Dynamic Analysis of Non-Linear Structures Using Higher Order Frequency Response Functions

A thesis submitted to *the* University *of* Manchester *for the degree of Doctor of* Philosophy *in the* Faculty *of Science.*

October 1991

David M Storer

Department of Engineering

University of Manchester

Contents

Abstract

Higher Order Frequency Response Functions [FRFs], defined from the Volterra Series, provide a firm basis for analysing the dynamics of a wide class of non-linear systems. In the past, methods have tended to concentrate on the use of random excitation but, as a consequence, have been limited in their application due to the difficulty in measuring the higher order FRFs. The more fundamental approach is to use a single sinewave input to measure higher order transfer functions, which are defined to approximate to the principal diagonals of the multi-dimensional, higher order FRFs. Employing sine excitation introduces simplifications in both the measurement and analysis stages, and estimates of FRFs up to third order can be obtained in practice from non-linear structures.

For non-linear structures which can be regarded as being governed by polynomial restoring force functions, mathematical models can be constructed by extending the conventional differential equations of motion to include non-linear terms. Recently, procedures have been developed for identifying the parameters of mathematical models of this type using time data measured from a structure. Using random excitation, it is possible either to identify a non-linear model directly from time data, or a linearised model starting with measured, conventional transfer functions. Since expressions can be derived which relate the non-linear mathematical models to the higher order FRFs, the principal diagonals of the FRFs can be regenerated from the identified parameters and compared with the higher order TFs measured using sine excitation. In this way, non-linear structures identified in the time domain can be analysed, interpreted and validated in the frequency domain, and vice-versa.

Declaration

No part ***of the work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or institute of learning.***

Acknowledgements

The author would like to thank Professor Geof Tomlinson who has supervised the work and provided considerable advice and assistance, occasional criticism and frequent encouragement throughout the project. The Dynamics Research Group at Manchester which he leads has become an excellent environment for theoretical and experimental studies into the dynamics of structures and, throughout the project, members of the group have advised on technical and practical issues, in particular Drs. Steve Gifford and Keith Worden. Through his own research work, Steve provided the motivation and initiated many of the techniques which have been developed in this project, and Keith has contributed consistently good advice to answer many of the theoretical problems which have arisen. Their help is greatly appreciated.

During the experimental stages of the project, the advice of Dr. Jan Wright and Roy Chng, and the skills of the technicians Paul Townsend and Phil Waters, who manufactured 3 experimental test rigs for this project, contributed a great deal. Roy also accepted responsibility for maintaince of the computer facility, and succeeded in making necessary updates to both software and hardware with minimum inconvenience to the users.

The author would also like to thank Dr. Mike Youngman of LMS Ltd. who assisted in the installation of the commercial software, Brian Jarvis who demonstrated non-linear finite element modelling using PAFEC software, and Colin Skingle and Duncan Eckford of the Royal Aerospace Establishment, Farnborough, who supervised and administered the project on behalf of the Ministry of Defence and provided advice and encouragement during their frequent visits.

The research described in this thesis has been supported by the Procurement Executive of the Ministry of Defence.

Details about the Author

David Storer graduated from the Department of Mechanical Engineering at the University of Southampton in July 1987 with first class honours and, after completing a Ministry of Defence sponsorship and training scheme in the aerospace industry, commenced appointment as research associate at Heriot-Watt University in Edinburgh for one year before transferring to the University of Manchester for the remainder of the three year research agreement.

Introductiori

This thesis describes a relatively new approach to the study of dynamical systems, and attempts to provide a new perspective on the vibration behaviour of mechanical structures (without stipulating that the structure must behave linearly.) Although the techniques which are discussed here are all well established mathematically, only few references have been found which address applications to physical systems. Perhaps the principal reason for this is that the equations which model the dynamics of non-linear systems cannot be solved using conventional mathematical techniques. In fact, exact closed-form solutions seldom exist. As a result, in order to address non-linear systems, methods of analysis are often required which are complicated and involved whilst revealing very little about the underlying nature of the system which is of particular interest to the dynamicist.

In general, the greater the mathematical content, the less attractive the analysis methods appear to the practicing engineer. Often, the main concern is to be able to understand the dynamic behaviour of a structure or component, and provide reassurance that the device will behave safely as designed, perhaps after modification or integration into a larger structure. In this light, one role of the research engineer may be to develop tools, often analytical and computer based, which can be used to assist with the task. A particularly difficult problem would be to predict the response of the system, a task which inevitably entails performing some test to identify the system, usually in terms of a mathematical model. During this project, the author has attempted to re-interpret some of the established methods for studying non-linear systems used predominantly in the engineering fields of electronics and control, and to introduce simplifications in an attempt to apply the techniques to the task of identifying parametric models of mechanical structures.

To introduce the thesis, it is appropriate to start by explaining the motivation behind research into non-linear structures. This is followed by an outline of the development of the Volterra series, the method which has formed the basis of many of the techniques described in this text, together with a brief historical survey of some of the literature published on the subject to date. The next stage is to list the research proposal, as agreed by the supervisors and the sponsoring organisation at the outset of the project, before proceeding to outline the organisation of the text. In the conclusion to the thesis, an account is included of the extent to which the original proposals have been fulfilled. The progress reports, compiled regularly throughout the project, have aimed at addressing directly the details of these proposals. On the other hand, the aim of this thesis is to discuss in general terms how the conventional techniques valid for linear systems may be extended to investigate non-linear structures.

Many physical structures behave in a way which cannot be fully described by a linear model, and can exhibit very complex behaviour which linear systems do not. Significant non-linear effects can arise in built-up structures in which the linking together of sub-structures or appendages can result in clearances between the components. One example is the connection of a store to the wing of an aircraft, as illustrated; the clearance between the wing and the pylon, which supports the store, can give rise to non-linear stiffness effects and friction, a dissipative non-linearity. In certain structures, non-linearities of this kind may be unavoidable, causing significant changes in the behaviour of the structure at different excitation levels. This presents a problem to the dynamicist, particularly if the levels of excitation used in vibration tests are not representative of the conditions during operation.

In some structures, such as aircraft, the location of the source of the non-linearity may be detected by inspecting the components, or performing a relatively straightforward load-deflection test. In most structures, however, detection of non-linearity purely from a physical viewpoint or from a static test would not be sufficient to develop a realistic model.

In the past, a considerable amount of research has been directed at being able to detect from measurements made in a dynamic test whether or not the structure can be considered to be linear, and techniques such as the Hilbert transform have been successful in a wide range of applications. However, the prediction of the dynamic behaviour of non-linear structures has progressed only slowly, principally because of the inability to identify from measured data a reliable numerical model which includes terms to represent the non-linearity. Mathematical models can be constructed to describe the behaviour of a large class of non-linear systems. The objective of the research contained in this thesis has been to define models of this type to represent the dynamics of structures, and to develop practical techniques to identify these models in a dynamic test.

As with all subjects, the route to the present level of understanding of the techniques investigated during this project has been paved by many individuals working in broad range of disciplines, and to keep track of all contributors would be an impossible task. However, it is relatively easy to indicate throughout the text which publications have been traced and used, and to acknowledge some of the authors whose work has been important in establishing the theory related to the Volterra series in particular.

The original studies into functional series by Vito Volterra date back to the 1880s, developing the series by analogy with the Taylor expansion of a function [1]. Further generalisations were provided by Frechet around 1910, whilst maintaining the purely mathematical viewpoint [2].

In 1942 Norbert Wiener, working in the field of communication systems, recognised that the techniques have an application to the characterisation of non-linear systems, and used the Volterra series as a basis on which to define the Wiener series [3].

One of the best survey papers on the properties of these functional series was written by Barrett in 1963. This precised the development of both series and described potential applications to control systems [4]. Also contained in this paper is a description of the use of Fourier transforms to formulate the kernels of the Volterra and Wiener series in the frequency domain, a subject expounded by Bedrossian and Rice in 1971, first proposing the use of harmonic inputs to probe analytical expressions for the kernel transforms, and introducing their definition as Frequency Response Functions (FRFs)[5].

A practical method of measuring the kernels was first proposed by Lee and Schetzen in 1965 who, while working in the field of electronics, demonstrated that the Wiener kernels could be determined directly through cross-correlation with a Gaussian white noise input [6]. Schetzen, originally a student of Wiener, ultimately published an excellent book in 1980 describing the motivation behind Wieners work, comparing the restrictions on the use of the different functional series, but entirely omitting practical results [7].

In the 1970's, Panos and Vasilis Marmarelis published papers on the application of the cross-correlation methods to biological systems such neuron chains [8] and mechanical systems [9], and in 1978 together published perhaps the first book to address functional series from a practical standpoint [10]. In the section 'Dialogue for Epilogue', many of the questions which may be posed by engineers approaching the subject of system identification are answered, providing motivation for the specific application to mechanical structures.

The Wiener series has also found practical application in the general field of signal processing through the techniques of polyspectral analysis, a subject borne out of a short paper by Powers in 1980 which succeeded in summarising the survey paper published by Barrett some 17 years previously, serving to re-introduce the subject to engineers [11]. Now polyspectral methods are an active area of research in several groups seeking new applications of the technology to physical systems, including structures, and the recent book by Bendat (1990) reviews some of the important developments [12].

It should also be remembered that the analytical basis of the functional series has continued to be developed and interpreted by mathematicians. For example, Sandberg has been working since 1963 on the application of functional series in the solution of non-linear differential equations, and despite being concerned with applications to electronics and communications system, this subject is entrenched in mathematics [13],[14].

Of particular interest to this project is the application of the functional series approach to the identification of non-linear structures. Extensive work into this field is being conducted by several research groups including those headed by Billings, Vinh and Tomlinson, each working on different approaches to solving similar problems. Predominately interested in control systems, Billings has developed methods of fitting non-linear discrete difference equations to input and output time data, establishing the relationship between these models and the Volterra kernels [15],[16]. Recent results indicate that this approach is highly promising and may have a very wide application, although as a means of measuring the kernels it is indirect, and suffers the problem that the difference model fitted may not have a physical interpretation, which may affect any subsequent interpretation of the kernels [17].

For direct measurement, Vinh et al. attempted to use multiple impulse testing and met with considerable practical problems [18],[19]. In 1989, Gifford and Tomlinson, succeeded in measuring second order Wiener kernels from mechanical structures using the cross-correlation technique proposed by Schetzen although, as with almost all the practical techniques formerly developed, estimating third or higher order kernels appeared to be impracticable [20]. This work also established for the first time the relationship between the FRFs, defined from the Volterra series, and non-linear differential equation models and, by developing a higher order curvefitting algorithm, attempted to give a physical interpretation to the results [21]. More direct time-domain routes to the identification of models of this type have been investigated in parallel by Mohammad, Worden and Tomlinson [22].

From the original publications on the use of harmonic inputs to probe the FRFs, a subject consolidated in the 1980 book by Spina and Wiener [23], practical methods using sine excitation began to be proposed, and in 1983, Boyd, Tang and Chua published results of a second order FRF measured from a loudspeaker horn [24]. Although the quality of the results were poor, the potential for using multiple sine inputs for measuring FRFs had been established.

The work in the initial stages of this project also concentrated on the use of sinusoidal excitation as a means of simplifying the measurement of higher order FRFs, whilst developing higher order curvefitting methods. In later stages of the project, the direct parameter estimation method was also incorporated, providing a means of developing the higher order FRFs from an identified model, which can be compared with directly measured results. Throughout, the work has been supervised by representatives of the sponsoring organisation, who also advised on the direction of the project which broadly followed the lines of the original research proposal, summarised by the following list,

1. (a) *Review of current methods of structural identification*

   (b) *Implementation of digital simulation package for simulating non-linear systems*

   (c) *Detailed study using digital simulation of SDOF systems which incorporate common non-linearities*

2. (a) *Extension of simulation studies to multi-mode systems*

   (b) *Assessment and development of techniques to detect and identify the form of non-linearity for simple multi-mode systems with one or more spatial non-linearities*

   (c) Design and construct an experimental rig with controlled non-linear characteristics

3. (a) *Experimental evaluation of simple non-linear dissipative and non-dissipative devices for inclusion in the rig*

   (b) *Evaluation of the detection methods developed and the construction of models for predicting the dynamics of the rig*

The first stage of the project focussed on understanding the Volterra series and reviewing the work conducted previously (in particular by Gifford [25] and Worden [26].) Numerical simulation routines were written to be used initially to study a selection of single and multi-degree of freedom systems, and subsequently to assist in the validation of the identification algorithms, as described subsequently.

Effectively, this thesis can be considered to be in two parts. Chapters 1 to 4 outline the mathematical background and, by discussing the different approaches in general terms, an attempt is made to justify the direction of the research by arguing the case for the methods adopted. To illustrate the discussion, results from numerical simulations have been included although only few experimental examples have been considered in this first section. On the other hand, chapters 5 to 7 present results from three simple mechanical systems, the primary objective being to describe how the methods have been implemented in a testing framework based around a commercial, computer software based, data acquisition system. This second part also serves to place the results from the first part in a different context, in order to highlight both theoretical limitations and practical problems associated with the methods presented. A summary of the findings and suggestions for further research are included in chapter 8, together with final conclusions.

Chapter 1 begins by discussing differences between linear and non-linear systems, providing a definition of non-linearity and outlining qualitative tests to detect non-linearity in structures based on this definition. The differential equation models which are considered throughout the thesis are introduced in this section, together with details of the direct parameter estimation technique, before proceeding to discuss how the mathematical models and the parameter identification technique can be extended to describe a class of weakly non-linear structure. The powerful method of the restoring force surface, which can be used when the system is best represented by a single degree of freedom, is also outlined, and the results are used to motivate the investigation of non-linear systems. The chapter concludes with an account of the use of random excitation to 'linearise' a non-linear system along the lines of the original development by Goyder [27].

Starting with an introduction to the Volterra series, with reference to the Marmarelis book [10], chapter 2 proceeds to discuss the properties of the series which make it particularly attractive from a system identification standpoint, before continuing to describe the Wiener series. The chapter finishes by demonstrating how the ordinary coherence function can be used in some situations to detect non-linearity from measurements using random excitation, despite the tendency to linearise the system.

Chapter 3 uses the FRFs, defined from the Volterra series in chapter 2, to re-introduce the differential equation model and to establish the relationship which exists between them. The higher order FRFs are multi-dimensional in frequency, and this chapter includes a general interpretation of the forms of the second and third order FRFs in terms of the energy transfer mechanism of non-linear systems, and a detailed account of the principal diagonals of the FRFs which are single-dimensional in frequency, enabling the discussion to be extended to all higher orders.

The difficulties of measuring estimates of the Volterra and Wiener kernels, first introduced in chapter 2, are addressed again in chapter 4 where a method is proposed for estimating the principal diagonals of the higher order FRFs of structures using sinusoidal excitation. The chapter proceeds to account for the distortion patterns which are frequently exhibited by measured transfer functions, a phenomenon observed when an attempt is made to measure conventional FRFs using sine excitation when the structure happens to be non-linear. The discussion also focusses on the familiar technique of Hilbert Transforms which can be regarded as identifying non-linearity by detection of this distortion. The chapter concludes with a discussion of one approach to system identification.

Chapter 5 is the first which is primarily concerned with practical issues and describes an implementation of the direct parameter estimation procedure to a three degree of freedom structure which behaves linearly, and introduces a discussion on the physical interpretation of the identified model, proposing validity tests in both the frequency and time domains.

The extensions to the direct parameter estimation procedures for identifying non-linear structures are discussed in chapter 6, together with details of the implementation of the stepped sine testing routines to measure an estimate of the principal diagonals of the higher order FRFs. This approach is presented as a means of validating in the frequency domain the differential equation of motion model of a non-linear structure identified in the time domain, using data measured from a non-linear beam rig.

In chapter 7, details are given of a non-linear multi-degree of freedom structure with localised non-linearity, designed and made to meet the specification of the original contract proposal. The rig was intended represent a non-linear pylon-store assembly connected to an aircraft wing, and is a well-behaved structure on which to validate system identification routines. Apart from indicating the limitations of the routines, particularly with regard to locating the source of non-linearity in a structure, the results from this rig also served to introduce again the discussion concerning the physical interpretation of the parametric models identified.

The final chapter is concerned with summarising the conclusions drawn in the preceeding chapters, and by pointing out the problems which have arisen during the course of the project, it has been possible to list suggestions for future work.

Chapter 1

Linear and Non-Linear Structures

1.1 Linear Structures

1.1.1 Testing for Linearity

Ordinarily, the starting point of a structural test should be a linearity check. Most conventional approaches to dynamic analysis make the assumption that the structure is linear, or can be approximated by a linear differential equation model. Considerable advantages are afforded by this simplification, but justifying the assumption of linearity is not easy.

Most commonly, to test for linearity, use is made of the fact that a linear system must obey 'the principle of superposition', which can be understood in terms of two input waveforms applied simultaneously to a system. Given that a system responds to an input $x_1(t)$ with an output $y_1(t)$, and to $x_2(t)$ with $y_2(t)$, superposition holds if the input $ax_1(t) + bx_2(t)$ causes the response to be $ay_1(t) + by_2(t)$ for all constants a and b. Any system which does not obey superposition is non-linear by definition.

A practical test for superposition along these lines would entail first exciting the system with two different inputs, independently measuring the response for each, and then repeating the test with the inputs combined. In such a test, before comparing the sum of the individual responses with the combined response, it would be essential to ensure that, when measuring the signals, the triggering in the individual cases corresponds to the timing between the two input components when combined. Synchronising the measurements in this way is important, but is difficult to achieve in practice.

However, the principle of superposition embodies other properties of linear systems which can be used to specify alternative tests which can be implemented practically. In particular, linear systems are 'homogeneous' and consequently if the magnitude of the input were doubled, the resulting output would also double. Homogeneity can be tested in practice by exciting a structure by a single frequency sinewave at two different amplitudes, and then comparing the amplitudes of the sinusoidal responses. The ratio of output to input amplitudes would remain the same if the structure were linear over the input range tested, although to ensure the system were not governed by, for example, a bilinear stiffness non-linearity, the test would need to be performed over a wide range of amplitudes.

Throughout this chapter, results are included from a test performed on an electronic circuit to illustrate the discussions. The only information known before the test was that the circuit (no.1 in the 'ETH box'[1]) was designed to represent a non-linear SDOF system, which was 'well-behaved', the behaviour approximating to that of a linear structure at low levels of excitation. Therefore this was considered to be a reasonable system to test initially. Also, the practical problems associated with a structural test are largely avoided since the signals into and out of the circuit are voltages; a range of voltage waveforms can be easily generated and applied to the system without requiring an electrodynamic shaker, and the measured response signals do not require conditioning by charge amplifiers. Nevertheless, since this project has been directed at developing analysis procedures which can be applied to physical structures, the practical problems which are often encountered are important, and are addressed in subsequent chapters where tests on three different structures are discussed.

In an attempt to test for non-linearity via non-homogeneity, a sinusoidal signal at amplitudes of $0.5V$ and $5.0V$ was input to the electronic circuit, and the responses measured for each case. The first response signal was then scaled by multiplying by 2, and the second divided by 5, to correspond with an arbitrary datum input of $1V$, enabling a direct comparison of the output waveforms to be made. Before taking the measurements, transients were allowed to die away, so that the comparison was made when the output had attained steady-state. Sections of the scaled time data are shown in figure 1.1 for excitation at 150 Hz, close to the resonant frequency of the system. By comparing the responses, the waveforms are seen to be different, both in terms of amplitude and phase. These results are consistent with the system being non-linear.

[1] A series of electronic circuits with response characteristics analogous to non-linear structures was designed and built for the purpose of validation of identification algorithms by the ETH Zentrum, Institute fiir Werkzengmaschimenban, Zurich, Switzerland.

One of the main themes running through this thesis is how the response of structures to sinusoidal forcing can be used to detect and identify non-linear behaviour. Sinusoidal excitation is particularly useful when identifying non-linearity in structures because it is known that if a linear system is subjected to a sinusoidal input, the output will also be sinusoidal. In practice, deviation of the response signal from a sinusoid can be detected in the frequency domain using a spectrum analyser. Assuming the excitation signal does approximate closely to a sinewave, the presence of components in the response spectra at frequencies other than the fundamental excitation frequency would indicate non-linearity if they occur at the higher *harmonics*, integer multiples of the fundamental frequency.

To investigate the harmonic content of the response signals from the homogeneity test performed on the electronic circuit, the spectra of the scaled sections of time data were also computed and are shown in figure 1.2, displayed on both linear and logarithmic scales. On log. scales, fundamental and higher harmonic components of the response at frequencies 150, 300 and 450 Hz are clearly visible, particularly at the higher amplitude, again indicating non-linearity. However, even at the higher amplitude of 5V, the distortion on the time signal is not particularly obvious, and on the linear scales, the fundamental term in the spectra is seen to remain dominant in both cases, with only the second order harmonic becoming significant at the higher excitation level.

Further tests were conducted to investigate this system in more detail, as described later in this chapter.

1.1.2 Linear Parametric Models

Having attemped to establish that the structure under test behaves nominally linearly, conventionally one of the objectives may be to obtain a mathematical model to provide an estimation of the natural frequencies and damping ratios. In recent years, much of the work devoted to the analysis of structures has concentrated on developing numerical algorithms which curvefit to measured time or frequency response data to provide this information [28]. In this section, a linear mathematical model of a structure is introduced from which the conventional curvefitting procedures have been developed, and which can be extended to describe the dynamics of a non-linear structure, as discussed in section 1.2.1.

Although physical structures would be best represented by a 'distributed parameter' model, the spatial model usually formulated is in terms of idealised mass, spring and damper 'lumped parameter' elements. Equations of motion are developed using an energy balance approach, in which the excitation may be regarded as the source of energy to the structure which stores kinetic energy in the inertia of the mass elements, and potential energy with the stiffnesses. Damping is included to account for energy dissipation and in the mathematical models considered throughout this project, viscous damping has been used which is proportional to velocity. An N degree of freedom linear structure is written as a set of N equations of motion of the form,

$$\begin{aligned} m_s.\ddot{y}_s(t) &+ c_{ss}.\dot{y}_s(t) + k_{ss}.y_s(t) \\ &+ \sum_{\substack{p=a \\ p\neq s}}^{N} [c_{sp}.(\dot{y}_s(t) - \dot{y}_p(t)) + k_{sp}.(y_s(t) - y_p(t))] = x_s(t) \end{aligned} \tag{1.1}$$

where,

s represents one of the N response locations on the structure,

p any other one of the N response locations on the structure,

c_{ss}/k_{ss} denotes the linear damping/stiffness between s and *ground*,

c_{sp}/k_{sp} denotes the linear damping/stiffness between s and p.

For all the response locations, the N equations of motion can be written together in matrix form,

$$[m]\{\ddot{y}(t)\} + [c]\{\dot{y}(t)\} + [k]\{y(t)\} = \{x(t)\} \tag{1.2}$$

Representing structures in terms of a mass matrix which has zero off diagonal terms, and damping and stiffness matrices which are symmetrical, is the usual starting point for identifying linear models of structures.

To emphasise that the electronic circuit considered throughout this chapter is analogous to a mechanical structure, a different notation has been introduced. Despite the system being non-linear, for low levels of excitation the dynamics may be approximated by a linear SDOF model of the form,

$$\mu.\ddot{y}(t) + \eta.\dot{y}(t) + \kappa.y(t) = x(t) \tag{1.3}$$

Clearly, μ corresponds to the mass element, η to the coefficient of viscous damping, and κ to the stiffness.

1.1.3 Harmonic Probing

The equations of motion can be used to represent the dynamic behaviour of a structure in the time domain. An equivalent model can be formulated in the frequency domain by considering the response of the structure to a single harmonic input applied at point r,

$$x_r(t) = X.e^{j\omega t} = X_r(\omega) \tag{1.4}$$

(In this expression, the notation for the 'complex phasor' is introduced, eg. $X_r(\omega)$ is defined as the single harmonic, magnitude X at point r.)

A linear system is characterised as having harmonic output to a harmonic input. In terms of acceleration, velocity and displacement, the frequency of each response signal at each location would have the same frequency as the harmonic forcing, although usually the phase of the signals would be different. Representing the displacement response at location s,

$$y_s(t) = Y.e^{j\omega t} = Y_s(\omega) \tag{1.5}$$

The phase difference between the force and the displacement, and different amplitudes of the two signals, are embodied in the complex phasors.

By differentiating the displacement signal, expressions for the velocity and acceleration responses can be written as $\dot{y}_s(t) = (j\omega).Y_s(\omega)$ and $\ddot{y}_s(t) = (j\omega)^2.Y_s(\omega)$ respectively. Formulating the time responses of a system in this way can be used to establish the relationship between the time and frequency domain models by direct substitution into the equations of motion and this is referred to as *harmonic probing, eg.* equation 1.1,

$$\left((j\omega)^2.m_s + (j\omega).c_{ss} + k_{ss}\right).Y_s(\omega) - \sum_{\substack{p=a \\ p\neq s}}^{N} \left((j\omega).c_{sp} + k_{sp}\right).\left(Y_p(\omega) - Y_s(\omega)\right)$$

$$= \begin{cases} 0 & \text{if } s \text{ is not at the point of excitation} \\ X_r(\omega) & \text{if s coincides with } r \end{cases}$$

$$\left((j\omega)^2.m_s + (j\omega).\sum_{p=a}^{N}.c_{sp} + \sum_{p=a}^{N}.k_{sp}\right).Y_s(\omega) - \sum_{\substack{p=a \\ p\neq s}}^{N}\left((j\omega).c_{sp} + k_{sp}\right).Y_p(\omega) = X_s(\omega) \tag{1.6}$$

This is one of N similar equations which describe mathematically the linear dynamics of the structure in the frequency domain. Each equation incorporates implicitly the response at each of the locations of interest. Note that to maintain generality, the term X,(w) has been included, although this would be zero unless s is at the 'direct point' r, the point of excitation. To obtain a set of scaled modal parameters from curvefitting, or physical mass, stiffness and damping matrices from parameter estimation, it is always necessary to incorporate responses measured at the location of the direct point.

1.1.4 Frequency Response Functions

The method of harmonic probing can be used to relate the time domain model, which was developed with regard to the physics of a structure, to the frequency domain model, which can be used to interpret its resonant behaviour by defining Frequency Response Functions (FRFs). In this section, conventional FRFs are defined for linear systems, whereas a similar approach is used subsequently in chapter 2 to define higher order FRFs for non-linear systems.

For response measured at location s to an excitation applied at point r, the displacement FRF, or ***receptance*** function, is written as,

$$H_{rs}(\omega) = \frac{Y_s(\omega)}{X_r(\omega)} \tag{1.7}$$

The FRF is a complex quantity conveying both the relative magnitude and phase difference between the harmonic displacement to the harmonic force at frequency w.

Simple tests for measuring an estimate of the FRF from a linear structure can be proposed directly from this formulation. (In this text, the function which is ***measured*** is referred to as the Transfer Function (TF), which may only approximate to the FRF which is defined. The distinction is particularly important if the system is non-linear.) Usually a structure can be excited with a harmonic force component by an electrodynamic shaker for example. The spectrum of sampled sections of the time signals can be computed at discrete frequencies using the standard technique of the Fourier transform, denoted $\mathcal{F}[.]$, and the complex value of the TF at the excitation frequency calculated as the ratio of the spectral components at the same frequency line, eg. the receptance function,

$$\mathrm{TF}_{rs}(\omega) = \frac{\mathcal{F}[y_s(t)](\omega)}{\mathcal{F}[x_r(t)](\omega)} \tag{1.8}$$

Similarly, if the response were measured in terms of velocity of acceleration, the mobility function can be written,

$$\mathrm{TF}'_{rs}(\omega) = \frac{\mathcal{F}[\dot{y}_s(t)](\omega)}{\mathcal{F}[x_r(t)](\omega)} \tag{1.9}$$

and the inertance function,

$$\mathrm{TF}''_{rs}(\omega) = \frac{\mathcal{F}[\ddot{y}_s(t)](\omega)}{\mathcal{F}[x_r(t)](\omega)} \tag{1.10}$$

To improve the estimate, the TF can be calculated for several sections of time data, usually arranging for each section to incorporate an integer number of periods of data, and the

results averaged. The procedure can be repeated for a range of excitation frequencies, and the value of the TF computed at each frequency increment over any bandwidth of interest.

Although stepped sine is traditional testing technique, the TFs measured in this way are often of a very high quality in practice. Furthermore, the method can be extended directly to the study of non-linear systems, as discussed in chapter 4. When the structure is linear, the main disadvantage with stepped sine testing is that it can be time consuming, since across a particular frequency range the TFs must be built up one frequency at a time, waiting for the transients to die away after each change. However alternative, faster tests can be proposed which make use of the properties of linearity. Since in a linear system there is no mechanism for energy transfer from one frequency to another, other excitation waveforms can be employed to excite a range of frequencies simultaneously.

To illustrate this, consider the response of a linear structure to random excitation. A broad-band random signal, which may be considered to be a summation of harmonics covering a range of frequencies, can be generated and used to excite a system. Computing the spectra of sections of measured force and response time data, and dividing corresponding spectral components at the same frequency, the TFs can be calculated for each spectral line over the bandwidth of the excitation. Because a linear system obeys superposition, each spectral line in the input can be considered individually, and by matching each with the component in the output at the same frequency, the TF can be constructed across the frequency range of the excitation. This avoids the need to vary the excitation and hence speeds up the test.

By repeating the computations for several sections of time data, the estimate of the FRF, provided by the measured TF, can be improved by averaging. In practice, the quantities which are averaged are the products $Y(\omega).X^*(\omega)$ and $X(\omega).X^*(\omega)$, where * denotes complex conjugate since, as can be shown statistically, the ratio of these quantities converges more quickly to the FRF through averaging,

$$\frac{\overline{Y_s(\omega).X_r^*(\omega)}}{\overline{X_r(\omega).X_r^*(\omega)}} \rightarrow H(\omega)$$

(The overbar denotes the quantities which are averaged.)

Applying the procedure described above to the electronic circuit, the TF was measured with broadband random excitation set at a level of 0.5V RMS, and is shown in figure 1.3. As can be seen, the resonance of the system is close to 150 Hz, and the amplification at this frequency is approximately 2.2, but since both the input and output signals are voltages, the TF is dimensionless.

