+5
0	0	0	0	0	0	-2.0E+5	4.0E+5

Table 6.2 - Analytical Stiffness matrix (N/m)

Based on equation (6-8a), FRF data for 10 different frequency points (which were randomly chosen in the 'measured' frequency range) were used each time to construct $[C(\omega)]$ and $\{B(\omega)\}$ and as expected, the updated model is always exact regardless of the number of modes which are included in the 'measured' frequency range ('measured' FRF data covering the first mode, the first two modes and the first three modes were considered and the mathematical proof will be given later to illustrate this point). Fig.6.3 shows the analytical, measured and regenerated (using the updated model) point receptances at coordinate x_1 in the case the measured FRF data only covering the first mode. As shown in Fig.6.3, not only the exact mass and stiffness matrices of the 'experimental' system are recovered, but also the exact damping matrix of the 'experimental' system.

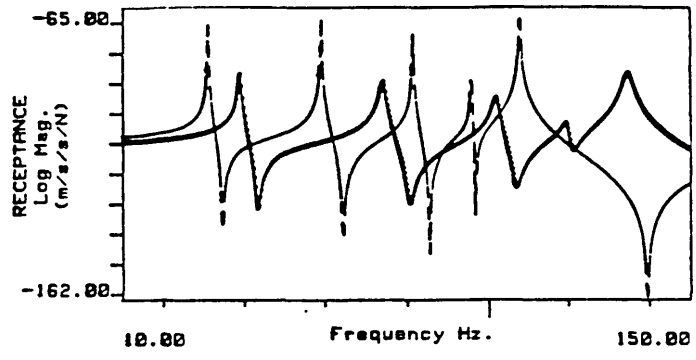


Fig.6.3 - Analytical, Measured and Regenerated Point Receptances at Coordinate x_1
— — — — — analytical, ····· measured and ———— regenerated

In order to consider the practical situation where the measured frequency response data are contaminated by noise, 1% uniformly distributed noise was added to the 'measured' frequency response function data and the above calculations were repeated. In all three cases ('measured' FRF data covering the first mode, the first two modes and the first three modes), it was found that the first iteration gave an estimation of errors to within 10% and after 2 or 3 iterations, the estimation error can be brought within 1% (for the stiffness and mass matrices but not damping). No further improvement is possible because of the presence of noise in the measured FRF data. The iteration strategy was introduced because in this case of data contaminated by noise, the effect of noise on the accuracy of estimation becomes less severe when the two models become closer. Again, the analytical, measured and regenerated point receptances of coordinate x_1 are shown in Fig.6.4, from which it is clear that the damping of the experimental system is underestimated because the existence of noise has drowned the effects of damping on off-resonance frequency response data. This demonstrates the difficulties involved in the damping property investigation of practical structures since the influence of damping on the frequency response data (off resonance) is always of second order. This problem of noise effecting the damping estimation can be overcome by smoothing the measured FRF data first or by using the data points around system resonances during the updating process.

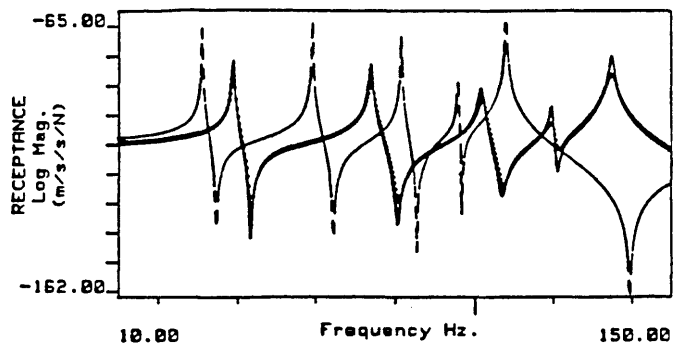
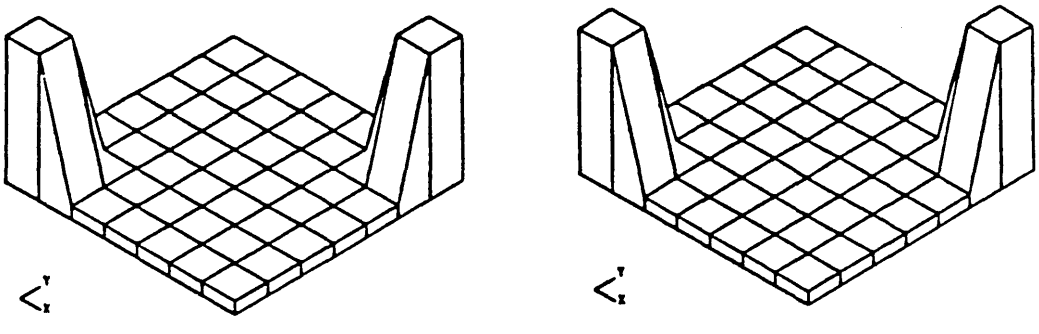


Fig.6.4 - Analytical, Measured and Regenerated Point Receptances at Coordinate x_1(1% noise)
——— analytical, ····· measured and ——— regenerated

When the measured coordinates are incomplete, the updating procedure is the same as for the case of complete coordinates except that during the formulation of equation (6-11), the unmeasured parts of the receptance $\{\alpha_x(\omega)\}_i$ are replaced by their analytical counterparts $\{\alpha_a(\omega)\}_i$. In this numerical case study, 4 coordinates x_1, x_3, x_6 and x_8 are supposed to be measured. The analytical model is the same as that shown in Tables 6.1&6.2 and the 'experimental' model is the analytical one perturbed in such a way that 30% stiffness modelling errors and 2% localised damping are introduced between coordinates x_1-x_2 and x_7-x_8, as shown in Fig.6.5. The point receptances at coordinate x_1 for each of these two models are shown in Fig.6.6. The mass matrix is assumed to be correct in this numerical case study although the method is equally applicable for the case of mass matrix updating as will be shown in the case study of GARTEUR structure in §6.4.4.



Stiffness error matrix Damping error matrix

Fig.6.5 - Exact Stiffness Error and Damping Matrices

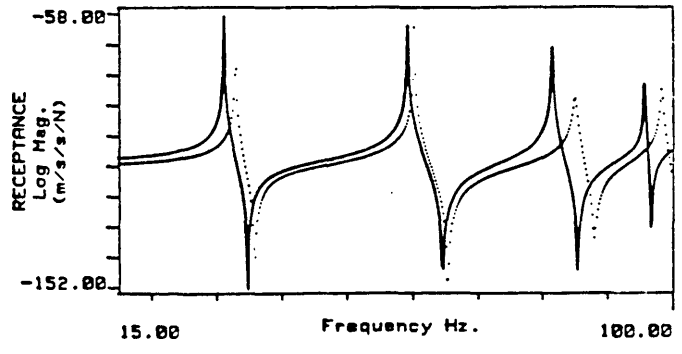


Fig.6.6 - Analytica and Experimental Point Receptances of Coordinate x1
——————— analytical, ... experimental

The program flow-chart for solving equation (6-11) iteratively is illustrated in Fig.6.7. As in the case of a complete set of measured coordinates, FRF data at 10 different frequency points were used in each iteration. Again, data covering a frequency range of just the first mode, the first two modes and the first three modes was investigated and it was found that in all these three cases, the error for the estimation (for the stiffness matrix) was less than 1% after 10 iterations. The iteration results for the case of data covering just the first mode are presented in Fig.6.8. The convergence criterion was chosen to be the relative norm (Euclidean) changes (see figure 6.7) of the stiffness and damping matrices of two successive iterations and it can be shown mathematically that under this criterion, the convergence is absolute (if it converges, it will converge to the true solution). As expected, the damping matrix converges more slowly than the stiffness matrix and only after the stiffness matrix has been obtained to some accuracy does the convergence of the damping matrix becomes faster.

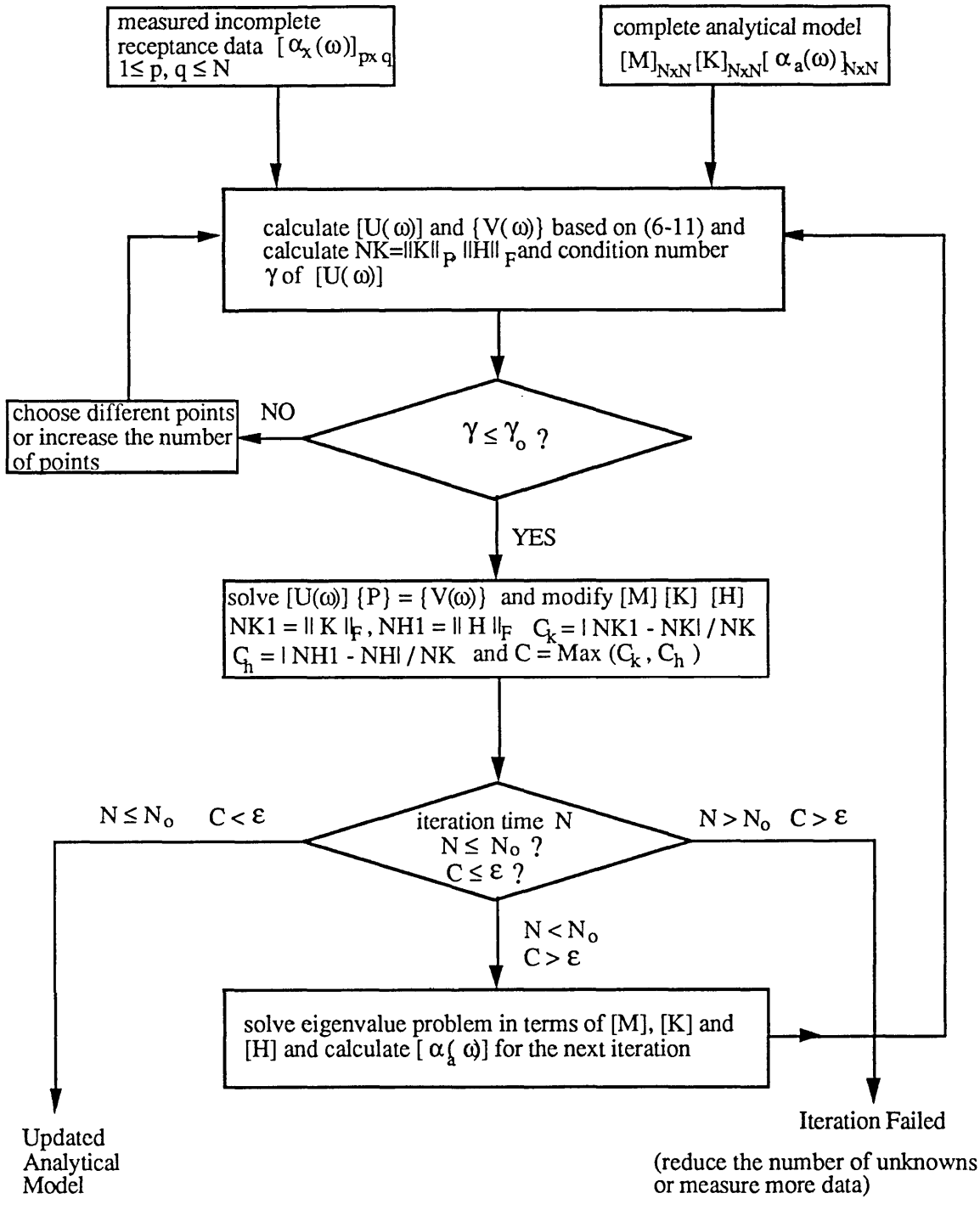
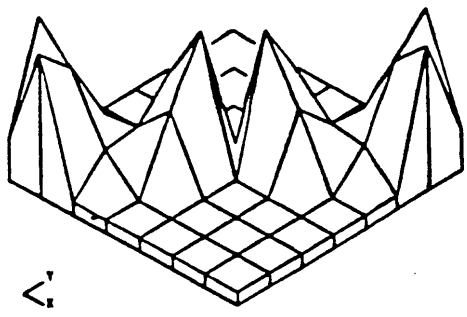
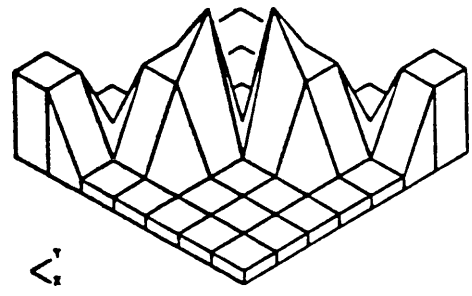


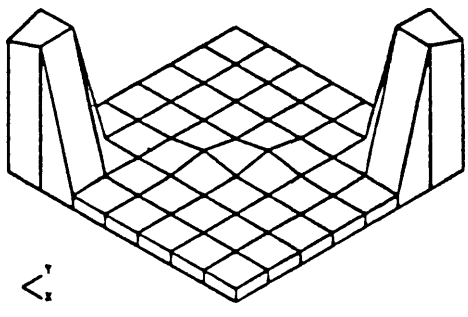
Fig.6.7 - Program Flowchart for Analytical Model Updating



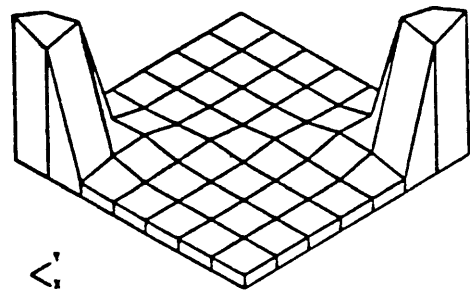
Stiffness Error Matrix (first iteration)



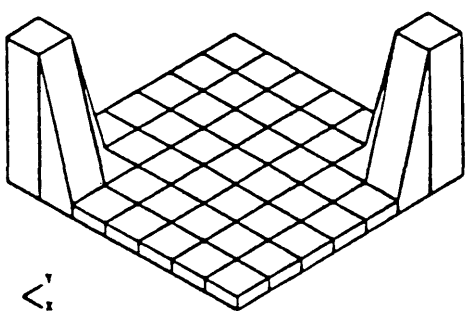
Calculated Damping Matrix (first iteration)



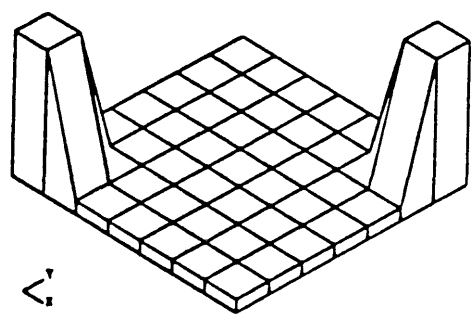
Stiffness Error Matrix (4th iteration)



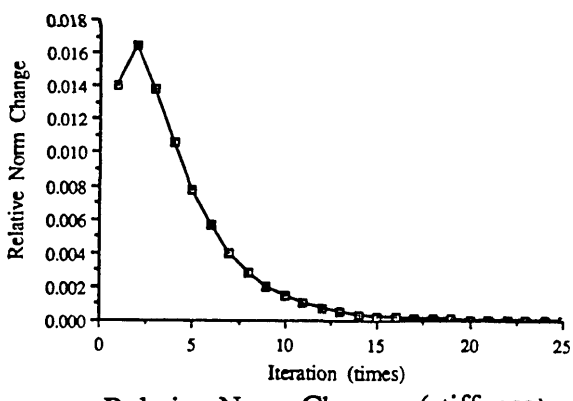
Calculated Damping Matrix (10th iteration)



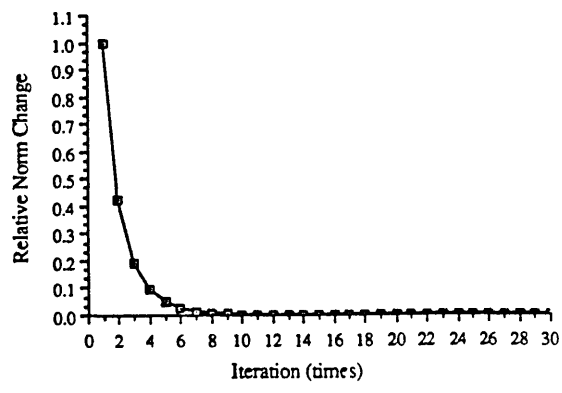
Stiffness Error Matrix (10th iteration)



Calculated Damping Matrix (20th iteration)



Relative Norm Changes (stiffness)



Relative Norm Changes (damping)

Fig.6.8 - The Iteration Results

During the calculation, it was found that when the stiffness modelling errors were greater than 35%, convergence became a problem (for this specific case). It is therefore considered that if the modelling errors are localised and there are some ways of locating them, then the number of unknowns involved in the updating process can be reduced and hopefully the restrictions imposed by the original assumption (the Euclidean norm of the modification should be of second order) can be relaxed. To illustrate this point, some 100% stiffness modelling errors are introduced to the same basic system mentioned above and the modelling errors are located in the ways which have been discussed in some detail in Chapter 5. After the location, only the unknowns which contain modelling errors are retained in the updating process. Even when measured FRF data covering only the first mode are used, convergence of the results was obtained.

From what is shown in Fig.6.7, a complete eigensolution of the analytical system is necessary during each iteration. For large practical problems, this could lead to a huge amount of calculation and so, if possible, this complete eigensolution should be avoided. This will be discussed in some detail later on in this Chapter.

6.4.3 EXTENSION OF THE METHOD TO THE MODELLING OF NONLINEAR STRUCTURES

The above developed method can be extended to the case where the structure is nonlinear. As mentioned before, when a structure to be modelled is nonlinear, its mathematical model has to be established on a mode by mode basis because of the inconsistency of measured data (even when response control is used to linearise the structure, the FRF data measured around different resonances could be the data from different linearised systems due the different modeshapes). For this reason, a mathematical model of a nonlinear structure cannot, in general, be established based on the use of measured modal data because, even when the nonlinearity location information is given, the data for one mode are, in most cases, insufficient for the problem to be solved. In order to establish a mathematical model of a nonlinear structure, it becomes essential that measured frequency response function data should be used instead of modal data.

It is assumed here that the nonlinearity of the structure is localised, as is usually the case in practice, and that the linearised frequency response functions corresponding to different response levels are measured. Also available is an analytical model of the structure which may be in error for the linear part (corresponding to very low response amplitude). What needs to be established is a mathematical model of the nonlinear structure corresponding

to these different response levels. Numerical cases of both complete and incomplete measured coordinates were investigated but only the results of the incomplete coordinates case are given here.

The analytical model used is that shown in Tables 6.1&6.2 and the 'experimental' model corresponding to the lowest response level is the same as the 'experimental' model used in the numerical case study of incomplete coordinates in §6.4.2 with stiffness and damping error matrices previously shown in Fig.6.5. Both cubic stiffness and quadratic damping nonlinearities are introduced between x_3-x_4 and are studied separately. Four coordinates are supposed to be measured x_1, x_3, x_6 and x_8 and the point receptances of coordinate x_1 corresponding to different response levels for the case of stiffness and damping nonlinearities are illustrated in Fig.6.9 in the frequency range of mode 3. First, the analytical model is updated using the measured frequency response functions corresponding to the lowest response level to obtain an accurate base-level ('zero' amplitude) linear model of the nonlinear structure. Then, based on this linear model and the measured FRF data at higher response amplitudes, the nonlinearity can be located in the ways discussed in Chapter 5. After the location has been made, only those unknowns corresponding to the nonlinear region are retained and, therefore, only the FRF data around one mode are necessary (mode 3 in this case study) to identify the mathematical model of the nonlinear structure (in fact, the only data available due to the inconsistency). The calculated stiffness and damping changes versus nondimensionalised response levels of coordinate x_1 are plotted in Fig.6.10. Clearly, the cubic stiffness and quadratic damping features of the nonlinearities are demonstrated.

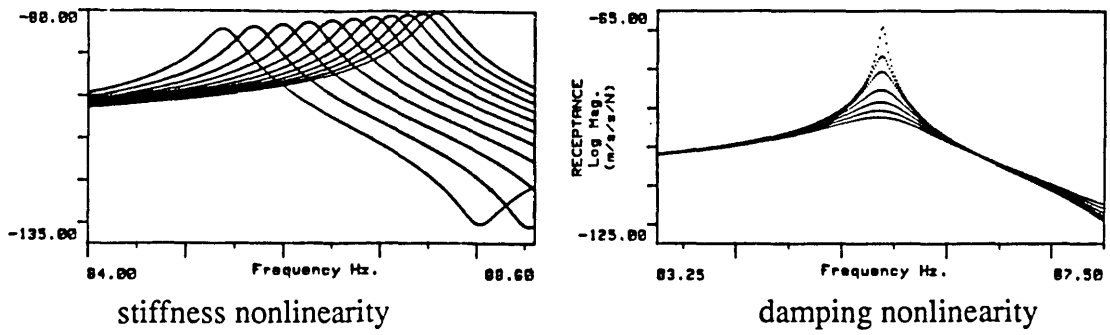


Fig.6.9 - FRFs of a Nonlinear Structure with Stiffness and Damping Nonlinearity

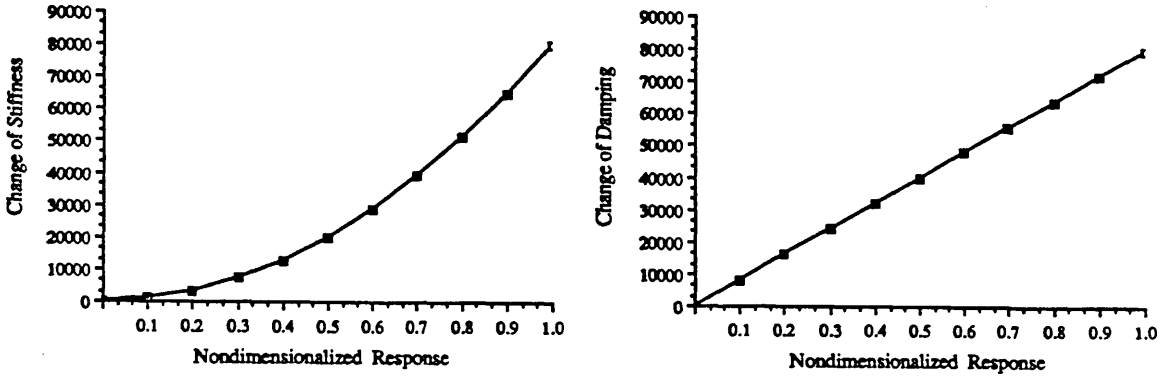


Fig.6.10 - Calculated Stiffness and Damping Changes Versus Response Levels

6.4.4 APPLICATION OF THE METHOD TO THE GARTEUR STRUCTURE

In the previous section, numerical case studies based on an 8DOF mass-spring system have been carried out to verify the new method and here an FE model of a more complex structure is considered. In general, as for the global mass and stiffness matrices, the mass and stiffness error matrices [ΔM] and [ΔK] can be expressed as linear combinations of element mass and element stiffness matrices $[m_s]$ and $[k_s]$ which have been appropriately (according to their positions in the global mass and stiffness matrices) expanded to the global dimension of the system, respectively, as:

$$[\Delta M] = \sum_{s=1}^{N_1} \beta_s [m_s] \quad \text{and} \quad [\Delta K] = \sum_{s=1}^{N_1} \gamma_s [k_s] \tag{6-14}$$

where N_1 is the number of elements, β_s and γ_s are the design variable changes associated with the s^{th} element, $[m_s]$ and $[k_s]$ are the s^{th} expanded element mass and stiffness matrices. In the case when the measured coordinates are incomplete, replacing the *unmeasured* receptance terms in $\{\alpha_x(\omega)\}_i$ (on the RHS of (6-6)) by their analytical counterparts and substituting $[\Delta Z(\omega)]=(-\omega^2[\Delta M]+[\Delta K])$ into the RHS of (6-6), with [ΔM] and [ΔK] being expressed in (6-14), the RHS of (6-6) becomes a vector with each element being a linear combination of all the unknown coefficients β_s and γ_s ($s=1,N_1$) and (6-6) becomes:

$$\begin{Bmatrix} {}_x\alpha_{1i}(\omega) - {}_a\alpha_{1i}(\omega) \\ {}_x\alpha_{2i}(\omega) - {}_a\alpha_{2i}(\omega) \\ \vdots \\ {}_x\alpha_{ni}(\omega) - {}_a\alpha_{ni}(\omega) \end{Bmatrix} = \begin{bmatrix} a_{11}(\omega)\ a_{12}(\omega) & \cdots & a_{12N_1}(\omega) \\ a_{21}(\omega)\ a_{22}(\omega) & \cdots & a_{22N_1}(\omega) \\ \vdots & \cdots & \vdots \\ a_{n1}(\omega)\ a_{n2}(\omega) & \cdots & a_{n2N_1}(\omega) \end{bmatrix} \begin{Bmatrix} \beta_1 \\ \vdots \\ \beta_{N_1} \\ \gamma_1 \\ \vdots \\ \gamma_{N_1} \end{Bmatrix} \quad (6\text{-}15)$$

Deleting the rows on both sides of (6-15) where the ${}_x\alpha_{ji}(\omega)$ terms have not been measured, the following reduced-order linear algebraic equations are obtained:

$$[A(\omega)]\ \{P\} = \{\Delta\alpha(\omega)\} \quad (6\text{-}16)$$

where $[A(\omega)]$ and $\{\Delta\alpha(\omega)\}$ are known for specific frequency ω. Again, equation (6-16) is based on FRF data at one frequency point. When data for a number of frequencies are used, (6-16) becomes overdetermined and the SVD technique can be used to solve for $\{P\}$ and then to reconstruct the updated analytical model. Since some approximation has been made during the formulation of (6-16) due to the incompleteness of measured coordinates, the updating problem has to be solved iteratively, as discussed in §6.4.1.

A plane truss structure as shown in Fig.6.11, which is a part of the GARTEUR structure as will be explained in next Chapter, is studied. In the formulation of FE model, 3DOFs (two translational and one rotational as shown figure 6.11) are considered. As compared with the previous mass-spring system, this example is larger in dimension and more typical because it is formulated based on a real structure. Modelling errors are introduced by overestimating the mass matrix for the 18^{th} element (nodes 17-18) and the stiffness matrices for the 11^{th} and 12^{th} element (nodes 10-12) by 100% (notice that in this case, the $\varepsilon = \|\Delta K\|_F / \|K_a\|_F = \sqrt{2}/\sqrt{19} = 0.324$). The exact mass and stiffness error matrices are shown in Fig.6.12. One incomplete column of the receptance matrix of the 'experimental' model - all the odd numbered nodes with their translational degrees of freedom (u,v) - is supposed to be measured over a frequency range covering just the first 5 modes and some of the measured and analytical frequency response functions are shown in Fig.6.13. Based on (6-16), FRF data at 20 frequency points were randomly chosen for each iteration in the measured frequency range to construct the coefficient matrices $[A(\omega)]$ and $\{\Delta\alpha(\omega)\}$ and the iteration results are shown in Fig.6.14.

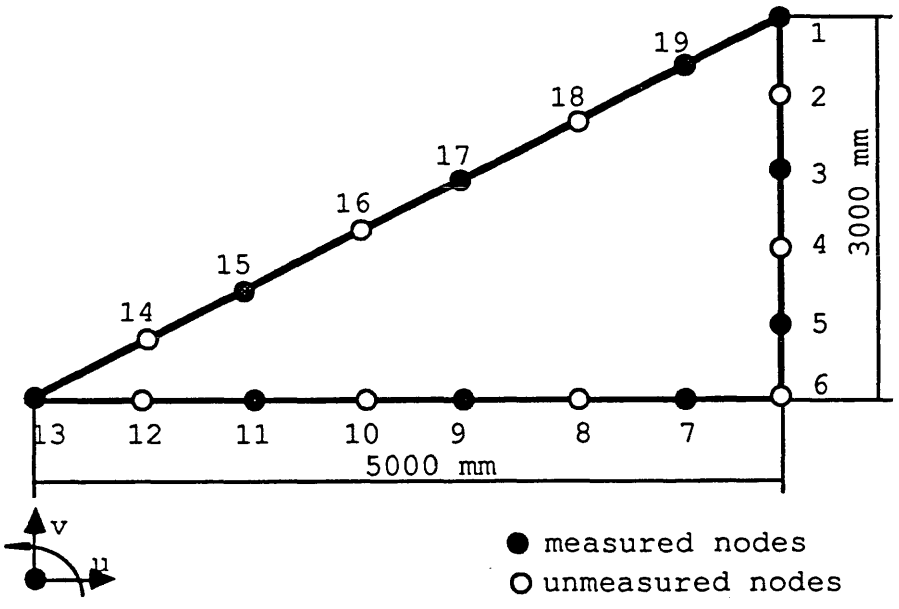
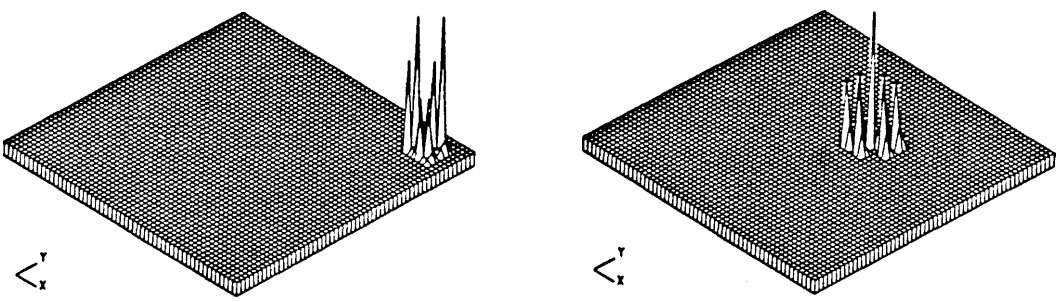
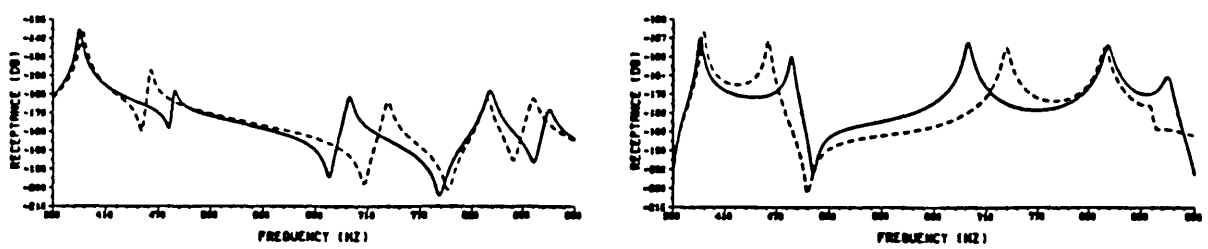


Fig.6.11 - A Free-free Frame Structure



Error Mass Matrix Error Stiffness Matrix

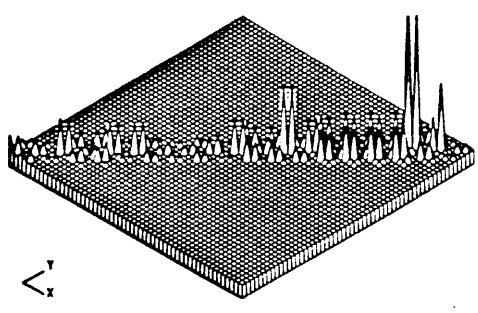
Fig.6.12 - Exact Mass and Stiffness Error Matrices



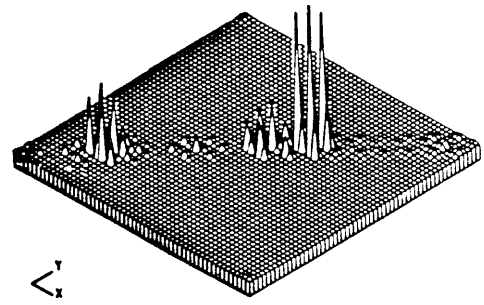
point receptances $\alpha_{11}(\omega)$ transfer receptances $\alpha_{1,25}(\omega)$

Fig.6.13 - FRFs of Analytical and 'Experimental' Models

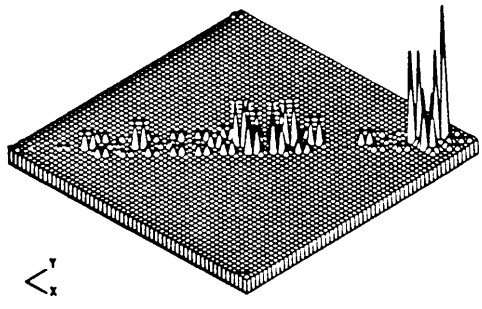
——— analytical, - - - - - - - 'experimental'



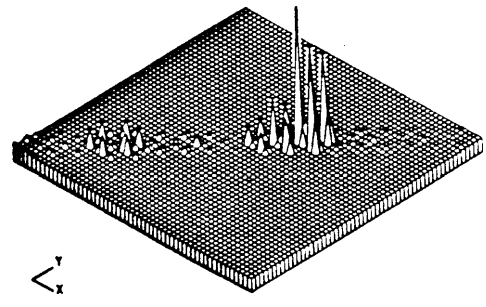
Error Mass Matrix (first iteration)



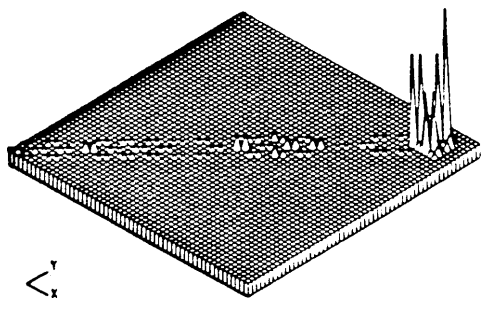
Error Stiffness Matrix (first iteration)



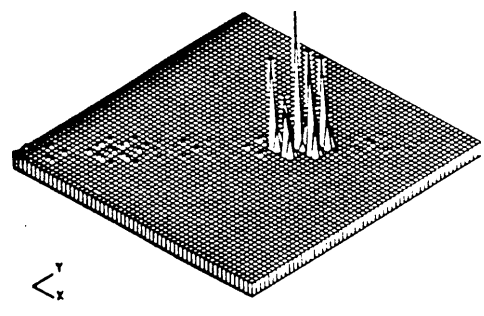
Error Mass Matrix (third iteration)



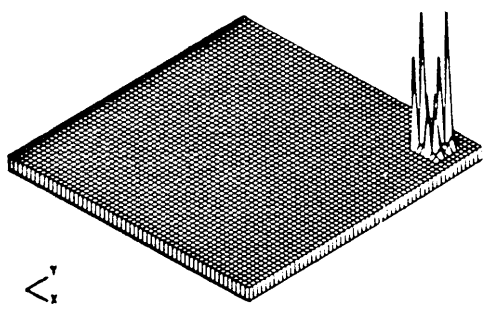
Error Stiffness Matrix (third iteration)



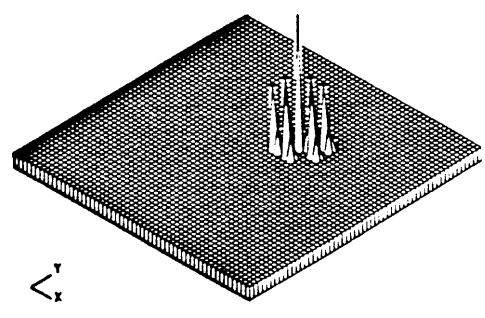
Error Mass Matrix (5^{th} iteration)



Error Stiffness Matrix (5^{th} iteration)



Error Mass Matrix (7^{th} iteration)



Error Stiffness Matrix (7^{th} iteration)

Fig.6.14 - Iteration Results

6.4.5 SOME CONSIDERATIONS ON THE COMPUTATIONAL ASPECTS

As shown in the numerical case studies, complete eigensolution of the updated system is required during each iteration when the measured coordinates are incomplete. Although computation is becoming cheaper as more powerful computers being produced at lower cost, this complete eigensolution is often computationally expensive, especially when systems with big dimensions are considered. It is therefore necessary to discuss some computational aspects of the eigenvalue problem so that computational effort involved can be minimised.