Although the system is non-linear, the measured TF resembles the characteristic receptance function of a SDOF linear system. There is, however, an important difference. At the low frequency end, and in particular below 10 Hz, the phase can be seen to distort away from the zero level which would be maintained by a receptance function of a SDOF system as the frequency approaches zero. Furthermore, the amplitude of the measured TF dips toward zero rather than converging to an asymptote parallel to the frequency axis. Often discrepancies in measured data are attributed to the presence of non-linearity in the system. However in this case, for reasons discussed in section 1.3, measuring the TF in this way tends to linearise the system. Consequently the spurious low frequency effect in the measured TF cannot have been caused by the non-linearity. The effect is consistent, however, with a high pass filter, cutoff frequency set at approximately 5 Hz, being connected in series with a SDOF system. Unfortunately, in this situation, the effectiveness of the electronic circuit at representing the conventional SDOF system is jeopardised because the conventional differential equation of motion model cannot be used to describe the combined system; the measured TF does not resemble the conventional FRF at low frequencies. Nevertheless at frequencies above 10 Hz, the effect is evidently insignificant, and subsequent tests performed on the circuit concentrated on the resonance region avoiding the 0 to 10 Hz range.

1.1.5 Conventional Curvefitting Model

In this section, equation 1.6 is used to introduce conventional modal analysis curvefitting algorithms, which determine the eigen-frequencies and damping factors from TFs measured from a structure.

The frequency domain formulation of the equations of motion relate *implicity* the response at N locations to the excitation at the direct point. To construct an expression for the FRF at each location, an alternative formulation is required to express the frequency response at any location ***explicity*** in terms of the excitation and the parameters of the system. Rewriting equation 1.6,

$$\left((j\omega)^2.m_s + (j\omega).\sum_{p=a}^{N}.c_{sp} + \sum_{p=a}^{N}.k_{sp}\right).Y_s(\omega) = X_s(\omega) + \sum_{\substack{p=a \\ p\neq s}}^{N}\left((j\omega).c_{sp} + k_{sp}\right).Y_p(\omega)$$

$$Y_s(\omega) = \frac{X_s(\omega) + \sum_{\substack{p=a \\ p\neq s}}^{N}\left((j\omega)._p\beta_1 + {}_p\beta_0\right).Y_p(\omega)}{(\alpha_1 - j\omega).(\alpha_0 - j\omega)} \qquad (1.11)$$

where,

$$\alpha_0 + \alpha_1 = \frac{-\sum_{p=a}^{N} .c_{sp}}{m_s} \; ; \; \alpha_0.\alpha_1 = \frac{\sum_{p=a}^{N} .k_{sp}}{m_s} \; ; {}_p\beta_1 = \frac{c_{sp}}{m_s} \; ; \; {}_p\beta_0 = \frac{k_{sp}}{m_s}$$

Although this formulation expresses the frequency response Y,(w) at any location s on the structure in terms of the parameters and the excitation, it includes the response of each of the other locations implicitly. However it is possible to formulate N similar equations for each response location and, by substitution, determine explicitly the response at s due to an excitation at r,

$$H_{rs}(\omega) = \frac{Y_s(\omega)}{X_r(\omega)} = \sum_{p=1}^{N} \frac{(j\omega).{}_p\gamma_1 + {}_p\gamma_0}{({}_p\lambda_1 - j\omega).({}_p\lambda_0 - j\omega)} \tag{1.12}$$

The coefficients ${}_q\gamma_0, {}_q\gamma_1, {}_q\lambda_0$ and ${}_q\lambda_1$ are related to the parameters of the model but the relationships become more complicated through substitution as the number of degrees of freedom increases. This is an important point since it restricts the ability to progress from the curvefitted modal model to the identification of the parameters in the differential equation model.

The FRF of each response location can be expressed in terms of a summation of single mode linear systems, where q denotes the mode number, and N becomes the total number of modes in the frequency range of interest,

$$H_{rs}(\omega) = \sum_{q=1}^{N} \frac{\omega_q.{}_qA_{rs} + (j\omega).{}_qB_{rs}}{\omega_q^2 - \omega^2 + (j\omega).2\zeta_q\omega_q} \tag{1.13}$$

where,

ω_q = the eigen frequency of mode q

ζ_q = the damping factor of mode q

${}_qA_{rs}, {}_qB_{rs}$ = the modal constants of mode q for excitation at r and response at s

Some modal analysis procedures use this formulation to identify ω_q and ζ_q by curvefitting to the TFs. Furthermore it is possible to show that the modal constants are related to the eigenvectors, and so the modeshapes can be determined by curvefitting to the TFs from each response location.

Using an established viscous-damping algorithm [29], a curvefit was performed on the TF data measured from the electronic circuit, yielding the coefficients,

$$A = 189.98 \; (s^{-1}), B = 0, \; \omega_q = 946.37 \; (rad/s) = 150.6 \; (Hz), \; \zeta_q = 0.0453$$

Since the system has only one degree of freedom, it is possible to work back easily to the parameters of the linearised model described by equation 1.3,

$$\begin{aligned} \mu &= \frac{1}{A.\omega_q} = 5.56 \text{ x } 10^{-6}(s^{-2}) \\ \eta &= \frac{2.\zeta_q}{A} = 4.77 \text{ x } 10^{-4}(s^{-1}) \\ \kappa &= \frac{\omega_q}{A} = 4.98 \end{aligned} \tag{1.14}$$

Using these parameters to regenerate the receptance FRF,

$$H(w) = \frac{1}{(j\omega)^2.\mu + (j\omega).\eta. + \kappa} \tag{1.15}$$

the quality of the **curvefit** can be assessed by comparing directly the regenerated function (figure 1.4) with the measured TF (figure 1.3). Above approximately 10 Hz, and particularly around the resonance, the agreement is seen to be excellent.

In this example, the conventional curvefitting model has been used to estimate the parameters of the equation of motion because the identification of parametric models is the aim of this research. In this way, it is possible to make a direct comparison with estimates provided by other identification procedures, as described subsequently. Since the parameters listed in equation 1.14 have been obtained from a **curvefit** to a linearised TF, they correspond to the optimum linear model for the random excitation signal used when the TF was measured, as described in section 1.3.

Modal models are effectively limited to linear systems because superposition is implicitly assumed in the construction of **FRFs** as a summation of SDOF modes: On the other hand, the equations of motion, from which the curvefitting algorithm has been derived, provide a more general platform since they can be extended to describe non-linear systems, as described in section 1.2.1. Correspondingly, the research in this project has concentrated on developing techniques for identifying parametric models which may include non-linear terms, rather than attempting to interpret differences between measured data and linear models to detect non-linearity. However, as described in chapter 4, the conventional approach to curvefitting could be employed *within* a frequency domain procedure for identifying a non-linear parametric model, if the modal model could be used to synthesise the 'full' FRF matrix, which is the complete set of FRF measurements in which the rows correspond to the response locations and the columns to different excitation points. The **TFs** measured for excitation applied at only one point on the structure provide a direct estimate of only one column of the matrix.

Only one frequency domain model based on viscous damping has been outlined above although alternative algorithms exist using different damping models, or one can fit similar linear modal models in the time domain using for example the complex exponential method [30]. However, because they are all formulations of the modal model, each method encounters the same problem when attempting to predict the full FRF matrix from curvefits to the measured TFs. This problem is related to the residual effect of modes outside the bandwidth selected for the test.

In the modal model, the natural frequency and damping factor of any mode in a structure are assumed to be constants, regarded as global properties which can be estimated from any one of the FRFs. Consequently, the poles of all the FRFs in the full matrix can be determined by curvefitting to the measured TFs. The algorithms also yield the modal constants for one column of the FRF matrix, and by using Maxwells' reciprocity principle, it should be possible in theory to synthesize the modal constants for each of the other columns in the matrix. In practice, however, the number of modes measured depends on the frequency range chosen for the test. Moreover real structures possess a large number of modes, although only a limited number may be of concern. This is important because there will always be higher or lower frequency modes which are not included in the bandwidth to which the measured TFs are limited. As such the residual effects of the out-of-band modes will be present within any bandwidth to some degree, and will affect the accuracy of the curvefit.

Traditionally, modal curvefitting algorithms have attempted to account for the residual effects by including two additional terms in the FRF expression: the effect of the modes at frequencies higher than the maximum cutoff frequency are approximated by a 'stiffness' residual S, and for measurements which are not baseband, a 'mass' residual term M may be included to take account of the lower frequency modes,

$$H_{rs}(\omega) = \frac{1}{(\omega)^2 M_{rs}} + \sum_{q=q_0}^{q_1} \frac{\omega_{q \cdot q} A_{rs} + (j\omega)_{\cdot q} B_{rs}}{\omega_q^2 - \omega^2 + (j\omega).2\zeta_q\omega_q} + \frac{1}{S_{rs}} \tag{1.16}$$

Whilst approximating the effects of out-of-band residuals in this way can improve the comparison between the TF and the modal curvefit, the problem encountered when generating the full FRF matrix remains, since the residual terms cannot be synthesised. The full FRF matrix would be liable to error if generated by omitting residual terms in situations in which the residual effects are significant.

1.1.6 Direct Parameter Estimation

Conventional curvefitting procedures have been largely successful in determining resonant frequencies, damping ratios and modeshapes from data measured from linear structures. Less successful has been the identification of the parameters in the equation of motion model (eqn. 1.1). Parameter identification has become an important area of research in recent years, and is partly motivated by the application to FE element model updating. Several solutions to this problem have been proposed [31],[32],[33], but one of the most straightforward and effective is the Direct Parameter Estimation (DPE) approach [22] in which the parameters are determined simply by substituting measurements of the displacement, velocity and acceleration and excitation force, into the equations of motion.

An outline of the development of the DPE method from the equations of motion is the subject of this section, whereas details of the application of the method to physical structures is described in later chapters, which also highlight practical limitations and extend the points of discussion introduced in this section. To outline the principles behind the DPE approach, consider the equation which represents the motion of the structure at the direct point,

$$m_r.\ddot{y}_r(t) + c_{rr}.\dot{y}_r(t) + k_{rr}.y_r(t) + \sum_{\substack{p=a \\ p\neq r}}^{N} [c_{rp}.(\dot{y}_r(t) - \dot{y}_p(t)) + k_{rp}.(y_r(t) - y_p(t))] = x_r(t) \quad \textbf{(1.17)}$$

The solution procedure assumes that the acceleration, velocity and displacement data can be determined at the same instants in time as the measured force signal. Denoting T as the total number of time instants over which the solution is determined, the data is arranged into a form which can be solved, for example using the established method of Singular Value Decomposition (SVD) [34],

$$[Y].\{b\} = \{x\} \quad \textbf{(1.18)}$$

Writing $\delta_{rp}(t)$ = y,(t) $- y_p(t)$ and $\dot{\delta}_{rp}(t) = \dot{y}_p(t) - \dot{y}_r(t)$,

$$[Y] = \begin{bmatrix} \ddot{y}_r(t_1) & \dot{y}_r(t_1) & \dot{\delta}_{ra}(t_1) & \cdots & \dot{\delta}_{rN}(t_1) & \cdots & y_r(t_1) & \delta_{ra}(t_1) & \cdots & \delta_{rN}(t_1) \\ \ddot{y}_r(t_2) & \dot{y}_r(t_2) & \dot{\delta}_{ra}(t_2) & \cdots & \dot{\delta}_{rN}(t_2) & \cdots & y_r(t_2) & \delta_{ra}(t_2) & \cdots & \delta_{rN}(t_2) \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \ddot{y}_r(t_1) & \dot{y}_r(t_T) & \dot{\delta}_{ra}(t_T) & \cdots & \dot{\delta}_{rN}(t_T) & \cdots & y_r(t_T) & \delta_{ra}(t_T) & \cdots & \delta_{rN}(t_T) \end{bmatrix}$$

$$\{b\} = \{\ m_r \quad c_{rr} \quad c_{ra} \quad \cdots \quad c_{rN} \quad \cdots \quad k_{rr} \quad k_{ra} \quad \cdots \quad k_{rN}\ \}^{\mathbf{T}}$$

$$\{x\} = \{\ x(t_1) \quad x(t_2) \quad \cdots \quad x(t_T)\ \}^{\mathbf{T}}$$

To solve this problem for the parameter vector $\{b\}$, the number of samples T must be at least equal to the maximum number of parameters which is 2N + 1 for the direct point.

For T > 2N + 1, the system of equations is overdetermined, but SVD produces a solution which is the best approximation in a least-squares sense.

Importantly, the parameters determined from the solution of the direct point equation are ***scaled*** with respect to the forcing signal. However the remaining N — 1 equations do not have a force term on the right hand side to provide scaling, and thus solving these equations independently would result in unscaled parameters. One solution to this problem, first proposed by Mohammad [35], is to make direct use of an assumption made when formulating the equations of motion, namely that the stiffness matrix should be symmetrical so that $k_{rp} = k_{pr}$. On this basis, parameters solved in the direct point equation can be passed to scale subsequent equations, for example,

$$m_s.\ddot{y}_s(t) + c_{ss}.\dot{y}_s(t) + k_{ss}.y_s(t) - \sum_{\substack{p=a \\ p\neq r,s}}^{N} [c_{sp}.(\dot{y}_p(t) - \dot{y}_s(t)) + k_{sp}.(y_p(t) - y_s(t))] - c_{rp}.(\dot{y}_p(t) - \dot{y}_r(t)) = k_{pr}.(y_p(t) - y_r(t)) \quad (1.19)$$

Progressing through each of the N equations enables all the parameters to be estimated.

Various schemes could be suggested for optimising the selection of scaling parameters. For example, the standard deviation of the parameters can be determined from the SVD procedure and used as a confidence criterion for the estimation. Alternatively, an iterative scheme could be considered whereby the errors generated by the selection of a certain set of parameters for scaling could be minimised to arrive at the optimal set. Since at each stage only one pre-determined parameter is needed to scale the solution, many options are possible, and the optimisation problem is the subject of further research [36]. On the other hand, the objective of this project has been to implement these procedures into a dynamic test framework and, in the implementation described subsequently, each stiffness parameter previously determined is passed down. In this way the stiffness matrix is constrained to be symmetrical. This option is unlikely to give the optimum solution, although it is adequate to assess in general terms the application of the method to physical structures.

Usually, large numbers of time data samples can be measured, and the requirement that $T \geq$ 2N + 1 is not difficult to meet in practice. However, as T is increased, the least-squares procedure does take longer to perform, particularly for large numbers of response locations. In modal analysis, the selection of the locations for the response measurement is made with regard to representing adequately the modeshapes in the bandwidth of interest, and usually the maximum number of locations is restricted only by the number of transducers available, or some other instrumentation limitation. However, the DPE procedure uses data from each location to solve each equation and, as the number of location increases, the network

For T > 2N + 1, the system of equations is overdetermined, but SVD produces a solution which is the best approximation in a least-squares sense.

Importantly, the parameters determined from the solution of the direct point equation are ***scaled*** with respect to the forcing signal. However the remaining N — **1** equations do not have a force term on the right hand side to provide scaling, and thus solving these equations independently would result in unscaled parameters. One solution to this problem, first proposed by Mohammad [35], is to make direct use of an assumption made when formulating the equations of motion, namely that the stiffness matrix should be symmetrical so that $k_{rp} = k_{pr}$. On this basis, parameters solved in the direct point equation can be passed to scale subsequent equations, for example,

$$m_s.\ddot{y}_s(t) + c_{ss}.\dot{y}_s(t) + k_{ss}.y_s(t) - \sum_{\substack{p=a \\ p\neq r,s}}^{N} [c_{sp}.(\dot{y}_p(t) - \dot{y}_s(t)) + k_{sp}.(y_p(t) - y_s(t))] - c_{rp}.(\dot{y}_p(t) - \dot{y}_r(t)) = k_{pr}.(y_p(t) - y_r(t)) \quad (1.19)$$

Progressing through each of the N equations enables all the parameters to be estimated.

Various schemes could be suggested for optimising the selection of scaling parameters. For example, the standard deviation of the parameters can be determined from the SVD procedure and used as a confidence criterion for the estimation. Alternatively, an iterative scheme could be considered whereby the errors generated by the selection of a certain set of parameters for scaling could be minimised to arrive at the optimal set. Since at each stage only one pre-determined parameter is needed to scale the solution, many options are possible, and the optimisation problem is the subject of further research [36]. On the other hand, the objective of this project has been to implement these procedures into a dynamic test framework and, in the implementation described subsequently, each stiffness parameter previously determined is passed down. In this way the stiffness matrix is constrained to be symmetrical. This option is unlikely to give the optimum solution, although it is adequate to assess in general terms the application of the method to physical structures.

Usually, large numbers of time data samples can be measured, and the requirement that $T \geq$ 2N + 1 is not difficult to meet in practice. However, as T is increased, the least-squares procedure does take longer to perform, particularly for large numbers of response locations. In modal analysis, the selection of the locations for the response measurement is made with regard to representing adequately the modeshapes in the bandwidth of interest, and usually the maximum number of locations is restricted only by the number of transducers available, or some other instrumentation limitation. However, the DPE procedure uses data from each location to solve each equation and, as the number of location increases, the network

of connections becomes complicated very quickly. Moreover, as the model becomes more generalised, the identified parameters are increasingly difficult to interpret.

In the example of the electronic circuit, measuring the output signal is analogous to measuring the displacement response of a SDOF structure. Consequently, to apply the DPE procedure to estimate the parameters of the approximate linear model, the output signal was differentiated to give velocity and acceleration time signals, sections of which are shown in figure 1.5 together with the frequency spectra. Using this data in the DPE procedure gave the following results,

$$\begin{aligned} \mu &= 5.56 \text{ x } 10^{-6}(s^{-2}) \\ \eta &= 4.69 \text{ x } 10^{-4}(s^{-1}) \\ \kappa &= 4.99 \end{aligned} \tag{1.20}$$

which agree very closely with the parameters identified from the modal curvefit (eqn. 1.14). The differences which do arise can be attributed to the fact that the previous identification was performed from linearised TFs, whereas in this case, since the parameters have been identified directly from the time data, no linearisation has taken place. Simply the model has been restricted to first order terms, the non-linearity which is present tending to bias the parameter estimates. (In section 1.2.2 the identified model is extended to include also non-linear terms.) The fact that there is very little difference between the parameters listed in equations 1.14 and 1.19 again indicates that the non-linearity in the system is only very weakly excited, particularly since a relatively low level of excitation (0.5V RMS) was used.

For reasons discussed in chapter 5, it is usually better to measure response in terms of acceleration from a structure directly using accelerometers, and then to integrate the data, rather than to measure displacement and differentiate. In this situation, there was no option. However a simple test for the quality of the numerically calculated data can be made by performing Fourier transforms on the displacement, velocity and acceleration signals, dividing by the spectrum of the excitation signal, and examining the resulting estimates of the receptance, mobility and inertance functions to see if they correspond to characteristic forms for a SDOF system (figure 1.6). Particularly important are the low and high frequency regions where errors from the integration or differentiation are most significant.

One conclusion which has been drawn in this project from the application of the DPE procedure to physical structures is that the DPE does provide a good method of estimating the full FRF matrix. At any frequency w, [H(w)] can be computed directly from the estimated parameters since,

$$[H(\omega)] = \left[(j\omega)^2.[m] + (j\omega).[c] + [k]\right]^{-1} \tag{1.21}$$

(This standard result is derived in chapter 3, where the method of harmonic probing is applied to the equations of motion.)

Using the estimated parameters to synthesise the full FRF matrix of a MDOF structure, problems of residuals can be circumvented. For example, if data from N response locations are included in the model, N lumped mass parameters will be estimated and consequently N modes would be predicted for each FRF. In practice, the DPE procedure uses random excitation which is bandlimited to a frequency range of interest in the same way as direct FRF measurements. However, the 'extra' modes included in the model act to take into account the residual effects within the bandwidth.

1.2 Non-Linear Structures

1.2.1 Non-Linear Parametric Models

In the preceeding discussion of linear systems, techniques were introduced for investigating linear structures, and a mathematical model of a linear structure in terms of the equations of motion was proposed. The aim of this section is to demonstrate how this model, developed from an energy balance approach, can be extended to describe a class of non-linear systems by adding stiffness and damping terms to generalise the representation of the kinetic energy and dissipation effects respectively.

In general, little may be known about a structure before testing, and it is advantageous to consider a non-linear model which does not over-restrict the type of non-linearity which can be represented. An N degree of freedom, n^{th} order non-linear structure is written as a set of N equations with the form,

$$m_s.\ddot{y}_s(t) + \sum_{i=1}^{n}[{}_{ss}c_i.\dot{y}_s^i(t) + {}_{ss}k_i.y_s^i(t)] - \sum_{i=1}^{n}\sum_{\substack{p=a \\ p\neq s}}^{N}\left[{}_{sp}c_i.(\dot{y}_p(t)-\dot{y}_s(t))^i + {}_{sp}k_i.(y_p(t)-y_s(t))^i\right] = x_s(t) \qquad (1.22)$$

where,
${}_{ss}c_i/{}_{ss}k_i$ denotes the i^{th} order non-linear damping/stiffness between s and ground,
${}_{sp}c_i/{}_{sp}k_i$ denotes the i^{th} order non-linear damping/stiffness between s and p.

The non-linear equations of motion are simply an extension of the linear equations, in which the stiffness and damping restoring forces are modelled as *continuous* differential functions. Note that in the cross-linking terms, eg. $(y_p(t)-y_s(t)), y_s(t)$ and $y_p(t)$ have been reversed, and a negative sign included to compensate for the change; the even order contributions to the restoring force, eg. ${}_{sp}k_2.(y_p(t)-y_s(t))^2$, give rise to a force which is always in the same (negative) sense. To correspond with this notation, it is sufficient to stipulate that,

$${}_{sp}c_i = {}_{ps}c_i \;;\; {}_{sp}k_i = {}_{ps}k_i \;: i = \text{an odd integer}$$
$${}_{sp}c_i = -{}_{ps}c_i \;;\; {}_{sp}k_i = -{}_{ps}k_i \;: i = \text{an even integer}$$

Starting with a generalised model of a structure of this type enables a wide class of structures to be represented, without 'a-priori' information. However, in this formulation, the non-linear functions are represented as polynomials. Consequently systems which are governed by

discontinuous differential functions such as piece-wise linear stiffness or Coulomb friction can only be approximated by this model. The reason for imposing the restriction of polynomial force functions in this way is that the models are used in conjunction with the Volterra series representation, as detailed in chapter 2, and systems governed by discontinuous functions do not have a Volterra series representation. In general, the approximation of a discontinuous function by a continuous polynomial would improve if more higher order terms were included. However, for MDOF systems, the model would become extremely complicated if terms higher than 3^{rd} order were included.

Although one of the advantages with this approach is that very little information is required before formulating a representative non-linear model, a-priori information can be used nevertheless to simplify the model before proceeding to the parameter identification stage. It might be known or intuitively obvious, for example, that the non-linearity is in the stiffness function rather than the damping, and correspondingly the higher order damping terms could be omitted from the model.

In the example of the non-linear electronic circuit, the only fact known before the identification tests was that the system is intended to be SDOF. Consequently, the possibility that non-linearity exist in both the stiffness and damping functions was maintained in the extended differential equation of motion,

$$\mu.\ddot{y}(t) + \eta_1.\dot{y}(t) + \eta_2.\dot{y}^2(t) + \eta_3.\dot{y}^3(t) + \kappa_1.y(t) + \kappa_2.y^2(t) + \kappa_3.y^3(t) = x(t) \qquad (1.23)$$

which was used as the model for the non-linear parameter estimation of the circuit as described in the next section.

1.2.2 Extended Direct Parameter Estimation

Having constructed the non-linear differential equation of motion model, it is possible to propose a straightforward extension of the DPE procedure to enable the mass and the stiffness and damping parameters of each order to be identified. Maintaining the notation introduced in section 1.1.6, the non-linear differential equations of motion (eq. 1.22) can be solved using SVD rearrangement in the form,

$$[Y].\{b\} = \{x\} \qquad (1.24)$$

where the matrix $[Y]$ and vector {b} have been extended to include non-linear functions,

$$[Y] = \begin{bmatrix} \ddot{y}_r(t_1) & \dot{y}_r(t_1) & \dot{y}_r^2(t_1) & \cdots & \dot{\delta}_{rN}(t_1) & \cdots & \dot{\delta}_{rN}^n(t_1) & y_r(t_1) & y_r{}^2(t_1) & \cdots & \delta_{rN}^n(t_1) \\ \ddot{y}_r(t_2) & \dot{y}_r(t_2) & \dot{y}_r^2(t_2) & \cdots & \dot{\delta}_{rN}(t_2) & \cdots & \dot{\delta}_{rN}^n(t_2) & y_r(t_2) & y_r{}^2(t_2) & \cdots & \delta_{rN}^n(t_1) \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \ddot{y}_r(t_T) & \dot{y}_r(t_T) & \dot{y}_r^2(t_T) & \cdots & \dot{\delta}_{rN}(t_T) & \cdots & \dot{\delta}_{rN}^n(t_T) & y_r(t_1) & y_r{}^2(t_T) & \cdots & \delta_{rN}^n(t_T) \end{bmatrix}$$

$$\{b\} = \{ \; m_r \;\; {}_{rr}c_1 \;\; {}_{rr}c_2 \;\; \cdots \;\; {}_{rN}c_1 \;\; \cdots \;\; {}_{rN}c_n \;\; {}_{rr}k_1 \;\; {}_{rr}k_2 \;\; \cdots \;\; {}_{rN}k_n \; \}^{\mathbf{T}}$$

$$\{x\} = \{ \; x(t_1) \;\; x(t_2) \;\; \cdots \;\; x(t_T) \; \}^{\mathbf{T}}$$

The vector of the force terms, $\{x\}$, remains the same as for the linear model.

Comments were made in section 1.1.6 regarding the complexity of the linear models for large numbers of response locations. When the models are non-linear, model complexity becomes an even more important consideration. For example, solving the direct point equation for the most general non-linearity would involve determining ***2.n.N*** **+** 1 parameters (where ***n*** is the order of the non-linearity, and N is the number of response locations.) Clearly, simplifying this very general model would be advantageous. Furthermore, the task of selecting the scaling parameters is broadened, since the non-linear parameters are also reciprocal, eg. ${}_{rp}k_3 = {}_{pr}k_3$. In the implementation of the extended DPE used in this project, only the first order parameters are passed down for scaling, and the symmetry of the higher order terms is not imposed directly.

In general, to identify the higher order terms, a system is tested with a relatively high level excitation since most continuous non-linearities have a greater influence on the behaviour of the system as the level is increased. Discontinuous non-linearities, such as friction, often have the opposite effect, although if they are to be approximated by a polynomial, a high level of excitation is again appropriate.

Clearly for SDOF systems such as the electronic circuit, the parametric models are not so complicated, and the problem of selecting scaling parameters does not arise since the system is still described by a single equation which involves the input signal. Using time data measured from the system at the higher level of excitation, 5V RMS, the following parameters were identified,

$$\mu = 5.630 \text{ x } 10^{-6}(\text{s}^{-2})$$

$$\eta_1 = 4.734 \text{ x } 10^{-4} \text{ (s-i)}, \quad \eta_2 = -1.384 \text{ x } 10^{-7}\,(\text{s}^{-1}\text{V}^{-1}), \quad \eta_3 = 3.137 \text{ x } 10^{-13}\,(\text{s}^{-1}\text{V}^{-2})$$

$$\kappa_1 = 5.052\ (-), \qquad \kappa_2 = 7.253 \text{ x } 10^{-2}(\text{V}^{-1}), \qquad \kappa_3 = -9.600 \text{ x } 10^{-3}(\text{V}^{-2})$$

(1.25)

Although it is often possible to identify the parameters in an extended model using a similar procedure as for a linear model, validation of the estimates is not as straightforward. For a linear system, it is possible to regenerate, for example, the conventional receptance FRF from the parameter estimates and make a direct comparison with the measured TF. In this example, it is only possible to make a comparison of μ, η_1 and κ_1 with the corresponding quantities when the identification was restricted to the linear model, and although the mass estimates should correspond, differences must be expected in the first order stiffness and damping parameters when the system is non-linear because restricting the model to be first order causes bias in the first order estimates (equation 1.19). It is assumed that, by extending the model to include non-linear terms, the estimates of the first order parameters will be more accurate. The fact that the estimates of μ are consistent, but that κ_1 differs from κ by 1 %, and η from η_1 by 1.2%, once again indicates that, at this level of excitation, the non-linear effects are still only relatively small.

Although it may be possible to validate first order parameters by comparison with a linear model, the problem of validating the higher order parameters remains. Nevertheless validating identified parameters is important and must be considered in conjunction with developing methods of identifying non-linear models. (As described subsequently, a method for comparing higher order TFs with regenerated higher order FRFs each of which exist only if the system is non-linear, has been developed during this project as a means of validating the higher order parameters.) Also, because they have different dimensions, it is not possible to determine the significance of the higher order parameters in the mathematical model simply by comparing the numerical values with the corresponding first order parameters. Perhaps the easiest way is to use the identified parameters to regenerate directly the polynomial functions, and to then compare the significance of each order in the polynomial. However, it is first necessary to decide on an appropriate range for the response since, for non-linear functions, the higher order terms in the polynomial can be greatly exaggerated if the range is either under- or over-estimated.

In the example of the electronic system, the appropriate response ranges for the damping and stiffness polynomials, regenerated from the parameters listed in equation 1.25 and shown in figures 1.7 and 1.8, were determined along the following lines. At the resonant frequency of 150 Hz, the magnitude of the measured receptance type TF was approximately 2.2 (figure 1.3). Consequently, a sinusoidal input of amplitude 2.5V would result in a output signal y(t) with an amplitude of approximately 5.5V at resonance. Hence $\pm$5V was chosen as the range for the stiffness polynomial. Also if the amplitude of the receptance function at ω_0 is 2.2, the amplitude of the corresponding mobility function would be 2.2 x ω_0.

Correspondingly a 150 Hz sinusoidal input of amplitude 2.5V would result in a signal j;(t) with amplitude 2.5 x 2.2 x $2\pi.150$ = 5184 (V/s). Hence f5000 (V/s) was chosen as the range for the damping polynomial.