If all the eigenvalues and eigenvectors of a matrix are of interest (the complete eigensolution), the LR and QR algorithms [91], which are the most effective of known methods for the general algebraic eigenvalue problem, can be used. Both methods use a reduction of the general matrix to triangular form by similarity transforms, but the reduction is achieved by non-unitary transform in the LR algorithm while it is achieved by unitary transform in the QR algorithm which is numerically more stable. On the other hand, if only some of the eigenvalues and their corresponding eigenvectors are of interest, iterative methods, which are often referred to as subspace iteration because only a subset of the whole eigensolution is of interest, can be employed. In fact, in practical Finite Element analysis, it is rare for all the modes of the system to be calculated because, in general, only the lower modes of the system are of interest or, even, valid. The computational cost of solving the eigenvalue problem is, in general, proportional to the number of modes which are required. The most effective algorithm used for partial eigensolution is the Inverse Iteration method [92]. In the following, it will be shown how Inverse Iteration method can be used effectively to reduce the computational effort involved in the eigensolution.

Let the system to be solved be described by matrix $[A]=[M]^{-1}([K]+\mathbf{i}[D])$ and suppose that only the first p modes are of interest. Then, for a given matrix $[Q_0] \in C^{nxp}$ with orthonormal columns (initial estimate of the required eigenvectors), the following Inverse Iteration generates a sequence of matrices $\{[Q_k]\} \in C^{nxp}$ (k=1,2,...) which will converge to the first p eigenvectors of interest

For $k = 1, 2, \ldots$

$$Z_k = [A]^{-1} Q_{k-1}$$

$Q_k R_k = Z_k$ (QR factorisation of Z_k to calculate Q_k for next iteration)

(6-17)

After the eigenvectors [ϕ] (which is the converged [Q_k]) are calculated, the eigenvalues [λ.] can be found easily based on the Rayleigh Quotient formulation:

$$[\lambda.] = [\phi]^T [A] [\phi] \tag{6-18}$$

The convergence rate of the sequence of (6-17), as shown in [92], is proportional to the ratio λ_p/λ_{p+1}. It should be noted that during the iteration process, only one complex inverse is required, that being $([K]+i[D])^{-1}$, and in the case of a free-free system in which [K] is singular, a shift μ becomes necessary so that $[A]^{-1}$ in (6-17) becomes $([A]-\mu[I])^{-1}$. It is easy to prove [91] that systems described by [A] and ([A-μ[I]) have the same set of eigenvectors and the correponding eigenvalues simply differ by a value of μ. Also, in the specific case of the Inverse Iteration method where only one eigenvalue and eigenvector of a system are of interest, μ should be so chosen such that it is the closest to the eigenvalue of interest [92].

In the model updating process, since [ΔM], [ΔK] and [D] are usually small in the sense of the Frobenius norm when compared with the original mass [M] and stiffness [K] matrices, the eigenvalue problem is actually reduced to the problem of dynamic reanalysis (often referred to as structural modification analysis) and the initial estimation of the eigenvectors [Q_0] can be very accurately chosen as the eigenvectors of the original system. Due to this accurate choice of initial conditions, it can be expected that the convergence will in general be very fast.

All this means that if only a partial eigensolution is required, the Inverse Iteration method is ideal for solving the eigenvalue problem. As shown in §6.4.1, the whole updated analytical receptance matrix is required in each iteration of the updating process and, theoretically, in order to calculate this receptance matrix, all modes should be available. However, experience shows that the receptance in the lower frequency range can be accurately approximated by using just a few of the lower modes and by considering the contribution of higher modes as constant residual terms. For example, for a reasonably-sized practical 500x500 system, if the receptances are of interest only up to the frequency of the fifth mode, then the estimation of $\alpha_{ij}(\omega)$ based on following equation should be very accurate when the first 20 modes are included

$$\alpha_{ij}(\omega) = \sum_{r=1}^{20} \frac{{}_r\phi_i \, {}_r\phi_j}{\omega_r^2 - \omega^2 + i\eta_r\omega_r^2} + R_{ij} \tag{6-19}$$

Here R_{ij} is a constant residual term representing the contribution of higher modes which can be determined as follows by using the $([K]+i[D])^{-1}$ which has been calculated before

$$([K]+i[D])^{-1} = [\alpha(0)] = [\alpha_c(0)] + [R] \tag{6-20}$$

where $[\alpha_c(\omega)]$ is the receptance matrix corresponding to the contribution of the calculated modes.

It has been shown that during model updating process, the problem of an eigensolution being involved in each iteration is actually reduced to the problem of dynamic reanalysis. Hence, performing the complete eigensolution *ab initio* is unnecessary. On the other hand, due to the fact that the contribution of the higher modes to the receptances in the lower frequency range decreases quadratically in terms of frequency separation, only a partial eigensolution is necessary in order to calculate accurately the receptance data needed in the model updating process. These two observations make the Inverse Iteration method as discussed above the most appropriate method for the specific problem addressed.

6.5 GENERALISATION OF MODEL UPDATING METHODS

A model updating method based on the correlation between analytical model and measured frequency response functions has been developed and it will be advantageous to examine the relationship between the present method and the existing methods which are based on the correlation between analytical and measured modal data. It can be recognised that model updating methods can be generalised if measured FRF data are used instead of the modal data. This argument is self-evident because modal data are effectively derived from measured FRF data at the resonance frequencies but a brief mathematical proof is given here.

When the measured coordinates are complete, the present method of updating is formulated based on equation (6-6). Since model updating methods based on the use of modal data rely only on the data points where $\omega=\omega_r$ (at the natural frequencies), it is interesting to know to what equation (6-6) will degenerate when $\omega=\omega_r$. Unfortunately, $\{\alpha_x(\omega)\}_i$ in (6-6) is not defined when $\omega=\omega_r$ (the damping of the experimental model is made to be zero in order to permit a comparison with the methods based on modal data which are supposed to be from an undamped system), therefore what needs to be

discussed is what will equation (6-6) degenerate to as $\omega \rightarrow \omega_r$ and the damping matrix $[D] \rightarrow [O]$.

It is not difficult to see that when $\omega \rightarrow \omega_r$, $\{\alpha_x(\omega)\}_i$ can be expressed as

$$\lim_{\omega \rightarrow \omega_r} \{\alpha_x(\omega)\}_i = \gamma \{\phi\}_r + \{\phi^0\} \tag{6-21}$$

where $\{\phi^0\}$ is due to the influence of the other modes and can be considered as a constant vector when there is slight change in frequency and change in damping. However, when the damping matrix $[H] \rightarrow [O]$, the coefficient $\gamma \rightarrow \infty$. Therefore, when $\omega \rightarrow \omega_r$ and $[H] \rightarrow [O]$, equation (6-6) becomes

$$\lim_{\omega \rightarrow \omega_r} \gamma (- \{\phi\}_r + \frac{1}{\gamma} (\{\phi^0\} + \{\alpha_a(\omega_r)\}_i))^T =$$

$$\lim_{\omega \rightarrow \omega_r} \gamma (- \{\phi\}_r + \frac{1}{\gamma} \{\phi^0\})^T [\Delta Z(\omega_r)] [\alpha_a(\omega_r)] \tag{6-22a}$$

$$- \{\phi\}_r^T = \{\phi\}_r^T [\Delta Z(\omega_r)] [\alpha_a(\omega_r)] \tag{6-22b}$$

$$(- \omega_r^2 ([M_a] + [\Delta M]) + [K_a] + [\Delta K]) \{\phi\}_r = \{0\} \tag{6-22c}$$

From equation (6-22c), it can be seen that equation (6-6) degenerates to the simple eigendynamic equation when $\omega \rightarrow \omega_r$ and since all the model updating methods using modal data are derived from the basic equation (6-22c), it can be concluded that, in a broad sense, model updating methods based on modal data are discrete versions of the present generalised method based on frequency response data.

When the measured coordinates are incomplete, as mentioned earlier, the present method is derived from some sort of perturbation analysis and again, this perturbation analysis in the domain of frequency response functions can, in a broad sense, be regarded as a generalised version of the perturbation analysis in the modal domain on which some of the methods, such as the inverse sensitivity analysis as will be discussed in detail in Chapter 7, are based.

As discussed above, model updating methods based on modal data are discrete versions of the present generalised method because modal data are virtually the data of frequency response functions at resonance frequencies. Furthermore, in the case of model updating using measured FRF data, each data point contains information from *all* the modes of the

system and, in general, the data points measured are plentiful. Thus, the updating problem, as discussed below, is in most cases uniquely defined while in the case of model updating using modal data, the updating problem, in most cases, is underdefined since on the one hand, the data points are so badly chosen (FRF data at resonance frequencies) that the contribution of other modes has been deliberately eliminated and, on the other, the number of measured modes is usually quite limited.

6.6 UNIQUENESS OF UPDATING PROBLEM

The uniqueness of the model updating problem - that is, how much data should be measured so that the model updating problem can be solved uniquely (the true solution of problem can be obtained) - needs to be discussed when a new method is developed. Since the present method seeks to obtain an updated model which will reproduce the measured frequency response function data continuously on the frequency axis, as shown in numerical case studies, the measured data required (the number of measured modes) in order to solve the updating problem uniquely are much less as compared with other existing methods. In fact, it can be proven mathematically that when the measured coordinates are complete, a unique solution always exists regardless of the number of modes which are included in the measured frequency range. On the other hand, when the measured coordinates are *incomplete*, a unique solution can, in most cases, be obtained even when just few modes are included in the measured frequency range if the physical connectivity of the analytical model is employed as additional constraints.

Suppose one complete column $\{\alpha_x(\omega)\}_i$ of the receptance matrix is measured and its updated analytical counterpart $\{\alpha_{au}(\omega)\}_i$ is available (calculated based on the updated analytical model). Since the updated model should reproduce the measured data, it follows that $\{\alpha_{au}(\omega)\}_i \equiv \{\alpha_x(\omega)\}_i$ because, otherwise, it will still be possible to improve the model. Thus:

$$\sum_{r=1}^{N} \frac{{}_r\phi_i \, {}_r\phi_j}{\omega_r^2 - \omega^2 + i\eta_r\omega_r^2} = \sum_{r=1}^{N} \frac{{}_r\tilde{\phi}_i \, {}_r\tilde{\phi}_j}{\tilde{\omega}_r^2 - \omega^2 + i\tilde{\eta}_r\tilde{\omega}_r^2} \quad \text{for } j = 1, N \qquad (6\text{-}23)$$

where modal parameters with ~ are those of the experimental system and modal parameters without ~ are those of the updated analytical system. Since equation (6-23) is valid for any ω, then mathematically, we have

$$\omega_r^2 = \tilde{\omega}_r^2, \quad \eta_r\omega_r^2 = \tilde{\eta}_r\tilde{\omega}_r^2 \quad \text{and} \quad {}_r\phi_i\,{}_r\phi_j = {}_r\tilde{\phi}_i\,{}_r\tilde{\phi}_j \quad \text{for} \quad r = 1, n \ \text{and} \ j = 1, N \qquad (6\text{-}24)$$

It is not difficult to see that (6-24) is equivalent to following equations

$$[\omega_r.] = [\tilde{\omega}_r.], \quad [\eta_r.] = [\tilde{\eta}_r.] \quad \text{and} \quad [\phi] = [\tilde{\phi}] \qquad (6\text{-}25)$$

From (6-25), it is clear that the thus-updated analytical model is unique and is the true solution of the problem because the model reproduces all the eigenvalues and eigenvectors of the experimental system. Also, it is clear that the number of modes measured is irrelevant during the development of the above argument and, therefore, the uniqueness of the updating problem when measured coordinates are complete is independent of the number of modes measured.

On the other hand, when the measured coordinates are *incomplete*, a unique solution can also be obtained as shown in numerical case studies when the physical connectivity of the analytical model is employed. This process is intuitively explained as below. From an information point of view, although the measured coordinates are incomplete, every piece of measured data contains information about all the modifications (the difference between the analytical and experimental models) and since the measured data are plentiful in terms of frequency points and the physical connectivity of the analytical model is imposed as constraints, the updating problem can, in most cases, still be turned into an overdetermined one and thus be solved uniquely.

Mathematically, when the measured coordinates are incomplete, the physical connectivity of the analytical model becomes essential in order to have a unique solution. Suppose one incomplete column $\{\alpha_x(\omega)\}_i$ of the receptance matrix has been measured (for simplicity, assume coordinates x_1 to x_n have been measured), then, after the model has been updated, the $\{\alpha_x(\omega)\}_i$ values and their updated analytical counterparts $\{\tilde{\alpha}_{au}(\omega)\}_i$ should satisfy $\{\tilde{\alpha}_{au}(\omega)\}_i = \{\tilde{\alpha}_x(\omega)\}_i$, that is

$$\sum_{r=1}^{N} \frac{{}_r\phi_i\,{}_r\phi_j}{\omega_r^2 - \omega^2 + i\eta_r\omega_r^2} = \sum_{r=1}^{N} \frac{{}_r\tilde{\phi}_i\,{}_r\tilde{\phi}_j}{\tilde{\omega}_r^2 - \omega^2 + i\tilde{\eta}_r\tilde{\omega}_r^2} \quad \text{for } j = 1, n \qquad (6\text{-}26)$$

Equation (6-26) is equivalent to the following equations

$$\omega_r^2 = \tilde{\omega}_r^2, \quad \eta_r\omega_r^2 = \tilde{\eta}_r\tilde{\omega}_r^2 \quad \text{and} \quad {}_r\phi_i\,{}_r\phi_j = {}_r\tilde{\phi}_i\,{}_r\tilde{\phi}_j \quad \text{for} \quad r = 1, n \ \text{and} \ j = 1, n \qquad (6\text{-}27)$$

From (6-27), it is clear that the updated model has the same natural frequencies and damping loss factors as those of the experimental model, but the modeshape matrix is not completely the same due to incompleteness of the measured coordinates as shown in Fig.6.15 (only the submatrix of the whole eigenvector matrix corresponding to the measured coordinates has been determined). Hence, the updated model cannot, in general, be unique. However, if we impose the physical connectivity of the analytical model so that the updated model preserves the physical connectivity also, then the solution, will in most cases, become unique.

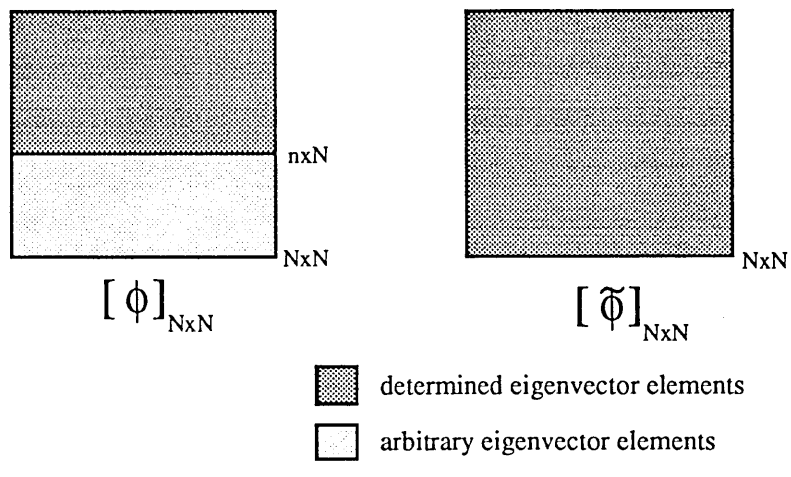


Fig.6.15 - Illustration of the Determination of Eigenvector Elements

6.7 BALANCE OF COEFFICIENT MATRICES

As mentioned in §6.2, when FRF data at j different frequency points are used, equations (6-8a) in the case of complete coordinates and (6-11) in the case of incomplete coordinates become

$$\begin{bmatrix} [C(\omega_1)] \\ [C(\omega_2)] \\ \cdot \\ \cdot \\ \cdot \\ [C(\omega_j)] \end{bmatrix} \{P\} = \begin{Bmatrix} \{B(\omega_1)\} \\ \{B(\omega_2)\} \\ \cdot \\ \cdot \\ \cdot \\ \{B(\omega_j)\} \end{Bmatrix} \qquad (6\text{-}8c)$$

$$\begin{bmatrix} [U(\omega_1)] \\ [U(\omega_2)] \\ \cdot \\ \cdot \\ \cdot \\ [U(\omega_j)] \end{bmatrix} \{P\} = \begin{Bmatrix} \{V(\omega_1)\} \\ \{V(\omega_2)\} \\ \cdot \\ \cdot \\ \cdot \\ \{V(\omega_j)\} \end{Bmatrix} \tag{6-11a}$$

Due to the different magnitudes of FRF data at different frequency points (e.g., data around resonance and anti-resonance frequencies), the RHS coefficient matrices of equations (6-8c) and (6-11a) could in some cases be poorly-conditioned in terms of its generalised inverse and this ill-conditioning could cause numerical difficulty in the updating process, especially when FRF data used are contaminated by measurement errors. Although such ill-conditioning problem can be overcome by properly choosing frequency points, an alternative numerical technique is presented here.

Consider equation (6-8c) and rewrite it as

$$\begin{bmatrix} \beta_1[C(\omega_1)] \\ \beta_2[C(\omega_2)] \\ \cdot \\ \cdot \\ \cdot \\ \beta_j[C(\omega_j)] \end{bmatrix} \{P\} = \begin{Bmatrix} \beta_1\{B(\omega_1)\} \\ \beta_2\{B(\omega_2)\} \\ \cdot \\ \cdot \\ \cdot \\ \beta_j\{B(\omega_j)\} \end{Bmatrix} \tag{6-8d}$$

The least-square solution of (6-8d) is

$$\{P\} = \left[\begin{bmatrix} \beta_1[C(\omega_1)] \\ \beta_2[C(\omega_2)] \\ \cdot \\ \cdot \\ \cdot \\ \beta_j[C(\omega_j)] \end{bmatrix}^T \begin{bmatrix} \beta_1[C(\omega_1)] \\ \beta_2[C(\omega_2)] \\ \cdot \\ \cdot \\ \cdot \\ \beta_j[C(\omega_j)] \end{bmatrix} \right]^{-1} \begin{bmatrix} \beta_1[C(\omega_1)] \\ \beta_2[C(\omega_2)] \\ \cdot \\ \cdot \\ \cdot \\ \beta_j[C(\omega_j)] \end{bmatrix}^T \begin{Bmatrix} \beta_1\{B(\omega_1)\} \\ \beta_2\{B(\omega_2)\} \\ \cdot \\ \cdot \\ \cdot \\ \beta_j\{B(\omega_j)\} \end{Bmatrix} \tag{6-28}$$

After some mathematical manipulation, equation (6-28) becomes

$$\{P\} = \left[\sum_{k=1}^{j} \beta_k^2 [C(\omega_k)]^T [C(\omega_k)] \right]^{-1} \left\{ \sum_{k=1}^{j} \beta_k^2 [C(\omega_k)]^T \{B(\omega_k)\} \right\} \tag{6-29}$$

By choosing β_k such that the maximum magnitude of the elements of the matrix $\beta_k^2[C(\omega_k)]^T[C(\omega_k)]$ (k=1,j) is unity, the accuracy of the solution could in some cases be improved.

6.8 CONCLUSIONS

In this Chapter, a new model updating method has been developed based on the correlation between the analytical model and measured frequency response function data. It has been mathematically demonstrated that existing methods based on the correlation between analytical and measured modal parameters are discrete versions of the new generalised method presented here and since only the frequency points which are the natural frequencies of the structure are used in those methods (these points are, in effect, badly chosen since they contain no information about other modes of the structure), the classical updating problem in most cases are underdefined. However, as shown in numerical examples, because of the direct use of measured response function data, the residual terms involved in the data are taken into account in the new method and the measured data are always of plentiful in terms of frequency points, and so, the present method turns the updating problem into an overdetermined one in most cases.

Considering the practical difficulty of measuring FRF data at all the coordinates which are specified in the analytical model, the present method has been extended to the case where measured coordinates are incomplete. Mathematically, in these circumstances, the method is based on a certain form of matrix perturbation analysis and, therefore, an iteration scheme has to be introduced during the updating process. Some computational aspects involved in the eigensolution during this iteration have also been discussed so that computational cost can be reduced.

As for the uniqueness of the updating problem, it is often asked how many modes are required in order to get a unique (true) solution? Of course, the number of modes required depends on which modes are chosen and what characteristics the structure possesses if the updating problem is to be solved based on measured modal data as will be discussed in next Chapter. However, based on the new method developed in this Chapter, it has been mathematically proven that in the case where measured coordinates are complete, the unique (true) solution always exists regardless of the number of modes measured. On the other hand, when the measured coordinates are incomplete, a unique solution can be obtained in most cases even when few modes are measured if the physical connectivity of the analytical model is employed.

Although experimental modal analysis has been highly developed, there still exist some problems. In the case where two modes are very close, accurate modal parameters can be difficult to obtain. Furthermore, measured modes are usually complex as discussed in Chapter 2 and although some investigation has been undertaken to devise ways of employing complex modes directly in the correlation [93], most of the correlation methods are based on the use of real modes. This means that the measured modes have to be realised first before they are used. The realisation process not only introduces numerical errors but at the same time, discards the damping information about the structure. These problems do not exist in the present generalised FRF method and, as a by-product of the updating process, the method reveals the damping information about the structure (damping model).

The method has been extended to the case where the structure under investigation is nonlinear. It has been shown that in this case, in order to establish the mathematical model, an accurate linear model as well as the location information of the nonlinearity are, in general, necessary because the model has to be established based on mode by mode basis. On the other hand, because of the inconsistency of measured data, it is essential to use measured FRF data in the correlation, rather than the measured modal data.

Numerical case studies based on an 8DOF mass-spring system as well as a 57DOF model of a frame structure are carried out to assess the practical applicability of the new method presented in this Chapter. Cases where measured coordinates are both complete and incomplete are considered and the results have shown marked advantages over the the existing methods based on the use of measured modal data. Since the method is developed for the practical case in which both measured modes and coordinates are incomplete, it is readily applicable to practical correlation analysis.

CHAPTER 7

POSSIBILITIES AND LIMITATIONS OF ANALYTICAL MODEL IMPROVEMENT

7.1 INTRODUCTION

As one of the major applications of modal analysis, analytical model improvement/updating using the measured dynamic properties of a structure has become a major research topic in the dynamic modelling of practical engineering structures and has generated many technical publications. A large number of different algorithms have been developed, some of which have been proven to be quite successful, such as the new method developed in Chapter 6. However, to the author's knowledge, there has been little technical discussion about what can be done (the possibilities) and what cannot (the limitations) in the updating of an analytical model when practical measurement cases in which both measured modes and coordinates are incomplete are considered. The purpose of this Chapter is to identify some of these possibilities and limitations with the objective of directing research in this subject towards more productive areas.

The discussion begins with a review of the limitations and difficulties of some of the recently developed methods based on full matrix updating. The mathematical underdeterminancy associated with these methods is explained. Then the possibility of updating a condensed (Guyan-reduced) model with error location based on Kidder's expansion method is examined. It is demonstrated that, when measured modes and/or coordinates are incomplete as they are in practice, updating of the analytical model using

full matrix updating method(s) or based on Guyan-reduced model with error location is very difficult, if not impossible. In order to solve the updating problem, it becomes clear that the physical connectivity of the analytical model should be respected during the updating process so that the number of unknowns involved can be reduced and the limited measured data available can have the possibility of solving the problem.

When the physical connectivity of the analytical model is imposed, the data required in order to update an analytical model are usually within the scope of practical measurements. As discussed in Chapter 6, by imposing the physical connectivity, measured FRF data covering few modes are, in most cases, enough to solve the updating problem even when the measured coordinates are incomplete. Nevertheless, in this present Chapter, criteria for the minimum measured data (modal data) required to solve the updating problem are established based on the Eigendynamic Constraint Method (where measured coordinates are complete) and the Inverse Eigensensitivity analysis (where measured coordinates are incomplete). Such criteria are important because they enable the analyst to judge whether an available set of measured modal data is able to obtain a unique solution of the updating problem. These criteria are then generalised by using the measured FRF data.

Basically, the discussions are illustrated using the analytical model updating exercise called 'GARTEUR' which is supposed to represent the true practical problem in terms of the incompleteness of both measured modes and coordinates. A mass-spring model is also used to illustrate the criterion developed.

7.2 REVIEW OF FULL MATRIX UPDATING METHODS

The term 'full matrix updating' is defined here to mean that all elements in the analytical mass and/or stiffness matrices are considered to be in error during the updating process and therefore should be corrected using the measured data if possible. Berman's method [94], the Error Matrix Method [90] and methods presented in references [83,95] belong to this category. Since the measured data are limited (both measured modes and coordinates), it will be shown that it is practically very difficult, if not impossible, to update the analytical model based on these methods. What these methods provide is an optimised solution which can reproduce the measured modal data to some accuracy, but which is not the true solution of the problem.

7.2.1 BERMAN'S METHOD

In reference [94], Berman developed a systematic method for improving analytical mass and stiffness matrices based on measured modal data. In general, the method is devised for the case where the analytical model is in the Guyan-reduced form in which only the measured degrees of freedom are retained as masters and, as a result, the analytical model itself becomes fully populated and the measured coordinates are compatible with the analytical model. In this case, it will be shown that only when all the modes have been measured, can exact updating become feasible based on Berman's method. Otherwise the solution obtained is an optimised one. A brief summary of the method is introduced in this section for the convenience of discussion and interested readers are referred to the original paper [94] for details.

During the optimisation procedure of Berman's method, an 'improved' analytical mass matrix is calculated first and then, based on this 'improved' mass matrix, an optimised stiffness matrix is obtained. Although it is difficult to provide a physical interpretation, the method is developed based on the minimisation of weighted Euclidean norms of the differences between the analytical and experimental mass matrices and between the analytical and experimental stiffness matrices respectively as:

$$\varepsilon_m = \| [M_a]^{-1/2} ([M_x] - [M_a]) [M_a]^{-1/2} \| \tag{7-1}$$

$$\varepsilon_k = \| [M_x]^{-1/2} ([K_x] - [K_a]) [M_x]^{-1/2} \| \tag{7-2}$$

Introducing orthogonality and symmetry properties into the experimental mass and stiffness matrices as physical constraints in the form of Lagrange multipliers, the final 'improved' mass and stiffness matrices can be derived from:

$$[M_x] = [M_a] + [M_a] [\phi_x] [m_a]^{-1} ([I] - [m_a]) [m_a]^{-1} [\phi_x]^T [M_a] \tag{7-3}$$

$$[K_x] = [K_a] + [\Delta_k] + [\Delta_k]^T \tag{7-4}$$

where $[m_a] = [\phi_x]^T [M_a] [\phi_x]$

and $[\Delta_k] = \frac{1}{2} [M_x] [\phi_x] ([\phi_x]^T [K_a] [\phi_x] + [\dot{\lambda}_x]) [\phi_x]^T [M_x] - [K_a] [\phi_x] [\phi_x]^T [M_x]$

From equations (7-3) and (7-4), it can be shown that the thus estimated mass matrix $[M_x]$ and stiffness matrix $[K_x]$ can never be the exact mass and stiffness matrices of the structure unless *all* the modes have been measured and this point will be further explained later on.

7.2.2 THE ERROR MATRIX METHOD (EMM)

The basic theory of EMM [90] can be summarised as shown below. Define stiffness error matrix $[\Delta K]$ to be $[\Delta K]=[K_x]-[K_a]$ and assume that the $[\Delta K]$, when compared with $[K_a]$, is of second order in the sense of the Euclidean norm. Then, based on matrix perturbation theory, $[\Delta K]$ can be expressed, to a first order approximation, as:

$$[\Delta K] \approx [K_a] ([K_a]^{-1} - [K_x]^{-1}) [K_a] \tag{7-5}$$

Again, $[K_x]^{-1}$ and $[K_a]^{-1}$ can be approximated using the m ($m \leq N$ where m is the number of measured modes and N is the number of degrees of freedom specified in the analytical model) corresponding measured and analytical modes and then substituted into (7-5) to obtain $[\Delta K]$ as:

$$[\Delta K] = [K_a] [\phi_a] [\lambda_a]^{-1} [\phi_a]^T [K_a] - [K_a] [\phi_x] [\lambda_x]^{-1} [\phi_x]^T [K_a] \tag{7-6}$$

As for Berman's method, the EMM requires *complete* measured coordinates and, even then, the estimation of the stiffness error matrix (the estimation of $[\Delta M]$ based on EMM is shown in [74]) is approximate even when *all* the experimental modes have been included. From this point, Berman's method is mathematically superiorer to the EMM because, at least when all the modes and all the coordinates are measured, Berman's formulation gives an exact solution (although the solution is trivial in this case). On the other hand, although the number of modes required can be reduced when an iteration procedure is introduced by imposing physical constraints (physical connectivity) as shown in [74], the situation cannot be improved significantly since, during the iteration, the total number of unknowns cannot be reduced and it is therefore expected that even when the physical constraints have been introduced, the number of modes necessary for the EMM to work properly will be in general beyond the capability of practical measurement. However, one could argue here that although it is strictly not possible to update the analytical model based on this method in practice, when few measured modes are available, is it not possible to pinpoint the modelling errors (the errors are supposed to be localised as they usually are in practice) using the first iteration results based on the EMM and then to reduce the number of unknowns in order to facilitate the solution of the

problem? This sounds reasonable but is mathematically contradictory because if the errors were located in the first iteration, that means something about the solution has been obtained and if the iteration is carried on, the condition will be improved and the true solution will, in most cases, be achieved. Since the solution cannot be achieved based on successive iteration, the location using first iteration results will be, in general, meaningless.

7.2.3 MATHEMATICAL ILLUSTRATIONS

It has been demonstrated that in the full matrix updating methods, *all* coordinates which are specified in the analytical model, and *all* modes (or at least a large proportion in the case of iterative EMM) should be measured in order to obtain the exact mass and stiffness matrices of the system. This point is to be illustrated here from a linear algebraic equation point of view based on the Eigendynamic Constraints Method, which will be discussed further in detail later in this Chapter. Based on the eigendynamic equation and mass-normalisation property of the i^{th} measured mode:

$$- (\lambda_x)_i [\Delta M] \{\phi_x\}_i + [\Delta K] \{\phi_x\}_i = (\lambda_x)_i [M_a] \{\phi_x\}_i - [K_a] \{\phi_x\}_i \qquad (7\text{-}7)$$

$$\{\phi_x\}_i^T [\Delta M] \{\phi_x\}_i = 1 - \{\phi_x\}_i^T [M_a] \{\phi_x\}_i \qquad (7\text{-}8)$$

it can be seen that (7-7) and (7-8) provide (N+1) linear algebraic equations in terms of the unknowns in the error mass and stiffness matrices. Therefore, when all the N modes are measured, the total number of linear algebraic equations is N(N+1). At the same time, since [ΔM] and [ΔK] are symmetric, the total number of unknowns involved in [ΔM] and [ΔK] is also N(N+1). This means that if all the elements of both mass and stiffness matrices are considered to be in error, the updating problem is just *deterministic* (the number of equations is equal to the number of unknowns) when all the modes and coordinates have been measured. Any measured piece of information (modes and/or coordinates) which is missing will cause the problem to be *underdetermined* and only optimised solutions become possible. This demonstrates the limitations of full matrix updating methods.

7.3 CONDENSED MODEL BASED ON GUYAN REDUCTION

So far, the difficulties and limitations of full matrix updating have been discussed. The possibilities of updating an analytical model when it is in the Guyan-reduced form (usually fully populated) with localised modelling errors now need to be investigated.

The modelling errors are located first in the ways as discussed in some detail in Chapter 5 and then the possibility of further updating of the model is investigated. The numerical results shown are calculated based on the free-free GARTEUR structure (Fig.7.1).

7.3.1 THE GARTEUR STRUCTURE

The structure being studied is a truss structure as shown in Fig.7.1. Each elemental segment shown in Fig.7.1 is the superposition of an axial bar element and a bending beam element. Young's modulus is assumed to be $E=0.75x10^{11}N/m^2$ and the density to be $\rho=2800kg/m^3$. For the bar element, the cross section areas are S_h(horizontal)$=0.004m^2$, S_v(vertical)$=0.006m^2$ and S_d(diagonal)$=0.003m^2$. For the bending beam element, the second moment of inertia is assumed to be the same for all the bending beam elements and is $I=0.0756m^4$.

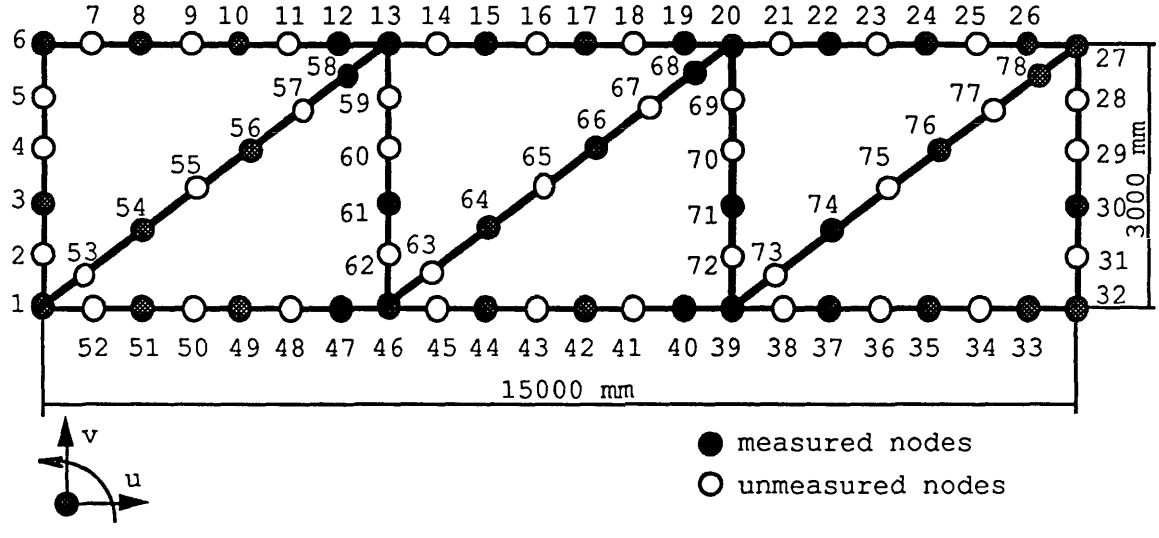


Fig. 7.1 - The Free-free GARTEUR Structure

7.3.2 THE GUYAN REDUCTION TECHNIQUE

In some cases, the full analytical model obtained using FE analysis needs to be reduced for computational reasons. The most commonly used technique for undertaking such a process is the Guyan reduction technique [96]. If the full analytical mass and stiffness matrices are partitioned according to master and slave degrees of freedom, then the frequency response function matrix corresponding to the master degrees of freedom at frequency ω can be calculated as follows:

$$[\alpha(\omega)] = \begin{bmatrix} [\alpha_{mm}(\omega)] & [\alpha_{ms}(\omega)] \\ [\alpha_{sm}(\omega)] & [\alpha_{ss}(\omega)] \end{bmatrix} = [Z(\omega)]^{-1} = \begin{bmatrix} [Z_{mm}(\omega)] & [Z_{ms}(\omega)] \\ [Z_{sm}(\omega)] & [Z_{ss}(\omega)] \end{bmatrix}^{-1} \quad (7\text{-}9)$$

From (7-9), the $[\alpha_{mm}(\omega)]$ can be calculated as

$$[\alpha_{mm}(\omega)] = ([Z_{mm}(\omega)] - [Z_{ms}(\omega)]\,[Z_{ss}(\omega)]^{-1}\,[Z_{ms}(\omega)]^{T})^{-1} \quad (7\text{-}10)$$

In order to preserve exactly the responses at the master degrees of freedom $[\alpha_{mm}(\omega)]$ at any frequency, the reduced mass and stiffness matrix must satisfy

$$[K^R] - \omega^2 [M^R] = [\alpha_{mm}(\omega)]^{-1} = [Z_{mm}(\omega)] - [Z_{ms}(\omega)]\,[Z_{ss}(\omega)]^{-1}\,[Z_{ms}(\omega)]^{T} \quad (7\text{-}11)$$

The reduced stiffness matrix can be obtained by setting $\omega = 0$ in (7-11) as

$$[K^R] = [K_{mm}] - [K_{ms}]\,[K_{ss}]^{-1}\,[K_{ms}]^{T} \quad (7\text{-}12)$$

Upon substitution of $[K^R]$ into (7-11), the reduced mass matrix $[M^R]$ is obtained as

$$[M^R] = [M_{mm}] - \frac{1}{\omega^2}[K_{ms}]\,[K_{ss}]^{-1}\,[K_{ms}]^{T} + \frac{1}{\omega^2}([K_{ms}] - \omega^2 [M_{ms}])\,[K_{ss}]^{-1}\,[M_{ss}]$$
$$([I] - \omega^2 [K_{ss}]^{-1}\,[M_{ss}])^{-1}\,[K_{ss}]^{-1}\,([K_{ms}] - \omega^2 [M_{ms}])^{T} \quad (7\text{-}13a)$$

Although, thus reduced, the mass matrix is nonlinear and nonunique, because it is a function of forcing frequency, the reduction based on (7-12) and (7-13a) is exact in the sense that the reduced model would predict exactly the steady state response of the structure at all master degrees of freedom at any frequency. In the static case, when $\omega=0$ (in fact, $\omega \to 0$, the limit of (7-13a) exists), the general reduction formula reduces to that of the well-known Guyan reduction and the reduced mass matrix becomes:

$$[M^R]_{(\omega \to 0)} = [M_{mm}] - [K_{ms}]\,[K_{ss}]^{-1}\,[M_{ms}]^{T} - [M_{ms}]\,[K_{ss}]^{-1}\,[K_{ms}]^{T}$$
$$+ [K_{ms}]\,[K_{ss}]^{-1}\,[M_{ss}]\,[K_{ss}]^{-1}\,[K_{ms}]^{T} \quad (7\text{-}13b)$$

What should be discussed here are some of the characteristics of the Guyan reduction technique from a modelling error location point of view, using appropriate numerical case studies. Since, as shown above, Guyan reduction is a static approximation, the reduced model cannot preserve the total energy of the original model, and this is demonstrated by

differences of the natural frequencies between these two models. Furthermore, during the reduction, there are energy exchanges from one part of the structure to another which are illustrated by the spreading of the localised modelling errors in the original model. To explain these two points, numerical case studies (case 1) based on the free-free GARTEUR structure shown in figure 7.1 are carried out. The full analytical model has 234 DOFs by considering 3 DOFs at each node as shown in figure 7.1, while the Guyan reduced model has 156 DOFs after condensing out the rotational degrees of freedom. The first 6 non-zero natural frequencies of the full and condensed models are shown in Table 7.1. For the lower modes, the natural frequencies of these two models are quite similar but become substantially different when higher modes are concerned. To illustrate how localised modelling errors in the full analytical model spread during the condensation process, the element stiffness matrices of element 12 (nodes 12-13) and element 45 (nodes 45-46) are doubled to make a modified version of the structure. In this case, the stiffness modelling errors on the full coordinate basis are localised in coordinates 37-42 (nodes 12-13) and 133-138 (nodes 45-46). However, after the condensation, these localised modelling errors are seen to have spread into neighbouring coordinates, depending on the connectivity, as shown in Fig.7.2&7.3. Although the dominant errors are still where they should be, some errors have been spread with considerable amplitudes. Taking a close look at where the dominant errors are as shown in Fig.7.3, it can be seen that more than 15 coordinates have clearly been contaminated although only the stiffness of one single element has been changed. The effect of this spreading of modelling errors on the updating of such analytical models can already be anticipated and will be discussed later.

Mode No.	1	2	3	4	5	6
ω_f	145.441	226.784	283.965	397.222	427.044	442.724
ω_c	145.244	225.737	282.556	393.635	422.128	436.142

ω_f: natural frequency of full analytical model (Hz)

ω_c: natural frequency of condensed analytical model (Hz)

Table 7.1 Natural Frequencies of Full and Condensed Analytical Models

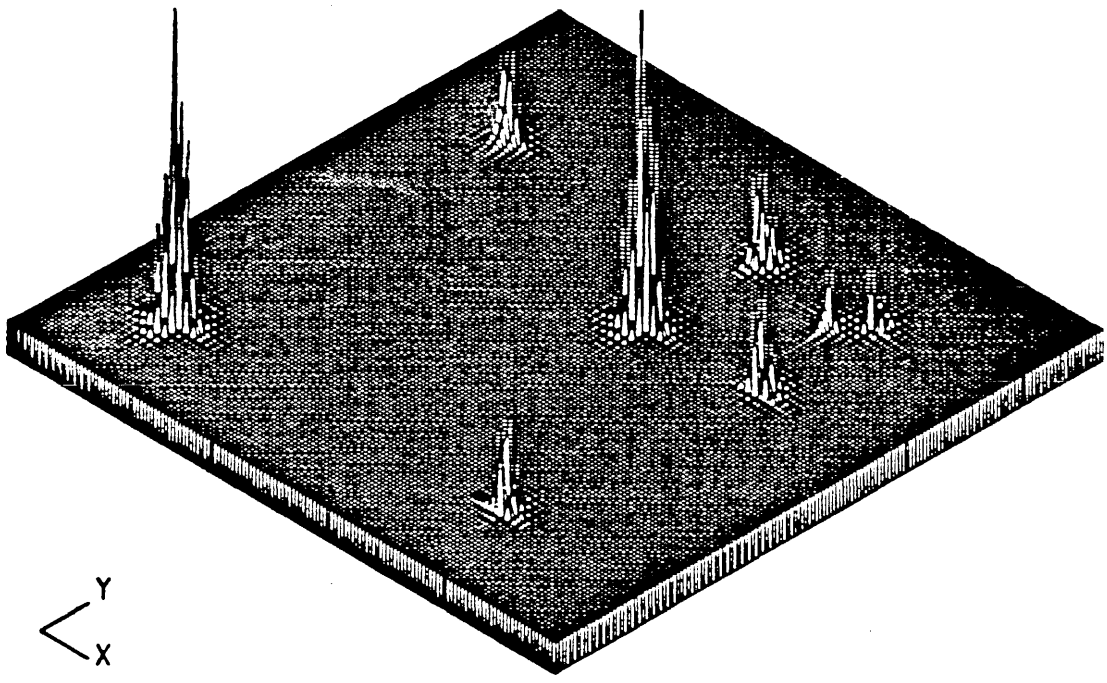


Fig.7.2 - Stiffness Modelling Errors (Guyan-reduced)

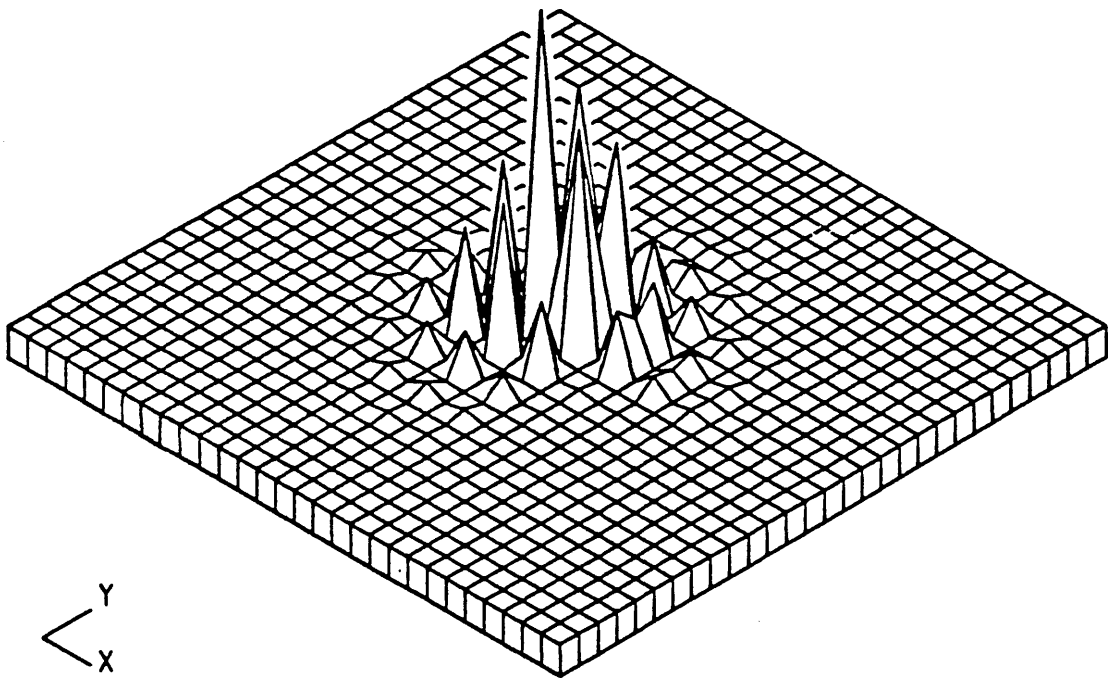


Fig.7.3 - Stiffness Errors between Coordinates 22-52 (Guyan-reduced)

7.3.3 LOCATION OF MODELLING ERRORS

As discussed above, since the physical connectivities of the original analytical model have been destroyed during dynamic condensation (Guyan reduction), exact updating of the condensed analytical model based on the full matrix updating methods such as Berman's method or the Error Matrix method becomes impossible unless all the modes and all the coordinates (specified in the condensed model) have been measured. Here, this problem is looked at from a different angle by locating the major modelling errors first and then by concentrating on those areas where the major modelling errors are believed to be, trying to reduce the mathematical difficulties involved and thus to solve the updating problem using the limited measured data available.

In the following example (case 2), stiffness modelling errors are introduced between nodes 59-62 and nodes 69-72 by doubling the values of the 6 element stiffness matrices between these nodes. The stiffness error matrix (condensed) is shown in Fig.7.4 after the condensation. Only the hatched nodes (figure 7.1) with their translational degrees of freedom (u,v) are supposed to have been measured. The unmeasured coordinates need to be interpolated before the location process can be undertaken.

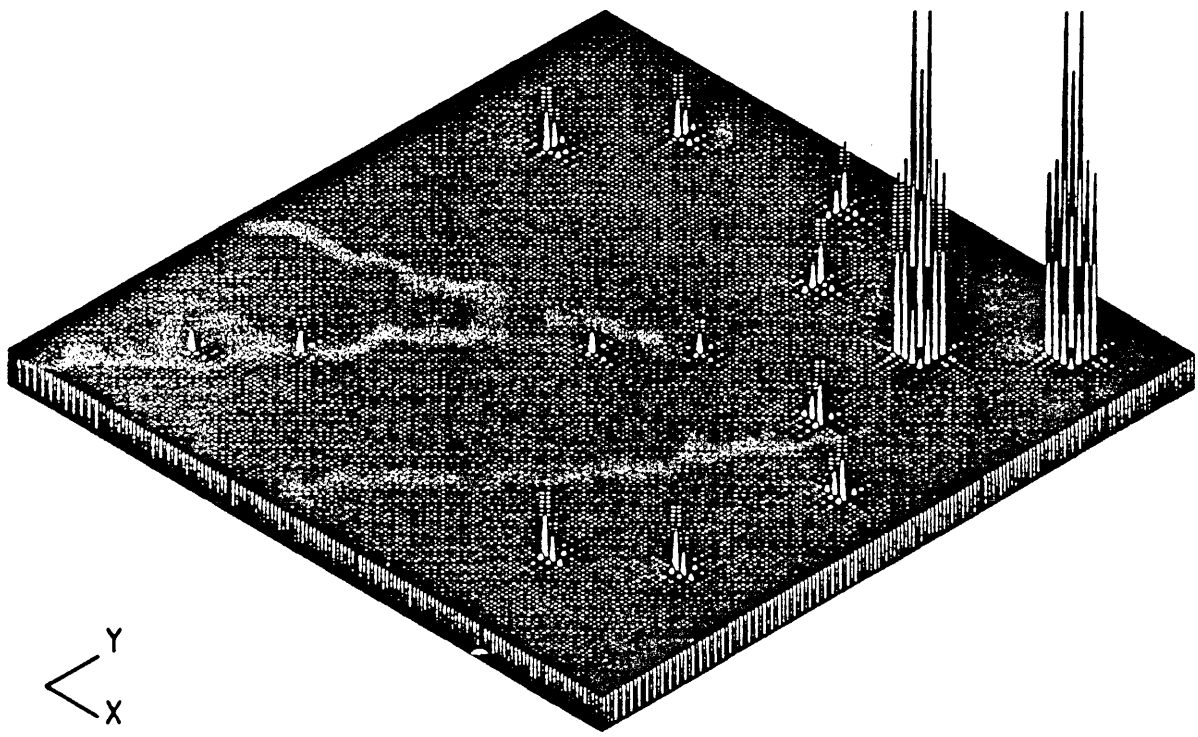


Fig.7.4 - Stiffness Errors for case 2 (Guyan-reduced)

This interpolation of the unmeasured coordinates can be achieved by using the analytical model itself as discussed in some detail in Chapter 5. Partitioning the analytical model according to measured and unmeasured degrees of freedom and using the eigendynamic equation, the unmeasured subvector $\{\phi_s\}$ of r^{th} mode can be interpolated in terms of the analytical model and measured subvector $\{\phi_m\}$ of r^{th} mode as:

$$\{\phi_s\} = (-\omega_r^2 [M_{ss}] + [K_{ss}])^{-1} (\omega_r^2 [M_{sm}] - [K_{sm}]) \{\phi_m\} \qquad (7\text{-}14)$$

The first 5 non-rigid body modes have been expanded based on (7-14) and the relative modeshape errors which are defined as the difference between the true modeshape and the expanded modeshape scaled by the maximum absolute value of the elements of the true modeshape, are shown in Figs.7.5-7.9. From these figures, it can be seen that the expanded modeshapes based on (7-14) are quite accurate except where the localised modelling errors are.

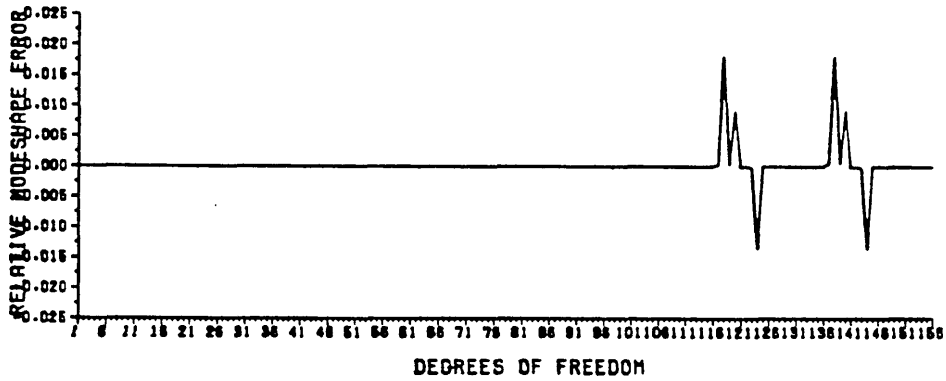


Fig.7.5 - Relative Modeshape Error of Mode 1 (case 2)

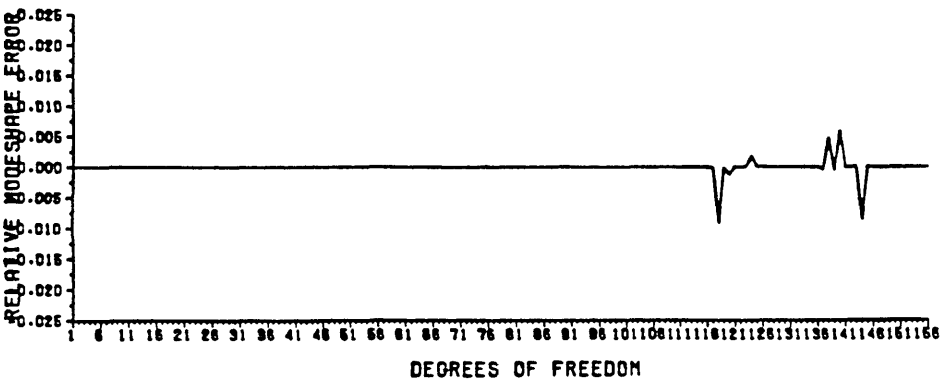


Fig.7.6 - Relative Modeshape Error of Mode 2 (case 2)

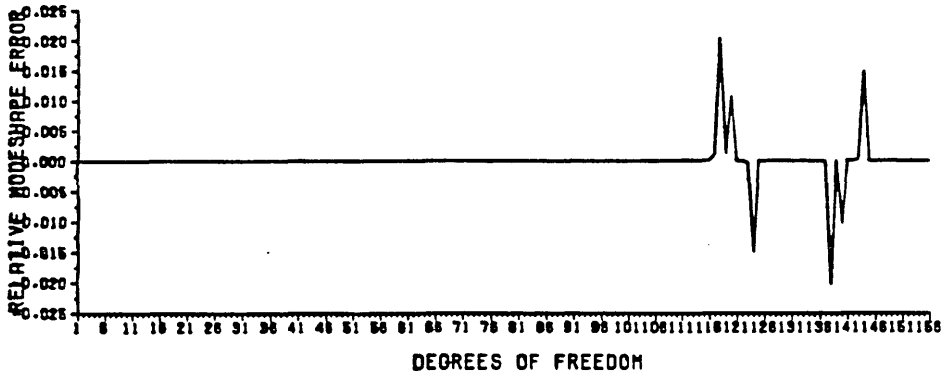


Fig.7.7 - Relative Modeshape Error of Mode 3 (case 2)

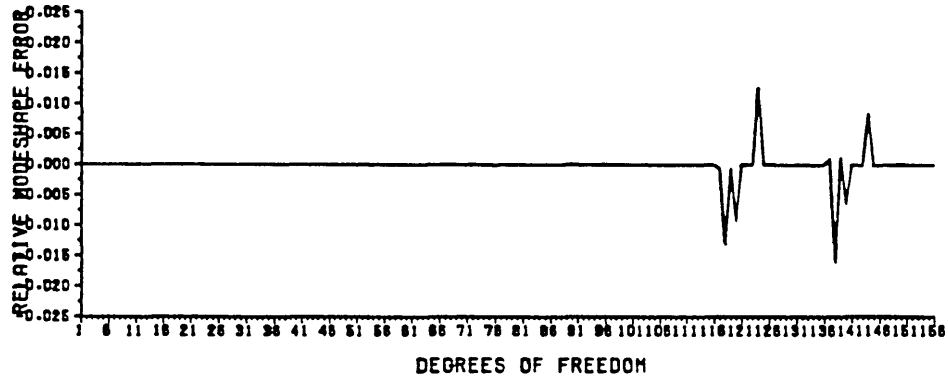


Fig.7.8 - Relative Modeshape Error of Mode 4 (case 2)

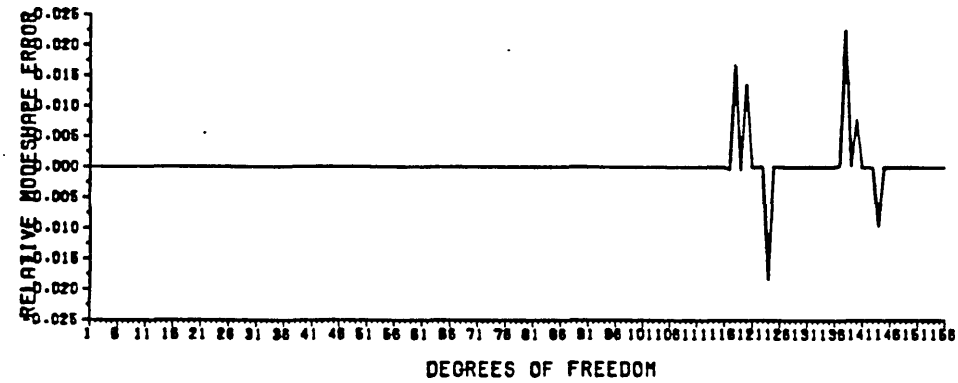


Fig.7.9 - Relative Modeshape Error of Mode 5 (case 2)

These expanded modes are then used to find the localised modelling errors based on the location methods discussed in Chapter 5. All 5 modes are used respectively and the location results are shown in Figs.7.10-7.14 (these are not strictly stiffness modelling errors [ΔK] but are $[\Delta K]\{\phi\}_r\{\phi\}_r^T$ as explained in Chapter 5). When compared with the stiffness error matrix of Fig.7.4, it can be seen that the location is quite successful since for every mode, all the areas where the major modelling errors are placed, are clearly located. However, it is difficult to identify the different amplitudes of these errors by looking at these location results even if, in fact, the errors are considerably different in amplitudes (figure 7.4). This adds difficulties for the next-step updating.

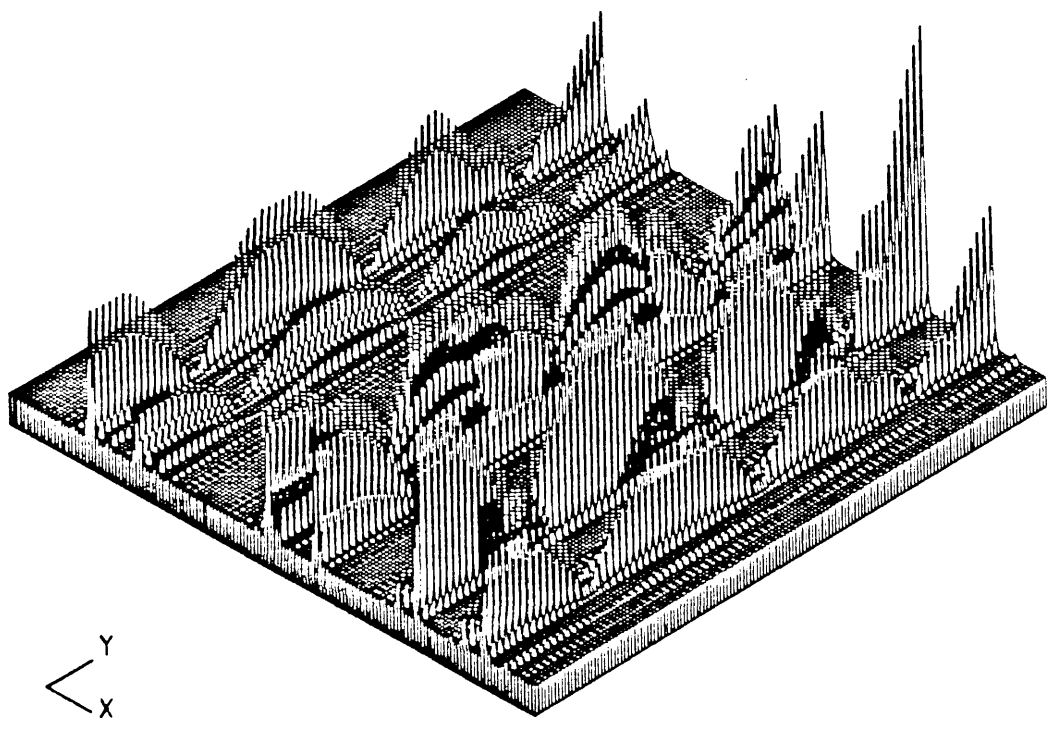


Fig.7.10 - Located Stiffness Errors Using Mode 1 (case 2)

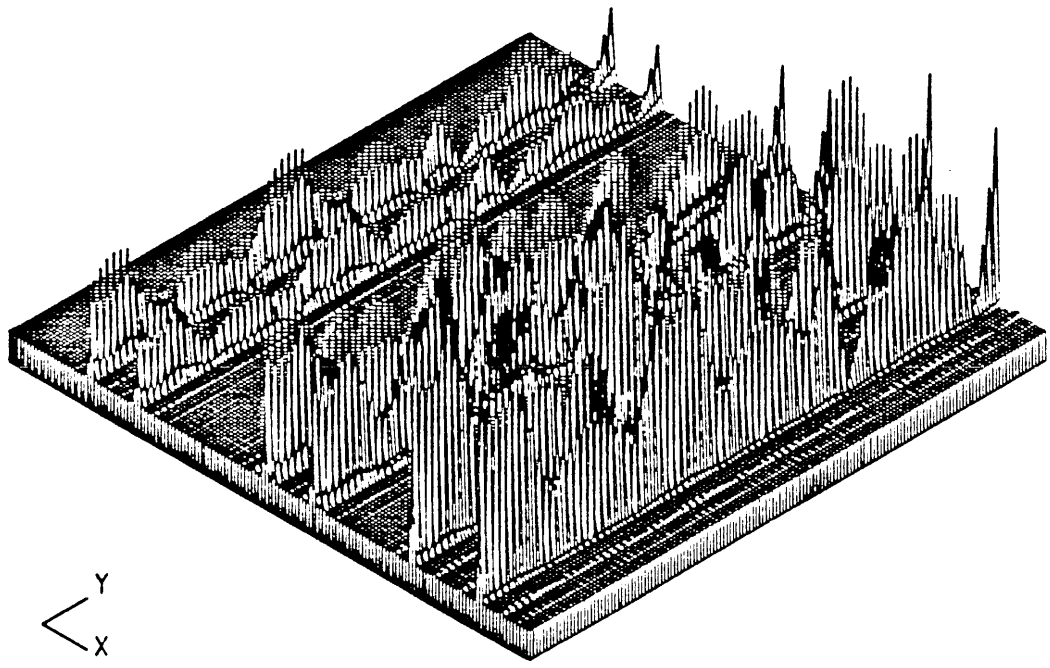


Fig.7.11 - Located Stiffness Errors Using Mode 2 (case 2)

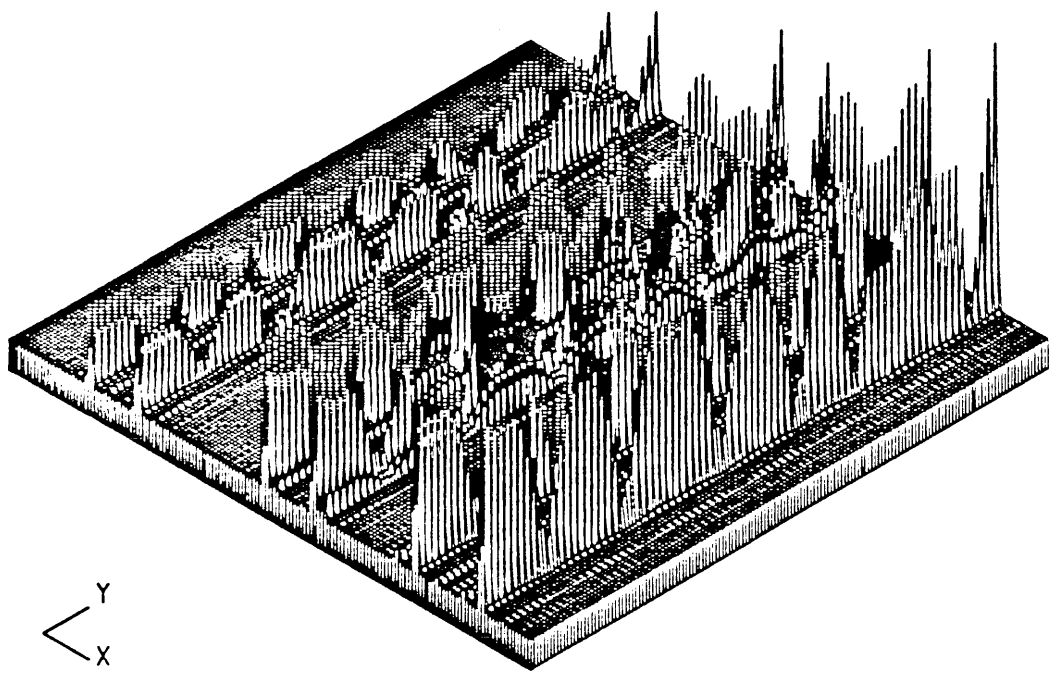


Fig.7.12 - Located Stiffness Errors Using Mode 3 (case 2)