From the polynomials regenerated for the electronic circuit, the damping function is apparently very non-linear, with a very significant quadratic effect but a negligable cubic term over the range selected. However this may also be a result of the unusual low frequency phase characteristic evident in the measured TF, which will also affect the measured time data, and hence the first and higher order parameter estimates, in particular the damping. The stiffness function is evidently non-linear also, incorporating both second and third order terms, but over this input range, the first order term appears dominant.

1.2.3 Restoring Force Surfaces

The method of regenerating the damping and stiffness functions from the identified parameters is useful in determining how the non-linear terms influence the dynamics of the system. In a technique first proposed by Masri and Caughey [37], it is possible to combine both damping and stiffness restoring force functions together to form a 3-dimensional plot known as the 'restoring force surface', and in this way make a qualitative and quantitive assessment of the non-linearity. Moreover, it is possible to compare the non-linearity in the damping and stiffness functions, and determine which effect is more dominant over the response range selected.

The restoring force method was developed simply by reformulating the equations of motion using Newton's second law; for a SDOF system,

$$m.\ddot{y}(t) + f(y,\dot{y}) = x(t) \tag{1.26}$$

where,

$$f(y,\dot{y}) = \sum_{i=1}^{n} c_i.\dot{y}^i(t) + \sum_{i=1}^{n} k_i.y^i(t) \tag{1.27}$$

$f(y,\dot{y})$ is called the 'internal restoring force function' and combines both the damping and stiffness effects together above the phase plane, (y, $\dot{y}$). Having identified all the stiffness and damping parameters, $f(y,\dot{y})$ can be calculated for any selected range of $\dot{y}(t)$ and y(t).

The method of regenerating the restoring force surface using the identified parameters can be applied similarly for MDOF systems since, from equation 1.22,

$$f((y_p - y_s),(\dot{y}_p - \dot{y}_s)) = -\sum_{i=1}^{n} \left[{}_{sp}c_i.(\dot{y}_p(t) - \dot{y}_s(t))^i + {}_{sp}k_i.(y_p(t) - y_s(t))^i \right] \tag{1.28}$$

A separate restoring force surface could be regenerated for each link in the model, eg. between s and p, in terms of the differences between the velocity and displacement response signals at the two locations.. Clearly constructing restoring force surfaces for each of the N links between each of the N locations and ground, and between the $\frac{1}{2}N(N-1)$ links between the locations would be time consuming if a large number of response locations were to be used.

Apart from simply regenerating the restoring force function, the method finds a particularly important application for SDOF systems, since measured data can be plotted directly,

$$f(y,\dot{y}) = x(t) - m.\ddot{y}(t)$$

if an estimate of the effective mass m can be determined beforehand, without requiring estimates of the damping and stiffness parameters. In this way, the restoring force function $f(y,\dot{y})$ can be determined at each point in the phase plane (y, $\dot{y}$) from the measured force signal, $x(t)$, and the acceleration $\ddot{y}(t)$ time histories, and the stiffness and damping parameters can be identified subsequently from the surface.

When applied in practice to SDOF structures, eg. an automotive shock absorber [38], it is usual that the response is measured either as acceleration or velocity or displacement. Whichever form is selected for measurement, signals in each form, corresponding to exactly the same points in time, are required to construct the force surface. Consequently, the most pragmatic approach is to measure acceleration and perform numerical integration in post-processing to estimate the velocity and displacement signals. In this way it is possible to calculate the triplet $(y(t_k),\dot{y}(t_k),f(y(t_k),\dot{y}(t_k)))$ at each sampling instant t_k, and plot the restoring force function for discrete points across the phase plane.

In order to reveal any non-linearity in the system, it is also important that the restoring force function cover as large an area of the phase plane as possible. Clearly, this is governed by the type and level of the excitation selected for the test. A wide variety of waveforms are commonly used to excite systems in tests and extensive investigations have been undertaken independently by Worden [39] and Al-Hadid [40], to determine which are the most suitable for this method. The deductions were made from tests on numerically simulated systems, using most of the commonly used excitation waveforms, including impulse and 'chirp', and the conclusion was that random and sine each have certain advantages over the alternatives.

Random excitation has the advantage over sine testing in that a large area of the phase plane can be covered in a single test and, as a consequence, is less time consuming. The random signal is usually bandlimited to the frequency range of interest, often around the resonance, to avoid any spurious low or high frequency effects. Numerical integration or

differentiation necessary for the response signals can also introduce spurious effects on the data, as described subsequently, and it is necessary also to numerically filter the data in post-processing to remove these effects before constructing the force surface.

Figure 1.9 shows the restoring force surface measured from the electronic circuit using random excitation bandlimited to between 10 and 300 Hz in order to encompass the resonance, but to avoid the low frequency region in which the phase of the measured TF (figure 1.3) appeared spurious. The relatively high level of excitation, 5V RMS, was selected to encourage the non-linear behaviour, but it is apparent from the distribution of the data points in the phase plane that the energy is nevertheless concentrated in the low level region.

With sine testing it is possible to cover a large area of the phase plane by conducting a series of tests incrementing either the amplitude of excitation, or the frequency so that the system passes through resonance for example. After the test, when the data has been processed, the results are combined to form a single restoring force surface.

Figure 1.10 illustrates the restoring force surface constructed from data measured from the electronic circuit, maintaining the excitation frequency close to the system resonance, but incrementing the input amplitude in eight approximately equal $0.5\mathrm{V}$ steps up to 4V. In this way, it is possible to sustain the vibration of a system at a high level and thus promote the non-linearity in the response. In the phase plane, the individual tests are evident as closed loops; since the system is non-linear and the response signals are not sinusoidal but include higher harmonics (as illustrated in figure 1.1 and 1.2), each loop deviates from the ellipsoid (or circle) expected for a linear system.

Both of these force surfaces were constructed using μ = 5.630 x $10^{-6}(\mathrm{s}^{-2})$, previously determined using DPE (equ. 1.25). In practice the need for an accurate mass estimate may present a problem particularly since it has been found that the force surface is very sensitive to errors in the value used. In the absence of a-priori information, it may be necessary to use the DPE initially on the data set to estimate the mass parameter before the restoring force function is calculated. However, if a mass estimate is known, the restoring force method can be used to identify the first and higher order stiffness and damping parameters by curvefitting to cross-sections of the surface. Along the displacement axis $\dot{y}(t)$ = 0,

$$f(y, \dot{y} = 0) = \sum_{i=1}^{n} k_i . y^i(t) \qquad (1.29)$$

By fitting a polynomial function to the values of the restoring force as it crosses the i(t) = 0 plane, the stiffness parameters can be estimated. Similarly the damping parameters can be obtained by curvefitting to $f(y = 0, \dot{y})$.

The damping and stiffness cross-sections of the force surfaces measured from the electronic circuit are shown in figures 1.11 and 1.12 respectively corresponding to cases of random and sine excitation. Both sets of results are consistent and have very similar forms to the damping and stiffness polynomials (figs. 1.7 and 1.8) regenerated from parameters identified using DPE. The principal difference between the two forms of excitation is that with sine it is possible to sustain the excitation at the higher levels. Consequently, the range of the functions is wider and any non-linearity becomes more apparent. In addition, the periodic signals can be numerical filtered and processed more accurately to remove the effects of measurement noise, so both the force surface and the cross-sections are less polluted. The fact that for both excitation waveforms, the stiffness cross-section is less noisy than the damping indicates that the stiffness effects dominate the restoring force function, an observation which is also confirmed by the overall shape of the force surfaces.

In order to estimate the parameters of the electronic circuit, polynomials were fitted to the cross-sections of the restoring force function measured using sine excitation to yield the parameters,

$$\eta_1 = 4.958 \times 10^{-4}(s^{-1}), \quad \eta_2 = -1.077 \text{ x } 10^{-7}(s^{-1}V^{-1}), \quad \eta_3 = -6.137 \text{ x } 10^{-13}(s^{-1}V^{-2})$$
$$\kappa_1 = 5.159, \qquad \kappa_2 = 1.051 \text{ x } 10^{-1}(V^{-1}), \qquad \kappa_3 = -1.148 \text{ x } 10^{-2}(V^{-2}) \tag{1.30}$$

Despite using the same mass estimate, these results are not exactly the same as the parameters identified with DPE (equation 1.25). The differences reflect the fact that fitting polynomials to measured data is extremely sensitive, and small changes in the lower order parameters can cause significant changes in the higher order parameters when performing the least squares fit., Nevertheless the measured restoring force functions closely resemble the functions regenerated previously (figure 1.4) thus validating the results.

For SDOF systems, the principal advantage of the restoring force method over DPE is that measured data is plotted initially without imposing any particular model on the structure, and it is possible to identify stiffness or damping functions which are not of the continuous polynomial type, for example bilinear stiffness. The disadvantage is that a mass estimate is required, which may require resorting to DPE anyway to provide this information [41]. For MDOF systems, estimates of the stiffness and damping parameters would be required in addition to the mass elements before force surfaces between any of the links in the model can be constructed [42]. This necessarily involves performing a full parameter estimation, and it transpires that the restoring force surface method for MDOF systems is equivalent to direct parameter estimation.

1.3 Linearisation of Non-Linear Systems

One technique which has proved useful in the practical testing of non-linear structures, is the use of broadband random excitation to 'linearise' the non-linearity in a system. Whether or not the system is linear, the TF measured using random excitation (as described in section 1.1.4) has the characteristic form of the conventional FRF defined from the linear model. This is true for both SDOF and MDOF structures. Linearisation is useful since conventional curvefitting techniques can be applied to TFs measured from non-linear structures and, in this way, it may be possible to identify a linear parametric model which approximately represents the system.

In this section, a procedure is outlined to demonstrate that the conventional TF measured using,

$$\mathrm{TF}(\omega) = \frac{\overline{Y(\omega).X^*(\omega)}}{\overline{X(\omega).X^*(\omega)}} \tag{1.31}$$

can be regarded as being an estimate of the FRF of the 'optimum linear model' of the non-linear system, defined for the particular input signal $x(t)$ used when the TF was measured. Theoretically, as the number of averages increases, the TF converges to this optimum, linear FRF.

The block diagrams, figures 1.13 and 1.14, have been included to illustrate the discussion which starts by considering how the output y'(t) of a linear model would differ from the response y(t) of the non-linear system when subjected to the same input $x(t)$.

Denoting y,(t) = y(t) − y'(t), the problem of determining the particular linear model which is the optimum, is equivalent to selecting the particular linear model which minimises y,(t) for the input $x(t)$. From the point of view of the linear model, y,(t) can be regarded as an error signal which arises from the non-linearity, and the optimisation can be formulated as minimising the mean square error of this signal,

$$\mathrm{MSE} = \int_{-\infty}^{\infty} |y_e|^2(t).dt \tag{1.32}$$

To proceed, it is convenient to consider the frequency domain representation of the system. representing the transfer characteristics of the linear model in terms of the FRF H'(w),

$$\mathrm{Y'(U)} = H'(\omega).X(\omega) \tag{1.33}$$

The output of the non-linear system can be written as,

$$Y(\omega) = H'(\omega).X(\omega) + Y_e(\omega) \tag{1.34}$$

where for example,

$$Y_e(w) = \mathcal{F}[y_e(t)] = \int_{-\infty}^{\infty} y_e(t).e^{j\omega t}.dt \tag{1.35}$$

The minimisation of the MSE can also be reformulated in the frequency domain using the definition of the Fourier transform,

$$\begin{aligned} \text{MSE} &= \int_{-\infty}^{\infty} |y_e|^2(t).dt = \int_{-\infty}^{\infty} y_e(t).y_e(t)^*.dt \\ &= \frac{1}{2\pi}.\int_{-\infty}^{\infty} y_e(t).\left[\int_{-\infty}^{\infty} Y_e(\omega).e^{-j\omega t}.d\omega\right].dt \\ &= \frac{1}{2\pi}.\int_{-\infty}^{\infty} Y_e(\omega).\left[\int_{-\infty}^{\infty} y_e(t).e^{-j\omega t}.dt\right].dw \\ &= \frac{1}{2\pi}.\int_{-\infty}^{\infty} Y_e(\omega).Y_e^*(\omega).d\omega \end{aligned} \tag{1.36}$$

In practice, the spectra are computed using DFT routines on digitised data, and over the bandwidth the spectral estimates are limited to discrete frequencies spaced 1/T apart. Consequently, the integral is replaced by the 'expected value of an ensemble average', denoted by an overbar, which must also be scaled by the factor 1/T to correspond with the integral in the limit. Correspondingly, the criterion for optimisation can be formulated as the minimisation of the discretised frequency domain MSE,

$$\text{MSE} = \frac{1}{2\pi T}.\overline{Y_e(\omega).Y_e^*(\omega)} \tag{1.37}$$

Having established that an optimum FRF exists, denoted by $H_0(\omega)$, it is always possible to write H'(w) as $H_0(\omega)$ plus a second, arbitrary, linear FRF scaled by the parameter ϵ,

$$H'(\omega) = H_0(\omega) + \epsilon.H_\epsilon(\omega) \tag{1.38}$$

The error signal can be written,

$$Y_e(\omega) = Y(\omega) - (H_0(\omega) + \epsilon.H_\epsilon(\omega)).X(\omega) \tag{1.39}$$

The auxiliary function, $H_\epsilon(w)$, has been introduced so that optimisation can be formulated in terms of the new parameter ϵ which tends to zero as the optimum conditions are reached. Moreover, the partial derivative of the MSE with respect to ϵ also becomes zero at the minimum,

$$\begin{aligned} \frac{\partial[\text{MSE}]}{\partial\epsilon} &= \frac{1}{2\pi T}\frac{\partial}{\partial\epsilon}\left[\overline{(Y(w) - H_0(\omega).X(\omega) - \epsilon.H_\epsilon(\omega).X(\omega))}\right. \\ &\left.\cdot\ \overline{(\mathbf{Y^*(w)} - H_0^*(\omega).X^*(\omega) - \epsilon.H_\epsilon^*(\omega).X^*(\omega))}\right] = 0 \end{aligned} \tag{1.40}$$

Performing the partial differentiation, and setting $\epsilon = 0$,

$$\overline{H_\epsilon(\omega).X(\omega).H_0^*(\omega).X^*(\omega)} - \overline{H_\epsilon(\omega).X(\omega).Y^*(\omega)} + \overline{H_\epsilon^*(\omega).X^*(\omega).H_0(\omega).X(\omega)} - \overline{H_\epsilon^*(\omega).X^*(\omega).Y(\omega)} = 0 \tag{1.41}$$

Observing that the terms are in conjugate pairs, the solution to this equation is given by,

$$\overline{H_\epsilon^*(\omega).X^*(\omega).H_0(\omega).X(\omega)} = \overline{H_\epsilon^*(\omega).X^*(\omega).Y(\omega)} \tag{1.42}$$

Hence,

$$H_0(\omega) = \frac{\overline{Y(\omega).X^*(\omega)}}{\overline{X(\omega).X^*(\omega)}} \tag{1.43}$$

Thus, the TF measured from a non-linear structure using random excitation in the usual way, does have the same form as a FRF of a linear system, ie. there is none of the distortion which is characteristic of the TF measured from a non-linear structure with sine excitation, as described in chapter 4.

From the viewpoint of the optimum linear model, the non-linearity of the system can be considered as contaminating the TF. In this approach, the error signal y,(t) may be regarded as a source of 'noise' on, the output of the optimum linear model, since there can be no linear correlation of y,(t) with the input signal $x(t)$, ie. y,(t) cannot be obtained from $x(t)$ by means of a linear operation. However y,(t) is correlated with $x(t)$, and some non-linear operation must exist to relate y,(t) with $x(t)$. In particular, as the level of the input $x(t)$ is varied, the behaviour of the non-linearity in the system will also vary. The effect will be manifest as a change in both the error signal $y_e(t)$, and the output of the system y(t). Consequently the optimum linear model identified would also be dependent on the level of the input signal.

A test was performed on the electronic circuit to verify that the excitation level does affect the noise content on the measured TF when the system is non-linear. The TF was measured using random excitation at RMS levels of $0.5V$ and 5V (a relatively high level), without averaging (figs. 1.15 and 1.16). Clearly, at the higher level of excitation, the contamination due to the non-linearity is significant. However, repeating the tests with 10 averages, it was observed that the contamination reduced to such an extent that hardly any difference between the two TFs could be detected, each approaching the function shown in figure 1.3 (which was measured with 50 averages).

As has been remarked previously, the measured TFs appear to deviate from the conventional receptance type FRF at low frequencies, particularly the phase. Often discrepancies in measured TFs are attributed to the effects of non-linearity in a system. However it is clear that this effect cannot be attributed to the non-linearity, despite the system being non-linear, because by measuring the TFs in this way, the system has been linearised.

For a SDOF system, it is often useful to compute the inverse of the TF to validate the linearisation, for example, since the real and imaginary parts of the inverse FRF have characteristic forms when plotted against frequency. The linearised TF of the electronic circuit has been assumed to approximate to a receptance function,

$$\frac{1}{H(\omega)} = (j\omega)^2.\mu + (j\omega).\eta. + \kappa = (\kappa - \omega^2.\mu) + j(\omega.\eta) \qquad (1.44)$$

The real part of the inverse receptance function has the form of a polynomial function with zero order coefficient κ and second order coefficient $-\mu$, whereas the imaginary part is a straight line with gradient η.

In figure 1.17, the real and imaginary parts of the inverted TF, measured with 50 averages, are shown plotted against frequency. From the value of the real part of the inverse TF when w $\rightarrow$ 0, κ = 5.0. Also, at the resonance ω_0 for which the real part becomes zero, $\kappa - \omega_0^2.\mu = 0$, giving μ = 5.6 x 10^{-6} (s^{-2}). These are consistent with the previous estimates of the mass and stiffness parameters (eg. equation 1.14).

Although the real part does have approximately the same form as the inverse receptance FRF, the imaginary part of the inverse TF can be seen to differ considerably from a straight line. This deviation corresponds with the distortion in the phase which is particularly apparent at the lower frequencies. However, from the inverse function, it is evident that this effect, which appears to be caused by a high pass filter connected in series with the system, actually extends across the entire frequency range to some degree. Consequently, the apparent non-linearity in the damping function detected by the restoring force surface method is likely to be influenced by this effect.

In the event, the electronic circuit did not behave exactly like a SDOF mechanical system, and therefore is not ideal for validating the identification techniques. However, the results from this system have assisted in introducing methods of detecting non-linearity in structures and direct parameter estimation, a technique which is discussed in a practical context throughout this thesis.

Chapter 2

The Volterra and Wiener Series

2.1 Introduction

In chapter 1, equations of motion were introduced as a means of modelling the dynamics of both linear and non-linear structures. These differential equations express the response of the system *implicitly* in terms of the input. Although it is usually possible to determine solutions for ordinary linear differential equations, exact solutions for non-linear differential equations exist only for a few specific cases. This prevents the response of a general non-linear system being formulated *explicitly* in terms of the input, except in the form of an infinite series. If such a series could be found, truncating to a finite number of terms would correspond to determining an approximate solution to the non-linear differential equation. The accuracy of such an approximation would be determined by the order of the truncation and the number of terms maintained in the series.

One such approximation method, called the Volterra series, has been used in several applications to investigate the behaviour of physical non-linear systems. The principal reason why the Volterra series may have an important role to play in the characterisation of non-linear structures is because it has a form which enables familiar concepts, such as convolution integrals and frequency response functions, which can only be used to characterise linear structures, to be extended and redefined for a wide class of non-linear structures. These extensions form a natural progression. The first term in the series is related to the solution of a linear differential equation and reflects the fact that a linear system is a special Volterra system. Furthermore, the Volterra series is now well established mathematically, a fact which lends confidence to the approach in general.

2.2 The Volterra Series

2.2.1 Developing the Volterra Series

In the time domain, a linear structure is usually characterised either as differential equations of motion, or as a convolution integral. As described in chapter 1, the equations of motion can be extended to describe non-linear structures. However, since the convolution integral embodies the principle of superposition, it cannot be used to characterise any non-linear system. The aim of this section is to demonstrate how the principles behind the convolution integral, and the definition of impulse response functions for linear systems, can be used to develop the Volterra series and define higher order impulse response functions for non-linear systems.

For a linear system, the response y(t) to an input $x(t)$ can be represented by y(t) = $\mathbf{L}[x(t)]$ where L[.] is a linear operator. One specific input which is of theoretical importance is the Dirac function $\delta(t)$, defined as a rectangular pulse of ***unit area*** such that $\int_{-\infty}^{\infty} \delta(t).dt = 1$.

The impulse response function of a linear system $h(t)$ is defined as the response to the Dirac impulse, $h(t) = \mathbf{L}[\delta(t)]$, and the convolution integral can developed by considering the response to this idealised input, and approximating any arbitrary input signal by a series of impulses:

1. The response of a linear system at any time t to a Dirac pulse applied τ seconds previously can be written,

$$y(t) = h(\tau).\delta(t-\tau)$$

and the response to an impulse width At and amplitude X can be written,

$$y(t) = h(\tau).\delta(t-\tau).X.\Delta t \tag{2.1}$$

2. Any arbitrary signal $x(t)$ can be approximated by a series of impulses,

$$x(t) = \sum_{i=1}^{\infty} x(t-\tau_i).\Delta\tau \tag{2.2}$$

In this 'staircase approximation', each impulse has equal width AT, and $x(t-\tau)$ is the instantaneous amplitude of the signal $x(t)$ at time $(t-\tau)$.

3. Since a linear system obeys the principle of superposition, the response to an input composed of a series of impulses must be equal to the sum of the responses to the individual

impulses,

$$y(t) = \mathbf{L}[x(t)] = \sum_{i=1}^{\infty} h(\tau_i).x(t-\tau_i).\Delta\tau \tag{2.3}$$

4. As the width of each impulse is reduced, $\Delta\tau \rightarrow 0$, the approximation of the signal $x(t)$ approaches the exact representation and in the limit,

$$\mathbf{y(t)} = \int_0^{\infty} h(\tau).x(t-\tau).d\tau \tag{2.4}$$

This general approach can be extended to non-linear systems making specific use of the fact that superposition does not hold. By understanding how the response would differ from a linear system as a train of impulses is applied, a general representation of the response to any input can be developed.

As an example, consider a SDOF system with a polynomial stiffness non-linearity of order n,

$$\ddot{y}(t) + \alpha.\dot{y}(t) + \beta_1.y(t) + \sum_{i=2}^{n} \beta_i.y^i(t) = x(t) \tag{2.5}$$

where β_1 is the first order stiffness parameter,
and for i $\geq$ 2, β_i is the i^{th} order non-linear stiffness parameter.
For each non-linear system a related linear *system* can be conceived which, in this example, is approached as,

$$\sum_{i=2}^{n} \beta_i.y^i(t) \rightarrow 0 \tag{2.6}$$

This condition either necessitates that the response level becomes so small that the contribution of the first order term dominates the stiffness function, or that β_i = 0 for all i $\geq$ 2, in which case the system would be linear by definition. The related linear system is characterised by the *first* orderimpulse response function $h_1(t)$ via the convolution integral,

$$y(t) = \int_0^{\infty} h_1(\tau).x(t-\tau).d\tau \tag{2.7}$$

If $\beta_i \neq 0$ for any i $\geq$ 2, the system is non-linear and representing the response using a convolution integral must be inexact. Using ϵ_1 to denote the difference between the responses of the non-linear system and the related linear system, the non-linear system response to a single Dirac impulse at time $(t-\tau_1)$ can be written as,

$$y(t) = h_1(\tau_1).\delta(t-\tau_1) + \epsilon_1 \tag{2.8}$$

Clearly the magnitude of the error ϵ_1 would depend on the parameters, and would be some function of the amplitude of the impulse.

Similarly, if two impulses were applied, one at $t = \tau_1$ and the other at $t = \tau_2$, the response of the non-linear system could be written as,

$$y(t) = h_1(\tau_1).\delta(t - \tau_1) + h_1(\tau_2).\delta(t - \tau_2) + \epsilon_2 \tag{2.9}$$

The magnitude of the difference term ϵ_2 would be a function of the system parameters and the amplitude of each impulse, and would be larger than ϵ_1 in general.

Since the action of the non-linear term in the differential equation is to multiply output terms together, it may be assumed that at least part of the difference term ϵ_2 will be a function of the ***product*** of the two impulse amplitudes, and that the response of the non-linear system can be written as,

$$y(t) = h_1(\tau_1).\delta(t - \tau_1) + h_1(\tau_2).\delta(t - \tau_2) + h_2(\tau_1, \tau_2).\delta(t - \tau_1).\delta(t - \tau_2) + \epsilon_2' \tag{2.10}$$

The term $h_2(\tau_1, \tau_2)$ is called the second order impulse response function which depends on the delays τ_1 and τ_2 of each of the two impulses, and it is assumed that $\epsilon_2' < \epsilon_2$.

The same argument can be extended to three impulses,

$$\begin{aligned} y(t) &= h_1(\tau_1).\delta(t - \tau_1) + h_1(\tau_2).\delta(t - \tau_2) + h_1(\tau_3).\delta(t - \tau_3) + \epsilon_3 \\ \epsilon_3 &= h_2(\tau_1, \tau_2).\delta(t - \tau_1).\delta(t - \tau_2) \\ &+ h_2(\tau_1, \tau_3).\delta(t - \tau_1).\delta(t - \tau_3) \\ &+ h_2(\tau_2, \tau_3).\delta(t - \tau_2).\delta(t - \tau_3) \\ &+ h_3(\tau_1, \tau_2, \tau_3).\delta(t - \tau_1).\delta(t - \tau_2).\delta(t - \tau_3) + \epsilon_3' \end{aligned} \tag{2.11}$$

where $\epsilon_3' < \epsilon_3$, and $h_3(\tau_1, \tau_2, \tau_3)$ is the ***third*** orderimpulse response function in the expression,

$$\begin{aligned} y(t) &= \sum_{i=1}^{3} h_1(\tau_i).\delta(t - \tau_i) \\ &+ \sum_{i_1=1}^{3}\sum_{i_2=1}^{3} h_2(\tau_{i_1}, \tau_{i_2}).\delta(t - \tau_{i_1}).\delta(t - \tau_{i_2}) \\ &+ h_3(\tau_1, \tau_2, \tau_3).\delta(t - \tau_1).\delta(t - \tau_2).\delta(t - \tau_3) + \epsilon_3' \end{aligned} \tag{2.12}$$

At each stage, the representation of the response is assumed to be continually improving as the higher order terms are introduced. The justification of this assumption is simply that, by expressing the response in a more general way, the approximation must either be improving, or staying the same. In fact, if the system happens to be linear, terms higher than first order each give a zero contribution and, in this situation, all the higher order impulse response functions are said to be 'null'.

For systems in which the non-linearity has odd order symmetry, ie. each of the even order parameters are zero, all the even order impulse response functions are also null. However, for a wide class of non-linear systems, it is possible to prove that the approximation of the response will be closer to the exact representation when the second order terms are included in addition to the first order terms in the expression of the response.

To proceed, consider the interaction of n impulses,

$$\begin{aligned} \mathbf{y(t)} \;=\; & \sum_{i=1}^{n} h_1(\tau_i).\delta(t-\tau_i) \\ & + \sum_{i_1=1}^{n}\sum_{i_2=1}^{n} h_2(\tau_{i_1},\tau_{i_2}).\delta(t-\tau_{i_1}).\delta(t-\tau_{i_2}) \\ & + \sum_{i_1=1}^{n}\sum_{i_2=1}^{n}\sum_{i_3=1}^{n} h_3(\tau_{i_1},\tau_{i_2},\tau_{i_3}).\delta(t-\tau_{i_1}).\delta(t-\tau_{i_2}).\delta(t-\tau_{i_3}) + \ldots \\ & + h_n(\tau_{i_1},\ldots,\tau_{i_n}).\delta(t-\tau_{i_1})\ldots\delta(t-\tau_{i_n}) + \epsilon_n' \end{aligned} \qquad \textbf{(2.13)}$$

The input to the system which has been considered to this stage has consisted only of a series of similar, discrete idealised impulses each occurring with arbitrary delay. However, the response to any arbitrary input signal $x(t)$ can be developed from this formulation by approximating $x(t)$ with an infinite, continuous series of impulses of width AT, equation 2.2. In the limit, as $\Delta\tau \rightarrow 0$ and n $\rightarrow\infty$, the response to any arbitrary input can be represented by the infinite series,

$$\begin{aligned} y(t) \;=\; & \int_0^{\infty} h_1(\tau_1).x(t-\tau_1).d\tau_1 \\ & + \int_0^{\infty}\int_0^{\infty} h_2(\tau_1,\tau_2).x(t-\tau_1).x(t-\tau_2).d\tau_1.d\tau_2 \\ & + \int_0^{\infty}\int_0^{\infty}\int_0^{\infty} h_3(\tau_1,\tau_2,\tau_3).x(t-\tau_1).x(t-\tau_2).x(t-\tau_3).d\tau_1.d\tau_2.d\tau_3 + \ldots \\ & + \int_0^{\infty}\ldots\int_0^{\infty} h_n(\tau_1,\ldots,\tau_n).x(t-\tau_1)\ldots x(t-\tau_n).d\tau_1\ldots d\tau_n + \ldots \end{aligned} \qquad \textbf{(2.14)}$$

This is called the Volterra series, and the higher order impulse response functions, $h_n(\tau_1,\ldots,\tau_n)$ are also known as the Volterra kernels of the system.

There are several restrictions on the applicability of the Volterra series to physical systems which are mentioned throughout this chapter. One of the main constraints is that the parameters of the system must not change with time. Although for most structural systems, the parameters either remain constant or vary only slowly with time, there are many physical situations in which, for example, the mass changes significantly within the time scale of interest. It would be necessary to use a different approach to analyse time varying systems of this kind.

2.2.2 Features of the Volterra Series

Causality

By developing the Volterra series representation in the way outlined in the previous section, causality of the system has already been assumed. A system is causal if, for any input, the response at any instant in time does not depend on the future of the input. On the basis of causality, the limits of the convolution integrals can be extended from $(0, \infty)$ to $(-\infty, \infty)$ by stipulating that the Volterra kernel $h_n(\tau_1, \ldots, \tau_n)$ is zero when any of the arguments $\tau_1, \ldots, \tau_n$ are less than zero, leading to the more familiar formulation of the Volterra series,

$$\begin{aligned} y(t) = & \int_{-\infty}^{\infty} h_1(\tau_1).x(t-\tau_1).d\tau_1 \\ & + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h_2(\tau_1, \tau_2).x(t-\tau_1).x(t-\tau_2).d\tau_1.d\tau_2 \\ & + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h_3(\tau_1, \tau_2, \tau_3).x(t-\tau_1).x(t-\tau_2).x(t-\tau_3).d\tau_1.d\tau_2.d\tau_3 + \ldots \\ & + \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} h_n(\tau_1, \ldots, \tau_n).x(t-\tau_1) \ldots x(t-\tau_n).d\tau_1 \ldots d\tau_n + \ldots \end{aligned} \tag{2.15}$$

The assumption of causality does not limit the practical application of the Volterra series since all known physical systems are causal.

Homogeneity

The convolution integrals of the Volterra series are defined to be ***homogeneous.*** For example, consider the n^{th} order component of the response,

$$y_n(t) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} h_n(\tau_1, \ldots, \tau_n).\prod_{i=1}^{n} x(t-\tau_i).d\tau_1 \ldots d\tau_n \tag{2.16}$$

If the amplitude of the input $x(t)$ were changed, the amplitude of all the response components would also change, but there would be no variation in the ***shape*** of each component waveform.

All linear systems are homogeneous. By definition, a system which is not homogeneous is non-linear, and the shape of the output waveform would distort if the amplitude of the input were varied. On this basis, the Volterra series can be regarded as a means of describing the response of a non-homogeneous system as a parallel combination of homogeneous subsystems, and this is illustrated in figure 2.1.

If the Volterra kernels were defined in a different way, the outputs of each operator would not be homogeneous for all inputs. As a consequence of homogeneity, the Volterra kernels form a unique description of a time-invariant system, and each kernel is independent of the type or level of the input and can be regarded as a property of the system.

Using the Volterra series representation, it is possible to consider the response of a non-linear system y(t) as a summation of individual components which are each the result of some operation on the input signal,

$$y(t) = y_1(t) + y_2(t) + y_3(t) + \ldots + y_n(t) + \ldots = \sum_{i=1}^{\infty} y_i(t) \tag{2.17}$$

in which,

$$y_n(t) = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} h_n(\tau_1, \ldots, \tau_n) . \prod_{i=1}^{n} x(t-\tau_i) . d\tau_1 \ldots d\tau_n = H[x(t)] \tag{2.18}$$

For an input $x'(t) = ax(t)$, the response $y'(t)$ is,

$$y'(t) = ay_1(t) + a^2 y_2(t) + a^3 y_3(t) + \ldots + a^n y_n(t) + \ldots = \sum_{i=1}^{\infty} a^i y_i(t) \tag{2.19}$$

Each of the higher order operators $\mathbf{H_n}[.]$ performs a different function and generate a different 'power' contribution to the response. In practice, measuring the Volterra kernels would necessitate separating each individual component in the output signal somehow. The most direct approach would be to measure the responses of the system at several excitation levels. Theoretically, by truncating the series to order *n* and approximating the representation of the response signal, the unknown output components $y_1(t), \ldots, y_n(t)$ could be identified by measuring the response at m different excitation levels and solving the equations,

$$\begin{Bmatrix} y(t)|_{a_1} \\ y(t)|_{a_2} \\ y(t)|_{a_3} \\ \\ y(t)|_{a_m} \end{Bmatrix} = \begin{bmatrix} a_1 & a_1^2 & a_1^3 & \cdots & a_1^n \\ a_2 & a_2^2 & a_2^3 & \cdots & a_2^n \\ a_3 & a_3^2 & a_3^3 & \cdots & a_3^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_m & a_m^2 & a_m^3 & \cdots & a_m^n \end{bmatrix} \cdot \begin{Bmatrix} y_1(t) \\ y_2(t) \\ y_3(t) \\ \\ y_n(t) \end{Bmatrix} \tag{2.20}$$

The main requirement to be met before a solution can be found using a least-squares procedure (such as SVD) is that $m \geq n$. In a structural test, this should not be too difficult to arrange for. One considerable disadvantage, however, is that any errors in the estimation of the excitation level factors $a_1, \ldots, a_m$ would have a significant effect, particularly on the accuracy of the higher order components identified.

Using different excitation levels to identify a non-linear system does seem to be a logical approach, exploiting the basic fact that the operator which relates *y(t)* to $x(t)$ is not homogeneous. This technique has not been used in this project simply because an alternative route to system identification has been adopted which uses the DPE procedure to identify a parametric model to develop the Volterra kernels, rather than vice-versa, as described subsequently.

The property of homogeneity, possessed by all Volterra kernels, can be used to develop a very interesting result which holds for all Volterra systems; each of the higher order operators $\mathbf{H_n}[.]$ can be considered in terms of a combination of first order operators $\mathbf{H_1}[.]$.

To illustrate this, consider a system which has a 3^{rd} order polynomial stiffness non-linearity,

$$\ddot{y}(t) + \alpha.\dot{y}(t) + \beta_1.y(t) + \beta_2.y^2(t) + \beta_3.y^n(t) = x(t) \tag{2.21}$$

For the input $ax(t)$, $\mathbf{y(t)} = \sum_{i=1}^{\infty} a^i y_i(t)$ and,

$$\sum_{i_1=1}^{\infty} a^{i_1}.\ddot{y}_{i_1}(t) + \alpha.\sum_{i_1=1}^{\infty} a^{i_1}.\dot{y}_{i_1}(t) + \beta_1.\sum_{i_1=1}^{\infty} a^{i_1}.y_{i_1}(t) + \beta_2.\sum_{i_1=1}^{\infty}\sum_{i_2=1}^{\infty} a^{i_1+i_2}.y_{i_1}(t).y_{i_2}(t) + \beta_3.\sum_{i_1=1}^{\infty}\sum_{i_2=1}^{\infty}\sum_{i_3=1}^{\infty} a^{i_1+i_2+i_3}.y_{i_1}(t).y_{i_2}(t).y_{i_3}(t) = ax(t) \tag{2.22}$$

An expression for the i^{th} order component of the response can be determined from this equation simply by equating coefficients of a^i,

$$a^1 \quad : \quad \ddot{y}_1(t) + \alpha.\dot{y}_1(t) + \beta_1.y_1(t) = x(t) \tag{2.23}$$

$$a^2 \quad : \quad \ddot{y}_2(t) + \alpha.\dot{y}_2(t) + \beta_1.y_2(t) + \beta_2.y_1(t)^2 = 0 \tag{2.24}$$

$$a^3 \quad : \quad \ddot{y}_3(t) + \alpha.\dot{y}_3(t) + \beta_1.y_3(t) + 2.\beta_2.y_1(t).y_2(t) + \beta_3.y_1(t)^3 = 0 \tag{2.25}$$

$$y_1(t) = \mathbf{H_1}[x(t)] = [s^2 + \alpha.s + \beta_1]^{-1} \tag{2.26}$$

$$y_2(t) = \mathbf{H_2}[x(t)] = -\mathbf{H_1}[\beta_2.y_1(t)^2] \tag{2.27}$$

$$y_3(t) = \mathbf{H_3}[x(t)] = -\mathbf{H_1}[2.\beta_2.y_1(t).y_2(t) + \beta_3.y_1(t)^3] \tag{2.28}$$

Although only the first three terms in the response have been written down, for this system the Volterra series representation of the response has an infinite number of terms. This is despite the polynomial having terms to third order only.

From these equations it can be seen that any higher order output component y;(t) can be expressed in terms of the non-linear parameters, β_2 and β_3, and lower order components using only the first order operator $\mathbf{H_1}[.]$ by substituting for all the higher order operators. Furthermore it would be possible to relate each of the higher order output components back to the input signal $x(t)$ by arranging first order operators in a network, as shown in figure 2.2.

The fact that a relationship does exist between $\mathbf{H_n}[.]$ and $\mathbf{H_1}[.]$ is fundamental to the Volterra series. As a consequence, it is also possible to show that the higher order impulse response functions can always be written as a combination of first order impulse response functions. This result has important implications when extended to the frequency domain as described in chapter 3, since it can be used to explain the mechanism for energy transfer characteristic of non-linear systems.

From this example, it is clear that the ability to represent the response of the system by a Volterra series (with an infinite number of terms) is tied to the fact that the non-linear functions are *continuous* **and have a polynomial form. Strictly systems governed by discontinuous functions do not have a Volterra series representation, although the series may give an approximation in the same way as a discontinuity can be approximated by a polynomial by including more higher order terms.**

In the previous example, the non-linearity function was formulated implicitly as a polynomial relating the stiffness restoring force function f(y) to the level of the response y,

$$f(y) = \beta_1.y + \beta_2.y^2 + \beta_3.y^3 = \sum_{i=1}^{3} \beta_i.y^i$$

A polynomial is a specific example of the general Taylor series expansion of a function. For example, a Taylor series for k independent variables $y_a, y_b, \ldots, y_k$ can be written,

$$f(y_a, y_b, \ldots, y_k) = \gamma_0 + \sum_{i_1=a}^{k} \gamma_{i_1}.y_{i_1} + \sum_{i_1=a}^{k}\sum_{i_2=a}^{k} \gamma_{i_1 i_2} y_{i_1}.y_{i_2} + \sum_{i_1=a}^{k}\sum_{i_2=a}^{k}\sum_{i_3=a}^{k} \gamma_{i_1 i_2 i_3} y_{i_1}.y_{i_2}.y_{i_3} + \ldots$$

The restoring force is a function of the *instantaneous* **response and is independent of previous response levels. In the same way, the Taylor series expansion of a function is said to have no 'memory'. For a system described by differential equations, determining the Volterra series entails inverting the equations and expanding the expression into a 'functional' series, in order to express the response** *y(t)* **explicitly in terms of the excitation $x(t)$. The functional series expansion can be regarded as a generalisation of a Taylor series representation of a function, in which the arguments are the values of the continuous function $x(t)$ in some interval. Because the response at any instant is related to the present and previous amplitudes of the excitation, the Volterra series is said to be a 'generalised Taylor series with memory'.**

To characterise a non-linear system with a Volterra series, it is also important that the system memory is not infinite, and that the effect of an instantaneous change in the input signal will tend to diminish with time. This fact is reflected in the interpretation of the Volterra kernels as (higher order) impulse response functions,

$$h_n(\tau_1, \ldots, \tau_n) \to 0 \quad \text{as} \quad \tau_1, \ldots, \tau_n \to \infty$$

This is always true for linear systems and is an important statement for this class of non-linearity since it implies the system will eventually reach a steady-state condition which is not governed by the initial conditions.

Some 'strongly' non-linear systems do have infinite memory. A specific example of this is a system which behaves chaotically [43]. In such a system, the response at any point in time has, by definition, 'sensitive dependence on the initial conditions', and therefore cannot be termed a Volterra system. On this basis, the Volterra series is usually said to characterise 'weakly' non-linear systems which have a single, stable solution.

One of the limitations of the Taylor series expansion is that, to approximate a function with a finite number of terms, it is necessary that the series be convergent. In the example of the polynomial stiffness function, the limit of convergence would correspond to a maximum response level. The Volterra series also suffers the problem of convergence, and the limit of convergence corresponds to a maximum level of excitation. In general, some level of excitation would exist to cause the Volterra series to diverge. This prevents the Volterra series from characterising a general non-linear system for all possible inputs.

Assuming that a Volterra series can be used to characterise a system, and that the strength of the non-linearity is such that at some level of excitation the Volterra series does converge, one possible approach to system modelling may be to identify the Volterra kernels. For a system which does comply with all the necessary conditions, each Volterra kernel can be shown to be a property of the system, independent of the excitation, and unique. Unique Volterra kernels are also symmetrical, for example $h_2(\tau_1, \tau_2) = h_2(\tau_2, \tau_1)$.

The uniqueness property is also useful from the point of view of modelling the system since, if it were possible to measure the Volterra kernels by testing the system with an input of a particular type and level, the Volterra series could be used to predict the response to any other input. It would be advantageous to exploit this feature by applying the Volterra series to the task of identifying a structure, since it is usually possible to control the conditions of a structural test, and to select the excitation which yields information regarding the behaviour of the structure in the most direct way. In this context, the experimental procedure would be to perform some test which would enable the Volterra kernels to be determined from the measurements, and then relate the kernels to the parameters of the model. The application of the Volterra series to the identification of structures has been one of the principal topics of this project, and the subject is addressed throughout this thesis.

2.2.3 Fourier Transforms of the Volterra Kernels

For a linear system, a Frequency Response Function (FRF) can be defined to characterise a system using the fact that the harmonic is an eigenfunction. A single harmonic cannot be generated although a single sinusoid can be. Because of superposition, a sinusoid, which is the sum of two harmonics, is also an eigenfunction of a linear system. Furthermore, since the FRF can be interpreted as quantifying the relationship of the system response to sinusoidal excitation both in magnitude and phase, an FRF can be measured directly from a linear system using sine excitation.

On the other hand, the Impulse Response Function (IRF) cannot be measured directly since the ideal impulse cannot be generated in practice. Intuitively, since both the IRF and FRF can be used to characterise the same linear system, it is clear that there must be some relationship between the IRF and the FRF, and each must embody the same information regarding the system.

The motivation for studying the relationship between the IRF and FRF in the context of the Volterra series is provided by the ability to interpret the Volterra kernels as higher order IRFs[44], and by the fact that the FRF defined for a linear system does not characterise a non-linear system. The objective of this section is to show how higher order FRFs can be defined from the higher order IRFs by extending the relationship which exists for a linear system between these two equivalent representations.

Consider the response of a linear system to the sinewave $x(t) = Xe^{j\omega t} + Xe^{-j\omega t}$ as represented by the convolution integral,

$$\begin{aligned} y(t) &= \int_{-\infty}^{\infty} h(\tau).x(t-\tau).d\tau \\ &= \int_{-\infty}^{\infty} h(\tau).\left[Xe^{j\omega(t-\tau)} + Xe^{-j\omega(t-\tau)}\right].d\tau \\ &= Xe^{j\omega t}.\int_{-\infty}^{\infty} h(\tau).e^{-j\omega\tau}.d\tau + Xe^{-j\omega t}.\int_{-\infty}^{\infty} h(\tau).e^{j\omega\tau}.d\tau \end{aligned} \qquad (2.29)$$

in which h(r) denotes the impulse response function.

The response y(t) to a sinewave at frequency w, would be sinusoidal at the same frequency with amplitude Y and relative phase angle θ,

$$Ye^{j(\omega t+\theta)} + Ye^{-j(\omega t+\theta)} = Xe^{j\omega t}.\int_{-\infty}^{\infty} h(\tau).e^{-j\omega\tau}.d\tau + Xe^{-j\omega t}.\int_{-\infty}^{\infty} h(\tau).e^{j\omega\tau}.d\tau \qquad (2.30)$$

By introducing the complex phasors X(w) and Y(w) which have a relative phase of θ,

$$Y(w) + Y^*(w) = X(\omega).\int_{-\infty}^{\infty} h(\tau).e^{-j\omega\tau}.d\tau + X^*(\omega).\int_{-\infty}^{\infty} h(\tau).e^{j\omega\tau}.d\tau \qquad (2.31)$$

The complex number $H(\omega)$ can be used to define the FRF of the linear system,

$$H(\omega) = \frac{Y(\omega)}{X(\omega)}, = \frac{Y^*(\omega)}{X^*(\omega)} : |H(\omega)| = \frac{Y}{X} \text{ and } \angle H(\omega) = \theta \tag{2.32}$$

$$H(\omega) = \int_{-\infty}^{\infty} h(\tau).e^{-j\omega\tau}.d\tau \tag{2.33}$$

and correspondingly,

$$h(\tau) = \frac{1}{2\pi}.\int_{-\infty}^{\infty} H(\omega).e^{j\omega\tau}.d\omega \tag{2.34}$$

These two single dimensional **functionals** are known as the Fourier transform pair. In this context they have been used to relate the FRF, H(w), of a linear system to the IRF, $h(\tau)$. It is useful to think of $h(\tau)$ and $H(\omega)$ as being two different representations of the same function, in the time and frequency domains respectively, and that the equations of the Fourier transform define how to go from one domain to the other.

The Fourier transformation is an established analytical technique which has found applications throughout signal processing [45]. Consequently, efficient numerical routines exist for computing the transform from streams of measured data [34]. Since the calculation must be performed over a finite number of sampled points in either time or frequency domain, these numerical procedures are often referred to as Discrete Fourier Transforms [DFTs].

Since DFT procedures can be used to estimate h(r) from H(w), and vice-versa, theoretically the identification of a linear structure can take place equivalently in either the time or frequency domain. For example, from the definition of the FRF,

$$H(\omega) = \frac{Y(\omega)}{X(\omega)} = \frac{Y(\omega).X^*(\omega)}{X(\omega).X^*(\omega)} \tag{2.35}$$

By definition, the auto-spectrum, $S_{xx}(\omega) = X(\omega).X^*(\omega)$, is the Fourier transform of the auto-correlation function,

$$s_{xx}(\tau) = \int_{-\infty}^{\infty} x(t).x(t-\tau).d\tau \tag{2.36}$$

and the cross-spectrum, $S_{yx}(\omega) = Y(\omega).X^*(\omega)$, is the transform of the cross-correlation function,

$$s_{yx}(\tau) = \int_{-\infty}^{\infty} y(t).x(t-\tau).d\tau \tag{2.37}$$

For a linear system, the FRF can be estimated from the measured signals $x(t)$ and y(t) either in the frequency domain, by using the DFT first to estimate the spectra X(w) and $Y(\omega)$ and hence compute the auto- and cross- spectra and the FRF, or alternatively in the time domain by performing discretised correlations and then using the DFT to compute the FRF.

The conventional Fourier transform is single-dimensional in that the integral is expressed for a single argument of the function. For functions of more than one variable, multi-dimensional Fourier transforms can be defined in terms of multiple integrals on the function. In addition, using these multi-dimensional transformations, it is possible to define higher order FRFs from the higher order IRFs, the kernels of the Volterra series.

Since the first term in the Volterra series has exactly the same form as the convolution integral, the first order FRF, $H_1(j\omega)$, is defined from the first order IRF, $h_1(\tau)$, in the same way as the FRF of a linear system,

$$H_1(\omega) = \int_{-\infty}^{\infty} h_1(\tau).e^{-j\omega\tau}.d\tau \tag{2.38}$$

$$h_1(\tau) = \frac{1}{2\pi}.\int_{-\infty}^{\infty} H_1(\omega).e^{j\omega\tau}.d\omega \tag{2.39}$$

However, for each of the higher order FRFs, it is necessary to make use of the multi-dimensional Fourier transforms defined as,

$$H_n(\omega_1,\ldots,\omega_n) = \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} h_n(\tau_1,\ldots,\tau_n).\prod_{i=1}^{n} e^{-j\omega_i\tau_i}.d\tau_1\ldots d\tau_n \tag{2.40}$$

$$h_n(\tau_1,\ldots,\tau_n) = \frac{1}{(2\pi)^n}.\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} H_n(\omega_1,\ldots,\omega_n).\prod_{i=1}^{n} e^{j\omega_i\tau_i}.d\omega_1\ldots d\omega_n \tag{2.41}$$

where $H_n(\omega_1,\ldots,\omega_n)$ denotes the n^{th} order multi-dimensional FRF.
Fortunately numerical procedures also exist for performing multi-dimensional DFTs[34].

Just as for the linear system, the existence of general Fourier transforms routines may suggest that the measurement of the Volterra kernels may be achieved either directly by performing multi-dimensional correlations and using the transforms to compute the FRFs or, if a method can be found for measuring a multi-dimensional FRF, the corresponding Volterra kernel may be computed using an inverse transformation. Although this is true, certain practical limitations arise with both approaches, as highlighted in the next section. For example, an estimate of the conventional FRF may be measured using a sine excitation from a structure which is assumed to be linear, but when the impulse response function is computed using the inverse Fourier transform, the function is found to be non-causal, indicating that assumption of linearity was invalid. Similarly, although it may be possible to estimate the multi-dimensional higher order FRFs, performing the Fourier transformation to estimate the corresponding IRF may serve to indicate discrepancies in the measured data. This fact is used in the detection of non-linearity via Hilbert transforms, as described in chapter 4.

2.2.4 Response Characteristics of the Volterra Series

The aim of this section is to demonstrate how the Volterra series can be used to represent the response to four characteristic inputs: a Dirac impulse, a single harmonic, a sinewave, and a broadband random signal. The first two of these signals are 'ideal' in that they cannot be generated in practice, but are particularly useful from a theoretical viewpoint since they can be used introduce simplifications, and assist with the interpretation of the Volterra kernels and the higher order FRFs respectively. On the other hand, the latter two signals are useful from a practical viewpoint, since they can be generated and are often used to test a structure.

Response to an Impulse

From the Volterra series formulation (equation 2.15), it is possible to write down an expression for the response of any system in the Volterra class to an impulse $x(t) = X\delta(t)$,

$$\begin{aligned} y(t) = & \int_{-\infty}^{\infty} h_1(\tau_1).X\delta(t-\tau_1).d\tau_1 + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h_2(\tau_1,\tau_2).X\delta(t-\tau_1).X\delta(t-\tau_2).d\tau_1.d\tau_2 \\ & + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h_3(\tau_1,\tau_2,\tau_3).X\delta(t-\tau_1).X\delta(t-\tau_2).X\delta(t-\tau_3).d\tau_1.d\tau_2.d\tau_3 + \ldots \\ & + X^n.\int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty} h_n(\tau_1,\ldots,\tau_n).\prod_{i=1}^{n}\delta(t-\tau_i).d\tau_1\ldots d\tau_n + \ldots \end{aligned} \tag{2.42}$$

Interpreting the Volterra kernels as impulse response functions, this expression reduces to,

$$y(t) = h_1(t).X + h_2(t,t).X^2 + h_3(t,t,t).X^3 + \ldots + h_n(t,\ldots,t).X^n + \ldots \tag{2.43}$$

For a non-linear system which is described by an infinite Volterra series, the response to a single impulse includes the main diagonals of all the kernels, but does not include any off-diagonal terms. Consequently, the response of a non-linear system to a single impulse cannot be used to generally characterise a non-linear system. To relate the multi-dimensional n^{th} order kernel to the output, it would be necessary to consider an input composed of ***n*** or more impulses at different time delays, eg. $x(t) = \sum_{i=1}^{n} X\delta_i(t)$.

Measuring the response of a system to one or more controlled impulses has been suggested as a possible method of determining Volterra kernels of a structure [46]. However, the ideal impulse cannot be generated and measuring transient oscillations is difficult in practice. Moreover, the method is very sensitive to errors in the measurement of the impulse amplitude. A better approach would be to measure the ***steady-state*** response of a structure when excited by, for example, a single sinewave[47].

Response to a Harmonic Input

Whereas the time domain response of a system is usually investigated using the impulse function, $X\delta(t)$, the harmonic, $Xe^{j\omega t}$, can used to interpret the frequency domain response. The representation of the response to a single harmonic can be written directly,

$$\begin{aligned} y,(t) &= \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} h_n(\tau_1,\ldots,\tau_n).\prod_{i=1}^{n} x(t-\tau_i).d\tau_1 \ldots d\tau_n \\ &= \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} h_n(\tau_1,\ldots,\tau_n).\prod_{i=1}^{n} Xe^{j\omega(t-\tau_i)}.d\tau_1 \ldots . d\tau_n \\ &= X^n e^{jn\omega t}.\int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} h_n(\tau_1,\ldots,\tau_n).\prod_{i=1}^{n} e^{-j\omega\tau_i}.d\tau_1 \ldots . d\tau_n \end{aligned} \quad \textbf{(2.44)}$$

Using the definition of the higher order FRF,

$$y,(t) = (Xe^{j\omega t})^n.H_n(\omega, \ldots,\omega) \quad (2.45)$$

From this expression, it may be seen that n^{th} order component of the response to a single harmonic input will be also a single harmonic, but at frequency $n\omega$,

$$Y(n\omega) = X(\omega)^n.H_n(\omega,\ldots,\omega) \quad (2.46)$$

The total time response' to a single harmonic corresponds to the infinite series,

$$y(t) = H_1(\omega).Xe^{j\omega t}+H_2(\omega,\omega).X^2e^{j2\omega t}+H_3(\omega,\omega,\omega).X^3e^{j3\omega t}+\ldots+H_n(\omega,\ldots,\omega).X^ne^{jn\omega t}+\ldots \quad (2.47)$$

Hence, terms on the principal diagonal of the FRFs quantify the energy transfer which may take place between the single input harmonic, and the fundamental and higher harmonic components of the output. Or, from an alternative standpoint, the single harmonic can be thought of as probing only the principal diagonal of each multi-dimensional FRF. To characterise fully the n^{th} order function $H_n(\omega_1,\ldots,\omega_n)$, it would be necessary to use series of at least *n* harmonics at different frequencies, as described in chapter 3.

If a system could be excited by a single harmonic, it would be possible to separate each of these higher harmonic terms from the spectra of the input and the output, and compute the FRF diagonals directly,

$$H_n(\omega,\ldots,\omega) = \frac{Y(n\omega)}{X(\omega)^n} \quad (2.48)$$

In theory, this would be a simple means of measuring the principal diagonal of each of the higher order FRFs. In practice, however, it is not possible to excite a system with an ideal harmonic. The closest physical signal is a single sinewave.

Response to a Sinusoidal Input

A sinewave can be regarded as a sum of two ideal harmonics, X. $\cos(\omega t) = \frac{X}{2}e^{j\omega t} + \frac{X}{2}e^{-j\omega t}$ In a linear system, it would be possible to predict the total output by considering each input harmonic individually and use the principle of superposition. In a non-linear system, the mechanism for energy transfer between frequencies means that two input harmonics will have a combined effect on the output [23],[48],

$$x(t) = X.\cos(\omega t) = \frac{X}{2}e^{j\omega t} + \frac{X}{2}e^{-j\omega t} \tag{2.49}$$

$$\begin{aligned} y(t) &= H_1(\omega).\left(\frac{X}{2}\right).e^{j\omega t} + H_1(-\omega).\left(\frac{X}{2}\right)e^{-j\omega t} \\ &+ H_2(\omega,\omega).\left(\frac{X}{2}\right)^2.e^{j2\omega t} + H_2(-\omega,-\omega).\left(\frac{X}{2}\right)^2 e^{-j2\omega t} \\ &+ H_2(\omega,-\omega).\left(\frac{X}{2}\right)^2 + H_2(-\omega,\omega).\left(\frac{X}{2}\right)^2 \\ &+ H_3(\omega,\omega,\omega).\left(\frac{X}{2}\right)^3.e^{j3\omega t} + H_3(-\omega,-\omega,-\omega).\left(\frac{X}{2}\right)^3 e^{-j3\omega t} \\ &+ H_3(\omega,\omega,-\omega).\left(\frac{X}{2}\right)^3.e^{j\omega t} + H_3(-\omega,-\omega,\omega).\left(\frac{X}{2}\right)^3 e^{-j\omega t} \\ &+ H_3(\omega,-\omega,\omega).\left(\frac{X}{2}\right)^3.e^{j\omega t} + H_3(-\omega,\omega,-\omega).\left(\frac{X}{2}\right)^3 e^{-j\omega t} \\ &+ H_3(-\omega,\omega,\omega).\left(\frac{X}{2}\right)^3.e^{j\omega t} + H_3(\omega,-\omega,-\omega).\left(\frac{X}{2}\right)^3 e^{-j\omega t} + \ldots \end{aligned} \tag{2.50}$$

Using the property of symmetry, this expression can be simplified to,

$$\begin{aligned} y(t) &= H_1(\omega).\left(\frac{X}{2}\right).e^{j\omega t} + H_1(-\omega).\left(\frac{X}{2}\right)e^{-j\omega t} \\ &+ H_2(\omega,\omega).\left(\frac{X}{2}\right)^2.e^{j2\omega t} + H_2(-\omega,-\omega).\left(\frac{X}{2}\right)^2 e^{-j2\omega t} \\ &+ 2.H_2(\omega,-\omega).\left(\frac{X}{2}\right)^2 \\ &+ H_3(\omega,\omega,\omega).\left(\frac{X}{2}\right)^3.e^{j3\omega t} + H_3(-\omega,-\omega,-\omega).\left(\frac{X}{2}\right)^3 e^{-j3\omega t} \\ &+ 3.H_3(\omega,\omega,-\omega).\left(\frac{X}{2}\right)^3.e^{j\omega t} + 3.H_3(\omega,-\omega,-\omega).\left(\frac{X}{2}\right)^3 e^{-j\omega t} + \ldots \end{aligned}$$

Using the property of conjugate symmetry, and observing that for a complex quantity z,

$$z.e^{j\omega t} + z^*.e^{-j\omega t} = |z|.e^{j(\omega t+\angle z)} + |z|.e^{-j(\omega t+\angle z)} = 2.|z|.\ \cos(\omega t + \angle z)$$

$$
\begin{aligned}
x(t) &= X.\cos(\omega t) \\
y(t) &= |H_1(\omega)|.2\left(\frac{X}{2}\right)\cos(\omega t + \angle H_1(\omega)) \\
&+ |H_2(\omega,\omega)|.2\left(\frac{X}{2}\right)^2\cos(2\omega t + \angle H_2(\omega,\omega)) \\
&+ H_2(\omega,-\omega).2\left(\frac{X}{2}\right)^2 \\
&+ |H_3(\omega,\omega,\omega)|.2\left(\frac{X}{2}\right)^3\cos(3\omega t + \angle H_3(\omega, w, w)) \\
&+ |H_3(\omega,\omega,-\omega)|.6\left(\frac{X}{2}\right)^3 cos(wt + \angle H_3(\omega, w, -\omega)) + \ldots
\end{aligned}
\qquad (2.51)
$$

From this expression it is evident that the response of a non-linear structure to a single frequency sinewave will be composed of a series of sinewaves at integer multiples of the fundamental frequency, with a possible D.C. shift. It is possible to confirm this in practice by performing one-dimensional Fourier transform on the measured response data; the spectrum would consist of a series of lines at the fundamental and discrete harmonic frequencies, and at zero frequency.