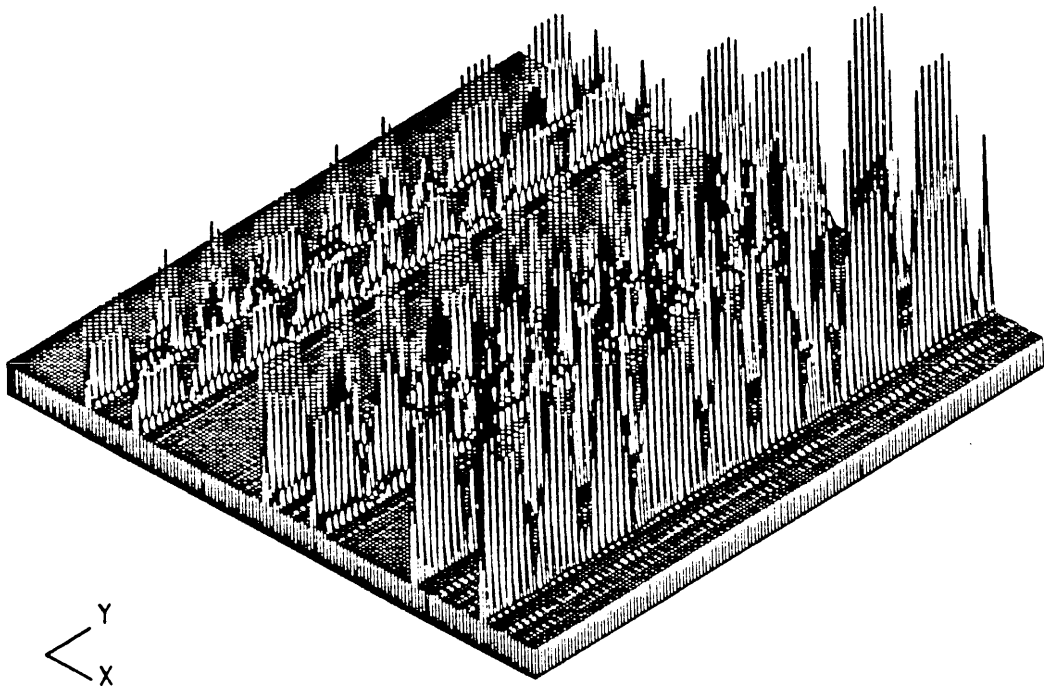


Fig.7.13 - Located Stiffness Errors Using Mode 4 (case 2)

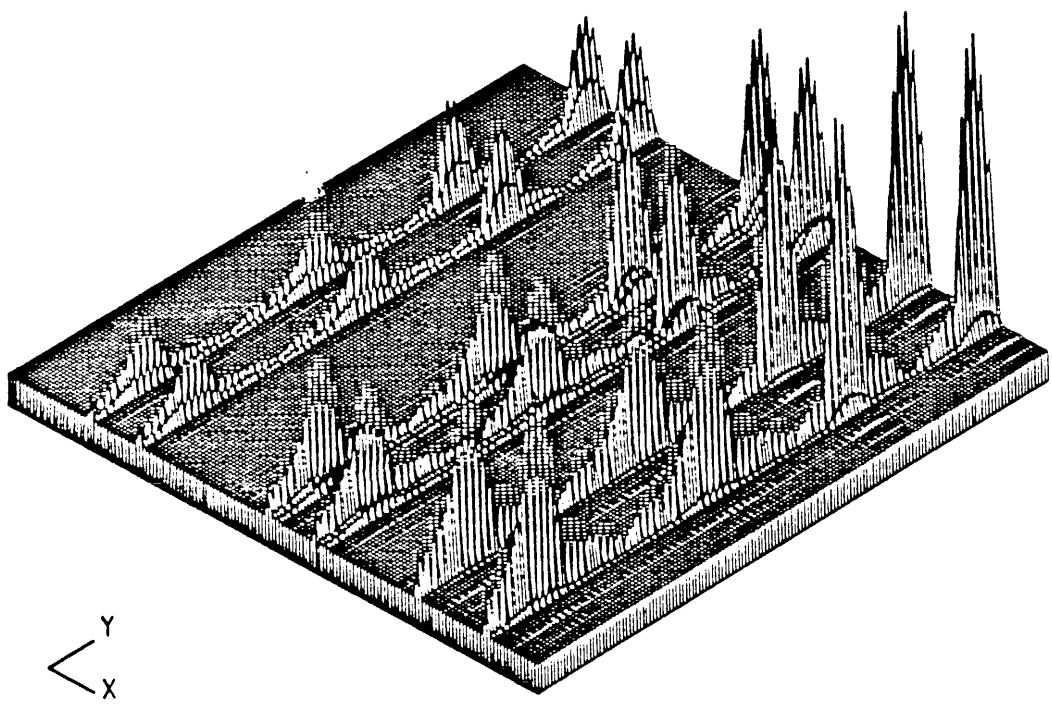


Fig.7.14 - Located Stiffness Errors Using Mode 5 (case 2)

7.3.4 THE UPDATING OF AN ANALYTICAL MODEL

Now, suppose the localised modelling errors are successfully located, what to do next? Obviously, the next step will be to update the analytical model since this is the ultimate target which is sought, but how? As mentioned in Chapter 6, the only way of updating an analytical model when the measured coordinates are incomplete, as in this case, is by some kind of perturbation or sensitivity analysis based on an iterative strategy. However, in this case, since the physical connectivities have been destroyed during the condensation process, every element of the mass (if mass modelling errors are considered) and stiffness matrices of the condensed analytical model in those located areas should be considered as an independent unknown in the updating process. Again, from the linear algebraic equation point of view, even for this simple numerical case, such a task as to update the analytical model will be beyond the capability of practical measurement (at least, as in the GARTEUR exercise, in which only 5 measured modes and a third coordinates (half the coordinates in terms of the condensed analytical model) are given, it is impossible to solve the problem addressed). On the other hand, although the major errors have been located and so considered in the updating process, the influence of those 'small' errors (which are not considered in the updating process) on the solution process probably cannot be simply overlooked because they are spread over the whole matrix.

7.4 STRUCTURAL CONNECTIVITY IN AN ANALYTICAL MODEL

So far, the mathematical difficulties involved in the full matrix updating process have been demonstrated and the limitations of updating an analytical model when it is in the Guyan-reduced form have been discussed. It becomes clear that in order to update an analytical model using limited measured data, it is essential to preserve the structural connectivity of the analytical model. As far as the stiffness properties are concerned, this suggests that any stiffness modelling errors which exist can only occur in the non-zero elements of the analytical stiffness matrix $[K_a]$. The mathematical validity of such physical constraints imposed in the updating process, which are essential, is discussed below.

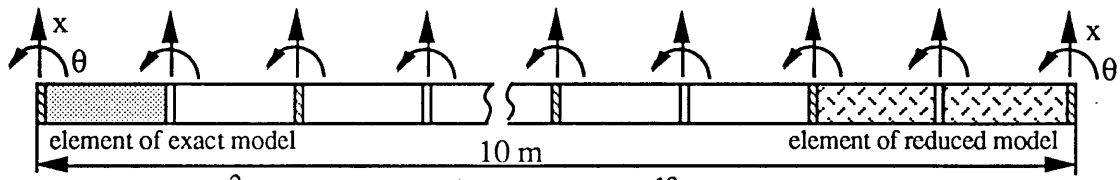
It is well known that practical structures are continuous and possess an infinite number of degrees of freedom while the mathematical models sought to represent their dynamic characteristics have a finite number of degrees of freedom. As shown before, this discretisation modelling process is mathematically a dynamic condensation process.

Suppose there exist unique mass and stiffness matrices $[M]_{N_s x N_s}$, $[K]_{N_s x N_s}$ where $N_s \rightarrow \infty$ which exactly describe the dynamic characteristics of the test structure, then according to dynamic condensation theory, the mathematical model ($[M^R(\omega)]_{NxN}$ and $[K^R]_{NxN}$ with $N<N_s$) with finite degrees of freedom which can exactly represent the test structure in terms of all these retained coordinates is not unique (the mass matrix is a function of frequency as can be seen from (7-13a)). On the other hand, during the condensation process, the physical connectivity of $[M]_{N_s x N_s}$, $[K]_{N_s x N_s}$ (which could themselves be fully populated matrices) has been destroyed, and in general, $[M^R(\omega)]_{NxN}$ and $[K^R]_{NxN}$ become fully populated. In the problem under consideration, $[M]_{N_s x N_s}$, $[K]_{N_s x N_s}$ are unknown and it is desired to identify constant $[M^R]_{NxN}$ and $[K^R]_{NxN}$ which can best represent the test structure. It is apparent therefore that such a constant coefficient model can only be an approximation with limits on the frequency range of applicability.

Although the exact model of the structure to be identified ($[M^R(\omega)]_{NxN}$ and $[K^R]_{NxN}$) is fully populated and the mass matrix is, in theory, a function of frequency, experience of Finite Element Analysis shows that as far as the lower modes of the structure, which are usually of practical interests, are concerned, a model with constant mass and stiffness matrices which preserves the physical connectivity can, in general, accurately describe the test structure. Also, as shown in the following numerical calculation, a fully populated model is not necessarily a better representation of the structure than a model with physical connectivity preserved.

To illustrate the above arguments, a free-free beam structure as shown in Fig.7.15 is considered. Suppose the 'true' model of the structure is the FE model formulated by discretising the beam into 50 elements and consider 2 DOFs at each node: this yields $[M]_{102x102}$, $[K]_{102x102}$. A reduced model of $[M^R]_{52x52}$, $[K^R]_{52x52}$ corresponding to 52 'measured' coordinates (hatched nodes) is to be identified which can best represent the structure. One choice of such a model is, naturally, the Guyan-reduced model with these measured coordinates considered as master degrees of freedom. A thus-obtained model is fully populated. Another choice of a suitable model is the FE representation of the structure composed of 25 elements with all the 'measured' coordinates being considered. This model preserves the physical connectivity (and is heavily banded). The question to be answered is: which model represents the structure better? In order to make this comparison, the first 10 natural frequencies of the 'true' model, the Guyan-reduced model and the model with 25 elements are tabulated in Table 7.1 from which it can be seen that although the Guyan-reduced model and the model with 25 elements are quite different in terms of their spatial forms (one is fully populated and another is heavily banded), the first 10 modes are almost the same (and, indeed, for all 52 modes, the

eigenvalues are the same up to 14 digits when double precision computation is considered). In addition, the eigenvectors of the true model and the model with 25 elements are the same (up to 14 digits) for all the corresponding modes and coordinates. The MAC values of the corresponding eigenvectors of the 'true' model and the Guyan-reduced model and of the true model and the model with 25 elements are shown in Fig.7.16. This suggests that it is possible for a constant coefficient model to represent the structure accurately while at the same time preserving its physical connectivity, a feature which is essential from the point of view of analytical model updating.



$S = 0.004\ (m^2)$ $I = 0.001\ (m^4)$ $E = 5.0 \times 10^{13}\ (N/m)$ $\rho = 2800\ (kg/m^3)$

Fig.7.15 - Free-free Beam Structure

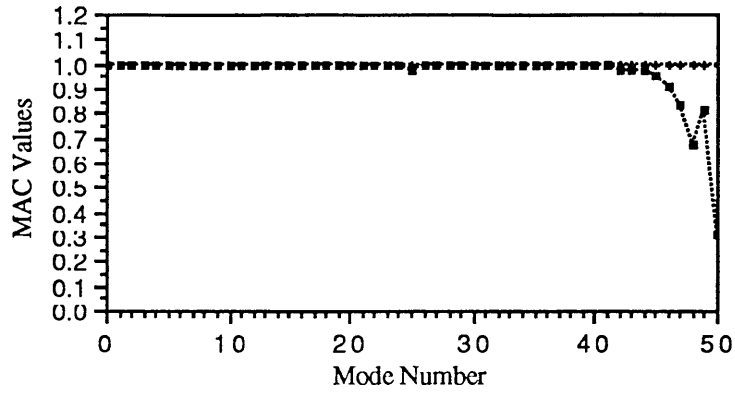
mode no.	1	2	3	4	5
ω_e	39.0937	107.763	211.259	349.224	521.683
ω_g	39.0938	107.764	211.265	349.247	521.759
ω_f	39.0938	107.764	211.265	349.247	521.759
mode no.	6	7	8	9	10
ω_e	728.639	970.096	1246.06	1556.54	1901.57
ω_g	728.842	970.570	1247.06	1558.46	1905.02
ω_f	728.842	970.570	1247.06	1558.46	1905.02

ω_e: exact natural frequency of the structure (Hz)

ω_g: natural frequency of the Guyan-reduced model (Hz)

ω_f: natural frequency of the model with 25 elements (Hz)

Table 7.1 - Comparison of Natural Frequencies of Three Different Models



+-+-+- 'true' and 25 element models, •–•–•– 'true' and Guyan-reduced models

Fig.7.16 - MAC values indicating the correlation of corresponding eigenvectors

7.5 MINIMUM DATA REQUIRED TO UPDATE AN ANALYTICAL MODEL

So far, it has been established that in order to update an analytical model using limited measured data, it is essential that the physical connectivity of the analytical model should be preserved. The mathematical validity of preserving such physical connectivity has been demonstrated. With this connectivity information available, it will be shown that although the measured data may be quite limited in practice, the updating problem can become overdetermined in most cases. In what follows, the criteria for the minimal measured data required to solve an updating problem uniquely are discussed.

Although many different methods have been developed to update an analytical model using measured data, there are no clear rules concerning just how much data should be measured (how many modes and coordinates) in order to solve the updating problem uniquely. Such a criterion is important not only because modal testing is costly, but more importantly because it is required for the analyst to judge whether an available set of modal data is enough to obtain a unique solution. In the following, a model updating method based on eigendynamic properties (Eigendynamic Constraint Method ECM) is developed and, based on this method, a criterion of how many modes should be measured in order to have a unique solution of the updating problem is established for the case when the measured coordinates are complete. When the measured coordinates are *incomplete* but the modelling errors are localised in the measured coordinates, a direct solution of the updating problem is still possible based on the Eigendynamic Constraint

Method. In general, however, when the measured coordinates are incomplete, a direct solution is not possible and some kind of perturbation or sensitivity analysis based on an iterative scheme has to be sought. In this case, an Inverse Eigensensitivity Analysis Method is employed to establish the criterion of how many modes and coordinates should be measured in order to solve the updating problem.

7.5.1 THE EIGENDYNAMIC CONSTRAINT METHOD

The problem of reconstructing system matrices from identified eigenvalues and eigenvectors has been considered by some authors [97-99]. Assuming a certain form of the modification matrix such as diagonal matrix, and using all the identified eigenvalues of the modified system, the modification matrix can be calculated [97-98]. Applications have been found for such studies in the solution of inverse Sturm-Liouville problems and nuclear spectroscopy [98]. In reference [99], Gladwell introduced and extended the theory of [97-98] to the identification of vibrational systems by using both measured eigenvalues and eigenvectors to reconstruct the mass and stiffness matrices. However, his analysis is restricted to his specific fixed-free mass-spring chain system because during the development of his arguments, he assumed that the system's eigenvalue problem is in the form of Jacobian matrices [99]. Based on this simple system, he established the necessary and sufficient conditions for a given vector to be one of the eigenvectors of the system and pointed out that it is possible to reconstruct the mass and stiffness matrices of the system by using two identified modes which satisfy certain conditions although the thus reconstructed system is not unique in the sense that it can be scaled by an arbitrary factor. Ibrahim [100] later extended Gladwell's theory into analytical model updating of dynamic structures. The Eigendynamic Constraint Method described below is similar to Ibrahim's method. However, since it makes use of the mass-normalisation properties of measured modes, the problem of uniqueness of the identified system is resolved.

The method is formulated based on the eigendynamic equation and the mass normalisation properties of measured modes as mentioned in §7.2.3,

$$- (\lambda_x)_i [\Delta M] \{\phi_x\}_i + [\Delta K] \{\phi_x\}_i = (\lambda_x)_i [M_a] \{\phi_x\}_i - [K_a] \{\phi_x\}_i \tag{7-7}$$

$$\{\phi_x\}_i^T [\Delta M] \{\phi_x\}_i = 1 - \{\phi_x\}_i^T [M_a] \{\phi_x\}_i \tag{7-8}$$

As discussed before, the physical connectivity of the analytical model should be preserved during the updating process, and so the updated model should have the same

connectivity as that of the analytical model. When the connectivity information is employed, (7-7) and (7-8) can be combined and turned into standard linear algebraic equations in terms of an unknown vector $\{P\}_{Lx1}$ consisting of all the design variable changes in the mass and stiffness matrices in a similar way to that discussed in Chapter 6, as

$$[A_i]_{(N+1)xL}\ \{P\}_{Lx1} = \{b_i\}_{(N+1)x1} \tag{7-15}$$

where N is the number of degrees of freedom specified in the analytical model, L is the number of total independent design variables in the mass and stiffness matrices, $[A_i]_{(N+1)xL}$ and $\{b_i\}_{(N+1)x1}$ are the coefficient matrix and vector formed using the analytical model and i^{th} measured mode properties.

To see how the coefficient matrix $[A_i]$ and $\{b_i\}$ can be obtained, again consider the 2DOF mass-spring system as mentioned in Chapter 6. When the first mode is used, equation (7-7) becomes

$$-(\lambda_x)_1\begin{bmatrix}\Delta m_1 & 0\\ 0 & \Delta m_2\end{bmatrix}\begin{Bmatrix}{}_x\phi_{11}\\ {}_x\phi_{21}\end{Bmatrix}+\begin{bmatrix}\Delta k_1+\Delta k_2 & -\Delta k_2\\ -\Delta k_2 & \Delta k_2+\Delta k_3\end{bmatrix}\begin{Bmatrix}{}_x\phi_{11}\\ {}_x\phi_{21}\end{Bmatrix}=$$

$$\begin{bmatrix}(\lambda_x)_1 m_1-k_1-k_2 & k_2\\ k_2 & (\lambda_x)_1 m_2-k_2-k_3\end{bmatrix}\begin{Bmatrix}{}_x\phi_{11}\\ {}_x\phi_{21}\end{Bmatrix}=$$

$$\begin{Bmatrix}((\lambda_x)_1 m_1-k_1-k_2)\ {}_x\phi_{11}+k_2\ {}_x\phi_{21}\\ k_2\ {}_x\phi_{11}+((\lambda_x)_1 m_2-k_2-k_3)\ {}_x\phi_{21}\end{Bmatrix}=\begin{Bmatrix}b_1\\ b_2\end{Bmatrix} \tag{7-7a}$$

After further mathematical manipulation, (7-7a) becomes

$$\begin{bmatrix}-(\lambda_x)_1\ {}_x\phi_{11} & 0 & {}_x\phi_{11} & {}_x\phi_{11}-{}_x\phi_{21} & 0\\ 0 & -(\lambda_x)_1\ {}_x\phi_{21} & 0 & {}_x\phi_{21}-{}_x\phi_{11} & {}_x\phi_{21}\end{bmatrix}\begin{Bmatrix}\Delta m_1\\ \Delta m_2\\ \Delta k_1\\ \Delta k_2\\ \Delta k_3\end{Bmatrix}=\begin{Bmatrix}b_1\\ b_2\end{Bmatrix} \tag{7-7b}$$

Similarly, (7-8) becomes

$$\{ {}_x\phi_{11} \;\; {}_x\phi_{11} \} \begin{bmatrix} \Delta m_1 & 0 \\ 0 & \Delta m_2 \end{bmatrix} \begin{Bmatrix} {}_x\phi_{11} \\ {}_x\phi_{21} \end{Bmatrix} = 1 - \{ {}_x\phi_{11} \;\; {}_x\phi_{11} \} \begin{bmatrix} m_1 & 0 \\ 0 & m_2 \end{bmatrix} \begin{Bmatrix} {}_x\phi_{11} \\ {}_x\phi_{21} \end{Bmatrix} = b_3 \qquad (7\text{-}8a)$$

Further, (7-8a) can be written as:

$$\{ {}_x\phi_{11} \, {}_x\phi_{11} \;\; {}_x\phi_{21} \, {}_x\phi_{21} \;\; 0 \;\; 0 \;\; 0 \} \begin{Bmatrix} \Delta m_1 \\ \Delta m_2 \\ \Delta k_1 \\ \Delta k_2 \\ \Delta k_3 \end{Bmatrix} = b_3 \qquad (7\text{-}8b)$$

Combining (7-7b) and (7-8b), coefficient matrix $[A]_{3x5}$ and $\{b\}_{3x1}$ can be obtained.

Equation (7-15) is obtained based on the i^{th} measured mode; when m measured modes are available, it is not difficult to see that the dimension of the coefficient matrix [A] becomes m(N+1)xL and {b} becomes m(N+1)x1, that is:

$$[A]_{m(N+1)xL}\{P\}_{Lx1} = \{b\}_{m(N+1)x1} \qquad (7\text{-}16)$$

In general, when $m(N+1) \geq L$, equation (7-16) becomes overdetermined and the SVD technique can be used to solve the unknown vector {P}. After {P} has been calculated, the updated mass and stiffness matrices can be reconstructed by using the physical connectivity of the analytical model.

7.5.2 INVERSE EIGENSENSITIVITY ANALYSIS

As mentioned previously, when the measured coordinates are incomplete, direct solution of the updating problem is generally not possible and some kind of perturbation or sensitivity analysis has to be employed based on an iterative scheme. In this case, in order to establish a criterion concerning how many modes and coordinates should be measured in order to solve uniquely the updating problem, the Inverse Eigensensitivity Analysis method [80] can be employed to calculate the design variable changes given differences between analytical and measured natural frequencies and modeshapes. The method was first introduced to analytical model improvement by Collin et al [80]. Later, Chen/Garba [81] modified the procedure described in [80] by introducing matrix perturbation to avoid the eigensolution required in every iteration. Lallement [82] recently extended the method to pinpoint where the significant modelling errors are first and then

to reduce the number of unknowns to improve the solution condition. Derivation of the eigenvalue and eigenvector derivatives which are required in the formulation of updating problem is explained in the Appendix II of this thesis.

From the theory of the algebraic eigenvalue problem, a system's eigenvalues and eigenvectors are implicit functions of its design variables. Hence, based on the Taylor series representation, the relationship between the change of modal parameter δ (δ can be the change of any eigenvalue or of any eigenvector element) and the vector of design variable change $\{P\}$ can be expressed as

$$\delta = \sum_{i=1}^{N} s_i P_i + \sum_{i=1}^{N} \sum_{j=1}^{N} v_{ij} P_i P_j + \ldots \tag{7-17}$$

where s_i and v_{ij} are the first- and second-order sensitivity coefficients. Suppose n out of N coordinates have been measured for the i^{th} mode then, based on (7-17), to a first order approximation, we have:

$$\begin{bmatrix} \frac{\partial\{\phi_a\}_i}{\partial P_1} & \frac{\partial\{\phi_a\}_i}{\partial P_2} & \cdots & \frac{\partial\{\phi_a\}_i}{\partial P_L} \\ \frac{\partial(\lambda_a)_i}{\partial P_1} & \frac{\partial(\lambda_a)_i}{\partial P_2} & \cdots & \frac{\partial(\lambda_a)_i}{\partial P_L} \end{bmatrix} \begin{Bmatrix} P_1 \\ P_2 \\ \cdot \\ \cdot \\ P_L \end{Bmatrix} = \begin{Bmatrix} \{\phi_x\}_i - \{\phi_a\}_i \\ \\ (\lambda_x)_i - (\lambda_a)_i \end{Bmatrix} \tag{7-18}$$

$$[S_i]_{(n+1)xL} \{P\}_{Lx1} = \{\delta_i\}_{(n+1)x1} \tag{7-18a}$$

where $[S_i]$ is the sensitivity matrix for n measured coordinates and one measured eigenvalue of the i^{th} mode, and can be calculated in a way as illustrated in Appendix II of this thesis, and $\{\delta_i\}$ is the difference vector between the measured and analytical eigenvector and eigenvalue of the i^{th} mode. If m measured modes are available, (7-18a) can be re-written as

$$[S]_{m(n+1)xL} \{P\}_{Lx1} = \{\delta\}_{m(n+1)x1} \tag{7-19}$$

When $m(n+1) \geq L$, (7-19) becomes a set of overdetermined linear algebraic equations and the SVD technique can be used to solve $\{P\}$. Since (7-19) is formulated based on first-order approximation, the exact solution of $\{P\}$ cannot be obtained directly and an iterative procedure has to be introduced as illustrated in Fig.7.16. Again, after $\{P\}$ has been calculated, the updated model can be reconstructed using the the analytical model.

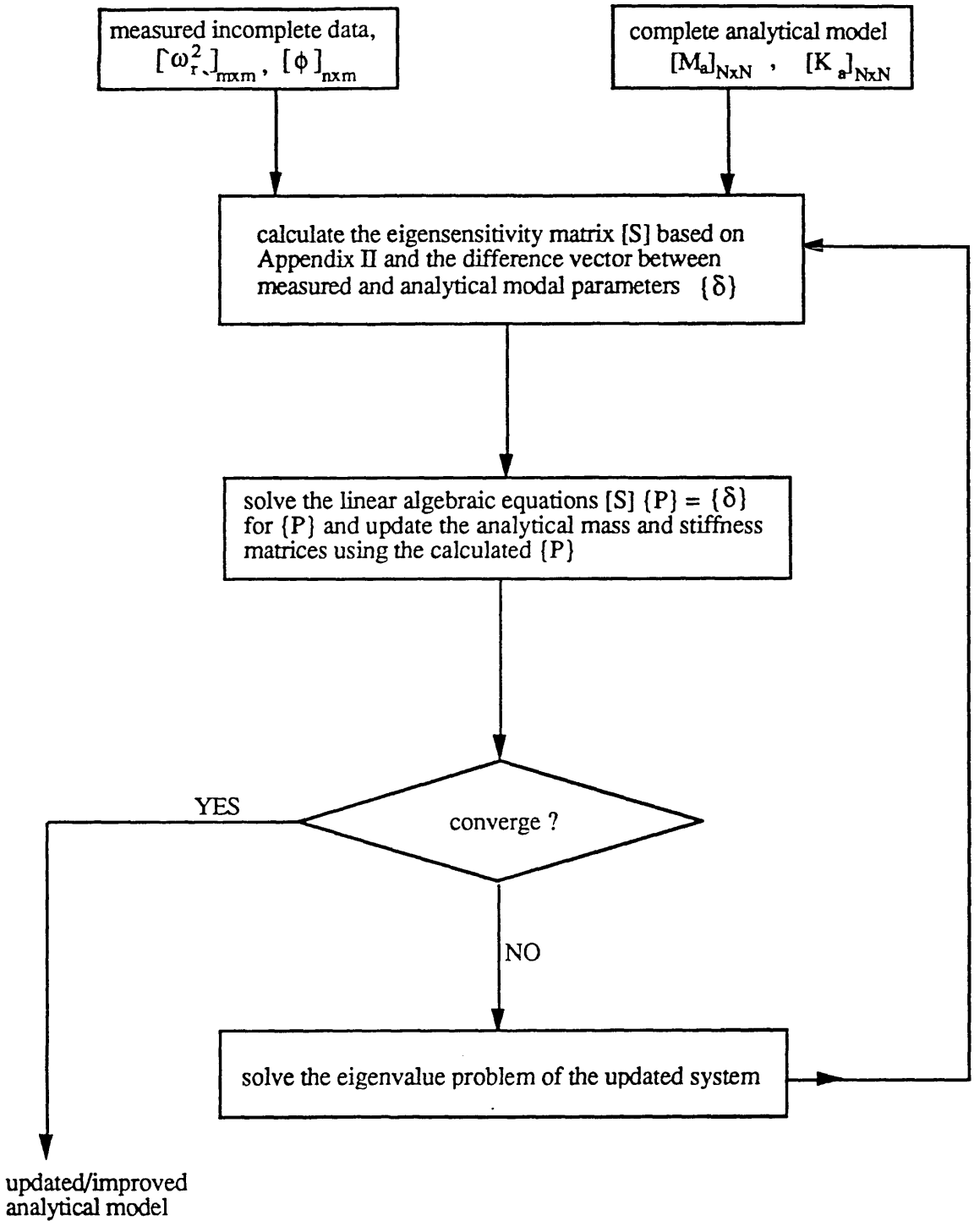


Fig.7.16 - Updating Process Using Inverse Eigensensitivity Analysis

7.5.3 NUMBER OF MODES AND COORDINATES REQUIRED TO UPDATE AN ANALYTICAL MODEL

When the measured coordinates are complete, it can be seen from equation (7-16) that in order to have a unique solution of the updating problem, the following condition has to be satisfied:

$$m(N+1) \geq L \tag{7-20}$$

where, as mentioned, m is the number of measured modes, N is the number of degrees of freedom specified in the analytical model and L is the number of independent design variables. Therefore the number of modes required (m) in order to solve the updating problem is $m \geq L/(N+1)$ or the minimum number of modes is $m=E(L/(N+1))$ where E(x) is a mathematical function such that when it is applied to a real number x, it returns an integer value which is $x \leq E(x) < x+1$.

When the measured coordinates are *incomplete*, the Eigendynamic Constraint Method cannot generally be applied directly. However, in the special case in which all the modelling errors are localised in the measured coordinates, the unmeasured coordinates can be interpolated exactly based on the analytical model itself using Kidder's expansion method, as discussed before. The thus-expanded modes can then be used in the Eigendynamic Constraint Method to update the analytical model. For this case, the criterion for the minimum number of measured modes is therefore the same as is given in (7-20).

In general, when the measured coordinates are incomplete, there exists a criterion concerning the number of modes and coordinates which should be measured in order to update the analytical model uniquely. From (7-19), in order to have a unique solution of the updating problem, the following criterion has to be satisfied:

$$m(n+1) \geq L \tag{7-21}$$

where n is the number of measured coordinates. Expression (7-21) is the necessary condition, and as will be shown in the numerical case study, it is also sufficient when only mass or stiffness modelling errors are considered. When the modelling errors are present in both the mass and stiffness matrices, the situation becomes more complicated and it is suggested that in this case an error location procedure be employed in order to

reduce the number of unknowns in {P} before the calculation can be carried out based on the Inverse Eigensensitivity Analysis.

7.5.4 NUMERICAL CASE STUDIES

In order to verify numerically the criteria presented above, a mass-spring system shown in Fig.7.17 is considered. The system has 10 degrees-of-freedom and consists of 21 design variables (L=21, 10 mass elements and 11 stiffness elements).

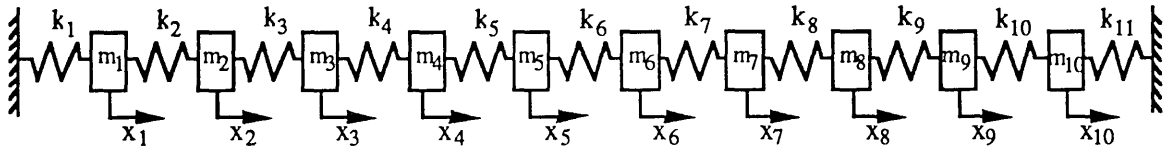


Fig.7.17 - A 10 DOF Mass-spring System

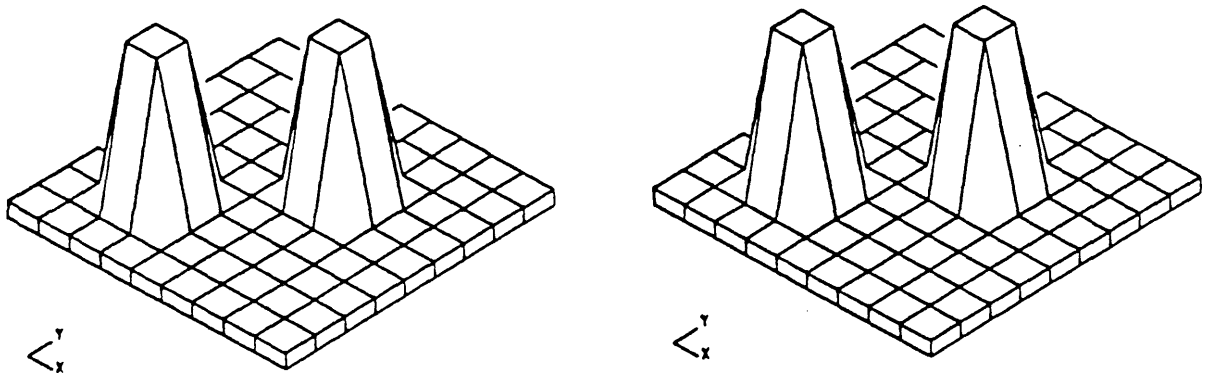
When the measured coordinates are complete, according to equation (7-20), m=E(21/11)=2. This means that for this specific mass-spring system, two measured modes with complete coordinates are in principle sufficient to identify all these 10 mass elements and 11 stiffness elements. Numerical results for the identification based on the Eigendynamic Constraint Method using the first and second 'measured' modes are shown in Table 7.2 (kg for mass and N/m for stiffness).

Variables	m_1	m_2	m_3	m_4	m_5	m_6	m_7
Exact	10.00000	11.00000	12.00000	13.00000	14.00000	14.00000	13.00000
Identified	10.00001	11.00000	12.00000	12.99999	14.00000	14.00000	12.99999
Variables	m_8	m_9	m_{10}	k_1	k_2	k_3	k_4
Exact	12.00000	11.00000	10.00000	100000.0	100000.0	300000.0	100000.0
Identified	12.00000	11.00000	10.00001	100000.1	100000.1	300000.2	100000.1
Variables	k_5	k_6	k_7	k_8	k_9	k_{10}	k_{11}
Exact	100000.0	300000.0	100000.0	100000.0	100000.0	100000.0	300000.0
Identified	100000.1	300000.2	100000.1	100000.1	100000.1	100000.1	300000.2

Table 7.2 - Identification Results

Although, in general, the ECM is not directly applicable to the case when the measured coordinates are incomplete, it has been noted that in the special case when the modelling

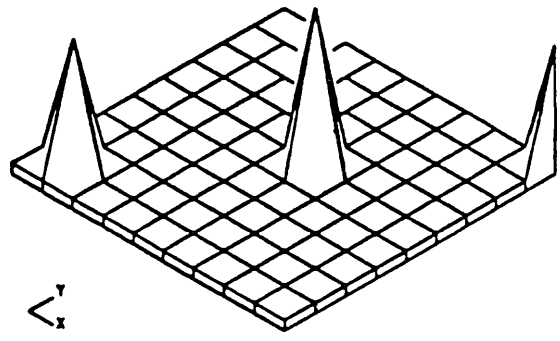
errors are localised in the measured coordinates, a direct solution of the updating problem is still possible because in this case, the unmeasured coordinates can be exactly interpolated based on the analytical model itself. In this numerical example, the exact stiffness modelling errors are shown in Fig.7.18(a) with a 50% stiffness modification in k_4 and k_7 (figure 7.17). Coordinates x_3, x_4, x_6, x_7 and x_9 are supposed to be measured, thereby including the stiffness errors introduced between coordinates x_3-x_4 and x_6-x_7. Again, the first two modes are used and the identified stiffness error matrix is exact, as shown in Fig.7.18(b).



(a) exact stiffness error matrix (b) identified stiffness error matrix

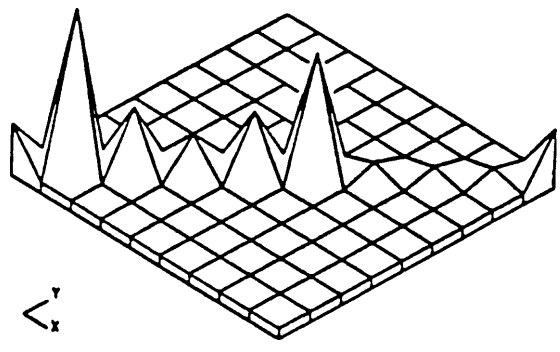
Fig.7.18 - Exact and Identified Stiffness Error Matrix (coordinates complete)

When mass modelling errors are considered, 10 design variables are taken into account in the Inverse Sensitivity Analysis (L=10 for 10 mass elements). Half the coordinates are supposed to be measured (all the odd-numbered coordinates). According to (7-21), the number of modes required in order to update the mass matrix can be calculated as m=E(10/6)=2. In the calculation, the first and second 'measured' modes are used. The mass modelling errors are introduced by modifying m_2, m_6 and m_{10} to 50% of their original values. The exact mass error matrix and the iteration results are illustrated in Fig.7.19(a) and Fig.7.19(b).

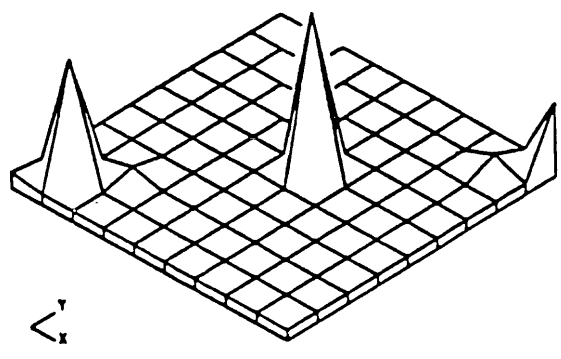


exact error mass matrix

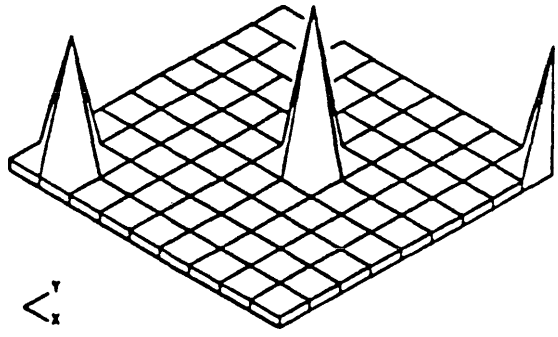
Fig.7.19(a) - Exact Mass Error Matrix



first iteration estimation



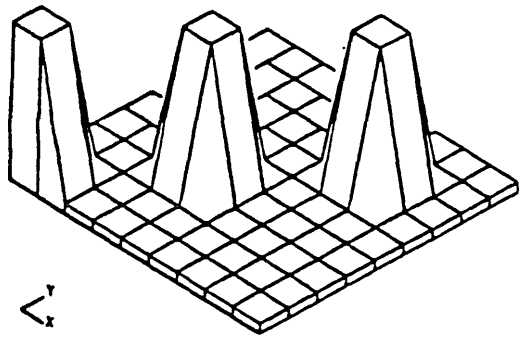
third iteration estimation



5th iteration estimation

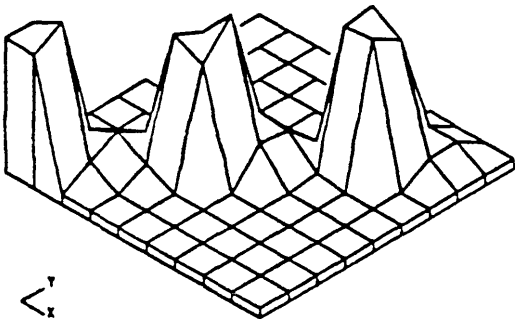
Fig.7.19(b) - Iteration Results (mass error case)

Similarly, in the case when stiffness modelling errors are considered, 11 design variables (11 stiffness elements) are taken into account in the Inverse Sensitivity Analysis. Again, all the odd numbered coordinates are assumed to be measured. According to (7-21), the number of modes required in order to update the stiffness matrix can be calculated as m=E(11/6)=2. The first two 'measured' modes are used in the calculation. The stiffness modelling errors are introduced by increasing k_2, k_5 and k_8 by 100% of the original values. The exact stiffness error matrix and the iteration results are shown in Fig.7.20(a) and Fig.7.20(b).

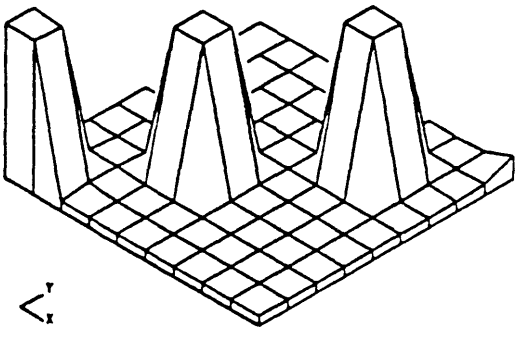


exact error stiffness matrix

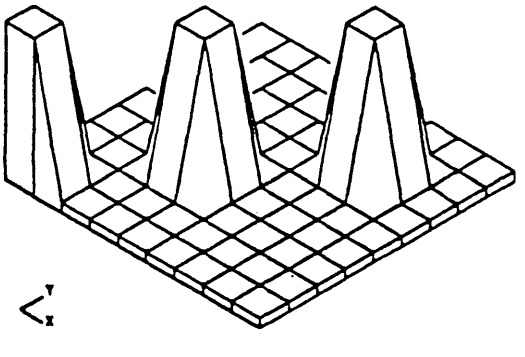
Fig.7.20(a) - Exact Mass Error Matrix



first iteration estimation



second iteration estimation



third iteration estimation

Fig.7.20(b) - Iteration Results (stiffness error case)

7.5.5 GENERALISATION OF THE CRITERION

The criteria concerning how much data should be measured in order to solve the updating problem uniquely have been developed based on the ECM and the Inverse Eigensensitivity Analysis method. Since, as discussed in detail in Chapter 6, the new

method developed based on the measured FRF data is a generalised version of the model updating methods based on measured modal data, it is therefore expected that such criteria can be generalised when measured FRF data are used in the updating process. In fact, as has been demonstrated in Chapter 6, a unique solution of the updating problem can always be obtained when measured coordinates are complete regardless of the number of measured modes, based on the method developed in Chapter 6. In the case when measured coordinates are incomplete, the method presented in Chapter 6 is based on a form of perturbation analysis and, as mentioned, such a perturbation analysis based the use of FRF data can be regarded as a generalisation of the Inverse Eiegensensitivity Analysis method presented in this Chapter. Therefore, the number of measured modes required to solve the updating problem using the method presented in Chapter 6 can be expected to be less than that required by the ECM in the case where measured coordinates are complete and, the Inverse Eigensensitivity Analysis in the case where measured coordinates are incomplete.

Also, it is perhaps worth mentioning that in the case when both mass and stiffness modelling errors exist, the criterion given in (7-21) based on the Inverse Eigensensitivity Analysis is not sufficient. It has been found that although (7-19) becomes largely overdetermined (e.g. the number of equations is twice as many as the number of unknowns), the condition of [S] in terms of its inverse is generally very poor when both mass and stiffness modelling errors are considered. This is probably because given modal parameter changes in certain modes might be achieved either by mass modification or stiffness modification and, as a result, the Inverse Eigensensitivity Analysis becomes mathematically uncertain. As shown in the numerical case studies of §6.4.4, this problem does not exist when FRF data are used in the method developed in Chapter 6.

7.5.6 APPLICATION OF THE METHODS TO THE GARTEUR STRUCTURES

As mentioned in Chapter 6, when a Finite Element model is considered, the mass and stiffness error matrices can in general be expressed as:

$$[\Delta M] = \sum_{s=1}^{N_1} \beta_s [m_s] \quad \text{and} \quad [\Delta K] = \sum_{s=1}^{N_1} \gamma_s [k_s] \tag{7-22}$$

where $[m_s]$ and $[k_s]$ are the s^{th} element mass and stiffness matrices which have been appropriately expanded to the global dimension of the system, N_1 is the total number of

elements and, β_s and γ_s are the design variable changes associated with the s^{th} element. With [ΔM] and [ΔK] being expressed as (7-22), the eigensensitivity coefficients $\frac{\partial\{\phi_a\}_i}{\partial\beta_j}$, $\frac{\partial\lambda_i}{\partial\beta_j}$, $\frac{\partial\{\phi_a\}_i}{\partial\gamma_j}$ and $\frac{\partial\lambda_i}{\partial\gamma_j}$ can be obtained using the analytical model (see Appendix II).

Based on the first order sensitivity analysis, the following relationship between the the changes of modal parameters and the changes of physical parameters can be established:

$$\begin{Bmatrix} \{\phi_x\}_1 - \{\phi_a\}_1 \\ (\lambda_x)_1 - (\lambda_a)_1 \\ \cdot \\ \{\phi_x\}_m - \{\phi_a\}_m \\ (\lambda_x)_m - (\lambda_a)_m \end{Bmatrix} = \begin{bmatrix} \frac{\partial\{\phi_a\}_1}{\partial\beta_1} & \cdot & \frac{\partial\{\phi_a\}_1}{\partial\beta_{N_1}} & \frac{\partial\{\phi_a\}_1}{\partial\gamma_1} & \cdot & \frac{\partial\{\phi_a\}_1}{\partial\gamma_{N_1}} \\ \frac{\partial\lambda_1}{\partial\beta_1} & \cdot & \frac{\partial\lambda_1}{\partial\beta_{N_1}} & \frac{\partial\lambda_1}{\partial\gamma_1} & \cdot & \frac{\partial\lambda_1}{\partial\gamma_{N_1}} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \frac{\partial\{\phi_a\}_m}{\partial\beta_1} & \cdot & \frac{\partial\{\phi_a\}_m}{\partial\beta_{N_1}} & \frac{\partial\{\phi_a\}_m}{\partial\gamma_1} & \cdot & \frac{\partial\{\phi_a\}_m}{\partial\gamma_{N_1}} \\ \frac{\partial\lambda_m}{\partial\beta_1} & \cdot & \frac{\partial\lambda_m}{\partial\beta_{N_1}} & \frac{\partial\lambda_m}{\partial\gamma_1} & \cdot & \frac{\partial\lambda_m}{\partial\gamma_{N_1}} \end{bmatrix} \begin{Bmatrix} \beta_1 \\ \cdot \\ \cdot \\ \cdot \\ \beta_{N_1} \\ \gamma_1 \\ \cdot \\ \cdot \\ \cdot \\ \gamma_{N_1} \end{Bmatrix} \quad (7\text{-}23)$$

where m experimental modes are supposed to have been measured. Again, the measured coordinates are assumed to be incomplete, i.e. $\{\{\phi_x\}_i - \{\phi_a\}_i\}_{nx1}$ where n<N. When a sufficient number m of modes are measured, (7-23) becomes overdetermined and the SVD technique can be used to solve for β_s and γ_s ($s=1,N_1$) iteratively.

To assess the practical applicability of the Inverse Eigensensitivity Analysis method, the fixed-free GARTEUR structure shown in Fig.7.20 was investigated. Only stiffness errors are introduced by overestimating the 10^{th} (nodes 9-10), 14^{th} and 15^{th} (nodes 14-16), 30^{th} (nodes 29-30) and 66^{th} and 67^{th} (nodes 62, 14, 63) element stiffness matrices by 100% (figure 7.21). The hatched nodes shown in Fig.7.21 with their translational degrees of freedom (u,v) are supposed to be measured. According to (7-21), the minimum number of modes required to update the analytical model is m=E(L/(n+1)) =E(78/(72+1))=2 (where L=78 is the total number and so the total number of design variables according to (7-22) when stiffness modelling errors are considered and n=72 is the number of measured coordinates (2 translational degrees at 36 nodes)). During the calculation, the first 4 modes are supposed to have been measured (twice as many as the minimum number required). The exact stiffness error matrix is shown in Fig.7.22 and the iteration results are shown in Fig.7.23-7.25.

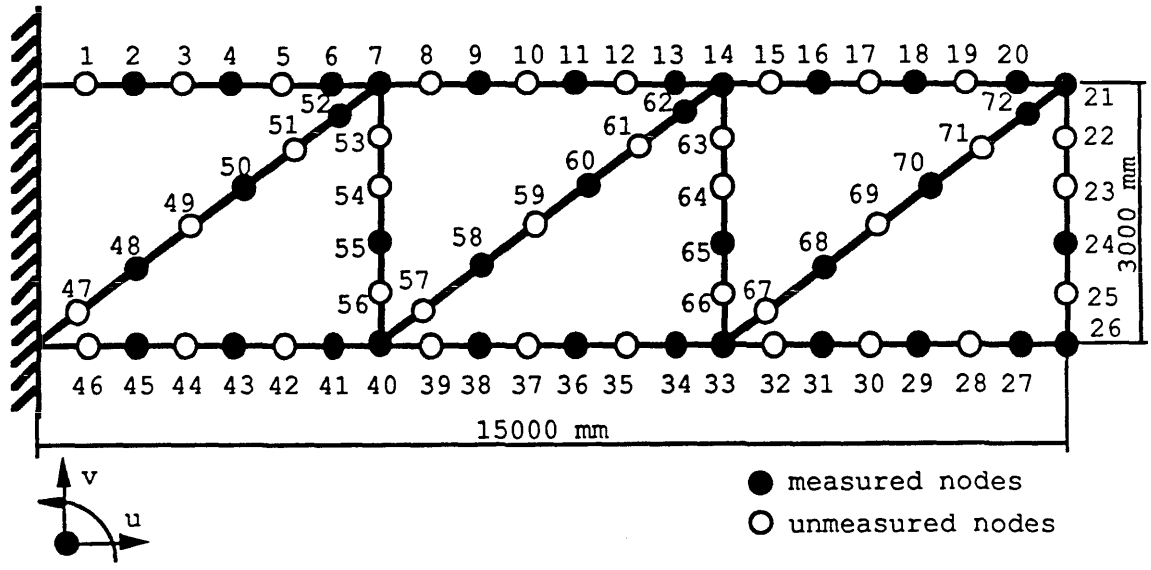


Fig.7.21 - The Fixed-free GARTEUR Structure

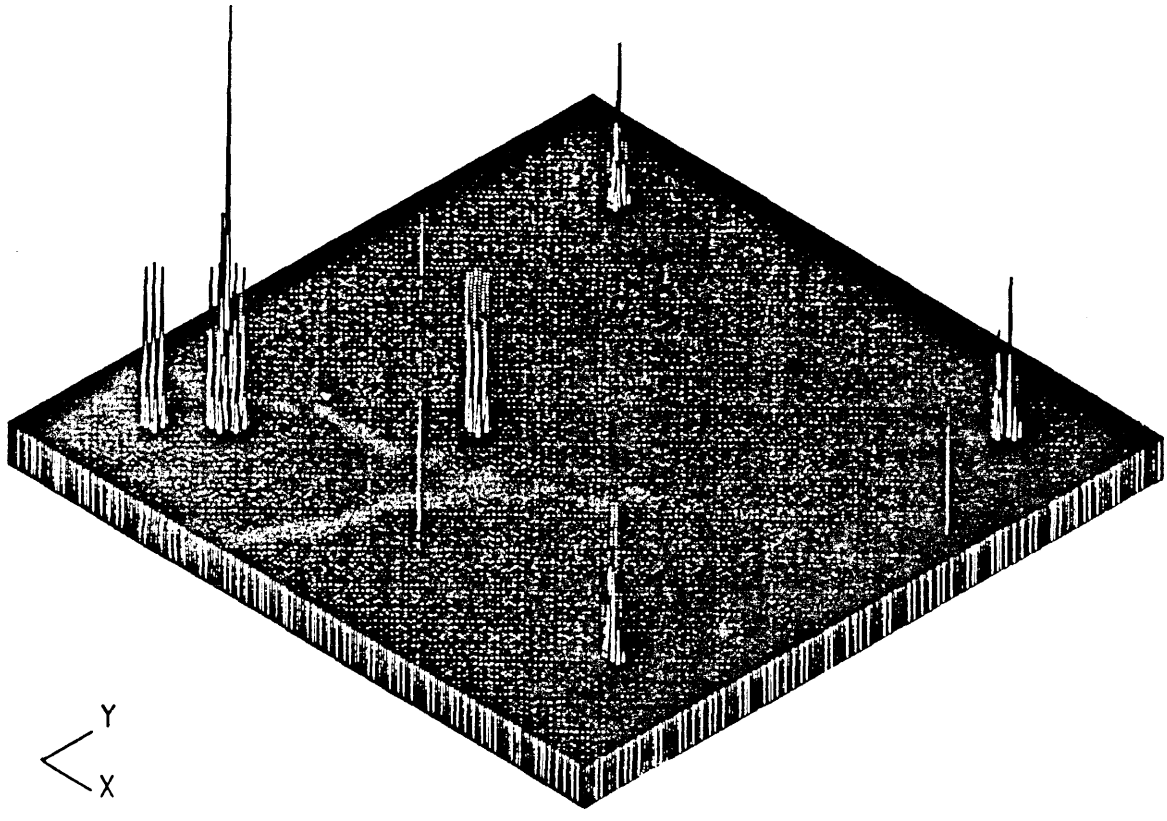


Fig.7.22 - Exact Stiffness Error Matrix

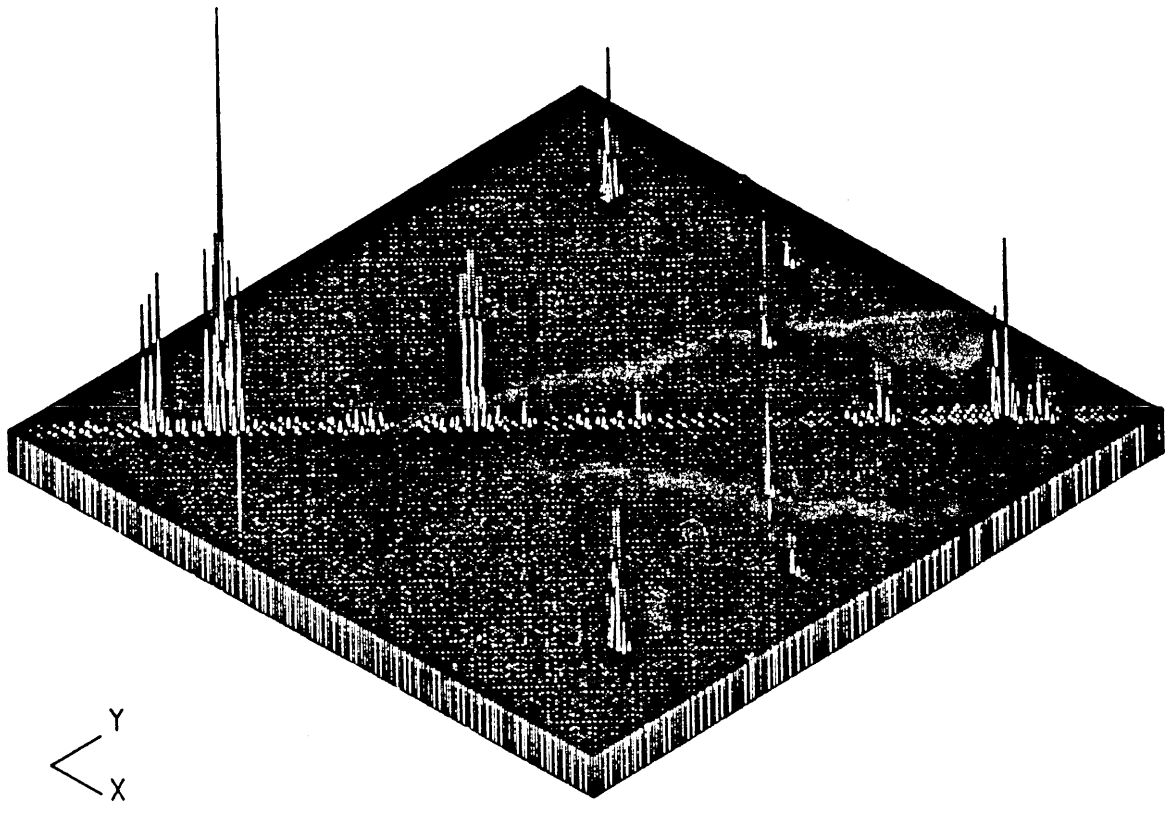


Fig.7.23 - First Iteration Results

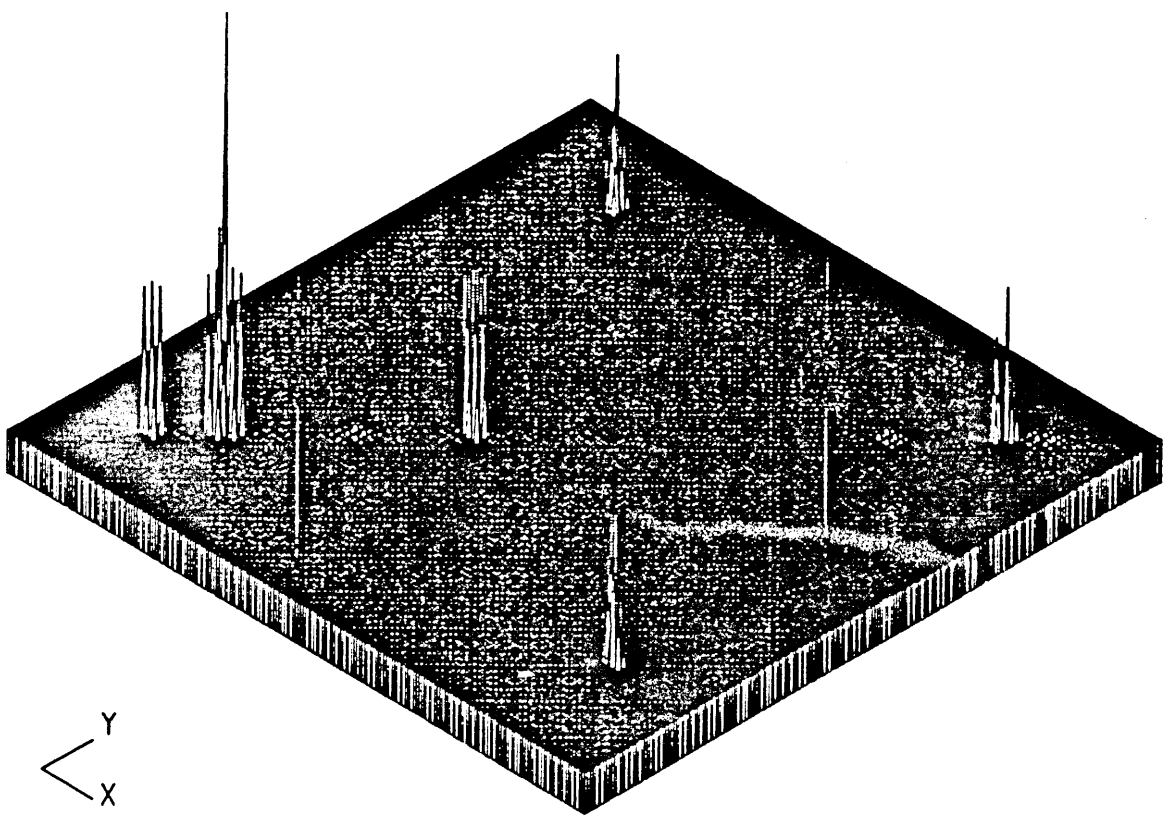


Fig.7.24 - Second Iteration Results

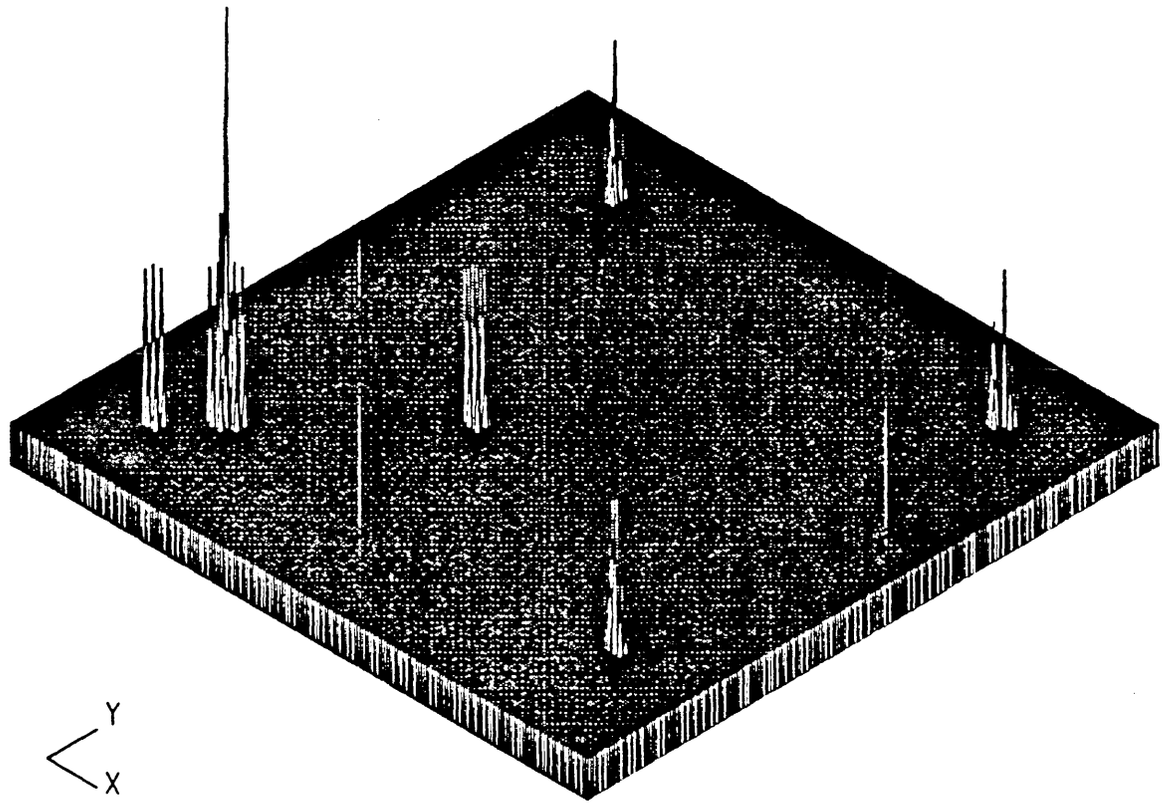


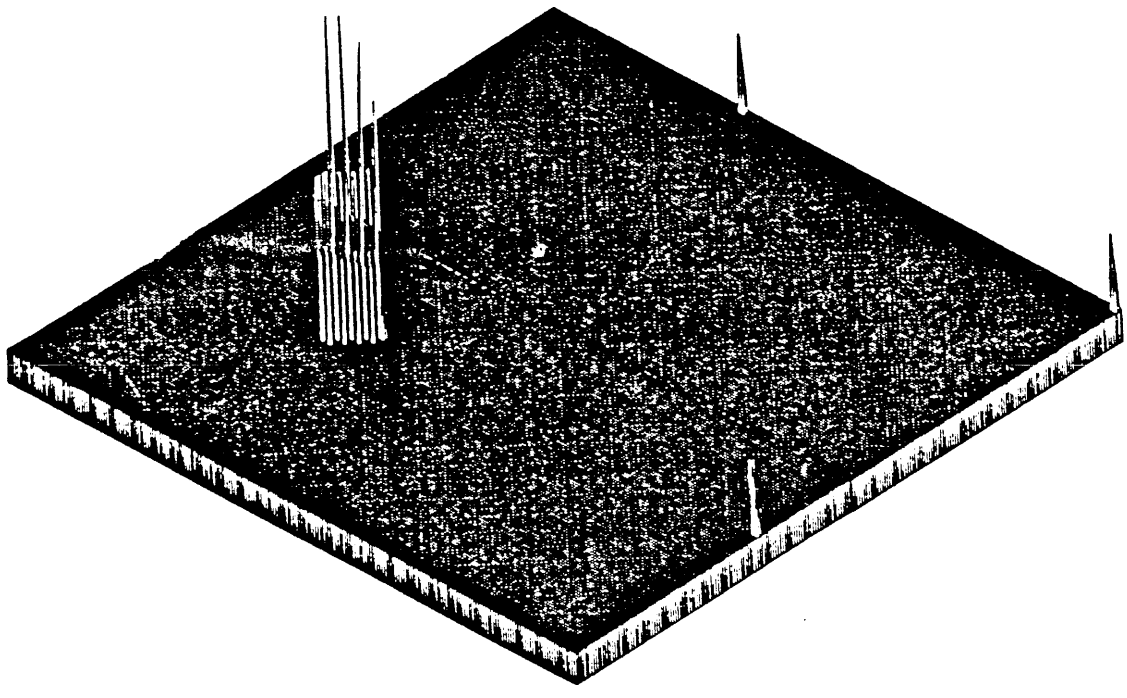
Fig.7.25 - Third Iteration Results

In some cases, it may not be appropriate to assume that the mass and stiffness modelling errors can be simply expressed as that of (7-22) and more independent physical design variables need to be considered. To illustrate this point, the free-free GARTEUR structure (figure 7.1) is also investigated. Every element shown in figure 7.1 is a superposition of an axial bar element and a bending beam element and these pairs of elements are considered to be independent of each other. Physically, this means that during the modelling, the cross-section area of the axial bar element and the second moment of area of the bending beam element are treated independently. As a result, the stiffness error matrix can in general be expressed as:

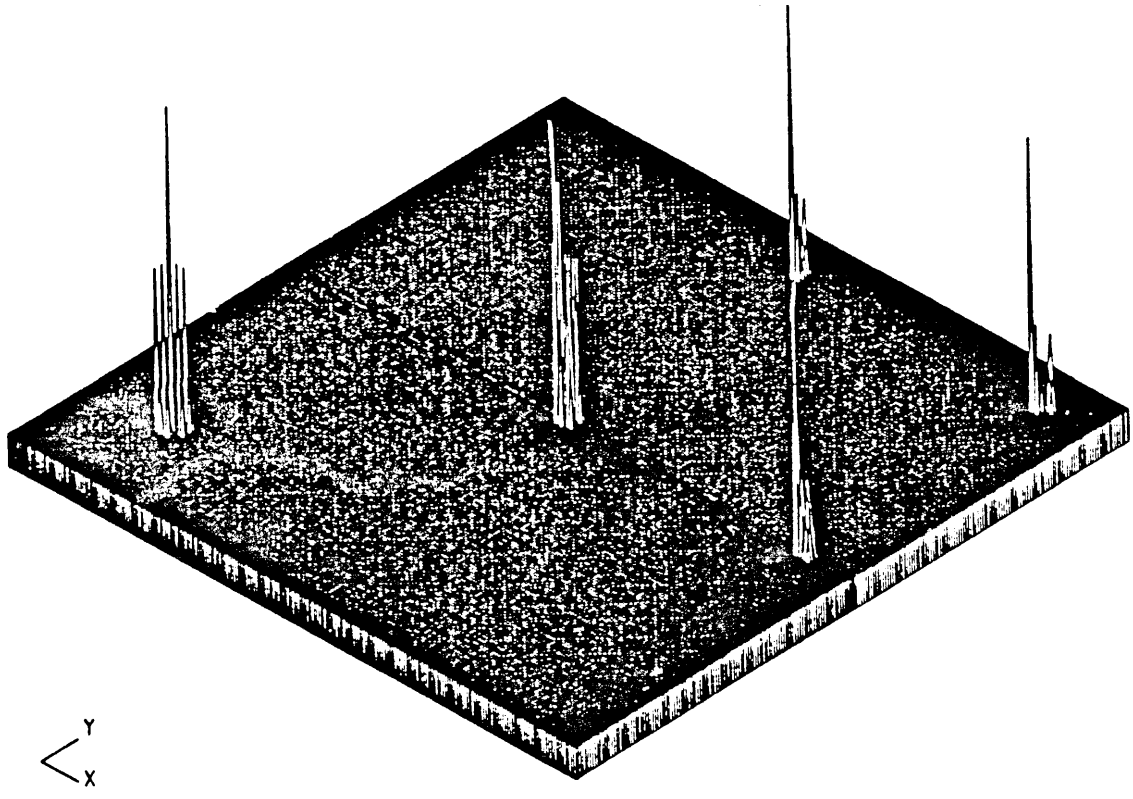
$$[\Delta K] = \sum_{r=1}^{N_l} \beta_s [k_s^e] + \gamma_s [k_s^b] \qquad (7\text{-}24)$$

where $[k_r^e]$ is the expanded extensional element stiffness matrix (which is proportional to the cross-section area) and $[k_r^b]$ is the expanded bending element stiffness matrix (which

is proportional to the second moment of inertia). Since, in this exercise, the number of elements is N_1=83, 166 unknowns need to be calculated in each iteration when only stiffness errors are considered. The stiffness errors are introduced by overestimating the 11st, 12nd, 39th, 76th and 77th bending element stiffness matrices and the 23rd, 24th, 25th, 26th and 83rd extensional element stiffness matrices by 100%. The exact stiffness error matrix is shown in Fig.7.26 (the stiffness errors for the bending and extension are of 100 times different in magnitudes and therefore they are plotted separately). Some of the 'experimental' and analytical receptance FRFs are shown in Fig.7.27. For comparison, the method developed in Chapter 6 is employed in this case. One incomplete column - all the hatched nodes with their translational degrees of freedom (u,v) - of the receptance matrix of the 'experimental' model with data covering the first 5 modes is supposed to have been measured. In each iteration, 20 FRF data points which are randomly chosen in the 'measured' frequency range are used and the iteration results are shown in Fig.7.28-7.31.