As mentioned in chapter 1, this procedure can be used to investigate whether or not a structure can be considered as linear. Assuming a structure can be excited by a high quality sinewave, if the structure were linear only the fundamental frequency line would be observed in the response spectra; each of the higher order FRFs would be null and there would be no mechanism for energy transfer to other frequencies. If, however, there were first and odd order harmonics in the response spectra, but no even harmonics, a further conclusion would be that all the even order FRFs are null and that the non-linear polynomial function which governs the dynamics of the structure in the structure possess odd order symmetry.

Response to a Random Input

A random time signal is the most general input for a system, and two systems are equivalent if and only if they respond identically to a random input [49]. For such an input, the Volterra series representation of the system is complicated since no simplifications can be introduced,

$$\begin{aligned} y_n(t) &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} h_n(\tau_1, \ldots, \tau_n).\prod_{i=1}^{n} x(t-\tau_i).d\tau_1 \ldots d\tau_n \\ &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} h_n(\tau_1, \ldots, \tau_n).\left(\prod_{i=1}^{n} \int_{-\infty}^{\infty} X(\omega_i).e^{j\omega_i(t-\tau_i)}.d\omega_i\right).d\tau_1 \ldots d\tau_n \\ &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} h_n(\tau_1, \ldots, \tau_n).\prod_{i=1}^{n} e^{-j\omega_i\tau_i}.d\tau_i\right).\prod_{i=1}^{n} X(\omega_i).e^{j\omega_i t}.d\omega_1 \ldots d\omega_n \\ y,(t) &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} H_n(\omega_1, \ldots, \omega_n).X(\omega_1) \ldots X(\omega_n).e^{j(\omega_1+\ldots+\omega_n)t}.d\omega_1 \ldots dw, \end{aligned} \quad \textbf{(2.52)}$$

The single dimensional Fourier transform, denoted $\mathcal{F}[.]$, can be used to reformulate the spectral content of each component of the output,

$$\mathcal{F}[y_n(t)] = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} H_n(\omega_1, \ldots, \omega_n).X(\omega_1) \ldots X(\omega_n).\mathcal{F}[e^{j(\omega_1+\ldots+\omega_n)t}].d\omega_1 \ldots \mathbf{dw}, \quad \textbf{(2.53)}$$

The Fourier transform of the harmonic is defined by the Delta function,

$$\mathcal{F}[e^{j\omega_0 t}] = \delta(\omega - \omega_0) = \begin{cases} 1 \text{ if } \omega = \omega_0 \\ 0 \text{ if } \omega \neq \omega_0 \end{cases} \quad (2.54)$$

$$\mathcal{F}[y_n(t)] = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} H_n(\omega_1, \ldots, \omega_n).X(\omega_1) \ldots X(\omega_n).\delta(\omega - \omega_1 - \ldots - \omega_n).d\omega_1 \ldots d\omega_n \quad (2.55)$$

Hence, using the property of the Delta function, each harmonic component at frequency w in the spectrum of the response can be written,

$$\begin{aligned} \mathbf{Y(w)} &= H_1(\omega).X(\omega) \\ &+ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} H_2(\omega_1, \omega_2).X(\omega_1).X(\omega_2).\delta(\omega - \omega_1 - \omega_2).d\omega_1.d\omega_2 \\ &+ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} H_3(\omega_1, \omega_2, \omega_3).X(\omega_1).X(\omega_2).X(\omega_3).\delta(\omega - \omega_1 - \omega_2 - \omega_3).d\omega_1.d\omega_2.d\omega_3 \\ &+ \ldots \\ &+ \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} H_n(\omega_1, \ldots, \omega_n).X(\omega_1) \ldots X(\omega_n).\delta(\omega - \omega_1 - \ldots - w_n).d\omega_1 \ldots d\omega_n \\ &+ \ldots \end{aligned} \quad (2.56)$$

Since no restrictions on the type of input have been imposed, this expression can be regarded as a general formulation of the Volterra series in the frequency domain [4],[11].

Polyspectra techniques make direct use of this frequency domain formulation [50]. The general approach is particularly suitable in applications when the input cannot be controlled, for example in the analysis of the dynamics of offshore structures [51]. Routines have been devised to identify non-linear behaviour from raw time data, and have been successful in explaining unusual behaviour. In particular, it is possible to understand the energy transfer between frequencies which occurs and, for example, how low frequency resonances of an offshore structure can be excited by predominantly higher frequency seawaves.

To investigate the energy transfer mechanism, the polyspectral approach is to reformulate the expression for discrete frequencies,

$$\begin{aligned} \mathbf{Y(U)} &= H_1(\omega).X(\omega) \\ &+ \sum_{\omega_{i_1}+\omega_{i_2}=\omega}^{\sum_{i_1}\sum_{i_2}} H_2(\omega_{i_1},\omega_{i_2}).X(\omega_{i_1}).X(\omega_{i_2}) \\ &+ \sum_{\omega_{i_1}+\omega_{i_2}+\omega_{i_3}=\omega}^{\sum_{i_1}\sum_{i_2}\sum_{i_3}} H_3(\omega_{i_1},\omega_{i_2},\omega_{i_3}).X(\omega_{i_1}).X(\omega_{i_2}).X(\omega_{i_3}) + \ldots \\ &+ \sum_{\omega_{i_1}+\ldots+\omega_{i_n}=\omega}^{\sum_{i_1}\cdots\sum_{i_n}} H_n(\omega_{i_1},\ldots,\omega_{i_n}).X(\omega_{i_1}).\ldots.X(\omega_{i_n}) + \ldots \end{aligned} \tag{2.57}$$

The properties of the $\delta(\omega-\omega_0)$ function have been used to prescribe conditions on the limits of the summations which have replaced the integrals. This approach is particularly appropriate with regard to implementation with a computer, since data processing algorithms usually employ digital numerical routines acting on discretised data.

The bispectrum approach is the simplest example of the application of polyspectral analysis, and makes the assumption that the frequency domain representation of the system can be approximated by the first and second order terms [52]. In general, this will be inexact since even a quadratic system would need an infinite number of higher order terms to describe the response accurately. However to begin to understand the energy transfer mechanism, the assumption is accepted,

$$Y(\omega) \approx H_1(\omega).X(\omega) + \sum_{\omega_{i_1}+\omega_{i_2}=\omega}^{\sum_{i_1}\sum_{i_2}} H_2(\omega_{i_1},\omega_{i_2}).X(\omega_{i_1}).X(\omega_{i_2}) \tag{2.58}$$

$$= H_1(\omega).X(\omega) + \sum_{\omega_i=w_1}^{w_N} H_2(\omega_i,\omega-\omega_i).X(\omega_i).X(\omega-\omega_i) \tag{2.59}$$

To develop the procedure, it is necessary to consider components at each individual frequency in the output spectrum, and to determine which input frequency components could interact and contribute to the output at this frequency.

The second order FRF describes interactions between pairs of input frequency components. With a baseband random input, any line ω_{i_1} in the digital spectrum from zero up to the high frequency cutoff could be 'paired' with a second component ω_{i_2} such that $\pm\omega_{i_1} \pm \omega_{i_2} = \omega_n$, the contributions from $H_2(\omega_{i_1}, \omega_{i_2})$ to w,. Throughout the discrete frequency range 0 to ω_q, the total contribution from the first and second order FRFs to the output component at frequency can be written,

$$\begin{aligned} Y(\omega_n) \approx H_1(\omega_n)X(\omega_n) + & \sum_{\omega_i=\omega_{n-1}}^{w_q} H_2(\omega_i, \omega_n - \omega_i).X(\omega_i).X(\omega - \omega_i) \\ & + H_2(\omega_n, 0).X(\omega_n).X(0) \\ & + \sum_{\omega_i=\omega_{n+1}}^{w_q} H_2(\omega_i, \omega_n - \omega_i).X(\omega_i).X(\omega - \omega_i) \end{aligned} \qquad (2.60)$$

The bispectrum approach, which has been applied to data measured from several physical systems [53], has been developed from this simplified approach. By rearranging into a matrix form, it is possible to solve for the complex quantity $H_1(\omega_n)$ at each spectral line ω_n, and the discrete points on $H_2(\omega_{i_1}, w;,)$ along each line $\omega_{i_1} + \omega_{i_2}$ = w,.

The main disadvantage with the bispectrum approach is that limiting the possible interactions to first and second order is far too restrictive to make the technique generally applicable to non-linear systems [54]. Required is a-priori knowledge that the non-linear interactions are limited to quadratic. Extending the procedure to include the third and higher order FRFs would give rise to a large number of interactions between components, making the solution procedure considerably more complicated [55].

2.3 Wiener Series

2.3.1 Defining the Wiener Series

During the 1940's, Norbert Wiener developed an alternative functional series using the the Volterra series as the basis [3]. Particular properties of the Wiener series enable the class of non-linear system which can be characterised to be broadened and, theoretically, would allow the Wiener kernels to be measured directly from a non-linear system [56]. In practice the measurement of the Wiener kernels remains fraught with difficulty, particularly for the higher orders [20].

The aim of this section is to introduce the Wiener series, compare important features with the Volterra series, and discuss the limitations with regard to identifying non-linear structures. This is relevant because to date most of the research conducted into the practical application of the Volterra series has been directed from the standpoint of the Wiener series.

Whereas the 3 terms of the Volterra series are written,

$$\begin{aligned} y(t) \approx & \int_{-\infty}^{\infty} h_1(\tau_1).x(t-\tau_1).d\tau_1 \\ & + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h_2(\tau_1,\tau_2).x(t-\tau_1).x(t-\tau_2).d\tau_1.d\tau_2 \\ & + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h_3(\tau_1,\tau_2,\tau_3).x(t-\tau_1).x(t-\tau_2).x(t-\tau_3).d\tau_1.d\tau_2.d\tau_3 \end{aligned} \quad (2.61)$$

the corresponding terms of the Wiener series can be written,

$$\begin{aligned} y(t) \approx & \int_{-\infty}^{\infty} k_1(\tau_1).x(t-\tau_1).d\tau_1 \\ & + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} k_2(\tau_1,\tau_2).(x(t-\tau_1).x(t-\tau_2) - \overline{x(t-\tau_1).x(t-\tau_2)}).d\tau_1.d\tau_2 \\ & + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} k_3(\tau_1,\tau_2,\tau_3).(x(t-\tau_1).x(t-\tau_2).x(t-\tau_3) \\ & \quad - x(t-\tau_1).\overline{x(t-\tau_2).x(t-\tau_3)} \\ & \quad - x(t-\tau_2).\overline{x(t-\tau_1).x(t-\tau_3)} \\ & \quad - x(t-\tau_3).\overline{x(t-\tau_1).x(t-\tau_2)}) \quad .d\tau_1.d\tau_2.d\tau_3 \end{aligned} \quad (2.62)$$

in which $k_n(\tau_1,\ldots,\tau_n)$ represents the n^{th} order Wiener kernel.

In the frequency domain, the Wiener series can be writ ten,

$$\begin{aligned} Y(\omega) \approx & K_1(\omega).X(\omega) \\ & + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} K_2(\omega_1,\omega_2).(X(\omega_1).X(\omega_2) - \overline{X(\omega_1).X(\omega_2)}).\delta(\omega-\omega_1-\omega_2).d\omega_1.d\omega_2 \\ & + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} K_3(\omega_1,\omega_2,\omega_3).(X(\omega_1).X(\omega_2).X(\omega_3) - X(\omega_1).\overline{X(\omega_2).X(\omega_3)} \\ & \quad - X(\omega_2).\overline{X(\omega_1).X(\omega_3)} \\ & \quad - X(\omega_3).\overline{X(\omega_1).X(\omega_2)}) \\ & \quad .\delta(\omega-\omega_1-\omega_2-\omega_3).d\omega_1.d\omega_2.d\omega_3 \end{aligned} \quad (2.63)$$

2.3.2 Features of the Wiener Series

The Wiener series was devised so that, if the output is considered to be made up of a series of individual components, each component is orthogonal to each other if and only if the input signal $x(t)$ is a white Gaussian time function; this is an ideal random signal defined to have a constant power spectrum for all frequencies. Writing the Wiener series,

$$y(t) = y_1'(t) + y_2'(t) + y_3'(t) + \ldots + y_n'(t) + \ldots \tag{2.64}$$

the condition of orthogonality can be written,

$$\overline{y_i'(t).y_j'(t)} = 0 \quad \text{if} \quad i \neq j \tag{2.65}$$

The principal difference between the Wiener and Volterra series is that, whereas the Volterra kernels are defined to be homogeneous and independent on the excitation, the Wiener kernels are non-homogeneous and rely on the excitation approximating to a white Gaussian time function to make use of its potential properties. When the input signal deviates appreciably from this ideal, for example in situations in which the excitation cannot be contolled, the Wiener series approach is usually abandoned in favour of the more general polyspectra techniques, as indicated in the previous section.

From the point of view of system identification, characterising non-linear systems in terms of the homogeneous Volterra kernels, rather than the inhomogeneous Wiener kernels, would appear to be more logical. However, the Wiener functional series has two features which are principally cited as being advantages over the Volterra series in the context of non-linear system characterisation [7].

Firstly, the class of systems which can be characterised by a Volterra series is a subset of the class of systems that possess a Wiener series representation, since any system that can be characterised in terms of a Volterra series can be characterised also by a Wiener series: The orthogonal Weiner series converges for a larger range of excitation levels.

Secondly, the kernel of each orthogonal functional in theory could be determined individually without interference from the presence of any of the other functionals. In addition, the truncated model could be extended to include higher order terms without affecting the terms already estimated. If the Wiener series is truncated after n terms, as it must in practice, the resulting approximation is the best in the mean error sense. In general, this is not true for the Volterra series representation.

To illustrate this latter point, consider the first order Wiener kernel transform, which is of particular interest since it can be related to the optimum linear model introduced in chapter 1. An expression can be developed by multiplying each term in the frequency domain expression by X*(w), and taking the average,

$$\begin{aligned}\overline{Y(\omega).X^*(\omega)} \approx\ & K_1(\omega).\overline{X(\omega).X^*(\omega)} \\ & + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} K_2(\omega_1,\omega_2).(\overline{X(\omega_1).X(\omega_2).X^*(\omega)} - \overline{X(\omega_1).X(\omega_2).X^*(\omega)}) \\ & \qquad .\delta(\omega-\omega_1-\omega_2).d\omega_1.d\omega_2 \\ & + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} K_3(\omega_1,\omega_2,\omega_3).(\overline{X(\omega_1).X(\omega_2).X(\omega_3).X^*(\omega)} \\ & \qquad -\overline{X(\omega_1).X^*(\omega).X(\omega_2).X(\omega_3)} \\ & \qquad -\overline{X(\omega_2).X^*(\omega).X(\omega_1).X(\omega_3)} \\ & \qquad -\overline{X(\omega_3).X^*(\omega).X(\omega_1).X(\omega_2)}) \\ & \qquad .\delta(\omega-\omega_1-\omega_2-\omega_3).d\omega_1.d\omega_2.d\omega_3\end{aligned} \tag{2.66}$$

and by observing that, if z(t) is defined to be a zero mean signal with Gaussian distribution, the spectrum of the signal, X(w), would also be Gaussian with zero mean, $\overline{X(w)} = 0$. The expression can be simplified by using the general properties of Gaussian random variables,

1. All odd order moments of zero-mean jointly Gaussian random variables are zero, eg.

$$\overline{X(\omega_1).X(\omega_2).X^*(\omega)} = \overline{X(\omega_1).X(\omega_2).X^*(\omega)} = 0$$

2. All even order moments (greater than 2) can be expressed as products of second order moments, eg.

$$\begin{aligned}\overline{X(\omega_1).X(\omega_2).X(\omega_3).X^*(\omega)} =\ & \overline{X(\omega_1).X^*(\omega).X(\omega_2).X(\omega_3)} \\ & + \overline{X(\omega_2).X^*(\omega).X(\omega_1).X(\omega_3)} \\ & + \overline{X(\omega_3).X^*(\omega).X(\omega_1).X(\omega_2)}\end{aligned}$$

With the exception of the terms incorporating the first order kernel $K_1(\omega_1)$, all the higher order terms on the left-hand side of the expression are found to have an estimated value of zero, with the result,

$$K_1(\omega) = \frac{\overline{Y(\omega).X^*(\omega)}}{\overline{X(\omega).X^*(\omega)}} \tag{2.67}$$

From this expression, it is clear that the linearised model of a non-linear system, obtained from the ratio of cross-spectrum to auto-spectrum, would correspond to the first order Wiener kernel for the specific case of Gaussian random input with zero-mean.

2.3.3 Measuring the Wiener Kernels

A similar procedure can be used to determine expressions for the Wiener kernel transforms of any order by multiplying the expression for the output $Y(\omega)$ with terms of the input spectrum to form higher order cross-spectral functions. For example, $K_n(\omega_1,\omega_2)$ is calculated from $\overline{Y(\omega).X^*(\omega_1)\ldots X^*(\omega_n)}$. Using conventional auto- and cross- correlation functions, and introducing notation for higher order cross-correlations functions,

$$S_{y'xx} = Y'(\omega_1+\omega_2).X^*(\omega_1).X^*(\omega_2) \tag{2.68}$$

$$S_{y''xxx} = Y''(\omega_1+\omega_2+\omega_3).X^*(\omega_1).X^*(\omega_2).X^*(\omega_3) \tag{2.69}$$

it is possible to show [11] that the first 3 Wiener kernel transforms have the form,

$$K_1(\omega) = \frac{S_{yx}(\omega)}{S_{xx}(\omega)} \tag{2.70}$$

$$K_2(\omega_1,\omega_2) = \frac{S_{y'xx}(\omega_1,\omega_2)}{2!S_{xx}(\omega_1).S_{xx}(\omega_2)} \tag{2.71}$$

$$K_3(\omega_1,\omega_2,\omega_3) = \frac{S_{y''xxx}(\omega_1,\omega_2,\omega_3)}{3!S_{xx}(\omega_1).S_{xx}(\omega_2).S_{xx}(\omega_3)} \tag{2.72}$$

where the signals y'(t) represent the output y(t) minus the mean level, and y"(t) is the output with the first order output component removed.

From the expressions above, the computation of the higher order cross-spectra and the Wiener kernels transforms would use only the one-dimensional Fourier transform. It is important to remember that the properties of Gaussian variables have been used throughout the development, and these relationships only strictly apply for a white Gaussian input. If the power level of the input is denoted X, the Wiener kernel transforms can be expressed in terms of higher order correlation functions with the higher order Fourier transforms,

$$K_1(\omega) = \frac{1}{X}\mathcal{F}_1\left(\overline{y(t)x(t-\tau)}\right) \tag{2.73}$$

$$K_2(\omega_1,\omega_2) = \frac{1}{2!X^2}\mathcal{F}_2\left(\overline{y'(t)x(t-\tau_1)x(t-\tau_2)}\right) \tag{2.74}$$

$$K_3(\omega_1,\omega_2,\omega_3) = \frac{1}{3!X^3}\mathcal{F}_3\left(\overline{y''(t)x(t-\tau_1)x(t-\tau_2)x(t-\tau_3)}\right) \tag{2.75}$$

In principle, expressions for the Wiener kernel transforms could be estimated from the measurements of the system input and output using either the correlation in the time domain followed by multi-dimensional Fourier Tranformation, a technique first proposed by Lee and Schetzen [6], or in the frequency domain using only single dimensional Fourier transforms [12].

Although it is evident from these formulations that the Wiener kernels could be estimated in theory from measurements of the input and output signals in either time or frequency domain, considerable practical problems are encountered with both approaches. The remainder of this section is devoted to highlighting these limitations with regard to implementing the Wiener series approach in order to justify the search for an alternative, more simple method of measuring estimates of the higher order FRFs defined using the Volterra series.

Even in a structural test in which the excitation source can be chosen, the idealised Gaussian input signal can never be realised in practice. All real signals have finite length and finite energy. Consequently the power density cannot be constant across all frequencies, but tends to decrease at both the low and high frequency ends of the bandwidth. To improve the approximation of a physical input signal to the ideal Gaussian signal, the bandwidth is usually selected to be considerably greater than that of the system under study. In addition, the magnitude of the power spectrum of the excitation signal is designed to be flat over the frequency range of interest. Unfortunately the interaction which frequently occurs between the shaker and structure can prevent the force input to a structure from having a flat spectral density, even over a zoomed bandwidth, particularly around the critical resonance regions. It has been noted by Gifford [25] that this can produce some strange results in the estimation of the Wiener kernel transforms.

Determination of the higher order kernel transforms using correlations in the time domain would be extremely time consuming and impractical for estimating kernels above second (or at most third) order. Because of the speed at which the single dimensional Fourier transform can be performed, calculation in the frequency domain is the more pragmatic approach.

As discussed previously, problems are encountered when computing Fourier transforms digitally on sets of discretised time data. In particular, the Fourier transform implicitly assumes the time function is periodic. Since a broadband random signal is indeterministic and not periodic, errors are introduced in each of the spectral estimates and are compounded in the estimate of the higher order kernels, eg. $Y'(\omega_1 + \omega_2).X^*(\omega_1).X^*(\omega_2)$

Another important fact is that the numerical DFT routines can only perform transformations over finite lengths of data. By ensuring adequate resolution, the transformations are usually reasonably accurate if, for example, the time signal is periodic over the interval selected. If, however, the signal is not periodic, or the sections do not contain only whole periods of the signal, the familiar problem of leakage occurs. Before performing the single dimensional Fourier transform, it is conventional to reduce the effect of leakage by multiplying each section of data by a 'window' which tends to profile the data to approach periodicity across

the sections [57]. However the inclusion of a window necessarily causes modification of the original data, and can cause differences between the computed spectra and the true spectra of the original signals. This discrepancy will be compounded when forming the higher order spectral functions from the single-dimensional spectra. On the other hand, if the Wiener kernel transforms are computed using multi-dimensional Fourier transforms, a multi-dimensional window would be required to reduce the effect of leakage. Although the concept is straightforward, the design of multi-dimensional windows to have minimal adverse effect on the spectra is a difficult task in itself [58].

Since the estimation of the higher order kernels necessitates processing sections of time data, for example to remove the mean, the standard approach is to measure digitally a large number points of time data from a system, and to compute the auto- and cross-spectra from sections of this database in post-processing. The length of each section is determined largely by the frequency resolution required in the spectra and hence the frequency selected to sample the data. The number of averages which can be performed is then limited by the size of the database.

The fact that the analysis is performed on raw time data is usually purported to be an advantage since measuring time data makes the smallest possible commitment to later analysis and preserves maximum flexibility [59]. Unfortunately, the ability to perform on-line tests to assess the quality of the data is restricted if the analysis is performed in post-processing. In addition, very large databases would be required to allow many averages in an attempt to ensure the results are of reasonable quality.

In theory, the expected values are only attained as a result of averaging an infinite population of possible terms. In practice, the averaging can only be performed using data samples of finite length, and averaging over a finite number implies that the expected values can be only estimated. However, the error between the estimates and expected values would be difficult to determine since it would depend on the variance of all possible finite samples formed from the infinite population. In addition because, before determining the second order Wiener kernel transform, it is necessary to first evaluate and then remove the mean from the output signal, the estimated kernel will be always in error to some degree since the mean of a signal cannot be reached by averaging a finite number of times. Similarly, the estimates of all higher order kernels would be polluted since each relies on low order components being removed from the output signal before calculation. Hence, for a non-linear system, even if the input does approximate well to Gaussian white noise, and assuming the signals are not significantly contaminated with measurement noise, the estimates the Wiener kernels from a finite number of data samples would still be polluted. Even if the system were linear,

the estimates of the second and higher order Wiener kernels would be non-zero. The only method of improving the estimates is to average over a larger number of data sets, and this is extremely time consuming particularly when estimates of the higher order Wiener kernels are required.

In independent tests conducted on simulated systems, one common observation was that the first order component usually dominates the output, and causes considerable pollution on the estimated second order Wiener kernel. In theory, the influence of the first order component on the second order kernel should approach zero as the averages approach infinity. In practice, too many averages are required to reduce the effect to a reasonable level. One modification which has been proposed to circumvent the problem is to estimate and remove the first order component from the output signal (in addition to the mean level) before computing the second order kernel.

A complete set of Wiener kernels, $k_1(\tau_1) \ldots k_n(\tau_1, \ldots, \tau_n)$, can be thought of as characterising a non-linear system, since knowledge of the kernels would allow the prediction of the system response to a given stimulus. However, this would require a measure of the power level X of the random input signal with which the Wiener kernels have been estimated, since the Wiener kernels of higher order than the first depend explicitly on X. Schetzen has shown that the n^{th} order Volterra kernel can be expressed as an infinite series of Wiener kernels of order higher than n but of the same type, ie. odd or even,

$$h_1(\tau_1) = k_1(\tau_1) - 3X.\int_{-\infty}^{\infty} k_3(\tau_1, \tau_2, \tau_2).d\tau_2 + \ldots \tag{2.76}$$

$$h_2(\tau_1, \tau_2) = k_2(\tau_1, \tau_2) - 6X.\int_{-\infty}^{\infty} k_4(\tau_1, \tau_2, \tau_3, \tau_3).d\tau_3 + \ldots \tag{2.77}$$

$$h_3(\tau_1, \tau_2, \tau_3) = k_3(\tau_1, \tau_2, \tau_3) - 10X.\int_{-\infty}^{\infty} k_5(\tau_1, \tau_2, \tau_3, \tau_4, \tau_4).d\tau_4 + \ldots \tag{2.78}$$

$$h_n(\tau_1, \ldots, \tau_3) = k_n(\tau_1, \ldots, \tau_3) - \frac{(n+2)!}{n!2!}X.\int_{-\infty}^{\infty} k_{n+2}(\tau_1, \ldots, \tau_{n+1}).d\tau_{n+1} + \ldots \tag{2.79}$$

In the context of identifying a model of a structure, the Wiener kernels are used as an estimate of the Volterra kernels. From this standpoint, the fact that Wiener kernels can be defined for systems which do not have a Volterra series representation is no longer an advantage. Furthermore, although truncating the orthogonal model to some order does give the optimum approximate representation of the output in a least squares sense, there are many difficulties associated with identifying the system kernels of second order and above. To date, results which have been published from either physical or simulated systems have attempted to identify only the first and second order Wiener kernels.

Clearly, a search for a different method of measuring an estimate the Volterra kernels of a structure is justified, to take advantage of the fact that, in a structural test, it is usually possible to choose the excitation signal. One approach is to use a deterministic input such as a single sinewave, and avoid needing to average away unwanted correlation products, a principal limitation of identifying the Wiener series using non-deterministic signals. Other problems, such as leakage when performing the Fourier transformation, can be avoided by using periodic signals.

There are several other advantages with using a single sinewave excitation to estimate the higher order FRFs defined by the Volterra series, and these are discussed in chapter 4.

2.4 Coherence Functions

One particularly useful application of the Wiener series is to demonstrate how the ordinary coherence function, often measured in a conventional structural test, can be used as a simple, practical way of detecting non-linearity, without identifying its precise nature.

Substituting the spectral functions for the Wiener kernels transforms (equations 2.70 to 2.72) into the frequency domain formulation of the Wiener series (equation 2.63),

$$\begin{aligned}\mathbf{Y(w)} \approx\ & \frac{S_{yx}(\omega)}{S_{xx}(\omega)}.X(\omega) \\ & + \frac{1}{2!}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{S_{y'xx}(\omega_1,\omega_2)}{S_{xx}(\omega_1).S_{xx}(\omega_2)}.(X(\omega_1).X(\omega_2) - \overline{X(\omega_1).X(\omega_2)}).\delta(\omega-\omega_1-\omega_2).d\omega_1.d\omega_2 \\ & + \frac{1}{3!}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{S_{y''xxx}(\omega_1,\omega_2,\omega_3)}{S_{xx}(\omega_1).S_{xx}(\omega_2).S_{xx}(\omega_3)}.(X(\omega_1).X(\omega_2).X(\omega_3) \\ & \qquad -X(\omega_1).\overline{X(\omega_2).X(\omega_3)} \\ & \qquad -X(\omega_2).\overline{X(\omega_1).X(\omega_3)} \\ & \qquad -X(\omega_3).\overline{X(\omega_1).X(\omega_2)}) \\ & \qquad .\delta(\omega-\omega_1-\omega_2-\omega_3)\ \ .d\omega_1.d\omega_2.d\omega_3\end{aligned} \tag{2.80}$$

An expression for the output auto-spectrum, S,,(w) = $Y(\omega).Y^*(\omega)$, can be determined by multiplying each side of the expression by its conjugate, Y*(w), and observing that for the orthogonal series, .

$$\begin{aligned}Y(\omega).Y^*(\omega) &= (Y_1(\omega)+ Y_2(\omega)\ \mathrm{t}\ Y_3(\omega)\ \mathrm{t}\ \ldots).(Y_1^*(\omega)\ \mathrm{t}\ Y_2^*(\omega)\ \mathrm{t}\ Y_3^*(\omega)\ \mathrm{t}\ \ldots) \\ &= Y_1(\omega).Y_1^*(\omega) + Y_2(\omega).Y_2^*(\omega) + Y_3(\omega).Y_3^*(\omega) + \ldots\end{aligned} \tag{2.81}$$

Using the properties of Gaussian random variables, it can be shown that,

$$\begin{aligned} Y(\omega).Y^*(\omega) &\approx \frac{|S_{yx}(\omega)|^2}{S_{xx}(\omega)} \\ &+ \frac{1}{2!}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{|S_{y'xx}(\omega_1,\omega_2)|^2}{S_{xx}(\omega_1).S_{xx}(\omega_2)}.\delta(\omega-\omega_1-\omega_2).d\omega_1.d\omega_2 \\ &+ \frac{1}{3!}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{|S_{y''xxx}(\omega_1,\omega_2,\omega_3)|^2}{S_{xx}(\omega_1).S_{xx}(\omega_2).S_{xx}(\omega_3)}.\delta(\omega-\omega_1-\omega_2-\omega_3).d\omega_1.d\omega_2.d\omega_3 \end{aligned} \quad \textbf{(2.82)}$$

This is a relatively simple expression for the output auto-spectrum, which is also known as the output power spectral density. It is also a power transfer relationship which can be used to define first and higher order coherence functions for a non-linear system.