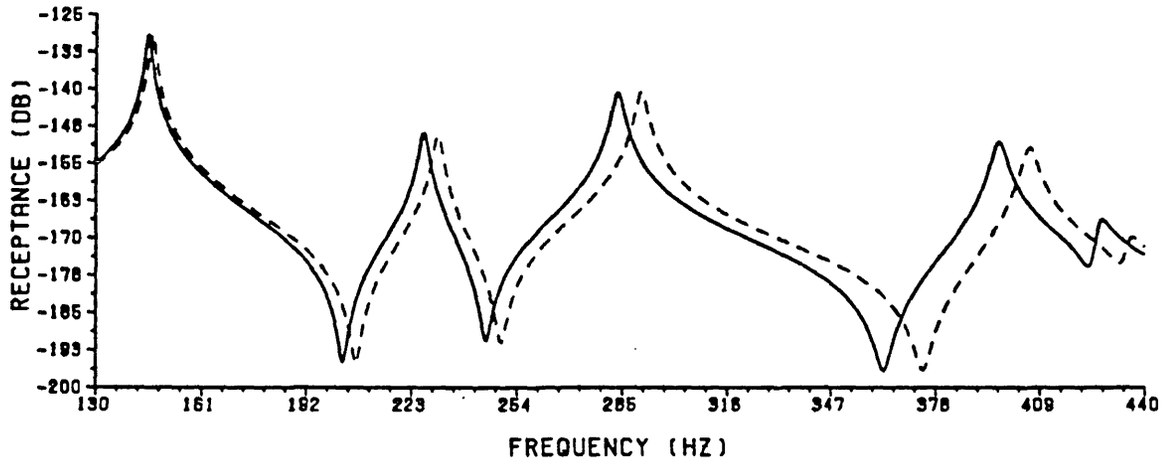


extensional stiffness

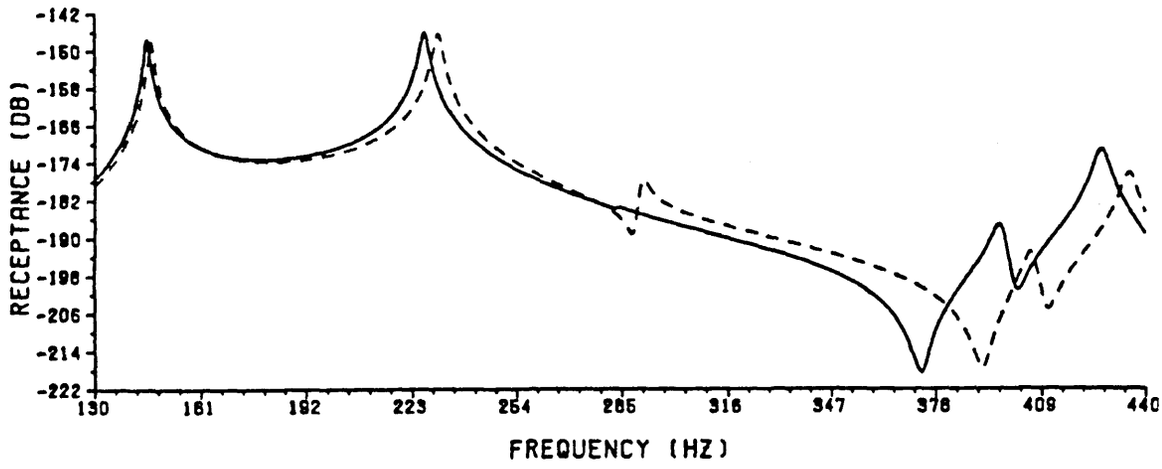


bending stiffness

Fig.7.26 - Exact Stiffness Error Matrix

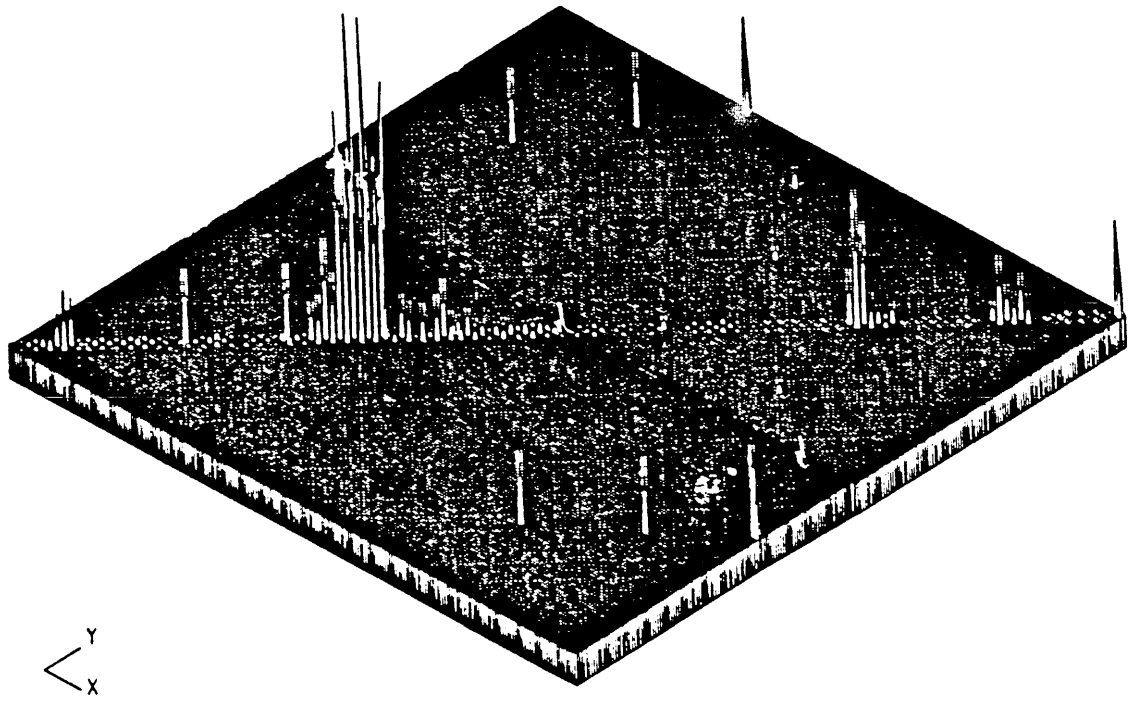


point receptance $\alpha_{11}(\omega)$

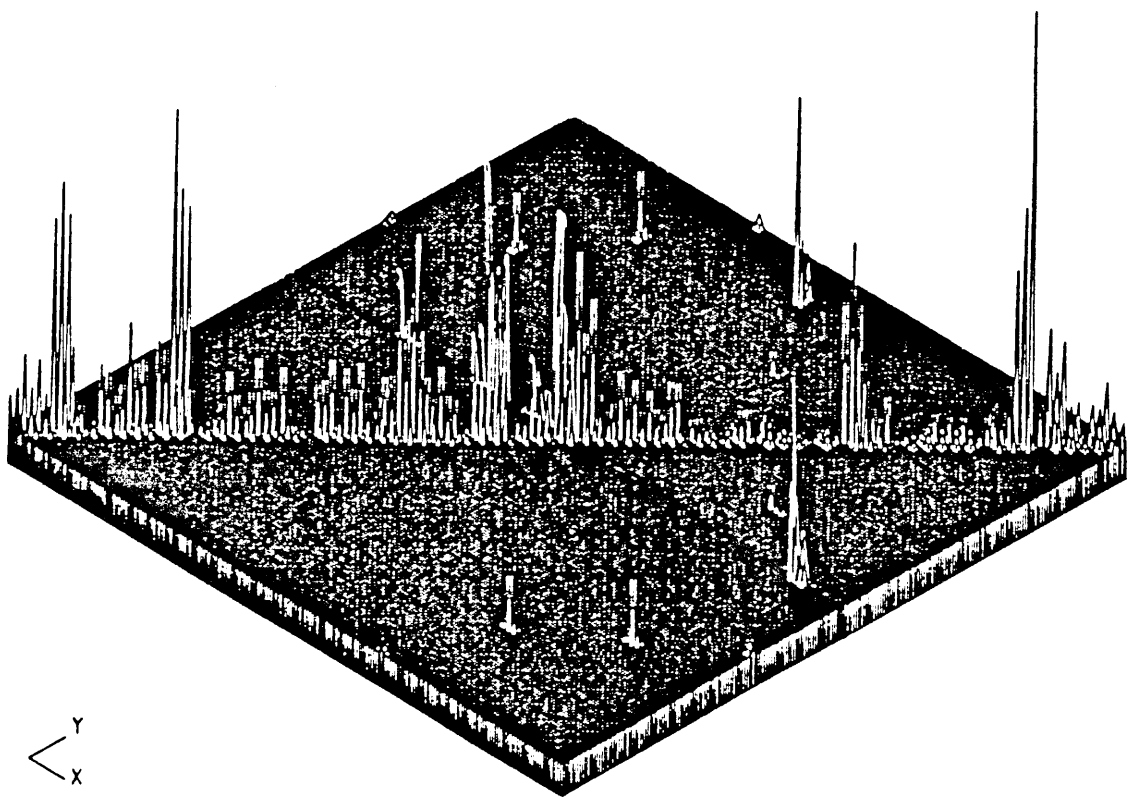


transfer receptance $\alpha_{1,50}(\omega)$

Fig.7.27 - FRFs of Analytical and 'Experimental' Models
——— analytical, - - - - - - - - - - 'experimental'

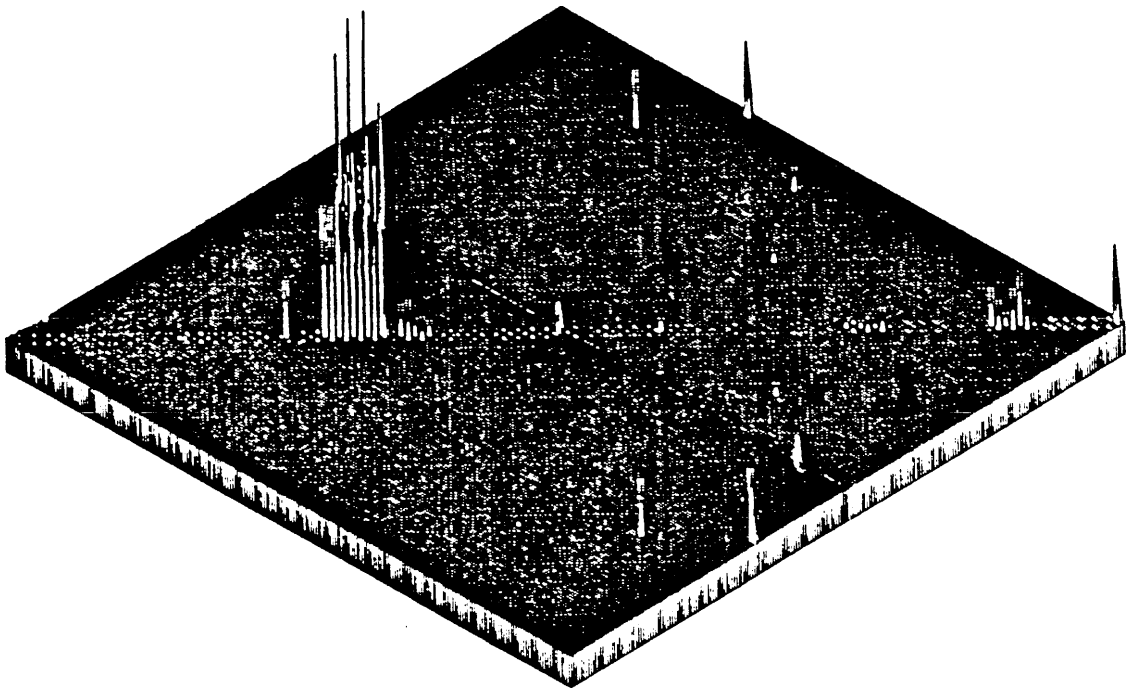


extensional stiffness

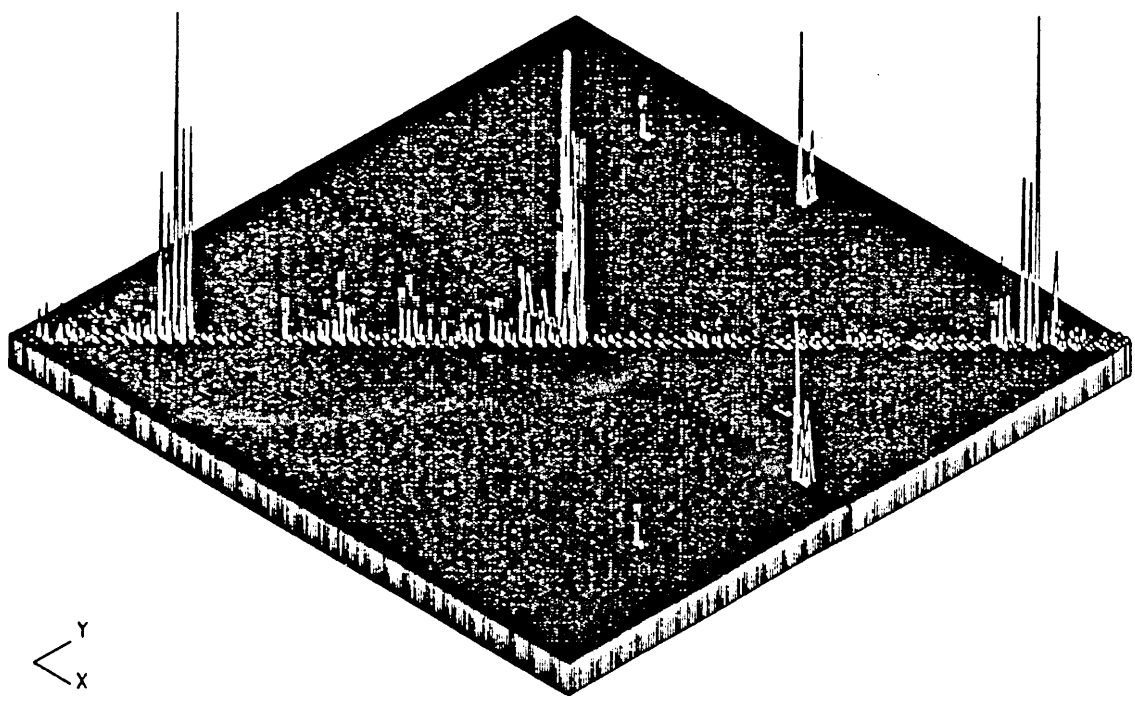


bending stiffness

Fig.7.28 - First Iteration Estimation

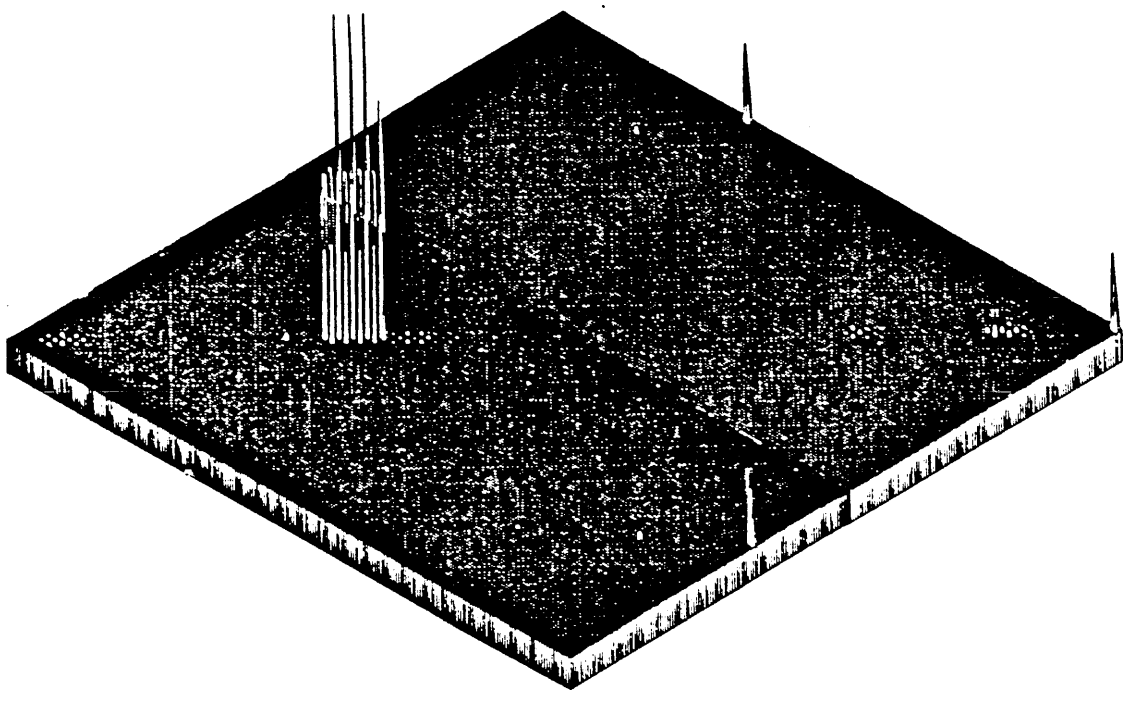


extensional stiffness

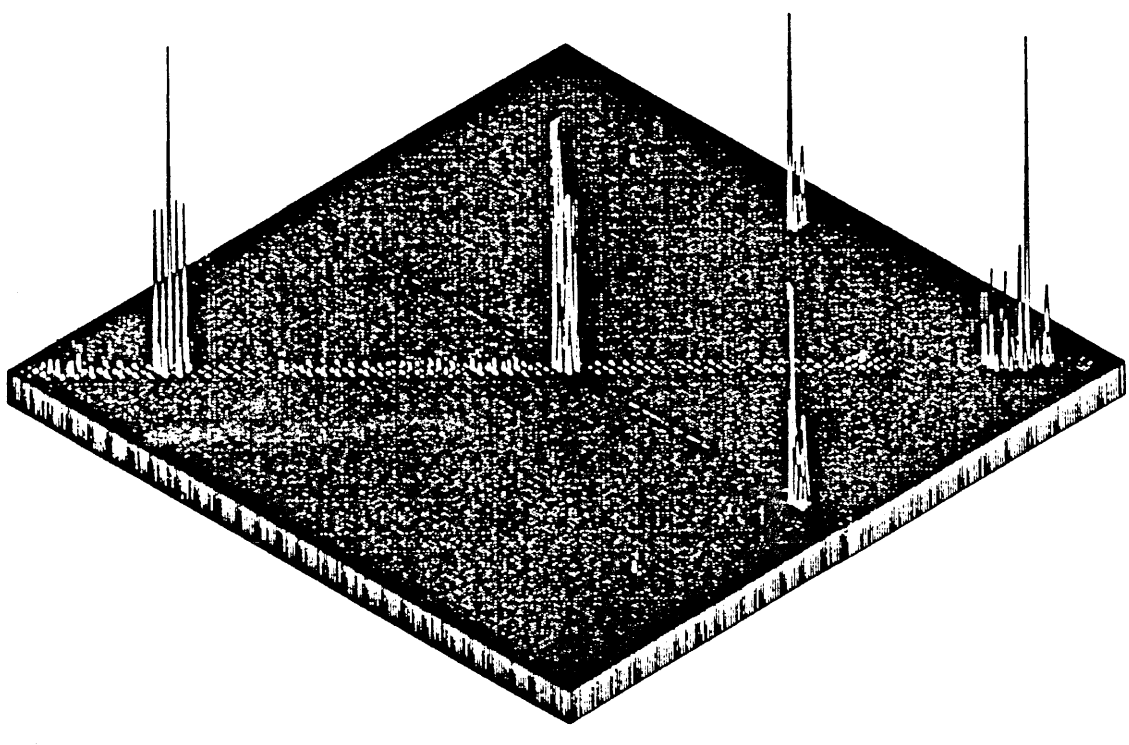


bending stiffness

Fig.7.29 - 4th Iteration Estimation

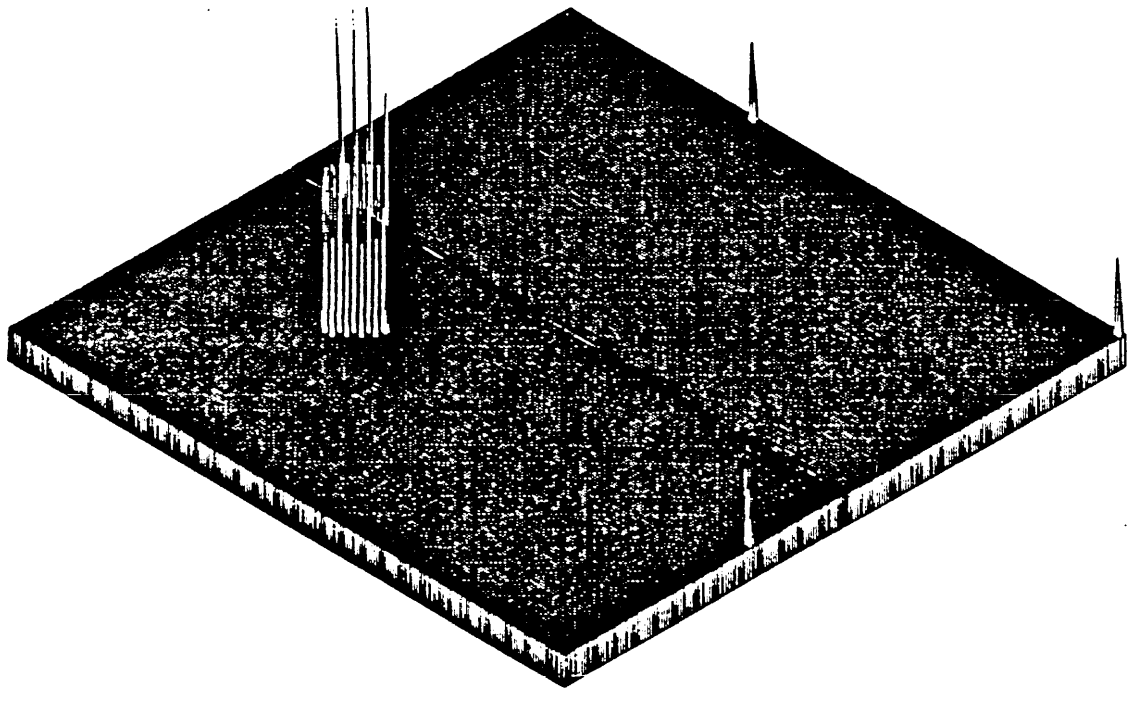


extensional stiffness

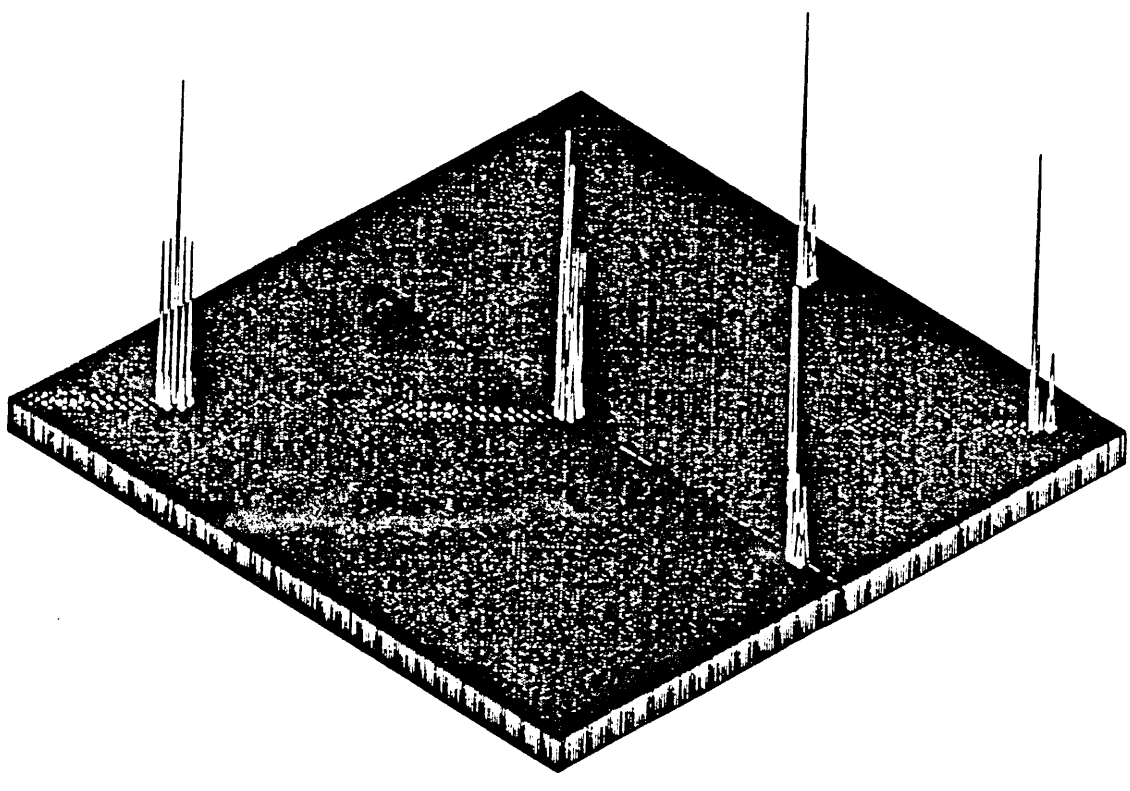


bending stiffness

Fig.7.30 - 7th Iteration Estimation



extensional stiffness



bending stiffness

Fig.7.31 - 12th Iteration Estimation

7.6 CONCLUSIONS

In this Chapter, the limitations and difficulties of some of the recently-developed methods based on full matrix updating such as Berman's method and the Error Matrix method have been discussed. The mathematical underdeterminacy associated with these methods is explained. Then, the possibility of updating a condensed (Guyan-reduced) model with error location based on Kidder's expansion method is examined. It has been demonstrated that when measured modes and/or coordinates are incomplete, as they are in practice, updating of the analytical model based on full matrix updating or the Guyan-reduced model with error location is truly difficult, if not impossible. Such a target as to update an analytical model by considering every element of mass and stiffness matrices to be (potentially) in error is overambitious and not necessarily appropriate when the inevitable limitations in measured data are considered. In order to solve the updating problem, it becomes clear that the physical connectivity of the analytical model should be respected during the updating process so that the total number of unknowns involved can be reduced and the limited measured data can have the possibility of solving the problem.

It has been illustrated that the modelling process - that is, to obtain a constant coefficient model of a continuous structure - is mathematically a dynamic condensation process and since the exact model is frequency-dependent (the mass matrix is a function of frequency), it is apparent that such a constant coefficient model can only be an approximation with limits on the frequency range of applicability. On the other hand, as has been demonstrated in the numerical example, it is quite possible for a constant coefficient model with well preserved physical connectivity (heavily banded) to represent the structure accurately as far as the lower modes of a structure are of interest (this is a limitation of any constant coefficient models - whether they are fully populated or heavily banded).

When the physical connectivity of the analytical model is applied, the measured data required in order to update an analytical model are usually within the capability of practical measurements. Hence, it is possible to establish the criteria on the minimum measured data required to solve the updating problem. Such criteria are important not only because modal testing is costly, but also because they enable the analyst to judge whether a set of measured data will have the potential to solve the updating problem so that 'blind' attempts can be avoided.

For those cases where measured coordinates are complete, the Eigendynamic Constraint method has been developed and employed in this Chapter to establish the criterion concerning the number of modes required to solve the updating problem. As shown in numerical case studies, this method requires less modes as compared with other existing model updating methods based on the use of modal data. When the measured coordinates are incomplete, a direct solution of the updating problem is, in general, not possible and some kind of iterative perturbation or sensitivity analysis has to be devised. In this case, the Inverse Eigensensitivity Analysis method has been employed to establish the criterion on the minimum data (measured modes and coordinates) required to update an analytical model. Furthermore, it has been shown that these criteria can be generalised when measured FRF data are used based on the method developed in Chapter 6.

The criteria developed have been verified numerically and the methods presented have been applied to the analytical model updating exercise called 'GARTEUR' which is intended to represent the realistic practical problem in terms of the incompleteness of both measured modes and coordinates.

CHAPTER 8

CONCLUSIONS

8.1 IDENTIFICATION OF STRUCTURAL NONLINEARITY

It is believed that most practical engineering structures possess a degree of nonlinearity. In some cases, they are treated as linear structures because the degree of nonlinearity is small and therefore insignificant in the response range of interest. In other cases, the effect of nonlinearity may become so significant that it has to be taken into account in the analysis of the structure's dynamic characteristics. In fact, for many engineering applications, structural nonlinearities need to be identified and, subsequently, nonlinear mathematical models must be established.

Unlike theoretical studies, where the vibration characteristics of nonlinear systems can be described by differential equations, a major problem in the identification of nonlinearity is to study unknown types of nonlinearity. In practice, not only the existence of nonlinearity needs to be detected, but more importantly, the degree of nonlinearity must be quantified and then the physical characteristics of the nonlinearity identified. A review and discussion of those methods currently used to investigate nonlinearity have shown their practical applicability. However, conclusive identification of practical structural nonlinearities is still problematic due to the existence of various different types of nonlinearity and numerous qualitatively different nonlinear phenomena.

For those structures whose nonlinearities are vibration amplitude dependent, the dynamic characteristics can be identified based on the analysis of measured first-order FRFs. It is found that one first-order FRF curve measured with a constant force level contains all the information of a series of FRF data with constant response controls. Therefore, a study of nonlinearity can be carried out using a single FRF curve measured with constant force. Based on this observation, a new nonlinearity analysis method has been proposed in this work. Instead of assuming that the mode to be analysed should be real, as in the case of inverse receptance method, this new method deals with the practical situation in which measured modes contain complexity due to the nonproportional distribution of structural damping. The final results of the analysis based on the proposed method are the response-amplitude-dependent eigenvalues $\lambda(\hat{X})$ and eigenvectors $\phi(\hat{X})$ which can be used not only to quantify the degree of nonlinearity but also to derive linearised spatial models of a nonlinear structure. The method has been effectively applied to the data measured from practical nonlinear structures even when the modes to be analysed are markedly complex. Also, it is found that the method does not require the condition of using constant force to measure the first-order FRF data as long as the force level is large enough to expose nonlinearity.

On the other hand, for structures whose nonlinearities are such that the measured first-order FRFs are apparently linear (the nonlinearity is of an nonsymmetrical type), a higher-order FRF analysis becomes necessary for the identification of such nonlinear structures. Fundamentally different from the analysis of first-order FRFs (in which only the fundamental frequency component is of interest), the analysis of higher-order FRFs takes into account the super-, sub- and combinational frequency components which, in some practical applications, are as important as the fundamental frequency component. By extending first-order FRF analysis to higher-order FRF analysis, it can be seen that linear system theory has been extended in a natural way to cover nonlinear systems. It has been found that many phenomena exhibited by nonlinear systems cannot be explained based on classical first-order FRFs, but can be interpreted with a series of FRFs of different orders. Furthermore, measured higher-order FRFs provide valuable information about the nature of system nonlinearities and can be used not only to identify structural nonlinearities, but also together with the first-order FRFs, to improve the response prediction of a nonlinear system due to known input.

Both first-order and higher-order FRF analysis techniques are essentially based on the classical assumption that the output of a nonlinear structure is periodic if the input is periodic. However, it has been recognised that for some nonlinear systems (chaotic

systems), a periodic input will result in an output of a random nature. Such a newly-discovered phenomenon is called chaos and is the most complicated dynamic behaviour of nonlinear systems. It is believed that the ability to identify such nonlinear behaviour is of practical importance. In Chapter 4, for the first time, the hidden chaotic behaviour of a mechanical backlash system with realistic system parameters has been revealed and, based on this system, qualitative as well as quantitative ways of identifying chaotic structures are presented. Both numerical studies and experimental investigations have been carried out. Indeed, there exist wide parameter regions, both in the system parameters and external forcing conditions, for which chaotic vibrations occur. Such nonlinear mechanisms as backlash stiffness nonlinearity represent an extensive group of engineering structures and it can be expected that many structural systems will exhibit chaotic behaviour under certain operating conditions. The anticipated engineering applications of the research work presented include, (i) design of mechanical control systems, (ii) statistical stress/fatigue analysis and, (iii) condition monitoring and diagnosis of machinery.

8.2 LOCATION OF STRUCTURAL NONLINEARITY

It is believed that structural nonlinearities, when they exist, are generally localised in terms of spatial coordinates as a result of the nonlinear dynamic characteristics of structural joints, nonlinear boundary conditions and nonlinear material properties. The ability to pinpoint a structure's localised nonlinearity(ies) thus has some important engineering applications. First, the information about where the structural nonlinearity is may offer opportunities to separate the structure into linear and nonlinear subsystems so that these can be analysed separately and efficiently using a nonlinear substructuring analysis. Second, since nonlinearity is often caused by the improper connection of structural joints, its location may give an indication of a malfunction or of poor assembly of the system. Third, from a materials property point of view, the stress at certain parts of the structure during vibration can become so high that the deformation of that part becomes plastic and the dynamic behaviour becomes nonlinear. In this case, location of the nonlinearity may offer the possibility of failure detection. Finally, the location information is essential if a nonlinear mathematical model of the structure is to be established.

Since structural nonlinearities cannot be foreseen and so cannot be analytically predicted, measurement becomes crucially important in locating and identifying them. Nevertheless, it is understood that such a task as to locate structure's localised nonlinearity is difficult to

achieve (if not impossible) based on measured data alone because they are in general quite limited and correlation between an analytical model and dynamic test data becomes essential. Based on such understanding, a nonlinearity location method was developed based on the correlation between an analytical model which contains modelling errors and modal test data which are measured at different response levels. The method was then extended to the practical case where measured coordinates are incomplete by interpolating the unmeasured coordinates based on the analytical model. Further, it has been shown that the method can be generalised when measured first-order FRF data are used. As illustrated in numerical case studies, a successful location can be achieved provided that the modelling errors are of second order in the sense of the Euclidean norm when compared with the analytical model itself and some coordinates around where the nonlinearities are located are measured.

The location technique has been demonstrated by a practical application. The structure used is a frame structure with a localised stiffness nonlinearity simulated using an electro-dynamic shaker and an analogue computer circuit based on feed-back system control theory. Only a third of the coordinates specified in the analytical (FE) model were measured. The second bending mode was found to the most sensitive one to nonlinearity and so was used in the location process. The unmeasured coordinates were first interpolated using the FE model and then the complete modeshape was used to locate the nonlinearity. The location process successfully indicates the location of the introduced nonlinearity.

8.3 MODELLING OF LINEAR AND NONLINEAR STRUCTURES

In the identification of dynamic characteristics of practical structures, analytical prediction and experimental modal testing are involved because both have their own advantages and shortcomings. Using modal testing results to improve an analytical model effectively makes use of the advantages of both while at the same time overcoming their shortcomings. It is therefore believed that by doing so, the most reliable mathematical model can be established.

Since practical structures are continuous, it is apparent that any constant coefficient models derived to represent their dynamic characteristics are approximate with limits on the frequency range of applicability. A constant coefficient model can be regarded as a *good* model if it will not only predict responses over the frequency range of interest, but will also be representative of the physical characteristics of the structure. Thus it must

have the capability to predict the effects of changes in physical parameters and to represent correctly the structure when it is treated as a component of a large system.

In effort to establish such a good model, many different methods have been developed in recent years to improve analytical models using modal testing results. Review and discussion of those recently-developed methods have suggested their limited practical applicability but at the same time, have revealed some of the existing problems. The full matrix updating methods such as the Berman's method and the Error Matrix method tackle the problem of analytical model improvement from a purely mathematical viewpoint, such as an optimisation one, rather than to consider structural characteristics as well, such as physical connectivity. As a result, the 'improved' model could be optimal in a mathematical sense, but physically unrealistic. Also, the completeness of measured coordinates is in general assumed in those methods and is critical to their success in most cases in spite of the fact that it is extremely difficult in practice to measure all the coordinates which specified in the analytical model. Furthermore, modal data are assumed to be used during analytical model improvement and since the number of measured modes are often quite limited. Consequently, the updating problem is usually mathematically underdetermined. Nevertheless, it is understood that the measured FRF data contain, in theory, the necessary information of all the modes of a structure. Based on such understanding, a new generalised model updating method has been developed which tackles the updating problem by directly using measured FRF data. The method allows the physical connectivity of the analytical model to be preserved and deals with the practical case in which measured coordinates are incomplete. It has been shown that model updating methods based on modal data are, in a broad sense, discrete versions of the present generalised method. Based on this method, the uniqueness of the updating problem has been discussed in some mathematical rigour. Numerical studies demonstrate the marked advantages of the new method as compared with other existing methods.

When the connectivity information of the analytical model is available, the model updating problem can in general be turned into an overdetermined problem even when, as in practice the measured data are limited. Therefore, it becomes possible and necessary to develop criteria concerning just how much data (modes and coordinates) need to measured in order to solve the updating problem uniquely. Such criteria are believed to be practically important because they enable the analyst to judge whether a set of measured data have the potential to solve the updating problem so that blind tries can be avoided.

The final target in the analysis of a nonlinear structure is perhaps to establish its nonlinear mathematical model. It is argued that such a mathematical model of a nonlinear structure

becomes possible only when, on the one hand, an accurate linear mathematical model (corresponding to very low response amplitude) is available and on the other, the location information of the localised nonlinearity is given because, unlike the modelling of a linear structure in which measured data are consistent, a nonlinear mathematical model has to be established based on mode by mode basis. Mathematically, structural nonlinearity can be considered as modelling errors and as a result, the problem of modelling nonlinear structures becomes the same as that of analytical model updating of linear structures except that in the former case, a series of linearised models are to be established. In the same way, the proposed method is ideally suited for the application of nonlinear structures for which FRF data can be measured while modal data are sometimes difficult to obtain.

8.4 SUGGESTION FOR FURTHER STUDIES

Whereas extensive research work on the identification of the dynamic characteristics of nonlinear structures has been carried out in this thesis, the study undertaken has revealed that some further development in the field may be necessary and of interest. Some general suggestions for possible further studies are summarised below.

Although extensive numerical simulations have been carried out in order to assess the practical applicability of the measurement and analysis of higher-order FRFs and it can be anticipated that similar results can be obtained in practical measurement as those of numerical simulation, no real measurement has been undertaken in this study due to the limited period of time available. Further research on the measurement of clean, consistent higher-order FRFs of practical engineering structures is recommended.

As discussed in Chapter 4, it can be envisaged that the existence of chaotic behaviour of such a simple nonlinear mechanical system with backlash stiffness nonlinearity will have important engineering applications such as fatigue analysis, condition monitoring and robotics design. Studies on these specific applications are recommended.

The nonlinearity location technique developed in Chapter 5 needs to be applied to more sophisticated practical nonlinear structures. In this study, the application has been made to a numerical study and a comparatively simple frame structure due to the difficulties involved in designing a nonlinear structure with known localised nonlinearity.

As for the mathematical modelling of nonlinear structures, it has to be mentioned that the mathematical models obtained based on the measured first-order FRFs are only the first order approximation of the true models of nonlinear structures. In order to improve the modelling accuracy, it is obvious that the measured higher-order FRFs should be incorporated into the modelling process. Some research work on this subject is necessary.

APPENDIX I

THE SINGULAR VALUE DECOMPOSITION

The singular Value Decomposition (SVD) is a mathematical tool which has proven to be very useful in many engineering applications and has accumulated many technical publications. It is not our purpose here to present a full and rigorous mathematical description but, instead, to give a simple introduction to the technique and to highlight its specific applications to the determination of the rank of a matrix and to the solution of a set of overdetermined linear algebraic equations which are necessary in the analytical model updating process as discussed in Chapters 6 & 7.

A review of the singular Value Decomposition of a matrix is given in [101] which includes a bibliography dealing with algorithms and applications. Improved algorithms can be found in [102-104]. Fortran subroutines are presented in [105] for real matrix cases and in [106] for complex matrices. In this appendix, we shall discuss the case where the matrix [A] to be decomposed is of dimension mxn with $m \geq n$, although the case $m \leq n$ can be treated by applying the analysis to the transpose matrix $[A]^T$.

The SVD of a real matrix [A] results in three component matrices as follows,

$$[A]_{mxn} = [U]_{mxm} [\Sigma]_{mxn} [V]^T_{nxn} \qquad \text{(A1.1)}$$

where the component matrices are described as:

$[U]_{mxm}$ orthonormal matrix with its columns called left singular vectors,

$$[\Sigma]_{mxn} = \begin{bmatrix} \sigma_1 & & & 0 \\ & \sigma_2 & & \\ & & \ddots & \\ 0 & & & \sigma_n \\ 0 & \cdots & & 0 \end{bmatrix}_{mxn}$$

a matrix which contains singular values of [A]. The singular values are ordered so that

$\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n \geq 0$,

$[V]_{nxn}$ orthonormal matrix with its columns called right singular vectors.

Since $[U]_{mxm}$ and $[V]_{nxn}$ are orthonormal matrices, they satisfy

$$[U]^T_{mxm}[U]_{mxm} = [U]_{mxm}\,[U]^T_{mxm} = [I]_{mxm} \tag{A1.2}$$

$$[V]^T_{nxn}[V]_{nxn} = [V]_{nxn}\,[V]^T_{nxn} = [I]_{nxn} \tag{A1.3}$$

Similarly, the SVD of a complex matrix results also in three component matrices as follows,

$$[A]_{mxn} = [U]_{mxm}\,[\Sigma]_{mxn}\,[V]^H_{nxn} \tag{A1.4}$$

where $[U]_{mxm}$ and $[V]_{nxn}$ are unitary matrices and $[V]^H_{nxn}$ is the complex conjugate transpose of $[V]_{nxn}$. Since $[U]_{mxm}$ and $[V]_{nxn}$ are unitary matrices, they satisfy

$$[U]^H_{mxm}[U]_{mxm} = [U]_{mxm}\,[U]^H_{mxm} = [I]_{mxm} \tag{A1.5}$$

$$[V]^H_{nxn}[V]_{nxn} = [V]_{nxn}\,[V]^H_{nxn} = [I]_{nxn} \tag{A1.6}$$

The singular values are the non-negative square-roots of the eigenvalues of the matrix $[A]^T[A]$, if [A] is real, and of $[A]^H[A]$, if [A] is complex. Because $[A]^T[A]$ is symmetric and non-negative definite and $[A]^H[A]$ is Hermitian, their eigenvalues are always real and non-negative and therefore, singular values are always real and non-negative. The left and right singular vectors [U] and [V] are the corresponding eigenvectors of $[A][A]^T$ and $[A]^T[A]$ ($[A][A]^H$ and $[A]^H[A]$ if [A] is complex).

In numerical calculations, the SVD of [A] is usually performed in two stages. First, [A] is reduced to upper bidiagonal form using Householder matrices [101]. Once the bidiagonalisation has been achieved, the next step is to zero the superdiagonal elements using QR algorithm [107]. The computational time depends upon "how much" of the SVD is required. For example, if only the information about the rank of a matrix (so the condition of the matrix) is required, then only the singular values are of interest and the computational time could be four times less than the case in which all the singular values and left and right singular vectors are required. In what follows next, we shall discuss the application of the SVD technique to the determination of the rank of a matrix and to the solution of a set of overdetermined linear algebraic equations.

In the analytical model updating process as discussed in Chapter 6 & 7, we encountered the problem of solving a set of linear algebraic equations with n unknowns and m equations ($m \geq n$). We called the set of linear algebraic equations overdetermined because the number of equations is greater than the number of unknowns. However, because of the linear dependence of some of the equations, the coefficient matrix of a set of 'overdetermined' linear algebraic equations is not necessarily of full rank and if this is the case, the solution obtained numerically is likely to be physically meaningless. Accordingly, it is necessary to check the rank of the coefficient matrix before solving the linear algebraic equations. Theoretically, the rank of a matrix is the number of linearly-independent rows (or columns) in the given matrix and it is generally believed that the SVD is the only reliable method of determining rank numerically. In SVD calculation, the rank of [A] can be determined by examining the nonzero singular values. Due to the numerical and/or experimental inaccuracies, it is most likely that none of the singular values of [A] will be zero and so [A] is 'full rank' according to the mathematical definition although, in fact, [A] has a rank of $r<n$. One way to circumventing the difficulties of the mathematical definition of rank is to specify a tolerance and say that [A] is numerically defective in rank if, within that tolerance, it is near to a defective matrix. A matrix [A] is said to have rank r if, for a given δ, the singular values of [A] satisfy

$$\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r \geq \delta \geq \sigma_{r+1} \geq \ldots \geq \sigma_n \qquad \text{(A1.7)}$$

The key quantity in rank determination is obviously the value σ_r. The parameter δ should be consistent with the machine precision ε in the case of numerical rounding error, e.g. $\delta = \varepsilon \|A\|_\infty$ and if the general level of relative error in the data is larger than ε, as in the case of experimental investigations, δ should be correspondingly bigger, e.g. $\delta = 10^{-2} \| A \|_\infty$, as suggested by Golub and Van Loan [92]. For our specific problem concerned with

model updating, we have to make sure that the coefficient matrix is of full rank within such tolerances. If this is not the case, different data sets or more data points should be used.

Once it has been established that the coefficient matrix [A] is of full rank, then the set of overdetermined linear algebraic equations can be reliably and efficiently solved based on the SVD of [A]. Suppose the problem to be solved is

$$[A]_{mxn}\{x\}_{nx1} = \{b\}_{mx1} \tag{A1.8}$$

Applying SVD to [A], we obtain

$$[U]_{mxm}[\Sigma]_{mxn}[V]^T_{nxn}\{x\}_{nx1} = \{b\}_{mx1} \tag{A1.9}$$

or

$$[\Sigma]_{mxn}[V]^T_{nxn}\{x\}_{nx1} = [U]^T_{mxm}\{b\}_{mx1} \tag{A1.10}$$

or

$$[\Sigma]_{mxn}\{z\}_{nx1} = \{d\}_{mx1} \tag{A1.11}$$

with

$$\{z\}_{nx1} = [V]^T_{nxn}\{x\}_{nx1} \tag{A1.12}$$

$$\{d\}_{mx1} = [U]^T_{mxm}\{b\}_{mx1} \tag{A1.13}$$

From (A1.11), $\{z\}_{nx1}$ can be easily solved. Having calculated $\{z\}_{nx1}$, the vector $\{x\}_{nx1}$ can be recovered from (A1.12) as

$$\{x\}_{nx1} = [V]_{nxn}\{z\}_{nx1} \tag{A1.14}$$

APPENDIX II

DERIVATION OF EIGENDERIVATIVES

Sensitivity analysis, the study of changes in system dynamic characteristics with respect to parameter variation, is being used in a variety of engineering disciplines ranging from automatic control theory to the analysis of large-scale physiological systems [108]. Some of the specific areas where sensitivity analysis has been applied include (i) system identification [80], (ii) development of insensitive control systems [109], (iii) use in gradient-based mathematical programming methods [109], (iv) assessment of design changes on system performance [109], and (v) approximation of system response to a change in a system parameter [110]. In this appendix, the derivation of eigenderivatives for eigenvalues and eigenvectors with respect to the system's design parameter changes is discussed. Eigenderivatives with respect to design variables have been found to be particularly useful in certain analysis and design applications, e.g., approximating a new vibration modeshape due to a perturbation in design variable, determining the effect of design changes on the dynamic behaviour of a structure, and tailoring modeshapes to minimise displacements at certain points on a structure. Here, the eigenderivatives are derived for the purpose of analytical model improvement as discussed in Chapter 7.

Methods for calculating eigenderivatives include the modal method [111], Nelson's method [112] and an improved modal method [113]. The modal method is straightforward in theory but requires all the modes of system which is sometimes computationally expensive especially when systems with large dimensions are considered. Nelson's method seeks to calculate the eigenderivatives of the r^{th} mode by just using the modal parameters of that mode. However, one matrix inverse of system dimension (in fact, of dimension (N-1) where N is the dimension of the system) is required for each mode in order to solve the linear algebraic equations involved. The improved modal

method aims to derive the eigenderivatives approximately (the eigenderivatives themselves are only the first-order approximation) by using the calculated lower modes and the flexibility matrix. In the following, all three methods are discussed and their advantages and disadvantages in terms of computational cost and numerical accuracy are examined.

However, before discussing the methods in detail, it is necessary to mention that although discussions have been made on the derivation of eigenderivatives of repeated modes [114], only the eigenderivatives of distinct modes are presented in this appendix. On the other hand, as it will be needed in later discussions, we state here that for any static structure whose mass and stiffness matrix are symmetric and whose mass matrix positive-definite (which is the case of interest here), the complete set of eigenvectors of the system forms a complete linearly-independent base and so any vector of the same dimension can be expressed as a linear combination of all these eigenvectors. This argument is briefly proved as below. The matrix representation of the vibration eigenvalue problem is

$$[K]\{\phi\}_i - \lambda_i [M]\{\phi\}_i = \{0\} \tag{A2.1}$$

Since [M] is symmetric and positive-definite, [M] can be decomposed as [92],

$$[M] = [L][L]^T \tag{A2.2}$$

where [L] is a non-singular lower triangular matrix. Upon substitution (A2.2) into (A2.1),

$$[K]\{\phi\}_i - \lambda_i [L][L]^T\{\phi\}_i = \{0\} \tag{A2.3}$$

Since [L] is non-singular, pre-multiply both sides of (A2.3) by $[L]^{-1}$, so that (A2.3) becomes

$$[L]^{-1}[K]\{\phi\}_i - \lambda_i [L]^T\{\phi\}_i = \{0\} \tag{A2.4}$$

Let $\{z\}_i = [L]^T\{\phi\}_i$ and substitute into (A2.4), then

$$[L]^{-1}[K][L]^{-T}\{z\}_i - \lambda_i \{z\}_i = \{0\} \tag{A2.5}$$

Since $[L]^{-1}[K][L]^{-T}$ is real and symmetric, the complete set of eigenvectors of $\{z\}_i$ forms a complete orthogonal base, regardless of the existence of repeated modes [92]. On the

other hand, since [L] is non-singular, $[\phi] = [L]^{-T}\,[z]$ forms a complete linearly-independent base.

THE MODAL METHOD

Differentiating (A2.1) with respect to the r^{th} design variable p_r:

$$([K] - \lambda_i [M])\frac{\partial\{\phi\}_i}{\partial p_r} + \left(\frac{\partial[K]}{\partial p_r} - \lambda_i \frac{\partial[M]}{\partial p_r} - \frac{\partial\lambda_i}{\partial p_r}[M]\right)\{\phi\}_i = \{0\} \qquad \text{(A2.6)}$$

Assume that $\{\phi\}_i$ is normalised such that

$$\{\phi\}_i^T [M] \{\phi\}_i = 1 \qquad \text{(A2.7)}$$

Multiply both sides of (A2.6) by $\{\phi\}_i^T$ and using (A2.1) and (A2.7), then

$$\frac{\partial\lambda_i}{\partial p_r} = \{\phi\}_i^T \frac{\partial[K]}{\partial p_r}\{\phi\}_i - \lambda_i \{\phi\}_i^T \frac{\partial[M]}{\partial p_r}\{\phi\}_i \qquad \text{(A2.8)}$$

From (A2.8), it can be seen that the eigenvalue derivative is determined by the mode itself. Since as discussed, the complete set of eigenvectors forms a complete linearly-independent base. Therefore, without any loss of generality, the i^{th} eigenvector derivative which is a vector can be expressed as a linear combination of all the eigenvectors of the system:

$$\frac{\partial\{\phi\}_i}{\partial p_r} = \sum_{j=1}^{N} \beta_{ij}\{\phi\}_j \qquad \text{(A2.9)}$$

In order to calculate the coefficients β_{ij}, substitute (A2.9) into (A2.6) and pre-multiply (A2.6) by $\{\phi\}_j^T$, then

$$\{\phi\}_j^T([K] - \lambda_i [M]) \sum_{k=1}^{N} \beta_{ik}\{\phi\}_k + \{\phi\}_j^T \left(\frac{\partial[K]}{\partial p_r} - \lambda_i \frac{\partial[M]}{\partial p_r} - \frac{\partial\lambda_i}{\partial p_r}[M]\right)\{\phi\}_i = 0 \qquad \text{(A2.10)}$$

In the case $j \neq i$, (A2.10) can be simplified to solve for β_{ij}

$$\beta_{ij} = \frac{\{\phi\}_j^T \left(\frac{\partial[K]}{\partial p_r} - \lambda_i \frac{\partial[M]}{\partial p_r}\right) \{\phi\}_i}{\lambda_i - \lambda_j}, \quad j \neq i \qquad \text{(A2.11)}$$

When j=i, β_{ii} can be computed from the mass-normalisation condition. Differentiating (A2.7)

$$\{\phi\}_i^T \frac{\partial[M]}{\partial p_r} \{\phi\}_i + 2\{\phi\}_i^T [M] \frac{\partial\{\phi\}_i}{\partial p_r} = 0 \qquad \text{(A2.12)}$$

Substitute (A2.9) into (A2.12), then β_{ii} becomes

$$\beta_{ii} = -\frac{1}{2} \{\phi\}_i^T \frac{\partial[M]}{\partial p_r} \{\phi\}_i \qquad \text{(A2.13)}$$

From (A2.11 and A2.13), the derivatives of i^{th} eigenvector are related to all the modes of the system. Also from (A2.11), it can be seen that the method breaks down when the eigenvector derivatives of repeated modes are to be calculated because for some j, the denominator of (A2.11) becomes zero.

NELSON'S METHOD

As discussed above, all the eigenvectors of a system are required in order to calculate the eigenderivatives of one single mode based on the modal method. As mentioned in §6.4.5, such complete eigensolution can be very expensive when systems of high order are concerned. In practice, usually only the few lowest modes which are of interest are available based on a partial eigensolution. In order to cope with such a practical situation, Nelson developed a method which seeks to derive the eigenderivatives of the r^{th} mode by using only those components of the mode itself.

The derivative of the eigenvalue is the same as expressed in (A2.8). After the eigenvalue derivative is calculated, the corresponding eigenvector derivatives can be determined as shown below. Define $\{f\}$ and [G] as

$$\{f\} \equiv \left(-\frac{\partial[K]}{\partial p_r} + \lambda_i \frac{\partial[M]}{\partial p_r} + \frac{\partial\lambda_i}{\partial p_r}[M]\right)\{\phi\}_i \quad \text{and} \quad [G] \equiv ([K] - \lambda_i [M]) \qquad \text{(A2.14)}$$

then equation (A2.6) becomes

$$[G]\frac{\partial\{\phi\}_i}{\partial p_r} = \{f\} \tag{A2.15}$$

Note that [G] has rank (N-1) (assuming λ_i is a single eigenvalue) and nullspace $\{\phi\}_i$ (nullspace of a matrix [A] is defined as the set of vectors $\{x\}$ such that $[A]\{x\}=\{0\}$). This means that for any $\{\sigma\}$ satisfying $[G]\{\sigma\}=\{f\}$, $\{\sigma\}+\gamma\{\phi\}_i$ is also a solution, where γ is any real number. The approach is first to find some particular solution $\{\sigma\}$, then to find γ so that $\{\sigma\}+\gamma\{\phi\}_i=\frac{\partial\{\phi\}_i}{\partial p_r}$. A suitable $\{\sigma\}$ can be found by setting one element of $\{\sigma\}$ to be zero (since [G] is rank 1 deficient) and solving the remaining elements. This will work for any $\{f\}$ in the range of [G] (the range of matrix [A] is defined as the set of vectors $\{y\}$ such that $\{y\}=[A]\{x\}$ for some $\{x\}$ and the necessary and sufficient condition for the existence of exact solution(s) of $[A]\{x\}=\{y\}$ is that $\{y\}$ is in the range of [A]) as long as the corresponding element of $\{\phi\}_i$ is nonzero. This suggests the following algorithm:

1) Find k such that $|_i\phi_k| = \|\{\phi\}_i\|_\infty \equiv \max_j |_i\phi_j|$.
2) Delete the k^{th} row and column of $[G]\rightarrow[\tilde{G}]$ and k^{th} element of $\{f\}\rightarrow\{\tilde{f}\}$ and let $\sigma_k=0$ and by removing it from $\{\sigma\}\rightarrow\{\tilde{\sigma}\}$.
3) Solve the reduced linear algebraic problem $[\tilde{G}]\{\tilde{\sigma}\} = \{\tilde{f}\}$.

Since [G] is rank 1 deficient and $\{f\}$ is in the range of [G] (since we know $\{\sigma\}$ exists), the resulting $\{\sigma\}$ automatically satisfies $[G]\{\sigma\}=\{f\}$. The reason for choosing such a k (the eigenvector element with maximum magnitude) is to make sure that the reduced coefficient matrix $[\tilde{G}]$ is rank full and numerically well-conditioned so that $\{\tilde{\sigma}\}$ can be uniquely and accurately determined. This is illustrated as follows. Defining $[A]\equiv([K]-\lambda_i[M])$ and partitioning the matrices, (A2.1) can be written as

$$\begin{bmatrix} [A]_{jj} & \{A\}_{jk} & [A]_{jl} \\ \cdots & \cdots & \cdots \\ \{A\}_{jk}^T & a_{kk} & \{A\}_{kl}^T \\ \cdots & \cdots & \cdots \\ [A]_{jl}^T & \{A\}_{kl} & [A]_{ll} \end{bmatrix} \begin{Bmatrix} \{\phi_j\}_i \\ \cdots \\ {}_i\phi_k \\ \cdots \\ \{\phi_l\}_i \end{Bmatrix} = \begin{Bmatrix} \{0\} \\ \cdots \\ 0 \\ \cdots \\ \{0\} \end{Bmatrix} \tag{A2.16}$$

where $[A]_{ij}$ are the partitioned matrices of $([K]-\lambda_i[M])$, a_{kk} is the k,k element of $([K]-\lambda_i[M])$, $\{\phi_j\}_i$ and $\{\phi_l\}_i$ are the partitioned subvectors, and ${}_i\phi_k$ the k^{th} component of $\{\phi\}_i$. Multiplying out (A2.16) and moving the ${}_i\phi_k$ term to the RHS, (A2.16) becomes

$$\begin{bmatrix} [A]_{jj} & \vdots & [A]_{jl} \\ \cdots & \cdots & \cdots \\ \{A\}_{jk}^T & \vdots & \{A\}_{kl}^T \\ \cdots & \cdots & \cdots \\ [A]_{jl}^T & \vdots & [A]_{ll} \end{bmatrix} \begin{Bmatrix} \{\phi_j\}_i \\ \cdots \\ \{\phi_l\}_i \end{Bmatrix} = -{}_i\phi_k \begin{Bmatrix} \{A\}_{jk} \\ \cdots \\ a_{kk} \\ \cdots \\ \{A\}_{kl} \end{Bmatrix} \qquad \text{(A2.17)}$$

If ${}_i\phi_k$ is nonzero and assigned a specific value, then subvectors $\{\phi_j\}_i$ and $\{\phi_l\}_i$ must be uniquely determined by solving (A2.17) whose coefficient matrix has rank (N-1) when λ_i is a single mode. This argument also shows that when ${}_i\phi_k$ is nonzero, the k^{th} column of the NxN matrix in (A2.16) is a linear combination of the remaining (N-1) columns and the remaining (N-1) columns are therefore linearly independent. Also, from a numerical point of view, ${}_i\phi_k$ should be chosen to be the the element with maximum magnitude so that the coefficient matrix of (A2.17) and so $[\bar{G}]$ is well-conditioned in terms of its inverse.

As can be seen from the above discussion, the process of deriving eigenvector derivatives based on Nelson's method preserves the sparseness and banded structure, if any, of [G], which allows for efficient solution of $[\bar{G}]\{\bar{\sigma}\}=\{\bar{f}\}$. It remains to find the scalar γ. Substitute $\{\sigma\}+\gamma\{\phi\}_i$ into (A2.12):

$$\{\phi\}_i^T \frac{\partial[M]}{\partial p_r} \{\phi\}_i + 2\{\phi\}_i^T [M] (\{\sigma\} + \gamma\{\phi\}_i) = 0 \qquad \text{(A2.18)}$$

$$\gamma = -\{\sigma\}^T [M] \{\phi\}_i - 0.5 \{\phi\}_i^T \frac{\partial[M]}{\partial p_r} \{\phi\}_i \qquad \text{(A2.19)}$$

This completes the solution for $\frac{\partial\lambda_i}{\partial p_r}$ and $\frac{\partial\{\phi\}_i}{\partial p_r}$ based on Nelson's method. Note that this method has the desirable properties of preserving the structure of [M] and [K] (allowing more efficient solution) and requiring the knowledge of only one eigenvalue-eigenvector pair. Both properties are important in realistic structural problems where [M] and [K] have very high order.