By dividing each side of equation 2.82 by the output auto-spectrum, a series of (squared) coherence functions can be defined,

$$1 = \gamma_1^2(\omega) + \gamma_2^2(\omega_1,\omega_2) + \gamma_3^2(\omega_1,\omega_2,\omega_3) + \ldots \quad \textbf{(2.83)}$$

in which,

$$\gamma_1^2(\omega) = \frac{|S_{yx}(\omega)|^2}{S_{yy}(\omega).S_{xx}(\omega)} \quad \textbf{(2.84)}$$

$$\gamma_2^2(\omega_1,\omega_2) = \frac{1}{2!}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{|S_{y'xx}(\omega_1,\omega_2)|^2}{S_{yy}(\omega).S_{xx}(\omega_1).S_{xx}(\omega_2)}.\delta(\omega-\omega_1-\omega_2).d\omega_1.d\omega_2 \quad (2.85)$$

$$\gamma_3^2(\omega_1,\omega_2,\omega_3) = \frac{1}{3!}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{|S_{y''xxx}(\omega_1,\omega_2,\omega_3)|^2}{S_{yy}(\omega).S_{xx}(\omega_1).S_{xx}(\omega_2).S_{xx}(\omega_3)} .\delta(\omega-\omega_1-\omega_2-\omega_3).d\omega_1.d\omega_2.d\omega_3 \quad (2.86)$$

The n^{th} order coherence function, $\gamma_n^2(\omega_1,\ldots,\omega_n)$, quantifies the energy coupling of the n^{th} term in the orthogonal model, and is positive and dimensionless. Whereas, for example, the bispectrum $S_{yxx}(\omega_1,\omega_2)$ depends on both the amplitude and phase coherence between the three spectral components involved, the bicoherence function $\gamma_2^2(\omega_1,\omega_2)$ removes the dependence upon the amplitude and detects phase coherence. Despite this fact, it is clear that the energy coupling due to the non-linear terms would be dependent on the input power level, and that the higher order coherence functions would be functions of the excitation level.

As an example, consider the information regarding the interactions which can interpreted, for the particular level of excitation used, from the first and second order coherence functions in several specific cases,

1. $\gamma_1 \approx 1$
   The first order terms dominate the model and the second and higher order terms do not contribute significantly, indicating that the system is behaving approximately linearly.

2. $0 < \gamma_1^2 < 1 \ ; \ 0 < \gamma_2^2 < 1 \ ; \ \gamma_1^2 + \gamma_2^2 \approx 0$
   The system is non-linear, with quadratic coupling. Higher order coupling is not detected, and a model with first and second order terms would characterise the system.

3. $0 < \gamma_1^2 < 1 \ ; \ 0 < \gamma_2^2 < 1 \ ; \ \gamma_1^2 + \gamma_2^2 < 1$
   The non-linear coupling is not limited to first and second order, and a model extended to include higher order terms would improve the representation.

4. $0 < \gamma_1^2 < 1 \ ; \ \gamma_2^2 \approx 0 \ ; \ \gamma_1^2 + \gamma_2^2 < 1$
   The system is non-linear, but the model would not be significantly improved by including second order terms. Higher order effects are considerable.

5. $\gamma_1^2 \approx 0$
   The system is so strongly non-linear that the first order effects have less influence on the behaviour of the system than the higher order effects.

Clearly, the expressions for coherence functions only strictly apply when the input is Gaussian. Consequently, the problems which would arise in the practical determination of the higher order coherence functions correspond to those encountered when attempting to measure the Wiener kernels. However, if a structure is excited with broadband excitation which exceeds the bandwidth of interest, it is possible to calculate at least an approximation to the first order coherence function,

$$\gamma_1^2(\omega) = \frac{|S_{yx}(\omega)|^2}{S_{yy}(\omega).S_{xx}(\omega)}$$

Most importantly, this is the ordinary coherence function which is usually measured by a spectrum analyser. Clearly, one reason for the ordinary coherence function falling significantly below unity may be attributed to the existence of non-linear energy coupling in the system.

It should be borne in mind, however, that there are other circumstances which could lead to a low coherence function measured from a structure. Noise on the measured signals would affect the power spectra computed and hence degrade the coherence function. If there is interaction between the shaker and the structure, the spectrum of the force signal may suffer 'drop-out' and be sensitive to noise particularly around the resonance regions. In addition the noise content on the response signals would have a greater effect in regions where the response signal is small, for example near an anti-resonance. Lack of resolution in the digitised spectra can also affect the coherence function, again around the resonances, in structures which are lightly damped, although usually this can be avoided by improving experimentation.

Chapter 3

Higher Order Frequency Response Functions

3.1 Introduction

Two approaches to modelling structures which may be non-linear have been introduced in chapters 1 and 2. Differential equations of motion can be used to represent a wide class of MDOF non-linear structures, and can be identified directly using Direct Parameter Estimation (DPE). A subset of this class can also be described by the Volterra series which is in effect the inverse of the equations of motion, expressing the response as an explicit operation on the system input. The multi-dimensional Fourier transforms of each of the kernels in the Volterra series form a set of higher order FRFs which can be used to represent and explain the resonance behaviour of many non-linear structures. The two alternative models are complementary since they can give insight into the behaviour of systems not provided by the other.

The objective of this chapter is to show how in this way the Volterra Series can be used to understand the frequency domain response characteristics of many non-linear structures, by deriving the general relationship which exists with the differential equation model. This relationship can be used in conjunction with system identification techniques since it provides a stepping-stone between frequency response measurements and the identified parametric model, a subject which is addressed in chapter 4. For a structure which has a Volterra series description, both the parameters of the equations of motion, and the higher order FRFs, can be regarded as properties of the structure in exactly the same way.

One approach to deriving the relationships is to use a method similar to one adopted in the solution of ordinary linear differential equations, which assumes that a harmonic input would give rise to a harmonic output. Clearly this assumption can be justified for a structure which behaves linearly. For a non-linear structure, the exact representation of the response to a harmonic input must be extended to include higher harmonic terms and, in this context, the most convenient formulation of the response signals is provided by the Volterra series description in the frequency domain. Each higher harmonic component in the response is considered to be a contribution from one or more of the higher order FRFs. By harmonically probing the equations of motion, substituting Volterra series expressions for each of the responses, the relationship between the parameters of the equations and the FRFs can be established [5].

In this chapter, an input consisting of only one harmonic is considered initially. This is the simplest way of probing the equations, since the response can be expressed using only the 'principal diagonals' of the FRFs, in effect reducing the multi-dimensionality of the higher order FRFs to a single-dimensional function of the fundamental input frequency. Apart from the relative ease of illustrating the FRFs of any order, another advantage of this simplification is that it is possible to develop the relationships between the parameters and the FRFs up to any order because a relatively simple pattern emerges as the harmonic probing is extended progressively to the higher orders.

The expressions which are derived using a single harmonic input are general in that, since each FRF is a property of the system, the relationships between the diagonals hold for all types of input waveform. However a single harmonic cannot be generated in practice, and the response to any physical input cannot be described simply in terms of the diagonals of the FRF, but would need to include additional points in the multi-dimensional FRF off the principal diagonal. Whereas probing with a single harmonic is relatively straightforward, the generalised procedure is considerably more complicated. For example, to deduce the multi-dimensional expression for $H_n(\omega_1, \ldots, \omega_n)$ it would be necessary to use an input of n harmonics, $\sum_{i=1}^{n} X_i e^{j\omega_i t}$. For any general input such as broadband random excitation, the Volterra description of the response signals must be in terms of the multi-dimensional FRFs, since all combinations of frequencies are possible.

An important fact to remember is that for a MDOF model, the expressions for the higher order FRFs are derived only for a single source of energy, ie. excitation at one location on the structure (and in one direction). However, an alternative family of higher order FRFs could be defined to represent simultaneous excitation at more than one location.

Whereas the equations of motion can be easily amended to express excitation at several locations, the higher order FRF representation of the response of the structure would be considerably more complex than the single-input, multi-output scenario. The entire harmonic probing technique can be regarded as an attempt to match each component in the response spectra with the input component which caused it, the higher order FRFs being used to quantify all the energy transfer patterns which can relate the input and output components. In general, the greater the number of harmonic input components, particularly if at different locations, the more numerous and complicated the possible interactions become.

3.2 The Principal Diagonals of the FRFs

For the class of structures which do have a Volterra series representation, a relationship can be established between the principal diagonals of the FRFs and the parameters of the equations of motion. The objective of this section is to derive these expressions for FRFs up to third order, and to establish a pattern to enable extension to any order.

The procedure uses the equation of motion model in which an N DOF; non-linear structure (with excitation at point r) is represented by a set of N equations each having the form,

$$
\begin{aligned}
& m_s.\ddot{y}_s(t) + {}_{ss}c_1.\dot{y}_s(t) + {}_{ss}k_1.y_s(t) \\
& \quad - \sum_{\substack{p=a \\ p\neq s}}^{N} [{}_{sp}c_1.(\dot{y}_p(t) - \dot{y}_s(t)) + {}_{sp}k_1.(y_p(t) - y_s(t))] \\
& \quad + {}_{ss}c_2.\dot{y}_s^2(t) + {}_{ss}k_2.y_s^2(t) \\
& \quad - \sum_{\substack{p=a \\ p\neq s}}^{N} \left[{}_{sp}c_2.(\dot{y}_p(t) - \dot{y}_s(t))^2 + {}_{sp}k_2.(y_p(t) - y_s(t))^2 \right] \\
& \quad + {}_{ss}c_3.\dot{y}_s^3(t) + {}_{ss}k_3.y_s^3(t) \\
& \quad - \sum_{\substack{p=a \\ p\neq s}}^{N} \left[{}_{sp}c_3.(\dot{y}_p(t) - \dot{y}_s(t))^3 + {}_{sp}k_3.(y_p(t) - y_s(t))^3 \right] \\
& \quad + \text{higher order terms} \ldots = \begin{cases} 0 & \text{if } s \neq r \\ x_r(t) & \text{if } s = r \end{cases}
\end{aligned}
\tag{3.1}
$$

When developing an expression for the n^{th} order FRF, interactions of up to n different components can play a part, and so terms in the equation to order n are of specific interest. However, as will be seen, terms of order higher than n do not play a part. The terms in equation 3.1 are limited to third order to correspond with the FRFs of interest.

The expressions for the principal diagonals can be probed by considering an input which is simply a single harmonic, applied at the direct point r in the equations of motion,

$$x_r(t) = Xe^{j\omega t} \tag{3.2}$$

For FRFs up to order n, the procedure is to substitute into each of the equations an expression for each of the response signals using the frequency domain expression of the Volterra series and then to equate coefficents of $X^n e^{jn\omega t}$. Using equation 2.47 from section 2.2.4, the response to a harmonic input can be expressed as an infinite series of terms incorporating the diagonals of each of the first and higher order FRFs,

$$\begin{aligned} y_s(t) \quad &= \quad {}_{rs}H_1(\omega).Xe^{j\omega t} + {}_{rs}H_2(\omega,\omega).X^2e^{j2\omega t} \\ &+ \quad {}_{rs}H_3(\omega,\omega,\omega).X^3e^{j3\omega t} + \ldots + {}_{rs}H_n(\omega,\ldots,\omega).X^ne^{jn\omega t} \bullet t \ldots \end{aligned} \tag{3.3}$$

By differentiating this representation of the displacement term by term, similar expressions for velocity and acceleration are obtained,

$$\begin{aligned} \dot{y}_s(t) &= (j\omega).{}_{rs}H_1(\omega).Xe^{j\omega t} + (j2\omega).{}_{rs}H_2(\omega,\omega).X^2e^{j2\omega t} \\ &+ (j3\omega).{}_{rs}H_3(\omega,\omega,\omega).X^3e^{j3\omega t} + \ldots + (jn\omega).{}_{rs}H_n(\omega,\ldots,\omega).X^ne^{jn\omega t} + \ldots \end{aligned} \tag{3.4}$$

$$\begin{aligned} \ddot{y}_s(t) &= (j\omega)^2.{}_{rs}H_1(\omega).Xe^{j\omega t} + (j2\omega)^2.{}_{rs}H_2(\omega,\omega).X^2e^{j2\omega t} \\ &+ (j3\omega)^2.{}_{rs}H_3(\omega,\omega,\omega).X^3e^{j3\omega t} + \ldots + (jn\omega)^2.{}_{rs}H_n(\omega,\ldots,\omega).X^ne^{jn\omega t} + \ldots \end{aligned} \tag{3.5}$$

To probe the diagonal of the n^{th} order FRF, all terms of order less than or equal to n must be included since, when the expressions of the response signals are substituted into the equations of motion, the polynomials cause the different components to interact. The interaction of two or more lower order terms may yield a coefficient of $X^ne^{jn\omega t}$ which should be included in the solution. However, terms of order higher than n cannot contribute to this particular coefficient. This is important because it enables the representation of each of the response signals to be truncated to n terms without causing any approximation in the ultimate expression of the principal diagonal of the n^{th} order FRF.

The analytical procedure is to probe each order sequentially, starting with $H_1(\omega)$. The expression for the first order FRF is central to the entire approach because it is used to simplify each of the higher order expressions. The pivotal role played by the first order terms in the Volterra series was also noted in chapter 2 and, because of its fundamental importance in understanding the energy transfer mechanism which takes place in weakly non-linear structures, the same subject is discussed in greater depth later in this chapter.

The first order FRF is probed by substituting expressions of the form,

$$\begin{aligned} x_r(t) &= Xe^{j\omega t} \\ y_s(t) &= {}_{rs}H_1(\omega).Xe^{j\omega t} + \text{higher order terms..} \end{aligned} \tag{3.6}$$

into the equations of motion (eqn. 3.1), and determining coefficients of $Xe^{j\omega t}$,

$$\begin{aligned} &((j\omega)^2.m_s + (j\omega).{}_{ss}c_1 + {}_{ss}k_1).{}_{rs}H_1(\omega) \\ &\quad - \sum_{i=1}^{n}((j\omega).{}_{sp}c_1 + {}_{sp}k_1).({}_{rp}H_1(\omega) - {}_{rs}H_1(\omega)) = \begin{cases} 0 & s \neq r \\ 1 & s = r \end{cases} \end{aligned} \tag{3.7}$$

Similar expressions for each of the N equations of motion can be written down and rearranged in matrix form,

$$\left[(j\omega)^2.[m] + (j\omega).[c_1] + [k_1]\right] . \begin{vmatrix} {}_{ra}H_1(\omega) \\ \\ {}_{rr}H_1(\omega) \\ {}_{rs}H_1(\omega) \\ \\ {}_{rN}H_1(\omega) \end{vmatrix} = \begin{vmatrix} 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{vmatrix} \tag{3.8}$$

in which the mass parameters are represented by a square matrix with zero off diagonal terms,

$$[m] = \begin{bmatrix} m_a & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & m_r & 0 & \cdots & 0 \\ 0 & \cdots & 0 & m_s & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & m_N \end{bmatrix}$$

and the first order damping and stiffness matrices each have the form,

$$[k_1] = \begin{vmatrix} \sum_{p=a}^{N} {}_{ap}k_1 & \cdots & -{}_{ar}k_1 & -{}_{as}k_1 & \cdots & -{}_{aN}k_1 \\ & \ddots & & & \ddots & \\ -{}_{ra}k_1 & \cdots & \sum_{p=a}^{N} {}_{rp}k_1 & -{}_{rs}k_1 & \cdots & -{}_{rN}k_1 \\ -{}_{sa}k_1 & \cdots & -{}_{sr}k_1 & \sum_{p=a}^{N} {}_{sp}k_1 & \cdots & -{}_{sN}k_1 \\ & \ddots & & & \ddots & \\ -{}_{Na}k_1 & \cdots & -{}_{Nr}k_1 & -{}_{Ns}k_1 & \cdots & \sum_{p=a}^{N} {}_{Np}k_1 \end{vmatrix}$$

The full first order FRF matrix can be related in a more general way to the mass, stiffness and damping matrices by considering the effect of moving the input to each of the N response locations in turn. In each case there is still only one input and N similar expressions involving the first order parameter matrices can be generated which, when combined into a matrix form, can be written,

$$\left[(j\omega)^2.[m] + (j\omega).[c_1] + [k_1]\right] . \begin{bmatrix} {}_{aa}H_1(\omega) & \dots & {}_{ra}H_1(\omega) & {}_{sa}H_1(\omega) & \dots & {}_{Na}H_1(\omega) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ {}_{ar}H_1(\omega) & \dots & {}_{rr}H_1(\omega) & {}_{sr}H_1(\omega) & \dots & {}_{Nr}H_1(\omega) \\ {}_{as}H_1(\omega) & \dots & {}_{rs}H_1(\omega) & {}_{ss}H_1(\omega) & \dots & {}_{Ns}H_1(\omega) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ {}_{aN}H_1(\omega) & \dots & {}_{rN}H_1(\omega) & {}_{sN}H_1(\omega) & \dots & {}_{NN}H_1(\omega) \end{bmatrix} = I$$

where I denotes the (N x N) Identity Matrix. Thus,

$$[H_1(\omega)]^{-1} = \left[(j\omega)^2.[m] + (j\omega).[c_1] + [k_1]\right] \tag{3.9}$$

Each of the terms ${}_{sp}H_1(\omega)$ describes the first order response of the system at point p due to a harmonic input at point s. Since the full first order matrix is related to only the first order stiffness and damping parameters without involving non-linear parameters, the result would be identical if all the higher order parameters were zero and the system were linear. In this case, all the higher order FRFs would be null, and the first order FRF matrix could be used to characterise the system.

Several simple but useful observations can be made at this stage:

1. The 'fundamental' resonances ω_0 of the non-linear system, corresponding to the resonances of the related linear system (as introduced in chapter 2), are determined by the mass and first order stiffness parameters via the familiar eigenvalue problem,

$$\left|[k_1] - \omega_0^2.[m]\right| = 0 \tag{3.10}$$

This will be important in the interpretation of the resonances of the higher order FRFs.

2. At low and high frequencies, $[H_1(\omega)]$ can be seen to approach $[k_1]^{-1}$ and $(j\omega)^2.[m]^{-1}$ respectively, the so-called mass and stiffness asymptotes of the FRF.
(These are referred to in chapter 4 regarding the application of the identified parameter matrices to the synthesis of the full first order FRF.)

3. From equation 3.9, an expression which will be used subsequently can be written directly,

$$\left[(jn\omega)^2.[m] + (jn\omega).[c_1] + [k_1]\right] = [H_1(n\omega)]^{-1} \tag{3.11}$$

To form an expression for the principal diagonal of the second order FRF, the representation of the responses (eqns. 3.3 to 3.5) are simply extended to include the first two terms, eg.

$$y_s(t) = {}_{rs}H_1(\omega).Xe^{j\omega t} + {}_{rs}H_2(\omega,\omega).X^2e^{j2\omega t} + \ldots \tag{3.12}$$

Substituting into the equation of motion (eqn. 3.1) and equating coefficients of $X^2e^{j2\omega t}$,

$$\begin{aligned}&((j2\omega)^2.m_s + (j2\omega).{}_{ss}c_1 + {}_{ss}k_1).{}_{rs}H_2(\omega,\omega)\\ &\quad -\sum_{\substack{p=a\\p\neq s}}^{N}((j\omega)^2.{}_{sp}c_1 + {}_{sp}k_1).({}_{rp}H_2(\omega,\omega) - {}_{rs}H_2(\omega,\omega))\\ &\quad +((j\omega)^2.{}_{ss}c_2 + {}_{ss}k_2).({}_{rs}H_1(\omega))^2\\ &\quad -\sum_{\substack{p=a\\p\neq s}}^{N}((j\omega)^2.{}_{sp}c_2 + {}_{sp}k_2).({}_{rp}H_1(\omega) - {}_{rs}H_1(\omega))^2 = 0\end{aligned} \tag{3.13}$$

This equation can be simplified by introducing 'T-notation' for the terms which include the second order parameters,

$${}_{ss}T_{11} = \left((j\omega)^2.{}_{ss}c_2 + {}_{ss}k_2\right).({}_{rs}H_1(\omega))^2 \tag{3.14}$$

$${}_{sp}T_{11} = \left((j\omega)^2.{}_{sp}c_2 + {}_{sp}k_2\right).({}_{rp}H_1(\omega) - {}_{rs}H_1(\omega))^2 \tag{3.15}$$

The right suffices on the T terms denote the order of the FRF terms in the expression. For example T_{11} indicates that the term is a function of the product of two first order FRFs.

$$\begin{aligned}&((j2\omega)^2.m_s + (j2\omega).{}_{ss}c_1 + {}_{ss}k_1).{}_{rs}H_2(\omega,\omega)\\ &\quad -\sum_{\substack{p=a\\p\neq s}}^{N}((j\omega)^2.{}_{sp}c_1 + {}_{sp}k_1).({}_{rp}H_2(\omega,\omega) - {}_{rs}H_2(\omega,\omega)) = -\left({}_{ss}T_{11} - \sum_{\substack{p=a\\p\neq s}}^{N}{}_{sp}T_{11}\right)\end{aligned} \tag{3.16}$$

By writing down N similar expressions, it is possible to write the matrix formulation,

$$\left[(j2\omega)^2.[m] + (j2\omega).[c_1] + [k_1]\right].\begin{Bmatrix} {}_{ra}H_2(\omega,\omega)\\ \\ {}_{rr}H_2(\omega,\omega)\\ {}_{rs}H_2(\omega,\omega)\\ \\ {}_{rN}H_2(\omega,\omega)\end{Bmatrix} = -\begin{Bmatrix} {}_{aa}T_{11} - \sum_{p=b}^{N}{}_{ap}T_{11}\\ \vdots\\ {}_{rr}T_{11} - \sum_{\substack{p=a\\p\neq r}}^{N}{}_{rp}T_{11}\\ \\ {}_{ss}T_{11} - \sum_{\substack{p=a\\p\neq s}}^{N}{}_{sp}T_{11}\\ \vdots\\ {}_{NN}T_{11} - \sum_{p=a}^{N-1}{}_{Np}T_{11}\end{Bmatrix} \tag{3.17}$$

Or in shorthand notation,

$$\{H_2(\omega,\omega)\} = -[H_1(2\omega)].\{T_{11}\} \tag{3.18}$$

For the diagonal of the third order FRF, the first three terms must be included in the response expressions, eg.

$$y_s(t) = {}_rs H_1(\omega).X e^{j\omega t} + {}_rs H_2(\omega,\omega).X^2 e^{j2\omega t} + {}_rs H_3(\omega,\omega,\omega).X^3 e^{j3\omega t} + \ldots \tag{3.19}$$

After substitution and equating coefficients of $X^3 e^{j3\omega t}$,

$$\begin{aligned}
&((j3\omega)^2.m_s + (j3\omega).{}_{ss}c_1 + {}_{ss}k_1).{}_rs H_3(\omega,\omega,\omega) \\
&\quad -\sum_{\substack{p=a \\ p\neq s}}^{N}((j\omega)^2.{}_{sp}c_1 + {}_{sp}k_1).({}_{rp}H_3(\omega,\omega,\omega) - {}_rs H_3(\omega,\omega,\omega)) \\
&\quad +2.((j\omega).(j\omega)^2.{}_{ss}c_2 + {}_{ss}k_2).{}_rs H_1(\omega).{}_rs H_2(\omega,\omega) \\
&\quad -2.\sum_{\substack{p=a \\ p\neq s}}^{N}((j\omega).(j\omega)^2.{}_{sp}c_2 + {}_{sp}k_2).({}_{rp}H_1(\omega) - {}_rs H_1(\omega)).({}_{rp}H_2(\omega,\omega) - {}_rs H_2(\omega,\omega)) \\
&\quad +((j\omega)^3.{}_{ss}c_3 + {}_{ss}k_3).({}_rs H_1(\omega))^3 \\
&\quad -\sum_{\substack{p=a \\ p\neq s}}^{N}((j\omega)^3.{}_{sp}c_3 + {}_{sp}k_3).({}_{rp}H_1(\omega) - {}_rs H_1(\omega))^3 = 0
\end{aligned} \tag{3.20}$$

Introducing the T-notation for the terms which incorporate the second and third order parameters,

$${}_{ss}T_{12} = ((j\omega).(j2\omega).{}_{ss}c_2 + {}_{ss}k_2).{}_rs H_1(\omega).{}_rs H_2(\omega,\omega) \tag{3.21}$$

$${}_{sp}T_{12} = ((j\omega).(j2\omega).{}_{sp}c_2 + {}_{sp}k_2).({}_{rp}H_1(\omega) - {}_rs H_1(\omega)).({}_{rp}H_2(\omega,\omega) - {}_rs H_2(\omega,\omega)) \tag{3.22}$$

$${}_{ss}T_{111} = \left((j\omega)^3.{}_{ss}c_3 + {}_{ss}k_3\right).({}_rs H_1(\omega))^3 \tag{3.23}$$

$${}_{sp}T_{111} = \left((j\omega)^3.{}_{sp}c_3 + {}_{sp}k_3\right).({}_{rp}H_1(\omega) - {}_rs H_1(\omega))^3 \tag{3.24}$$

$$\begin{aligned}
&((j3\omega)^2.m_s + (j3\omega).{}_{ss}c_1 + {}_{ss}k_1).{}_rs H_3(\omega,\omega,\omega) \\
&\quad -\sum_{\substack{p=a \\ p\neq s}}^{N}((j\omega)^2.{}_{sp}c_1 + {}_{sp}k_1).({}_{rp}H_3(\omega,\omega,\omega) - {}_rs H_3(\omega,\omega,\omega)) \\
&\qquad = -2.\left({}_{ss}T_{12} - \sum_{\substack{p=a \\ p\neq s}}^{N}{}_{sp}T_{12}\right) - \left({}_{ss}T_{111} - \sum_{\substack{p=a \\ p\neq s}}^{N}{}_{sp}T_{111}\right)
\end{aligned} \tag{3.25}$$

Writing down N similar equations, the matrix form can be written,

$$\{H_3(\omega,\omega,\omega)\} = -[H_1(3\omega)].\{2.\{T_{12}\} + \{T_{111}\}\} \tag{3.26}$$

Summarising, the principal diagonals of the first three receptance FRFs can be written,

$$[H_1(\omega)] = \left[(j\omega)^2.[m] + (j\omega).[c_1] + [k_1]\right]^{-1} \tag{3.27}$$

$$\{H_2(\omega,\omega)\} = -[H_1(2\omega)].\{T_{11}\} \tag{3.28}$$

$$\{H_3(\omega,\omega,\omega)\} = -[H_1(3\omega)].\{2.\{T_{12}\} + \{T_{111}\}\} \tag{3.29}$$

In these expressions, no restrictions have been imposed on either the highest order term in the governing polynomial function, or on the number of degrees of freedom. However to understand some of the features of the principal diagonals of the FRFs, it is useful to consider initially a SDOF system,

$$\ddot{y}(t) + \alpha.\dot{y}(t) + \beta_1.y(t) + \beta_2.y^2(t) + \beta_3.y^3(t) = x(t) \tag{3.30}$$

In this example, the non-linearity is in the stiffness restoring force which can be represented by a polynomial with terms up to order three. The diagonals of the FRFs simplify to,

$$H_1(\omega) = \frac{1}{(j\omega)^2 + (j\omega).\alpha + \beta_1} \tag{3.31}$$

$$H_2(\omega,\omega) = -H_1(2\omega).\beta_2.H_1(\omega)^2 \tag{3.32}$$

$$H_3(\omega,\omega,\omega) = -H_1(3\omega).(\beta_2.H_1(\omega).H_2(\omega,\omega) + \beta_3.H_1(\omega)^3) \tag{3.33}$$

One important observation can be made regarding the higher order FRFs at this stage:
The principal diagonals of the higher order FRFs can be expressed in terms of the diagonals of lower order FRFs, and parameters of the same and lower orders.
In the SDOF example, $H_2(\omega,\omega)$ is completely independent of the third order parameter β_3 but, through $H_1(\omega)$, does depend on β_1 and the mass and damping parameters. As a consequence, the FRF of any order can be expressed in terms of the parameters and only the first order FRF by substitution, for example,

$$\begin{aligned} H_3(\omega,\omega,\omega) &= -H_1(3\omega).(\beta_2.H_1(\omega).H_2(\omega,\omega) + \beta_3.H_1(\omega)^3) \\ &= -H_1(3\omega).(\beta_2^2.H_1(2\omega).H_1(\omega)^3 + \beta_3.H_1(\omega)^3) \\ &= -H_1(3\omega).(\beta_2^2.H_1(2\omega) + \beta_3).H_1(\omega)^3 \end{aligned} \tag{3.34}$$

As a direct consequence, the poles of the diagonals of the higher order FRF are determined only by the poles of the first order FRF.
A specific example of a system of this type was investigated using the parameters,

$$\alpha = 10\ (\mathrm{N/ms^{-1}}),\ \beta_1 = 10^4\ (\mathrm{N/m}),\ \beta_2 = 10^7\,(\mathrm{N/m^2})\ \text{and}\ \beta_3 = 10^{10}\,(\mathrm{N/m^3}) \tag{3.35}$$

The polynomial stiffness restoring force function is shown in figure 3.1 together with a block diagram of the system, and figure 3.2 illustrates the principal diagonals of the first three FRFs generated for this system.

In general, if ω_r is a pole of H(w), then $\frac{\omega_r}{m}$ is a pole of $H_n(\omega,\ldots,\omega)$, where $1 \leq m \leq n$.

For the SDOF system, since $\omega_0 = \sqrt{\beta_1}$ is the fundamental resonance, the 'primary resonance' of $H_2(\omega,\omega)$ will be at ω_0, with a 'secondary resonance' at $\frac{\omega_0}{2}$ due to $H_1(2\omega)$.

The phase change associated with the secondary resonance is 180" whereas for the primary resonance it is 360" because $H_1(\omega)$ is raised to the power 2 in the expression for $H_2(\omega,\omega)$.

In general, the third order FRF will also exhibit primary and secondary resonances at both ω_0 and $\frac{\omega_0}{2}$, and an additional 'tertiary resonance' at $\frac{\omega_0}{3}$ due to $H_1(3\omega)$.

In the numerical example, the fundamental resonance occurs at,

$$\omega_0 = \sqrt{\beta_1} = 100 \text{ rad/s} = 15.92 \text{ Hz}$$

and the second and tertiary resonances at 8.0 and 5.3 Hz respectively are clearly shown on the diagonals. Also, at the fundamental resonance,

$$|H_1(\omega_0)| = \left| \frac{1}{(j\omega_0).\alpha} \right| = 10^{-3} \quad (\text{m / N})$$

Hence for -1 (N) $\leq$ z(t), ≤ 1 (N), the maximum displacement response will be limited to fO.OO1 (m),which are the limits selected for the stiffness polynomial shown.