IMPROVED MODAL METHOD

The improved modal method seeks to calculate the eigenderivatives approximately using the lowest m (m«N) pairs of eigenvalues and eigenvectors obtained based on partial eigensolution. The eigenvalue derivatives are related to the modes themselves and they are expressed in (A2.8). In order to calculate the eigenvector derivatives, rewrite (A2.9) as

$$\frac{\partial\{\phi\}_i}{\partial p_r} = \beta_{ii}\{\phi\}_i + \{z\}_i \quad \text{where} \quad \{z\}_i = \sum_{\substack{j=1 \\ j\neq i}}^{N} \beta_{ij}\{\phi\}_j \tag{A2.20}$$

Using (A2.11), $\{z\}_i$ can be written as

$$\{z\}_i = \sum_{\substack{j=1 \\ j\neq i}}^{N} \beta_{ij}\{\phi\}_j = \sum_{\substack{j=1 \\ j\neq i}}^{N} \frac{\{\phi\}_j^T\{F\}_i}{\lambda_i - \lambda_j}\{\phi\}_j =$$

$$\sum_{\substack{j=1 \\ j\neq i}}^{m} \frac{\{\phi\}_j^T\{F\}_i}{\lambda_i - \lambda_j}\{\phi\}_j + \sum_{j=m+1}^{N} \frac{\{\phi\}_j^T\{F\}_i}{\lambda_i - \lambda_j}\{\phi\}_j \tag{A2.21}$$

where $\{F\}_i \equiv \left(\frac{\partial[K]}{\partial p_r} - \lambda_i \frac{\partial[M]}{\partial p_r}\right)\{\phi\}_i$ and we have assumed that $i \leq m$. If the eigenvalues are numbered according to their magnitude in ascending order, then for the class of problems with a large frequency gap is

$$\lambda_j - \lambda_i \approx \lambda_j \quad \text{for} \quad j > m \tag{A2.22}$$

It is clear that the above approximation is very accurate for $j \gg m$. Thus (A2.22) can be approximated as

$$\{z\}_i \approx \sum_{\substack{j=1 \\ j\neq i}}^{m} \frac{\{\phi\}_j^T\{F\}_i}{\lambda_i - \lambda_j}\{\phi\}_j + \sum_{j=m+1}^{N} \frac{\{\phi\}_j^T\{F\}_i}{-\lambda_j}\{\phi\}_j \tag{A2.23}$$

which can also be written as

$$\{z\}_i \approx \sum_{\substack{j=1 \\ j\neq i}}^{m} \frac{\{\phi\}_j^T\{F\}_i}{\lambda_i - \lambda_j}\{\phi\}_j - \sum_{j=1}^{N} \frac{\{\phi\}_j^T\{F\}_i}{\lambda_j}\{\phi\}_j - \sum_{j=1}^{m} \frac{\{\phi\}_j^T\{F\}_i}{-\lambda_j}\{\phi\}_j \tag{A2.24}$$

In equation (A2.24), only the middle term is unknown and it can be calculated using the flexibility matrix as shown below.

Since from the theory of algebraic eigenvalue problems, the flexibility matrix $[K]^{-1}$ can be calculated from the system's eigenvalue and eigenvector matrices as

$$[K]^{-1} = [\phi]\,[\lambda.]^{-1}\,[\phi]^T = \sum_{j=1}^{N} \frac{\{\phi\}_j \{\phi\}_j^T}{\lambda_j} \qquad \text{(A2.25)}$$

Using this spectral decomposition of (A2.25), the middle term of the RHS of (A2.24) becomes $[K]^{-1}\{F\}_i$ and (A2.24) can be written as

$$\{z\}_i \approx \sum_{\substack{j=1 \\ j \neq i}}^{m} \frac{\{\phi\}_j^T \{F\}_i}{\lambda_i - \lambda_j} \{\phi\}_j - [K]^{-1}\{F\}_i - \sum_{j=1}^{m} \frac{\{\phi\}_j^T \{F\}_i}{\lambda_j} \{\phi\}_j \qquad \text{(A2.26)}$$

We note that the $[K]^{-1}\{F\}_i$ term in (A2.26) approximates the effect of higher uncalculated modal components on eigenvector derivatives of lower modes. With $\{z\}_i$ being obtained and β_{ii} as expressed in (A2.13), the eigenvector derivatives of i^{th} mode can be simply calculated based on (A2.20).

REFERENCES

[1] **Ewins, D. J.**
"Modal Testing: Theory and Practice"
Research Studies Press, 1984

[2] **He, J.**
"Identification of Structural Dynamic Characteristics"
Ph.D thesis, Mech. Eng. Dept., Imperial College, 1983

[3] **Titchmarsh, W. C.**
"Introduction to the Theory of Fourier Integrals"
Oxford University Press, 1937

[4] **Simon, M and Tomlinson, G. R.**
"Use of the Hilbert Transform in Modal Analysis of Linear and Nonlinear Structures"
Journal of Sound and Vibration, 1984

[5] **He, J. and Ewins, D. J.**
"A simple Method of Interpretation for the Modal Analysis of Nonlinear Systems"
IMAC 5, 1987

[6] **Schetzen, M.**
"The Volterra and Wiener Theories of Nonlinear Systems"
John Wiley and Sons, 1980

[7] **Wiener, N.**
"Nonlinear Problems in Random Theory"
Wiley, 1958

[8] **Chouychai, T.**
"Dynamic Behaviour of Nonlinear Structures"
Ph.D thesis (in French), 1986, Sain-Ouen, Paris

[9] **Ibrahim, S. R. and Mikucik, E. C.**
"The Experimental Determination of Vibration Parameters from Time Response"
Shock and Vibration Bulletin, Vol. 46(5), 1976

[10] **Jordan, D. W. and Smith, P.**
"Nonlinear Ordinary Differential Equations"
Oxford University Press, 1987

[11] **Masri, S. and Caughey, T.**
"A Non-parametric Identification Technique for Nonlinear Dynamic Problems"
J. Appl. Mech., ASME **46**, 1979

[12] **Crawley, E. F. and Aubert, A. C.**
"Identification of Nonlinear Structural Elements by Force-State Mapping"
AIAA Journal, Vol. 24, No. 1, 1986

[13] **Crawley, E. F. and O'Donnel, K. J.**
"Force-State Mapping Identification of nonlinear Joints"
AIAA Journal, Vol. 25, No. 7, 1987

[14] **Worden, K. and Tomlinson, G. R.**
"Developments in the Force State Mapping for Nonlinear Systems"
IMAC 6, 1988

[15] **Pandit, S. M. and Wu, S. M.**
"Time Series and System Analysis with Applications"
Wiley, 1983

[16] **Billings, S. A. and Voon, W. S. F.**
"Structure Detection and Validity Tests in the Identification of Nonlinear Systems"
Proc. IEE, Vol. 130D, 1983

[17] **Billings, S. A. and Frazil, M. B.**
"The Practical Identification of Systems with Nonlinearities"
The 7th Symposium on Identification and System Parameter Estimation, New York, 1985

[18] **Billings, S. A. and Tsang, K. M.**
"Application of the NARMAX Method to Nonlinear Frequency Response Estimation"
IMAC 6, 1988

[19] **Coppolino, R. N.; Bendat, J. S. and Strout, R. C.**
"New Processor Integrates Dynamic Testing and Analysis"
Sound and Vibration, August 1988

[20] **Piszczek, K. and Niziol, J.**
"Random Vibration of Mechanical Systems"
PWN - Polish Scientific Publishers, 1986

[21] **Ewins, D. J. He, J. and Lieven, N.**
"A Review of the Error of Matrix Method (EMM) for Structural Dynamic Model Comparison"
Proceedings of StruCoMe, 1988

[22] **Roberts, J. B.**
"Response of Nonlinear Mechanical Systems to Random Excitation"
Part 2, Shock Vib. Dig. 13(5), 1987

[23] **Zhu, W.-Q. and Yu J.-S.**
"The Equivalent Nonlinear System Method"
Journal of Sound and Vibration, 129(3), 1989

[24] **Nayfeh, A. H. and Mook, T. D.**
"Nonlinear Oscillations"
John Wiley and Sons, 1979.

[25] **Hanagud, S. V.; Meyyappa, M. and Craig, J. I.**
"Method of Multiple Scales and Identification of Nonlinear Structural Dynamic systems"
J. AIAA, 23(5), 1985

[26] **Nayfeh, A. H. and Zavodney, L. D.**
"Modal Interaction in the Nonlinear Response of Structural Elements -- Theory and Experiment"
Proceeding of Recent Advances in Structural Dynamics, 1988

[27] **Siljak, D. D.**
"Nonlinear Systems"
John Wiley and Sons, 1969.

[28] **Bogoliubov, N. N. and Mitropolsky, J. A.**
"Asymptotical Methods in the Theory of Nonlinear Oscillations"
Hindustan Publishing Co., Delhi, India, 1963

[29] *"The Fundamentals of Signal Analysis"*
Application Note 243, Hewlett Packard

[30] **Vinh, T. and Chouychai, T.**
"Impact Testing of Nonlinear Structures"
IMAC 5, 1987

[31] **Rayleigh, L.**
"The Theory of Sound"
Dover, New York, 1945

[32] **Coughey, T. K.**
"Classical Normal Modes in Damped Linear Dynamic Systems"
Journal of Applied Mechanics, 1960

[33] **Coughey, T. K. and O'Kelly, M. E. J.**
"Classical Normal Modes in Damped Linear Dynamic Systems"
Journal of Applied Mechanics, 1965

[34] **Prater, G. Jr. and Singh, R.**
"Quantification of the Extent of Nonproportional Viscous Damping in Discrete Vibratory Systems"
Journal of Sound and Vibration, 104(1), 1986

[35] **Lin, R. M.**
"On the Relationship Between Viscous and Hysteretic Damping Models"
Internal Report No. 88002, Dynamics Section, Mech. Eng. Dept., Imperial College of Science, Technology and Medicine, London.

[36] **Beards, C. F.**
"Damping in Structural Joints"
Sound and Vibration Digest, Vol. 11, No. 9, 1979

[37] **Robb, D. A.**
"User's Guide to Program MODENT"
Dynamics Section, Mech. Eng. Dept., Imperial College of Science, Technology and Medicine, London.

[38] **Sidhu, J.**
"Reconciliation of Predicted and Measured Modal Properties of Structures"
PhD thesis, 1983, Mechanical Engineering Department, Imperial College of Science, Technology and Medicine, London.

[39] **He, J.; Lin, R. M. and Ewins, D. J.**
"Evaluation of Some Nonlinear Modal Analysis Methods"
13th International Seminar on Modal Analysis, 1988

[40] **Narayanan, S.**
"Transistor Distortion Analysis Using Volterra Series Representation"
Bell Syst. Tech. J., 46, 1967

[41] **Maurer, R. E. and Narayanan, S.**
"Noise Loading Analysis of a Third-order System with Memory"
IEEE Trans. Commun. Technol. COM-16, 1968

[42] **Narayanan, S.**
"Application of Volterra Series to Intermodulation Distortion Analysis of Transistor Feedback Amplifier"
IEEE, Trans. Circuit Theory, CT-17, 1970

[43] **Gilford, S. J. and Tomlinson, G. R.**
"A Functional Series Approach in the Identification of Nonlinear Structures"
IMAC 5, 1987

[43] **Gilford, S. J. and Tomlinson, G. R.**
"Higher-order Frequency Response Functions"
Short Course Note on *Vibration analysis and Identification of Structural Nonlinearities*, November, 1988

[45] **Vihn, T. and Liu, H.**
"Extension of Modal Analysis of Nonlinear Systems"
IMAC 7, 1989

[46] **Edward, B. and Stephen, O. R.**
"The Output Properties of Volterra System (Nonlinear Systems with Memory) Driven by Harmonic and Gaussian Input"
Proc. IEEE, Vol. 59, No. 12, 1971

[47] **Schetzen, M.**
"Measurement of Wiener Kernels of a Nonlinear System by Cross-correlation"
Int. J. Control, Vol. 2, 1965

[48] **Stephen, B.; Tang, Y. S. and Chua, L. O.**
"Measuring Volterra Kernels"
IEEE, Vol. CAS-30, No. 8, August 1983

[49] **Bendat, J. S. and Piersol, A. G.**
"Spectral Analysis of Nonlinear Systems Involving Square-law Operation"
J. Sound & Vibration, 81(2), 1982

[50] **Iwan, W. D.**
"A Generalisation of the Concept of Equivalent Linearisation"
Int. J. of Nonlinear Mechanics, Vol. 8, 1973

[51] **Oppenheim, A. V. and Schafer, R. W.**
"Digital Signal Processing"
prentice-Hall, Inc., 1975

[52] **Maia, N. M. M.**
"Extraction of Valid Modal Properties from Measured Data in Structural Vibrations"
Ph.D thesis, Mech. Eng. Dept, Imperial College of Science, Technology and Medicine, London, 1988

[53] **Dowell, E. H. and Pezeshki, C.**
"On the Understanding of Chaos in Duffing's Equation Including a Comparison with Experiment"
Journal of Applied Mechanics 53, 1986

[54] **Shaw, S. W. and Holmes, P. J.**
"A Periodically Forced Piecewise Linear Oscillator"
Journal of Sound and Vibration 90(1), 1983

[55] **Hénon, M.**
"A two dimensional mapping with a strange attractor"
Comm. Math. Phys. **50**, 1976

[56] **Mandelbrot, B. B.**
"Fractals: Form, Chance and Dimension"
San Francisco, CA, Freeman, 1977

[57] **Matsumoto, T.**
"A Chaotic Attractor from Chua's Circuit"
IEEE Transactions on Circuit and System CAS-31(12), 1984

[58] **Takens, F.**
"Detecting Strange Attractors in Turbulence", in Dynamical Systems and Turbulence, D. A. Rand and L.-S. Young (eds.), 1980, pp. 366-381. Springer Lecture Notes in Mathematics, Vol. 898, Springer-Verlag.

[59] **Epstain, I. R.**
"Oscillations and chaos in chemical systems"
Physica 7-D, 1983

[60] **Dooren, R. V.**
"On the Transition from Regular to Chaotic Behaviour in the Duffing's Equation"
Journal of Sound and Vibration 123(2), 1988

[61] **Pezeshki, C. and Dowell, E. H.**
"An Examination of Initial Condition Maps for the Sinusoidally Excited Buckled Beam Modelled by the Duffing's Equation"
Journal of Sound and Vibration 117(2), 1987

[62] **Dowell, E. H. and Pezeshki, C.**
"On the Necessary and Sufficient Condition for Chaos to Occur in Duffing's Equation: An Heuristic Approach"
Journal of Sound and Vibration 121(2), 1988

[63] **Tongue, B. H.**
"On the Existence of Chaos in One-degree-of-freedom System"
Journal of Sound and Vibration 110(1), 1986

[64] **Shaw, S. W.**
"The Dynamics of a Harmonically Excited System Having Rigid Amplitude Constraints"
Journal of Applied Mechanics 52, 1985 (part 2)

[65] **Minorsky, N.**
"Nonlinear Oscillations".
Huntington, New York: Krieger Publishing, 1962.

[66] **Maezawa, S.**
"Steady Forced Vibrations of Unsymmetrical Piecewise Linear Systems"
Bulletin of Japanese Society of Mechanical Engineering 4, 1961

[67] **Moon, F. C.**
"Chaotic Vibrations"
John Wiley & Sons, 1987.

[68] **Wolf, A.**
"Quantifying Chaos with Lyaponov Exponents", Nonlinear Science: Theory and Application. Ed. A. V. HOLDEN, Manchester University Press, 1984.

[69] **Lin, R. M.**
"nonlinear Coupling Analysis Based on a Describing Function Method"
Internal Report No. 88030, Dynamics Section, Mech. Eng. Dept., Imperial College of Science, Technology and Medicine, London.

[70] **Rosenberg, R. and Kuo, J.**
"Nonsimilar Normal Mode Vibration of Nonlinear Systems Having Two Degrees of Freedom"
J. Appl. Mech., 31(2), 1964

[71] **Rosenberg, R.**
"On Nonlinear Vibration of Systems with Many Degrees of Freedom"
Adv. Appl. Mech., 1966

[72] **Rand, R.**
"Nonlinear Normal Modes in Two-Degree-of-Freedom Systems"
J. Appl. Mech., 1971

[73] **Vakakis, A. and Caughey, T.**
"Some Topics in the Free and Forced Oscillations of a Class of Nonlinear Systems"
Dynamics Laboratory Report DYNL-89-1, CalTech, Pasadena, California

[74] **He, J.**
"Identification of Structural Dynamic Characteristics"
Ph.D Thesis, 1987, Mechanical Engineering Department,
Imperial College of Science, Technology and Medicine, London.

[75] **Berman, A.**
"Limitations on the Identification of Discrete Structural Dynamic Models"
Proceeding of Second International Conference on Recent Advances in Structural Dynamics, Southampton, England, 1984

[76] **Young, J. P. and On, F. J.**
"Mathematical Modeling Via Direct Use of Modal Data"
Preprint 690615, Oct. 1969, National Aeronautic and Space Engineering and Manufacturing Meeting, Society of Automotive Engineers, Los Angeles, Calif.

[77] **Berman, A. and Flannelly, W.**
"Theory of Incomplete Models of Dynamic Structures"
AIAA J, Vol. 9, 1971

[78] **Link, M.**
"Theory of a Method for Identifying Incomplete System Matrices from Vibration Test Data"
Z. Flugwiss. Weltraumforsch. 9 (1985), Heft 2

[79] **Luk, Y. W.**
"Identification of Physical Mass, Stiffness and Damping Matrices Using Pseudo-Inverse"
IMAC 5, 1987

[80] **Collins, J. D.; Hart, G. C.; Hasselman, T. K. and Kennedy, B.**
"Statistical Identification of Structures"
AIAA J, Vol. 12, 1974

[81] **Chen, J. C. and Garba, J. A.**
"Analytical Model Improvement Using Modal Test Results"
AIAA J, Vol. 18, 1980

[82] **Lallement, G.**
"Identification Methods IV - Adjustment of Mathematical Model by the Results of Vibration Tests"
To Be Presented to *Identification of Rotationary Structures*, October 1980 CISM, Italy

[83] **Baruch, M. and Bar Itzhack, I. Y.**
"Optimal Weighted Orthogonalisation of Measured Modes"
AIAA J, Vol. 16, 1978

[84] **Baruch, M.**
"Optimization Procedure to Correct Stiffness and Flexibilty Matrices Using Vibration Tests"
AIAA J, Vol. 16, 1978

[85] **Berman, A.**
"Mass Matrix Correction Using an Incomplete Set of Measured Modes"
AIAA J, Vol. 17, 1979

[86] **He, J. and Ewins, D. J.**
"Analytical Stiffness Matrix Correction Using Measured Vibration Modes"
The Intl J of Analytical and Experimental Modal Analysis, Vol. 1, No. 3, July 1986

[87] **Fritzen, C.-P.**
"Identification of Mass, Damping and Stiffness Matrices of Mechanical Systems"
Journal of Vibration, Acoustics, Stress, and Reliability in Design, Vol. 108, 1986

[88] **Mottershead, J. E. and Stanway, R.**
"Identification of Structural Vibration Parameters by Using a Frequency Domain Filter"
Journal of Sound and Vibration, 1986

[89] **Foster, C. D. and Mottershead, J. E.**
"A Method for Improving Finite Element Models by Using Experimental Data: Application and Implications for Vibration Monitoring"
Int. J. of Mech. Sci, Vol. 32, 1990

[90] **Sidhu, J. and Ewins, D. J.**
"Correlation of Finite Element and Modal Test Studies of a Practical Structure"
IMAC 2, 1984

[91] **Wilkinson, J. H.**
"The Algebraic Eigenvalue Problems"
Oxford University Press, 1965

[92] **Golub, G. H. and Van Loan, C. V.**
"Matrix Computations"
North Oxford Academic Publishing Co. Ltd., 1983

[93] **Xu, Q.-Y. and Ma, G.-F.**
"Correction of Structural Dynamic Model Using the Results of Complex Modal Parameter Identification"
IMAC 6, 1988

[94] **Berman, A. and Nagy, E. J.**
"Improvement of Large Analytical Model Using Test Data"
AIAA, 21(8), 1983

[95] **Wei, F. S.**
"Stiffness Matrix Correction from Incomplete Test Data"
AIAA J, Vol. 18, 1980

[96] **Guyan, R. J.**
"Reduction of Mass and Stiffness Matrices"
AIAA J, Vol. 3, 1965

[97] **Downing, A. C. Jr. and Householder, A. S.**
"Some Inverse Characteristic Value Problems"
J. Assoc. Comput. Math., Vol. 3, 1956

[98] **Friedland, S.; Nocedal, J. and Overton, M. L.**
"The Formulation and Analysis of Numerical Methods for Inverse Eigenvalue Problems"
SIAM J. Numer. Anal., No. 3, Vol. 24, 1987

[99] **Gladwell, G. M. L.**
"The Inverse Mode Problem for Lumped-mass System"
Quart. J. Mech. & Appl. Math., Vol. 39, 1986

[100] **Ibrahim, S. R.; Stavrinidis, C.; Fissette, E. and Brunner, O.**
"A Direct Two Response Approach for Updating Analytical Dynamic Models of Structures with Emphasis on Uniqueness"
IMAC 7, 1989

[101] **Golub, G. H. and Kahan, W.**
"Calculating the Singular Values and Pseudo-Inverse of a Matrix"
SIAM J. Numer. Anal. 2, 1965

[102] **Wilkinson, J. H. and Reinsch, C.**
"Linear Algebra"
Hand Book of Automatic Computation, Springer Verlag, 1971

[103] **Chan, T. F.**
"An Improved Algorithm for Computing the singular Value Decomposition"
ACM Trans. on Math. Software, Vol. 8, 1982

[104] **Chan, T. F.**
"Algorithm 581 - An Improved Algorithm for Computing the Singular Value Decomposition [F1]"
ACM Trans. on Math. Software, Vol. 8, 1982

[105] **Forsythe, G. E.; Malcolm, M. A. and Moler, C. B.**
"Computer Methods for Mathematical Computations"
Prentice-Hall, 1977

[106] **Businger, P. A. and Golub, G. H.**
"Algorithm 358 - Singular Value Decomposition of a Complex Matrix [F1, 4, 5]"
Comm. of the ACM, Vol. 12, 1969

[107] **Golub, G. H. and Reinsch, C.**
"Singular Value Decomposition and Least Squares Solutions"
Numer. Math. 14, 1970

[108] **Adelman, H. M. and Haftka, R. T.**
"Sensitivity Analysis of Discrete Structural Systems"
AIAA J, Vol. 24, 1986

[109] **Prasad, B. and Emerson, J. F.**
"A General Capability of Design Sensitivity for Finite Element Systems"
AIAA/ASME/ASCE/AHS 23rd Structures, Structural Dynamics and Material Conf., May, 1982

[110] **Inamura, T.**
"Eigenvalue Reanalysis by Improved Perturbation"
International Journal for Numerical Methods in Engineering, Vol. 26, 1988

[111] **Fox, R. L. and Kapoor, M. P.**
"Rates of Change of Eigenvalues and Eigenvectors"
AIAA J, Vol. 6, 1968

[112] **Nelson, R. B.**
"Simplified Calculation of Eigenvector Derivatives"
AIAA J, Vol. 14, 1976

[113] **Lim, K. B.; Junkins, J. L. and Wang, B. P.**
"Re-examination of Eigenvector Derivatives"
AIAA J, Vol. 10, 1987

[114] **Dailey, R. L.**
"Eigenvector Derivatives with Repeated Eigenvalues"
AIAA J, Vol. 27, 1988

[115] **Ewins, D. J.**
"Measurement and Application of Mechanical Impedance Data"
Journal of the Society of Environmental Engineers, Vol. 14(3), 15(1) and 15(2), 1975-1976.

[116] **Ewins, D. J. and Kirshenboim, J.**
"A Method for Recognising Structural Nonlinearities in Steady State Harmonic Testing"
J. of Vib., Acoustics, Stress, and Reliability in Design, 1984

[117] **Imregun, M.; Robb, D. A. and Ewins, D. J.**
"Structural Modification and Coupling Dynamic Analysis Using FRF Data"
IMAC 5, 1987

[118] **Imregun, M.**
"Dynamic Condensation: Theory and Application"
M.Sc. Thesis, Mech. Eng. Dept., Imperial College of Science, Technology and Medicine, London, 1980

[119] **Parker, T. S. and Chua, L. O.**
"Chaos: A Tutorial for Engineers"
Proceeding of the IEEE 75(8), 1987

[120] **Poddar, B.; Moon, F. C. and Mukherjee, S.**
"Chaotic Motion of Elastic Plastic Beam"
Journal of Applied Mechanics 55, 1986

[121] **Kapitaniak, T.**
"Combined Bifurcations and Transition to Chaos in a Nonlinear Oscillator with Two External Periodic Forces"
Journal of Sound and Vibration 121(2), 1988

[122] **Paîdousis, M. P. and Moon, F. C.**
"Nonlinear and Chaotic Vibrations of a Flexible Pipe Conveying Fluid"
Journal of Fluids and Structures 2, 1988

[123] **Ibrahim, S. R.**
"Computation of Normal Modes from Identified Complex Modes"
AIAA J, Vol. 21, 1983

[124] **O'Callaham, J. C. and Chou, C.-M.**
"Location of Errors in Optimised Mass and Stiffness Matrices Using Modal Test Data"
J. Modal Analysis, January 1989

[125] **She, S. H. and Liu, R. Y.**
"Sensitivity Analysis and Optimisation for the Modal Parameters of Structural Vibration"
IMAC 6, 1988

[126] **Zeischska, H.; Storrer, O.; Leuridan, L. and Vandeurzen, U.**
"Calculation of Modal Parameter Sensitivities Based on a Finite Element Proportionality Assumption"
IMAC 6, 1988

[127] **Bishop, R. E. D. and Johnson, D. C.**
"The Mechanics of Vibration"
Cambridge University Press, 1960

[128] **Timoshenko, S.; Young, D. H. and Weaver, W.**
"Vibration Problems in Engineering"
John Wiley and Sons, 1974

[129] **Zienkiewicz, O. C.**
"The Finite Element Methods in Engineering Science"
MacGraw-Hill, 1971

[130] **Guckenheimer, J. and Holmes, P.**
"Nonlinear Oscillations, Dynamic Systems, and Bifurcation of Vector Fields"
Springer-Verlag, 1983

[131] **To, W. M. and Lin, R. M.**
"A Criterion for the Location of Modification Sites"
IMAC 8, 1990

[132] **Lin, R. M.**
"On the Location of Structural Nonlinearity from Modal Testing"
IMAC 8, 1990

[133] **Lin, R. M.**
"Model Updating Using FRF Data"
15th International Seminar on Modal Analysis, 1990

[134] **Lin, R. M.**
"Identification of Nonlinearity from Analysis of Complex Modes"
Internal Report No. 88001, Dynamics Section, Mech. Eng. Dept., Imperial College of Science, Technology and, Medicine, London.

[135] **Lin, R. M. and Ewins, D. J.**
"Chaotic vibration of Mechanical Systems with Backlash"
Submitted to the Journal of Mechanical Systems and Signal Processing.

[136] **Lin, R. M.**
"On Minimum Data Required to Update An Analytical Model"
Internal Report No. 89010, Dynamics Section, Mech. Eng. Dept.,
Imperial College of Science, Technology and, Medicine, London.

[137] **Lin, R. M. and H. Jung**
"Possibilities and Limitations of Analytical Model Improvement -- A Review Based on the GARTEUR Exercise"
Internal Report No. 90001, Dynamics Section, Mech. Eng. Dept.,
Imperial College of Science, Technology and, Medicine, London.

APPENDIX III

DETAILS OF EXPERIMENTAL WORK

1. THE BEAM/ABSORBER STRUCTURE (USED IN CHAPTER 2)

The Beam/Absorber test rig, as referred in Chapter 2, is an existing apparatus used for standard projects for many years and was not specially designed for research. A photograph of the rig and its main components are shown Fig. A3.1.1 and Fig. A3.1.2. The physical dimension of the main beam and the centre mass are shown in Fig. A3.1.3 and Fig. A3.1.4.

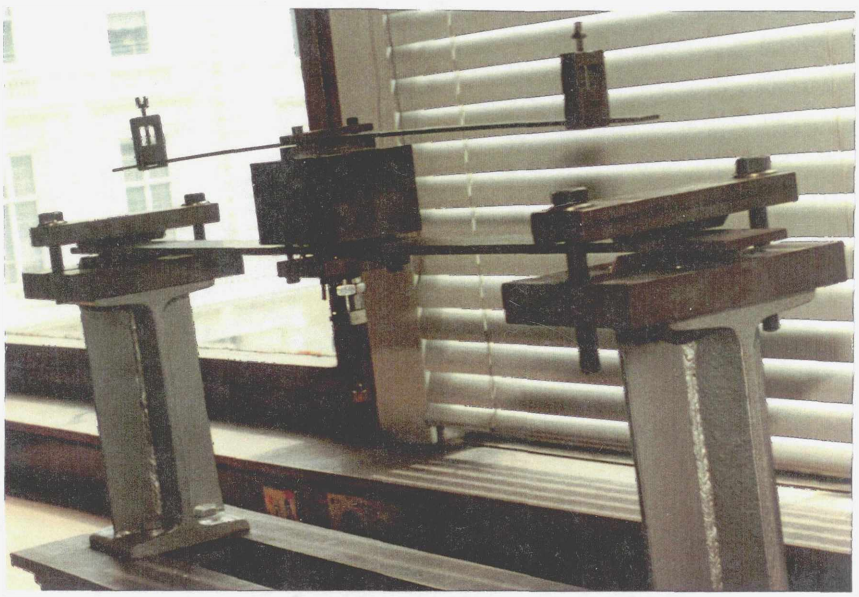


Fig. A3.1.1 - Photograph of the Test Rig

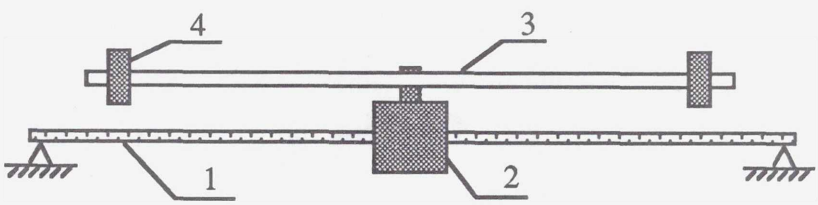


Fig. A3.1.2 - Main Components of the Beam/Absorber Rig

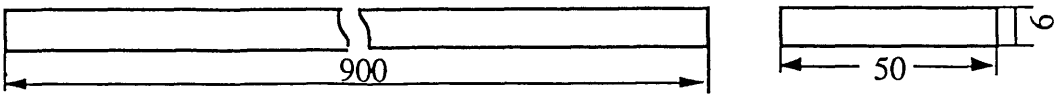


Fig. A3.1.3 - Dimension of Component No. 1 (Main Beam)

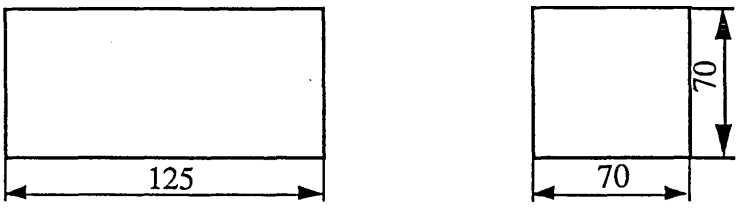


Fig. A3.1.4 - Dimension of Component No. 2 (Centre Mass)

Component No. 3 is a uniform beam with dimension (length x width x height) being 760x20x4 (all dimensions are in mm) and Component No. 4 is a rigid mass with dimension 80x40x40 as shown in Fig.A3.1.1.

2. DETAILS OF CONSTANT FORCE CONTROL

In some modal tests, such as for the measurement of structural nonlinearity, it is required that the amplitude of the excitation force should be kept constant. In cases where the damping of the structure to be tested is not very light (i.e., if the damping ratio is of the order of 1% or more), the amplitude of the excitation force can be controlled to be constant based on a feed-back control loop as shown in Fig.A3.2.1.

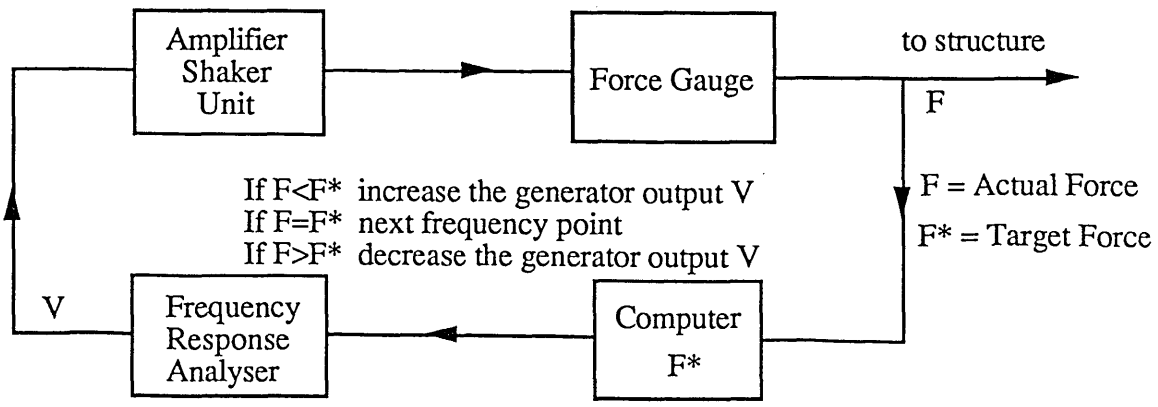


Fig.3.2.1 - Excitation Force Amplitude Control

In the case where the damping of the structure is very low, for the frequency points very near structural resonance, the response amplitude can become very high and the control of constant force becomes difficult, then the FRF data at these frequency points should be left out.