The low and high frequency asymptotes of each of the higher order diagonals can be written for this system,

$$\omega \to 0 \quad \begin{cases} H_1(\omega) \to \dfrac{1}{\beta_1} = 10^{-4} \ \text{(m/N)} \\ H_2(\omega,\omega) \to -\dfrac{1}{\beta_1}.\beta_2.\dfrac{1}{\beta_1^2} = -10^{-5} \ \text{(m/N}^2\text{)} \\ H_3(\omega,\omega,\omega) \to -\dfrac{1}{\beta_1}.\left(\dfrac{\beta_2^2}{\beta_1} + \beta_3\right).\dfrac{1}{\beta_1^3} = -2 \times 10^{-6} \ \text{(m/N}^3\text{)} \end{cases}$$

$$\omega \to \infty \quad \begin{cases} H_1(\omega) \to \dfrac{1}{(j\omega)^2} = -\omega^{-2} \ \text{(m/N)} \\ H_2(\omega,\omega) \to -\dfrac{1}{(j2\omega)^2}.\dfrac{\beta_2}{(j\omega)^4} = \dfrac{10^7}{4}\omega^{-6} \ \text{(m/N}^2\text{)} \\ H_3(\omega,\omega,\omega) \to -\dfrac{1}{(j3\omega)^2}.\left(\dfrac{\beta_2^2}{(j\omega)^2} + \beta_3\right).\dfrac{1}{(j\omega)^6} = \dfrac{10^{14}}{36}\omega^{-10} + \dfrac{10^{10}}{9}\omega^{-8} \ \text{(m/N}^3\text{)} \end{cases}$$

Clearly, the slopes of the diagonals at high frequencies have a steeper (negative) gradient as the order of the FRF increases.

The diagonals have been defined with respect to the displacement response and can therefore be thought of as higher order receptance functions. If the frequency domain response of the system to a single harmonic is written as

$$\mathcal{F}[y(t)] = Y_1(\omega) + Y_2(2\omega) + Y_3(3\omega) + \ldots + Y_n(n\omega) + \ldots \tag{3.36}$$

where the n^{th} harmonic component of the output $Y(n\omega) = X^n(\omega).H_n(\omega,\ldots,\omega)$, then it is possible to write the velocity and acceleration responses as,

$$\begin{aligned}\mathcal{F}[\dot{y}(t)] &= (j\omega).Y_1(\omega) + (j2\omega).Y_2(2\omega) + (j3\omega).Y_3(3\omega) + \ldots + (jn\omega).Y_n(n\omega) + \ldots \\ &= Y_1'(\omega) + Y_2'(2\omega) + Y_3'(3\omega) + \ldots + Y_n'(n\omega) + \ldots \qquad \textbf{(3.37)}\\ \mathcal{F}[\ddot{y}(t)] &= (j\omega)^2.Y_1(\omega) + (j2\omega)^2.Y_2(2\omega) + (j3\omega)^2.Y_3(3\omega) + \ldots + (jn\omega)^2.Y_n(n\omega) + \ldots \\ &= Y_1''(\omega) + Y_2''(2\omega) + Y_3''(3\omega) + \ldots + Y_n''(n\omega) + \ldots \qquad \textbf{(3.38)}\end{aligned}$$

Hence the diagonals of higher order mobility functions which can be defined from the spectrum of the velocity response as,

$$H_n'(\omega,\ldots,\omega) = \frac{Y_n'(nw)}{X(\omega)^n} \tag{3.39}$$

and inertance functions defined as,

$$H_n''(\omega,\ldots,\omega) = \frac{Y_n''(nw)}{X(\omega)^n} \tag{3.40}$$

can each be related to the diagonals of the receptance functions,

$$H_n''(\omega,\ldots,\omega) = (jn\omega).H_n'(\omega,\ldots,\omega) = (jn\omega)^2.H_n(\omega,\ldots,\omega) \tag{3.41}$$

The diagonals of the higher order mobility and inertance functions naturally have different asymptotes than the receptance functions. Both $H_n'(\omega,\ldots,\omega)$ and $H_n''(\omega,\ldots,\omega)$ approach zero as w $\rightarrow$ 0. For the numerical SDOF system, the high frequency asymptotes can be written down, for example,

$$\omega \rightarrow \infty \quad \begin{cases} H_1''(\omega) \rightarrow \dfrac{(j\omega)^2}{(j\omega)^2} = 1 \ (\mathrm{ms}^{-2}/\mathrm{N}) \\ H_2''(\omega,\omega) \rightarrow -\dfrac{(j2\omega)^2}{(j2\omega)^2}.\dfrac{\beta_2}{(j\omega)^4} = -10^7\omega^{-4} \ (\mathrm{ms}^{-2}/\mathrm{N}^2) \\ H_3''(\omega,\omega,\omega) \rightarrow -\dfrac{(j3\omega)^2}{(j3\omega)^2}.\left(\dfrac{\beta_2^2}{(j\omega)^2} + \beta_3\right).\dfrac{1}{(j\omega)^6} = -\dfrac{10^{14}}{4}\omega^{-8} + 10^{10}\omega^{-6} \ (\mathrm{ms}^{-2}/\mathrm{N}^3) \end{cases}$$

The FRFs have different dimensions in each of the three forms between the different orders.

The first, second and third order inertance diagonals generated from the parameters of the SDOF example are shown in figure 3.3.

In this example, a general asymmetrical polynomial has been considered. In practice, some information may be known a priori; for example the system may be governed by a polynomial which has odd order symmetry, in which all the even order parameters are zero. In the SDOF example, this would correspond to $\beta_2 = 0$,

$$H_1(\omega) = \frac{1}{(j\omega)^2 + (j\omega).\alpha + \beta_1} \tag{3.42}$$

$$H_2(\omega,\omega) = -H_1(2\omega).\beta_2.H_1(\omega)^2 = 0 \tag{3.43}$$

$$H_3(\omega,\omega,\omega) = -H_1(3\omega).\beta_3.H_1(\omega)^3 \tag{3.44}$$

Clearly the first order FRF is unaffected by changes to higher order parameters. However the diagonal (and all other off-diagonal points) of the second order FRF would be null. In general, if a non-linear system does have odd order symmetry in each of the polynomials, all of the even order FRFs throughout the infinite series would be null. Although all the odd order FRFs exist, the form of each does change if the polynomial is symmetrical. For example, it is clear from equation 3.44 that $H_3(\omega, w, w)$ would not exhibit a secondary resonance. Substituting the values for the parameters α, β_1 and β_3 as in equation 3.35, and setting $\beta_2 = 0$, the stiffness polynomial is shown in figure 3.4, and $H_3(\omega, w, w)$ in figure 3.5.

In practice, even if the governing stiffness polynomial were symmetrical, a slight preload or deformation would cause an offset in the origin of the polynomial and destroy the symmetry. (In chapter 6, this issue is raised again in relation to measurements from a physical structure.)

Throughout this thesis, one observation continually made is that the first order terms play a particularly important role in the Volterra series representation. In this example, to be able to define higher order FRFs at all using the Volterra series, it is necessary that the system has 'linear rigidity', ie. $\beta_1 \neq 0$. However the higher, odd order terms in the non-linear function may be very small or zero. Consider the SDOF example when $\beta_3 = 0$. Both the first and second order FRFs are unaffected by changes in the third order parameter. Importantly, the odd order FRFs do exist despite the odd order parameters being zero, for example,

$$H_3(\omega,\omega,\omega) = -H_1(3\omega).\beta_2^2.H_1(2\omega).H_1(\omega)^3 \tag{3.45}$$

because the presence of the first order term means that the polynomial does not have even order symmetry. Inserting the numerical values for β_1 and β_2 from equation 3.35, the stiffness polynomial was constructed (figure 3.6), and the $H_3(\omega, w, w)$ generated (figure 3.7).
Since the governing polynomial remains asymmetrical, all the odd order FRFs exist.
In the SDOF example, the only way to make all the odd order FRFs null would be to set both β_2 and β_3 to zero. Clearly in this situation the system would be linear, and only the first order would FRF exist, completely characterising the dynamics of the system.

The discussion of the SDOF example may appear rather laboured, but is appropriate since, for a general MDOF model, the diagonals of the FRF are related to the parametric model in a similar way. Moreover, the diagonals of higher order FRFs exhibit primary, secondary and tertiary resonances, etc. which again correspond to each of the fundamental resonances determined by the mass and first order stiffness parameters.

To illustrate the discussion, a numerical MDOF example is introduced, representing a structure with 3 degrees of freedom including third order non-linear stiffness functions to link each response location, except between locations 1 and 3. A diagram of this lumped parameter model is shown in figure 3.8.

For an input at location 1, this system is represented by 3 equations of motion,

$$\begin{aligned}
&\ddot{y}_1(t) \quad +10.\dot{y}_1(t) + 10^4.y_1(t) + 10^7.y_1^2(t) + 10^{10}.y_1^3(t) \\
&\qquad -10.(\dot{y}_2(t) - \dot{y}_1(t)) - 10^4.(y_2(t) - y_1(t)) - 10^7.(y_2(t) - y_1(t))^2 - 10^{10}.(y_2(t) - y_1(t))^3 = z(t) \\
&\ddot{y}_2(t) + 10.\dot{y}_2(t) + 10^4.y_2(t) + 10^7.y_2^2(t) + 10^{10}.y_2^3(t) \\
&\qquad -10.(\dot{y}_1(t) - \dot{y}_2(t)) - 10^4.(y_1(t) - y_2(t)) + 10^7.(y_1(t) - y_2(t))^2 - 10^{10}.(y_1(t) - y_2(t))^3 \\
&\qquad -10.(\dot{y}_3(t) - \dot{y}_2(t)) - 10^4.(y_3(t) - y_2(t)) - 10^7.(y_3(t) - y_2(t))^2 - 10^{10}.(y_3(t) - y_2(t))^3 = 0 \\
&\ddot{y}_3(t) \quad +10.\dot{y}_3(t) + 10^4.y_3(t) + 10^7.y_3^2(t) + 10^{10}.y_3^3(t) \\
&\qquad -10.(\dot{y}_2(t) - \dot{y}_3(t)) - 10^4.(y_2(t) - y_3(t)) + 10^7.(y_2(t) - y_3(t))^2 - 10^{10}.(y_2(t) - y_3(t))^3 = 0
\end{aligned} \tag{3.46}$$

Expressions for the diagonals of the first, second and third order FRFs can be formulated directly by inserting the parameters into the appropriate expressions. Also, if the parameters of the equations of motion of a structure had been identified, for example by using the DPE method, then it would be possible to determine the diagonals of the FRFs directly by substitution in the same way. For this system, the first, second and third order receptance functions for the direct location (1) have been generated and are shown in figure 3.9.

Since each mass parameter is set to unity, the fundamental resonances can be determined,

$$\left|[k_1] - \omega_0^2.I\right| = 0 \tag{3.47}$$

and are at ω_a = 15.92, ω_b = 22.51, ω_c = 31.83 Hz, as shown on the first order FRF.

In addition to the primary resonances, both the second and third order FRFs exhibit clear secondary resonances at 7.96 and 11.25 Hz corresponding to ω_a and ω_b respectively. However it is difficult to distinguish the secondary resonance from ω_c since it occurs at 15.91 Hz and is practically coincident with the more dominant primary resonance of w,.

Similarly, although the tertiary resonance of ω_a at 5.31 Hz is clear on the third order diagonal, the tertiary resonances at 7.50 and 10.61 Hz, corresponding to ω_b and ω_c, have a relatively small contribution and can be hardly distinguished from the primary and secondary resonances. The coincidence of resonance terms in this manner is an unavoidable feature of the diagonals of the higher order FRFs, and sometimes it presents difficulties when attempting to interpret the FRFs for systems with many fundamental resonances.

In general, although only expressions for the first three FRFs have been derived in this text, the comments made regarding the relationships between the principal diagonals and the parameters of the model can be shown to hold true for FRFs of any order. The harmonic probing procedure can be employed for any order n, and $H_n(\omega,\ldots,\omega)$ is found to be related to lower order diagonals and parameters. A pattern emerges for both the form and coefficient of each term involved, making it possible to write down the expression for $H_n(\omega,\ldots,\omega)$ directly from the *partitions* of n. As an example, consider the sixth order FRF diagonal: The partitions of 6 are $(0,6)$, which corresponds to the left hand side of the expression, and $(1,5)\,(2,4)\,(1,1,4)\,(3,3)\,(1,2,3)\,(2,2,2)\,(1,1,1,3)\,(1,1,2,2)\,(1,1,1,1,2)$ and $(1,1,1,1,1,1)$.

Using the T-notation for the terms which include parameters higher than first order, the partitions can be used to form the right hand side of the expression,

$$\{H_6(\omega,\ldots,\omega)\} = -[H_1(6\omega)]\cdot\left\{\tfrac{(1+1)!}{1!1!}\{T_{15}\} + \tfrac{(1+1)!}{1!1!}\{T_{24}\} + \tfrac{(2+1)!}{2!1!}\{T_{114}\} + \tfrac{2!}{2!}\{T_{33}\} + \tfrac{(1+1+1)!}{1!1!1!}\{T_{123}\} + \tfrac{3!}{3!}\{T_{222}\} + \tfrac{(3+1)!}{3!1!}\{T_{1113}\} + \tfrac{(2+2)!}{2!2!}\{T_{1122}\} + \tfrac{(4+1)!}{4!1!}\{T_{11112}\} + \tfrac{6!}{6!}\{T_{111111}\}\right\} \tag{3.48}$$

where, as before, the coefficient of each T term relates to the number of FRF terms of the same order in each product. Since the polynomial function of the SDOF example has an order of three, only the terms formed from the interaction of three FRFs appear in the expression for the diagonal of the sixth order FRF,

$$H_6(\omega,\ldots,\omega) = -H_1(6\omega).(2.H_1(\omega).H_5(\omega,\ldots,\omega) + 2.H_2(\omega,\omega).H_4(\omega,\ldots,\omega) + 3.H_1(\omega)^2.H_4(\omega,\ldots,\omega) + H_3(\omega)^2 + 6.H_1(\omega).H_2(\omega).H_3(\omega,\omega,\omega) + H_2(\omega,\omega)^3) \tag{3.49}$$

Once again it can be observed that, although the governing stiffness polynomial has terms only to third order, an expression can be formulated for the FRFs of any order. In general, the FRFs of all orders would exist. The response of this system could only be represented exactly by a Volterra series with an infinite number of terms, provided β_2 or $\beta_3 \neq 0$.

3.3 Multi-Dimensional Frequency Response Functions

Perhaps the easiest way of deriving expressions for the multi-dimensional FRFs in terms of the parameters of the equations of motion is to start by considering the SDOF system used in the previous section (equation 3.30),

$$\ddot{y}(t) + \alpha.\dot{y}(t) + \beta_1.y(t) + \beta_2.y^2(t) + \beta_3.y^3(t) = x(t) \tag{3.50}$$

The multi-dimensional second-order FRF can be used to quantify the interaction between any two harmonic components of the input signal. For the input,

$$x(t) = X_1 e^{j\omega_1 t} + X_2 e^{j\omega_2 t} \tag{3.51}$$

the displacement response of the system can be written,

$$\begin{aligned} y(t) = {} & H_1(\omega_1).X_1 e^{j\omega_1 t} + H_1(\omega_2).X_2 e^{j\omega_2 t} \\ & + H_2(\omega_1,\omega_1).X_1^2 e^{j2\omega_1 t} + H_2(\omega_2,\omega_2).X_2^2 e^{j2\omega_2 t} \\ & + 2.H_2(\omega_1,\omega_2).X_1.X_2.e^{j(\omega_1+\omega_2)t} + \text{higher order terms} \ldots \end{aligned} \tag{3.52}$$

An expression for $H_2(\omega_1,\omega_2)$ can be developed by substituting for y(t), $\dot{y}(t)$ and $\ddot{y}(t)$, and equating coefficients of $X_1.X_2.e^{j(\omega_1+\omega_2)t}$,

$$\left[(j(\omega_1+\omega_2))^2 + (j(\omega_1+\omega_2)).\alpha + \beta_1\right].H_2(\omega_1,\omega_2) + \beta_2.H_1(\omega_1).H_1(\omega_2) = 0 \tag{3.53}$$

which simplifies to,

$$H_2(\omega_1,\omega_2) = -H_1(\omega_1+\omega_2).\beta_2.H_1(\omega_1).H_1(\omega_2) \tag{3.54}$$

For the numerical example, $\ddot{y}(t) + 10.\dot{y}(t) + 10^4.y(t) + 10^7.y^2(t) + 10^{10}.y^3(t) = z(t)$ the 'principal' quadrant of the second order FRF, for which $0 < \omega_1, \omega_2$, is shown in figure 3.10.

Clearly $H_2(\omega_1,\omega_2)$ is symmetrical in the independent frequency variables ω_1 and ω_2. The axis of symmetry is the line $\omega_1 = \omega_2$, corresponding to the particular situation in which the two input harmonics have the same frequency, ie. can be considered as a single harmonic. Clearly the axis of symmetry is also the principal diagonal of the FRF.

The statement made in section 2.2.2 that the second order FRF also possesses conjugate symmetry can be demonstrated for this example using the fact that $H_1(\omega) = H_1^*(-\omega)$,

$$\begin{aligned} H_2(-\omega_1, -\omega_2) &= -\beta_2.H_1(-\omega_1-\omega_2).H_1(-\omega_1).H_1(-\omega_2) \\ &= -\beta_2.H_1^*(\omega_1+\omega_2).H_1^*(\omega_1).H_1^*(\omega_2) = H_2^*(\omega_1,\omega_2) \end{aligned} \tag{3.55}$$

and also $H_2(\omega_1,-\omega_2) = H_2^*(-\omega_1,\omega_2)$.

The axis of conjugate symmetry is defined by the line $\omega_1 + \omega_2 = 0$, which is orthogonal to the principal diagonal (the axis of symmetry), and consequently does not cross the principal quadrant.

Although the expression for the multi-dimensional second order FRF has been written only for a particular SDOF example, the technique of using two harmonics can be applied in the same way to the general SDOF system with polynomial non-linearity in both the damping and stiffness restoring force functions,

$$\begin{aligned} m.\ddot{y}(t) + c_1.\dot{y}(t) + c_2.\dot{y}^2(t) + c_3.\dot{y}^3(t) + \ldots \\ +k_1.y(t) \text{ t } k_2.y^2(t) + k_3.y^3(t) + \ldots \;=\; x(t) \end{aligned} \tag{3.56}$$

with the result,

$$H_2(\omega_1,\omega_2) = -H_1(\omega_1 + \omega_2).((j\omega_1).(j\omega_2).c_2 + k_2).H_1(\omega_1).H_1(\omega_2) \tag{3.57}$$

As mentioned in the previous section, parameters of higher orders do not affect the defined FRF at all. Also, the properties of symmetry and conjugate symmetry are possessed by the second order FRFs defined for all Volterra systems.

The general form of the second order FRF can also be interpreted from this formulation. The poles of $H_2(\omega_1,\omega_2)$ are determined by the poles of $H_1(\omega)$, in a similar way to the principal diagonal, but in the multi-dimensional case, the resonances are ridges rather than peaks. For the SDOF example, if ω_0 denotes the fundamental resonance at 15.92 Hz, three ridges can be observed in the principal quadrant of the magnitude plot along the lines $\omega_1 = \omega_0$, $\omega_2 = \omega_0$ and $\omega_1 + \omega_2 = \omega_0$ corresponding to the poles of $H_1(\omega_1), H_1(\omega_2)$ and $H_1(\omega_1 + \omega_2)$ respectively. Local peaks occur at the points in the plane only where two (or more) ridges are coincident, for example where both ω_1 and ω_2 are equal to ω_0. This point also lies along the principal diagonal and is the primary resonant peak described in the previous section.

In terms of the spectrum of the response signals, the second order FRF quantifies the interaction which can take place between pairs of harmonic components in the input. In particular, since the second order FRF exhibits a resonance at $\omega_1 + \omega_2 = \omega_0$, it is clear that if any two input harmonics are at frequencies which sum to the fundamental resonant frequency, then the resonance will be excited (provided the second order FRF exists). This is particularly important in some situations, for example when the input signal is broadband random. Clearly pairs of harmonics can always combine to excite the resonance to some degree, if $H_2(\omega_1,\omega_2)$ exists, even if the input signal is bandlimited to a frequency range which does not encompass the fundamental resonant frequency. $H_2(\omega_1,\omega_2)$ is not limited to the principal quadrant but, in theory, is defined for all frequencies and can quantify any interaction $(i_1.\omega_1 + i_2.\omega_2)$, where i_1 and i_2 are integers which can be positive or negative.

The conditions for existence of the multi-dimensional second order FRF are exactly the same as for the existence of the principal diagonal $H_2(\omega, w)$. In the general SDOF example (equation 3.57), the requirement is that either c_2 or k_2 are non-zero, ie. that the system is non-linear and the polynomials do not have odd order symmetry.

To determine expressions for higher order multi-dimensional FRFs, the general harmonic probing procedure can be extended. For $H_n(\omega_1, \ldots, w,)$, it is necessary to consider an input with at least n incommensurable harmonic terms, eg. $x(t) = \sum_{i=1}^{n} X_i.e^{j\omega_i t}$, **and after** substitution equate coefficients of $\prod_{i=1}^{n} X_i.e^{j\omega_i t}$. The number of interactions which can occur between these input components increases for the higher orders. This is **unfortunate in terms** of deriving the various expressions because it complicates the generalised harmonic probing procedure. For example, to determine the third order FRF, the expressions for the responses must include all terms which include FRFs up to third order,

$$\begin{aligned} y(t) = {} & H_1(\omega_1).X_1e^{j\omega_1 t} + H_1(\omega_2).X_2e^{j\omega_2 t} + H_1(\omega_3).X_3e^{j\omega_3 t} \\ & + H_2(\omega_1,\omega_1).X_1^2e^{j2\omega_1 t} + H_2(\omega_2,\omega_2).X_2^2e^{j2\omega_2 t} + H_2(\omega_3,\omega_3).X_3^2e^{j2\omega_1 t} \\ & + 2.H_2(\omega_1,\omega_2).X_1.X_2.e^{j(\omega_1+\omega_2)t} + 2.H_2(\omega_2,\omega_3).X_2.X_3.e^{j(\omega_2+\omega_3)t} \\ & + 2.H_2(\omega_1,\omega_3).X_1.X_3.e^{j(\omega_1+\omega_3)t} \\ & + H_3(\omega_1,\omega_1,\omega_1).X_1^3e^{j3\omega_1 t} + H_3(\omega_2,\omega_2,\omega_2).X_2^3e^{j3\omega_2 t} + H_3(\omega_3,\omega_3,\omega_3).X_3^3e^{j3\omega_1 t} \\ & + 3.H_3(\omega_1,\omega_1,\omega_2).X_1^2.X_2.e^{j(2\omega_1+\omega_2)t} + 3.H_3(\omega_1,\omega_2,\omega_2).X_1.X_2^2.e^{j(\omega_1+2\omega_2)t} \\ & + 3.H_3(\omega_2,\omega_2,\omega_3).X_2^2.X_3.e^{j(2\omega_2+\omega_3)t} + 3.H_3(\omega_2,\omega_3,\omega_3).X_2.X_3^2.e^{j(\omega_2+2\omega_3)t} \\ & + 3.H_3(\omega_1,\omega_1,\omega_3).X_1^2.X_3.e^{j(2\omega_1+\omega_3)t} + 3.H_3(\omega_1,\omega_3,\omega_3).X_1.X_3^2.e^{j(\omega_1+2\omega_3)t} \\ & + 6.H_3(\omega_1,\omega_2,\omega_3).X_1.X_2.X_3.e^{j(\omega_1+\omega_2+\omega_3)t} + \text{higher order terms} \ldots \end{aligned} \tag{3.58}$$

For the SDOF example considered previously, substituting this expression into the equation of motion (eqn. 3.50) and equating coefficients of $X_1.X_2.X_3.e^{j(\omega_1+\omega_2+\omega_3)t}$, the result which is eventually obtained can be written,

$$\begin{aligned} H_3(\omega_1,\omega_2,\omega_3) = -H_1(\omega_1+\omega_2+\omega_3).(&\beta_3.H_1(\omega_1).H_1(\omega_2).H_1(\omega_3) \\ &+\beta_2.H_1(\omega_1).H_2(\omega_2,\omega_3) \\ &+\beta_2.H_1(\omega_2).H_2(\omega_1,\omega_3) \\ &+\beta_2.H_1(\omega_3).H_2(\omega_1,\omega_2)) \end{aligned} \tag{3.59}$$

Correspondingly, for the general SDOF system (eqn. 3.56),

$$\begin{aligned} H_3(\omega_1,\omega_2,\omega_3) = -H_1(\omega_1+w2+\omega_3).(&((j\omega_1).(j\omega_2).(j\omega_3).c_3 + k_3).H_1(\omega_1).H_1(\omega_2).H_1(\omega_3) \\ &+((j\omega_1).(j(\omega_2+\omega_3)).c_2 + k_2).H_1(\omega_1).H_2(\omega_2,\omega_3) \\ &+((j\omega_2).(j(\omega_1+\omega_3)).c_2 + k_2).H_1(\omega_2).H_2(\omega_1,\omega_3) \\ &+((j\omega_3).(j(\omega_1+\omega_2)).c_2 + k_2).H_1(\omega_3).H_2(\omega_1,\omega_2)) \end{aligned} \tag{3.60}$$

The third order FRF quantifies the interaction between three input harmonics at the output frequency $(i_1.\omega_1 + i_2.\omega_2 + i_3.\omega_3)$, where i_1, i_2 and i_3 are positive or negative integers. Clearly $H_3(\omega_1, \omega_2, \omega_3)$ possesses symmetry in the independent frequency variables about the axis w = $\omega_1 = \omega_2 = \omega_3$ which is the principal diagonal, $H_3(\omega$, w, w).

Illustrating a function of three independent variables is difficult, but since $H_3(\omega_1, \omega_2, \omega_3)$ is symmetrical between ω_1, ω_2 and ω_3, a 'slice' across the plane of one of the frequencies may be considered, representing third order FRF as a function of two variables and keeping the third constant, eg. $\omega_3 = \omega_3'$. $H_3(\omega_1, \omega_2, \omega_3')$ is shown in figure 3.11 for the numerical SDOF example considered previously (eqn. 3.35), where ω_3' has been set arbitrarily to the value of approximately 2.5 Hz.

From the formulations it is clear that the ridges in the principal quadrant of the magnitude plot, and the corresponding phase changes, can be interpreted in terms of the poles of the the first and second order FRFs. If ω_0 denotes the fundamental resonance, ridges in the principal quadrant are possible at $\omega_1 = \omega_0$ and $\omega_2 = \omega_0$ from the terms $H_1(\omega_1)$ and $H_1(\omega_2)$ respectively, and at $\omega_1 + \omega_2 = \omega_0 \pm \omega_3'$ from $H_1(\omega_1 + \omega_2 + \omega_3')$. In addition, the second order FRF terms give the possibility of ridges at $\omega_1 = \omega_0 \pm \omega_3', \omega_2 = \omega_0 \pm \omega_3'$ and $\omega_1 + \omega_2 = \omega_0$.

To this stage, all the examples of multi-dimensional FRFs have been for SDOF systems. However, the extended harmonic probing technique can be applied also to the MDOF differential equations of motion. Rather than performing the substitutions and simplifying the complicated expressions which result, it is possible to write down the multi-dimensional FRFs directly from the diagonal formulation using the property of symmetry.

For example, since $H_2(\omega_1, \omega_2)$ is symmetrical in ω_1 and ω_2, expressions for the second order FRF of a MDOF system can be written in matrix form by generalising the T-notation,

$$\{H_2(\omega_1, \omega_2)\} = -[H_1(\omega_1 + \omega_2))] . \{T_{11}\} \qquad \textbf{(3.61)}$$

where,

$$_{ss}T_{11} = ((j\omega_1).(j\omega_2)._{ss}c_2 + {}_{ss}k_2)._{rs}H_1(\omega_1)._{rs}H_1(\omega_2) \qquad \textbf{(3.62)}$$

$$_{sp}T_{11} = ((j\omega_1).(j\omega_2)._{sp}c_2 + {}_{sp}k_2).(_{rp}H_1(\omega_1) - {}_{rs}H_1(\omega_1)).(_{rp}H_1(\omega_2) - {}_{rs}H_1(\omega_2)) \qquad \textbf{(3.63)}$$

Figure 3.12 shows the principal quadrant of $H_2(\omega_1, \omega_2)$ for the direct point of the 3DOF example considered in the previous section (equation 3.46). As can be seen from the magnitude plot, the poles of the $H_2(\omega_1, \omega_2)$ become ridges along the lines $\omega_1 = \omega_0, \omega_2 = \omega_0$ and $\omega_1 + \omega_2 = \omega_0$, where ω_0 now denotes either of the three fundamental resonant frequencies of the system. In total, nine ridges can be observed in the selected frequency range, $0 < \omega_1, \omega_2 < 37.5$ Hz, and at the points where two or more ridges coincide, local peaks occur.

The multi-dimensional third order FRF can be also written down directly from the diagonal formulation,

$$\{H_3(\omega_1,\omega_2,\omega_3)\} = -[H_1(\omega_1+\omega_2+\omega_3)].\{2.\{T_{12}\}+\{T'_{111}\}\} \tag{3.64}$$

where,

$$\begin{aligned}
{}_{ss}T_{111} &= \frac{1}{3}((j\omega_1).(j(\omega_2+\omega_3)).{}_{ss}c_2+{}_{ss}k_2).{}_{rs}H_1(\omega_1).{}_{rs}H_2(\omega_2,\omega_3)\\
&+\frac{1}{3}((j\omega_2).(j(\omega_1+\omega_3)).{}_{ss}c_2+{}_{ss}k_2).{}_{rs}H_1(\omega_2).{}_{rs}H_2(\omega_1,\omega_3)\\
&+\frac{1}{3}((j\omega_3).(j(\omega_1+\omega_2)).{}_{ss}c_2+{}_{ss}k_2).{}_{rs}H_1(\omega_3).{}_{rs}H_2(\omega_1,\omega_2)
\end{aligned} \tag{3.65}$$

$$\begin{aligned}
{}_{sp}T_{111} &= \frac{1}{3}((j\omega_1).(j(\omega_2+\omega_3)).{}_{sp}c_2+{}_{sp}k_2).({}_{rp}H_1(\omega_1)-{}_{rs}H_1(\omega_1))\\
&\qquad .({}_{rp}H_2(\omega_2,\omega_3)-{}_{rs}H_2(\omega_2,\omega_3))\\
&+\frac{1}{3}((j\omega_2).(j(\omega_1+\omega_3)).{}_{sp}c_2+{}_{sp}k_2).({}_{rp}H_1(\omega_2)-{}_{rs}H_1(\omega_2))\\
&\qquad .({}_{rp}H_2(\omega_1,\omega_3)-{}_{rs}H_2(\omega_1,\omega_3))\\
&+\frac{1}{3}((j\omega_3).(j(\omega_1+\omega_2)).{}_{sp}c_2+{}_{sp}k_2).({}_{rp}H_1(\omega_3)-{}_{rs}H_1(\omega_3))\\
&\qquad .({}_{rp}H_2(\omega_1,\omega_2)-{}_{rs}H_2(\omega_1,\omega_2))
\end{aligned} \tag{3.66}$$

$${}_{ss}T_{12} = ((j\omega_1).(j\omega_2).(j\omega_3).{}_{ss}c_3\ {}_{t}\ {}_{ss}k_3).{}_{rs}H_1(\omega_1).{}_{rs}H_1(\omega_2).{}_{rs}H_1(\omega_3) \tag{3.67}$$

$$\begin{aligned}
{}_{sp}T_{12} &= ((j\omega_1).(j\omega_2).(j\omega_3).{}_{sp}c_3+{}_{sp}k_3).({}_{rp}H_1(\omega_1)-{}_{rs}H_1(\omega_1))\\
&\qquad .({}_{rp}H_1(\omega_2)-{}_{rs}H_1(\omega_2)).({}_{rp}H_1(\omega_3)-{}_{rs}H_1(\omega_3))
\end{aligned} \tag{3.68}$$

The plot of $H_3(\omega_1,\omega_2,\omega_3)$, where ω_3 is once again maintained at the arbitrary frequency 2.5 Hz, is shown in figure 3.13. Interpretation is made particularly difficult by the number of ridges which arise and the interactions between the ridges which occur.

The expressions for the diagonals of the higher order FRFs can be formulated using the partition pattern established at the end of section 3.2. However the multi-dimensional FRFs become increasingly complicated when the order is extended beyond three, and since it is very difficult to represent visually a function of three or more independent variables, listing expressions above third order is not worthwhile.

3.4 Features of the Frequency Response Functions

The n^{th} order FRF quantifies the possible interactions between n harmonic components of the input signal at frequencies $\omega_1, \omega_2, \ldots, \omega_n$, which may generate components in the output signal at frequencies $(\pm\omega_1 \pm \omega_2 \ldots \pm \omega_n)$. In a non-linear system governed by a general polynomial, FRFs of all orders would usually exist, and the possible interactions between n harmonic components can generate output components at $(i_1.\omega_1 + i_2.\omega_2 \ldots i_n.\omega_n)$, where $i_1, i_2, \ldots, i_n$ can each take any integer value positive, negative or zero.

The response of a non-linear system to a broadband random signal can contain interaction components at all frequencies, with the possibility of exciting all of the resonances to some degree. For example, if the signal is bandlimited to a frequency range ω_l to ω_u, and a fundamental system resonance occurs at a frequency below the lower limit ω_l, the non-linear interactions can excite this resonance nevertheless. This explains the low-frequency drift oscillations of moored vessels which are subject to random excitation by sea-waves [51]. The energy of the waves is concentrated to a bandwidth considerably higher in frequency than the drift resonance. This specific example has been studied extensively using the polyspectral techniques outlined in chapter 2. In this application, of particular interest is the fact that by extending the conventional coherence function to include the second order term, it is possible to demonstrate that the mechanism for energy transferral is dominated by the 'linear' and 'quadratic' interactions. Terms in the Volterra series of order higher than second are neglected, and bispectral techniques are used for the analysis. The conclusion which is drawn is that the governing equations are non-linear with a strong quadratic term. Approximating the Volterra series representation to the first and second order terms is justified in this situation from an energy approach, since extensions to third order being prohibited by the complexity of the analysis. However for non-linear systems which are governed by polynomials with cubic terms, it is necessary to consider third order terms also.

Clearly, the Volterra series must always be limited to some finite order, and restricting the order of the series has important implications regarding the processing of the signals prior to analysis, filtering in particular. This subject is addressed in chapter 5 in the context of implementation of the identification techniques. Here it is appropriate to point out that, if a system is excited by a baseband signal with a low-pass frequency set at frequency ω_u, and if the Volterra series is to be approximated by first, second and third order terms, then the highest frequency component in the output signal which can be represented is at $3\omega_u$. Hence, the response signals should not be filtered within the range 0 to $3\omega_u$ if terms up to third order are to be identified from the data.

Interpreting the energy transfer mechanism using higher order FRFs can be used to understand the differences between a linearised TF, measured from a non-linear system using broadband random excitation, and the first order FRF. Each of the higher order FRFs can contribute to the response at a particular spectral line. In forming the cross-spectrum, $\overline{Y(\omega).X^*(\omega)}$, the conjugate of the input component at frequency w is multiplied by the output component at the same frequency, which will be the sum of contributions from each of the higher order FRFs. Clearly, if the system is non-linear, and higher order FRFs exist, then the cross-spectrum of the non-linear system will be different from the cross-spectrum $\overline{Y_1(\omega).X(\omega)}$ of the related linear system which is characterised by the first order FRF.

Another result from chapter 2 can be re-examined now that the relationships for the higher order FRFs have been established. In section 2.2.2, the property of homogeneity possessed by the Volterra kernels made it possible to express each of the higher order operators $\mathbf{H_n}[.]$ as a combination of first order operators $\mathbf{H_1}[.]$. To illustrate this, the SDOF system of equation 2.21 was considered as an example, and the first three terms in the Volterra series were written (equations 2.26 to 2.28),

$$\begin{aligned} y_1(t) &= \mathbf{H_1}[x(t)] \\ y_2(t) &= \mathbf{H_2}[x(t)] = -\mathbf{H_1}[\beta_2.y_1(t)^2] \\ y_3(t) &= \mathbf{H_3}[x(t)] = -\mathbf{H_1}[2.\beta_2.y_1(t).y_2(t) + \beta_3.y_1(t)^3] \end{aligned} \tag{3.69}$$

The principal diagonals of the second and third order FRFs have been written down for this system in section 3.2, (equations 3.32 and 3.33) and are repeated here,

$$\begin{aligned} H_2(\omega,\omega) &= -H_1(2\omega).\beta_2.H_1(\omega)^2 \\ H_3(\omega,\omega,\omega) &= -H_1(3\omega).(2.\beta_2.H_1(\omega).H_2(\omega,\omega) + \beta_3.H_1(\omega)^3) \end{aligned} \tag{3.70}$$

Clearly there is a direct correspondence between equations 3.69 and 3.70 which were both determined by substitution followed by equating coefficients of a particular order. Consequently it is possible to use the partition pattern established for $H_n(\omega,\ldots,\omega)$ at the end of section 3.2 to determine an expression for y,(t). For example, the sixth order diagonal for this system can be written (eqn. 3.49),

$$\begin{aligned} H_6(\omega,\ldots,\omega) = \\ -H_1(6\omega).(2.H_1(\omega).H_5(\omega,\ldots,\omega) + 2.H_2(\omega,\omega).H_4(\omega,\ldots,\omega) + 3.H_1(\omega)^2.H_4(\omega,\ldots,\omega) + \\ H_3(\omega)^2 + 6.H_1(\omega).H_2(\omega).H_3(\omega,\omega,\omega) + H_2(\omega,\omega)^3) \end{aligned}$$

and consequently,

$$\begin{aligned} y_6(t) = -\mathbf{H_1}.[2.y_1(t).y_5(t) + 2.y_2(t).y_4(t) + 3.y_1(t)^2.y_4(t) + \\ y_3(t)^2 + 6.y_1(t).y_2(t).y_3(t) + y_2(t)^3] \end{aligned} \tag{3.71}$$

Chapter 4

Higher Order Transfer Functions

4.1 Introduction

One of the principal objectives of this project has been to develop a practical method, which employs the Volterra series and makes use of its properties for characterising non-linear systems, to be applied to the identification of physical structures. The Volterra kernels cannot be measured directly from a test on a structure because, for any physical excitation waveform, the contributions of each kernel interact in the responses making it impossible to distinguish between the different components. In this project, the aim has been to determine the Volterra kernels indirectly, and in particular to find a means of measuring an estimate of the higher order FRFs, the Volterra kernel transforms.

In chapter 2, the Wiener series was introduced as a feasible method of estimating the FRFs from measurements. However, numerous problems associated with the implementation arise which render impractical extensions to third order and above. This motivated the need to look for an alternative approach for application to the identification of non-linear structures.

Usually the type of excitation can be selected for a structural test and one approach, which is described in this chapter, concentrated on the use of sinusoidal excitation to simplify the estimation of the higher order FRFs from measurements. Importantly these simplifications make possible the investigation of a wide class of weakly non-linear structures without needing to restrict the form or order of the non-linearity which can be considered, although when attempting to interpret the results in terms of the non-linear equations of motion presented in section 1.2.1, the restoring force functions are assumed to be continuous polynomials.

4.2 Defining the Transfer Functions

As shown in section 3.2, expressions for the principal diagonals of the higher order FRFs can be extracted from the equations of motion analytically using a single harmonic input. The aim of this section is to show that a practical technique can be conceived along similar lines, to estimate the diagonals of the FRFs from measurements using *sinusoidal* excitation.

To comply with earlier notation, the measured estimates of the FRFs are termed Transfer Functions (TFs). In the context of non-linear systems, the distinction is important. There are often marked differences between the FRFs defined from the Volterra series and the TFs measured using sine excitation, an observation which is also addressed in this section. The TFs which are defined here rely on the input being a sinewave, and consequently they can only find an application to situations in which the excitation can be generated, and any external disturbances, which would corrupt particularly the excitation signal, are assumed to be minimal,

In chapter 2, one observation made was that the single harmonic, $X.e^{j\omega t}$, is particularly useful for analysis, but cannot be generated in practice. The closest physical signal is a single sinusoid which.is simply a combination of two harmonics,

$$x(t) = X.\cos(\omega t) = \frac{X}{2}.e^{j\omega t} + \frac{X}{2}.e^{-j\omega t} \tag{4.1}$$

The action of any general polynomial-type non-linearity present is to cause these two harmonic terms to interfere and give rise to combination components in the output signals at all integer multiples of ω, as shown in section 2.2.4.

Higher order FRFs are used to quantify the energy transfer which takes place from the input harmonics to the fundamental and higher order harmonic frequencies. The harmonic components in the spectra of the input and output signals can be determined simply by performing single-dimensional Fourier transformations on sections of time data. The ability to determine an estimate of at least part of the higher order FRFs whilst avoiding using multi-dimensional Fourier transforms is one of the main advantages of this technique.

If the fundamental harmonic component in the spectrum of the input at the frequency of the sinewave w is denoted X(w), and the n^{th} order harmonic component in the spectrum of the output is denoted $Y(n\omega)$, it is possible to write down the following expressions directly from equation 2.51,

$$\frac{Y(\omega)}{X(\omega)} = H_1(\omega) + \frac{3X^2}{4}.H_3(\omega,\omega,-\omega) + \ldots + n_1.\frac{X^{n_1-1}}{2^{n_1-1}}.H_{n_1}(\omega,\ldots,-\omega) + \ldots \quad (4.2)$$

$$\frac{Y(2\omega)}{X(\omega)^2} = \frac{1}{2}.H_2(\omega,\omega) + \frac{3X^2}{4}.H_4(\omega,\omega,\omega,-\omega) + \ldots$$

$$+ \frac{n_0!}{2!(n_0-2)!}\frac{X^{n_0-2}}{2^{n_0-1}}.H_{n_0}(\omega,\ldots,-w) + \ldots \quad (4.3)$$

$$\frac{Y(3\omega)}{X(\omega)^3} = \frac{1}{4}.H_3(\omega,\omega,\omega) + \frac{5X^2}{8}.H_5(\omega,\omega,\omega,\omega,-\omega) + \ldots$$

$$+ \frac{n_1!}{3!(n_1-3)!}\frac{X^{n_1-3}}{2^{n_1-1}}.H_{n_1}(\omega,\ldots,-\omega) + \ldots \quad (4.4)$$

$$\frac{Y(n\omega)}{X(\omega)^n} = \frac{1}{2^{n-1}}.H_n(\omega,\ldots,\omega) + \frac{(n+2)!}{n!}.\frac{X^2}{2^{n+2}}.H_{n+2}(\omega,\ldots,-\omega) + \ldots \quad (4.5)$$

(n_0 denotes an even integer, and n_1 an odd integer.)

(Factors such as $\frac{1}{2^{n-1}}$ arise because of the convention which has been adopted. Since X denotes the amplitude of the input sinewave, the fundamental input harmonic component must be written X(w) = $\frac{X.e^{j\omega t}}{2}$.)

The first term on the right-hand side of each expression is the principal diagonal of an FRF. From the point of view of using the ratios on the left-hand side to estimate the diagonals of the FRFs, the estimates can be regarded as being corrupted by the other 'residual' terms on the right-hand side, which are subsequently referred to as the *degenerative* contributions from the higher order FRFs. Degenerative terms arise because of the interaction between the two harmonic components of the input sinewave. For example, if only one of these harmonics could be isolated and applied to the structure, degenerative terms would not appear. Unfortunately this cannot be realised in practice. All physical inputs when applied to a non-linear system will cause some interaction between two or more input harmonic components. The single sinewave can be thought of as the optimum physical signal to minimise the number of interactions possible.

On this basis, it is possible to define a series of higher order Transfer Functions,

$$\mathrm{TF}_n(\omega) = 2^{n-1}.\frac{\mathcal{F}[y(t)](n\omega)}{\mathcal{F}[x(t)](\omega)^n} \quad (4.6)$$

$$H_n(\omega,\ldots,\omega) + \frac{(n+2)!}{\ldots}.\frac{X^2}{2^3}.H_{n+2}(\omega,\ldots,-\omega) + \ldots \quad (4.7)$$

to be measured from a structure in practice.

In general, few problems are faced in generating a sinewave, using the signal to excite a structure, and performing single dimensional Fourier transforms on the measured signals. However, it is important to bear in mind that the higher order TFs defined by equation 4.6 only strictly apply when the input is a pure sinewave. In practice, the excitation signal applied to a structure will always be polluted to some degree by the interaction of a structure with the shaker, particularly when the structure is lightly damped. This is the main practical problem encountered with this approach, and is addressed again in section 6.4.3 where details of the experimental procedure for measured higher order TFs are provided.

From equation 4.7, the approximation provided by the n^{th} order TF to the principal diagonal of the n^{th} order FRF can be seen to improve as the amplitude X is reduced, and the degenerative terms have less influence. Remembering that these relationships only apply for polynomial type non-linearities, by definition the two functions are the same in the limit,

$$H_n(\omega, \ldots, \omega) = \mathrm{TF}_n(\omega)|_{X \to 0} \tag{4.8}$$

The higher order TFs approach the principal diagonals of the FRFs as the amplitude of the input sinewave approaches zero. (This corresponds with the approximation of the Wiener kernel transforms to the multi-dimensional FRFs which in theory would also improve as the level of the Gaussian input were reduced.)

From this viewpoint, the correspondence between the two functions in the limit may appear intuitive. However it is usual to regard a system with polynomial non-linearities as approaching linearity when the input level is reduced, and it is possible to see from the output spectra measured from systems of this type that the amplitude of the higher harmonic terms decrease as the input level is decreased (for example, figure 1.2). One might suppose that, with the exception of $\mathrm{TF}_1(\omega)$, the higher order TFs would approach zero in the limit rather than the diagonals of the higher order FRFs.

One way of understanding the limiting condition is by realising that all of the higher order TFs including the first are in the form of a simple polynomial function of the type,

$$f(x) = \alpha_0 + \alpha_2.x^2 + \ldots$$

The parameters $\alpha_0, \alpha_2, \ldots$ are properties of the function f(x) in the same way that the higher order FRFs are properties of the system, and are independent of the amplitude variable x. Clearly as x $\rightarrow$ 0, f(x) $\rightarrow \alpha_0$.

In terms of the spectral components, as the amplitude of the sinusoidal input is reduced, the output harmonic Y (nw) $\rightarrow$ 0 at the same rate as $X(\omega)^n \rightarrow 0$.

Consequently the ratio $\dfrac{Y(n\omega)}{X(\omega)^n}$ would not approach zero if the n^{th} order FRF exists.

In equation 4.6, the TFs were defined in terms of the displacement time signal y(t), and consequently can be regarded as higher order receptance functions. Correspondingly, higher order mobility functions $\mathrm{TF}'_n(\omega)$, and inertance functions $\mathrm{TF}''_n(\omega)$, can be defined respectively in terms of the velocity and acceleration response signals,

$$\mathrm{TF}'_n(\omega) = 2^{n-1}.\frac{\mathcal{F}[\dot{y}(t)](n\omega)}{\mathcal{F}[x(t)](\omega)} \quad ; \quad \mathrm{TF}''_n(\omega) = 2^{n-1}.\frac{\mathcal{F}[\ddot{y}(t)](n\omega)}{\mathcal{F}[x(t)](\omega)} \tag{4.9}$$

$$\mathrm{TF}''_n(\omega) = (nj\omega).\mathrm{TF}'_n(\omega) = (nj\omega)^2.\mathrm{TF}_n(\omega) \tag{4.10}$$

In practice, the minimum amplitude of the sinewave is determined primarily by the levels of noise on the measured signals. For weakly non-linear structures, the higher harmonic terms in the spectra of the response signals are usually very small compared with the fundamental term, particularly away from resonance. Although instrumentation is available to resolve very small harmonic components, the noise content always present on the measured signals remains a considerable problem.

As already mentioned, a good quality sinewave is not particularly difficult to generate in practice (even if it may become slightly distorted when applied via a shaker to a structure.) This is one of the principal advantages of the higher order TF approach over the use of Wiener kernels which requires a random excitation signal to be generated which is similar to the ideal Gaussian signal. Problems of shaker-structure interaction are encountered for all types of excitation waveform (except impulse), and for the random excitation case would cause further deviation of the generated signal from Gaussian. A further advantage of the higher order TF method is that the single dimensional Fourier transform can be easily implemented and is fast, particularly when compared to performing higher order correlations. Also, the need for windowing the time data before computing the spectra can be avoided by calculating the Fourier transforms over whole time periods of the fundamental sinewaves. Since the signals are periodic and deterministic, the TFs are formulated directly as ratios of spectral components, as opposed to expected values of an ensemble average which is appropriate for indeterministic signals. Averaging can be employed to reduce the effect of uncorrelated noise on each of the measured signals, hence improving the quality of the TFs. In addition, there is no need to measure and store vast amounts of time data, which would require time consuming post-processing before determining whether or not the data were of sufficient quality. Higher order TFs can be computed on-line and visualised making it easier to check the quality of the results during the stepped-sine test. (Several further practical advantages of the higher order TF approach are discussed in chapter 6, where details of implementation are described in conjunction with results from a physical structure.)

4.3 Relating TFs to FRFs

Transfer functions measured from physical systems using sine excitation frequently exhibit distortion when compared to the principal diagonal of the corresponding FRF. This distortion may be considerable around the resonances and usually becomes more significant at higher amplitudes. For levels of excitation small enough for the Volterra series to converge, the distortion can accounted for by the degenerative terms which in theory always occur for non-linear Volterra systems. This is the subject of this section.

As can be seen from equation 4.7, the degenerative contributions to the n^{th} order TF arise only from FRFs with order higher than n and of the same type, ie. odd or even. Again this is similar to the relationship between the Volterra and Wiener kernels (equations 2.76 to 2.79) noted at the end of section 2.3.3.

Of particular interest is the first order TF, the ratio of output to input spectral components at the fundamental frequency, since this is the function which is usually measured in a conventional stepped sine test. It is known that $TF_1(\omega)$ is affected by degenerative effects from each of the odd order FRFs, and most significantly by $H_3(\omega, \mathrm{w}, -\omega)$[17].

Expressions for the degenerative contributions can be written down directly from the multi-dimensional FRF relations derived in section 3.3. For example, consider the SDOF system used throughout chapter 3,

$$\ddot{y}(t) + \alpha.\dot{y}(t) + \beta_1.y(t) + \beta_2.y^2(t) + \beta_3.y^3(t) = x(t)$$

Using equation 3.59,

$$\begin{aligned} H_3(\omega_1,\omega_2,\omega_3) = & -\beta_3.H_1(\omega_1+\omega_2+\omega_3).H_1(\omega_1).H_1(\omega_2).H_1(\omega_3) \\ & -\beta_2.\frac{2}{3}.H_1(\omega_1+\omega_2+\omega_3).[H_1(\omega_1).H_2(\omega_2,\omega_3)+ \\ & \qquad H_1(\omega_2).H_2(\omega_1,\omega_3)+ \\ & \qquad H_1(\omega_3).H_2(\omega_1,\omega_2)] \end{aligned}$$

This expression can be used to describe any point in the multi-dimensional third order FRF, including $(\omega, \mathrm{w}, -\omega)$ which is along an axis orthogonal to the principal diagonal. Thus,

$$\begin{aligned} H_3(\omega,\omega,-\omega) = & -\beta_3.H_1(\omega+\omega-\omega).H_1(\omega).H_1(\omega).H_1(-\omega) \\ & -\beta_2.\frac{2}{3}.H_1(\omega+\omega-\omega).[H_1(\omega).H_2(\omega,-\omega) + H_1(\omega).H_2(\omega,-\omega) \\ & \qquad + H_1(-\omega).H_2(\omega,\omega)] \\ = & -\beta_3.H_1^*(\omega).H_1(\omega)^3 - \beta_2.\frac{2}{3}.H_1(\omega).[2.H_1(\omega).H_2(\omega,-\mathrm{w}) + H_1^*(\omega).H_2(\omega,\omega)] \end{aligned} \tag{4.11}$$

Hence, to a second approximation,

$$\begin{aligned} \mathrm{TF}_1(\omega) &\approx H_1(\omega) - \frac{3X^2}{4}H_3(\omega,\omega,-\omega) \\ &= H_1(\omega) - \beta_3.\frac{3X^2}{4}.H_1^*(\omega).H_1(\omega)^3 - \beta_2.X^2.H_1(\omega)^2.H_2(\omega, -\omega) \\ &\quad - \beta_2.\frac{X^2}{2}.H_1^*(\omega).H_1(\omega).H_2(\omega, \omega) \end{aligned} \tag{4.12}$$

In general, distortion will be exhibited by all the higher order TFs in addition to the first order function, because each is affected by degenerative contributions from higher order FRFs as described by equation 4.7.

To illustrate the distortion and the effect of increasing the input amplitude on the measured TFs, a numerical simulation routine was written to perform the stepped sine test emulating the procedure used in the testing of structures. The simulation model uses the lumped parameter equations of motion introduced in section 1.2.1, and the integration was performed using an adaptive Bulirsch-Stoer routine with Richardson extrapolation [34].
(This procedure was found to improve significantly on execution times over the adaptive fourth order Runge-Kutta routine, which is often used in applications of this kind [60].)

Importantly the practical problems of introducing transients when the conditions of the input are changed, ie. after each increment in frequency, arise in simulated systems in the same way as in a stepped sine test on a structure. Since the TFs should be calculated from the signals only after the system has attained steady-state, it is important to allow the transients to die away before determining the TFs. Applying a sinewave waveform at one 'location' in the numerical model, and then allowing transients to die away, the outputs $\ddot{y}(t), \dot{y}(t)$ and y(t) at each of the locations were computed. Employing a conventional DFT routine [34], the input and output spectra were used to calculate each of the complex valued higher order transfer functions at w_i, before the frequency of the input sinewave was incremented and the procedure repeated. Whereas in a structural test, measurement noise always presents a problem, the numerical noise caused by round-off errors which affects the simulated data is very small by comparison, and can be attenuated further by averaging the TFs over consecutive periods. In addition, the simulated signal-to-noise ratio does not necessarily decrease for smaller input amplitudes.

Using simulation, it is possible to estimate the FRF diagonals accurately using low level sinusoidal inputs with constant amplitudes, and to investigate how the presence of higher order degenerative effects give rise to distortion in the TFs of each order as the input amplitude is increased. To illustrate the dependence of the distortion on the amplitude, simulations were

performed using both the numerical SDOF and MDOF systems considered in the previous chapter (and shown in figures 3.1 and 3.8), and TFs up to order three were determined for a range of input amplitudes.

A stepped sine test on the numerical SDOF system described by the equation,

$$\ddot{y}(t) + 10.\dot{y}(t) + 10^4.y(t) + 10^7.y^2(t) + 10^{10}.y^3(t) = \text{x. } \cos(\omega t) \tag{4.13}$$

was simulated for the amplitudes X = 0.1, 0.3, 0.5, 0.7, 0.9 by applying a sinewave which was stepped-up in frequency each increment. The distortion which occurs as the amplitude is increased skews the effective resonances toward higher frequency, as shown in figure 4.1. This indicates that the stiffness is 'hardening', an effect which is consistent with the governing polynomial having an appreciable (positive) cubic term. In addition, at large input amplitudes, the system ceases to be weakly non-linear, as indicated by the bifurcation which occurs for input amplitudes exceeding approximately 0.6. Conversely, if instead the coefficient of the cubic term were negative, the non-linear stiffness would become 'softening' and the apparent resonances would skew toward lower frequency. As described in section 6.4.2, an apparent softening effect can occur when the quadratic term in the polynomial is dominant, whether it be positive or negative, even if there is also a large positive cubic term in the function tending to cause hardening. This effect has been observed in the pre-loaded flexible beam described in chapter 6.

When compared to the ideal FRF, which was generated previously for this system and is shown in figure 3.2, there is also a slight distortion in each of the TFs, occurring at $\frac{\omega_0}{2}$ on $\mathrm{TF}_1(\omega)$, at $\frac{\omega_0}{3}$ on $\mathrm{TF}_2(\omega)$, and at $\frac{\omega_0}{4}$ on $\mathrm{TF}_3(\omega)$, where ω_0 denotes the fundamental resonance at 15.92 Hz. This effect can be attributed to the degenerative terms which have additional poles corresponding to these frequencies.

The MDOF system described by equation 3.44 was also simulated, and the TFs for the location of the direct point are shown in figure 4.2 for amplitudes of 0.5, 1.0, 1.5, 2.0, 2.5, again stepping-up in frequency. The distortions occur predominately around the first and second resonances since at these frequencies the displacement response of the system approach maxima, and the non-linearity in the stiffness function has greatest effect. In terms of the degenerative effects, the influence on the TFs is most significant at the frequencies which correspond to the poles of the degenerative terms. Clearly this occurs at each of the fundamental resonances. Again the direction of the distortions indicates that the stiffness effectively hardens for higher amplitudes, with bifurcation occurring when the input amplitude exceeds approximately 1.6.

In section 3.2, it was observed that the third order FRF would usually exist for polynomial type non-linearities. In the SDOF example described, it is clear that there will be a degenerative contribution to $TF_1(\omega)$ from the third order FRF even if either β_2 or β_3 were non-zero, as mentioned in section 3.2. FRFs of order higher than three would also exist and give rise to additional degenerative terms which contribute to the distortion. The influence of these degenerative terms will increase with amplitude, and it is possible that the point of bifurcation corresponds to the amplitude at which degenerative terms become so large that the Volterra series diverges, although this has not been proved. Certainly the Volterra series cannot be used to characterise bifurcation.

Because in theory the distortion on TFs is some function of the amplitude of the input sinewave (eqn. 4.7), an observation which has been supported in practice by performing stepped sine tests on non-linear structures, it is evident that one method of reducing the distortion would be to decrease the amplitude of the input signal as the system approaches a resonance. Moreover, since the extent to which the non-linearity affects the dynamics of a system is a function of the level of the output, one possible scheme would be to adjust the amplitude of the excitation in order to maintain the response at nominally the same level. Clearly this would entail either profiling the excitation amplitude to dip near the resonances, or alternatively including some form of control loop, whereby at each frequency increment the response were measured, compared to the datum level, and then the amplitude of the generated input signal adjusted accordingly, before the TF measurements taken. This would greatly increase the time taken for a stepped sine test (which is time consuming anyway), and for a MDOF structure would entail either selecting one reference response location, or using the signals from all locations to determine a nominal average response level to be maintained.

In a lightly damped structure, shaker-structure interaction would introduce further complications in the regulation of the excitation by controlling the signal generation. More importantly, however, the excitation would need to be reduced to a very small level near the resonances in order to maintain the response at the same level as near the anti-resonances, where conversely the excitation would need to be relatively high. Because of the practical difficulties which would be encountered in attempting to control the excitation level to reduce distortion in the TFs, the method for measuring higher order TFs implemented in this project and described in section 6.4.3 used a constant amplitude sinewave to provide the excitation, although modifications to this procedure are suggested as an area for future development.

4.4 Hilbert Transforms

Since in theory higher order FRFs will always exist for non-linear systems governed by polynomial-type non-linear functions, the first order TF measured using sine excitation would usually exhibit distortion. Using the explanation of the distortion provided by the degenerative terms as established in section 4.3, it may be possible to reveal a new way of understanding traditional methods of detecting non-linearity from the first order TF, in particular the use of the Hilbert transform [61].

Single-dimensional FRFs can be used to characterise all structures which are linear. Furthermore, these FRFs can be always written in the form,

$$H(\omega) = \frac{\prod_{i=1}^{Q}(\omega - q_i)}{\prod_{i=1}^{P}(\omega - p_i)} \tag{4.14}$$

in which the zeros, q_i, and the poles, p_i, are usually complex.
The inverse Fourier transform of the FRF of a linear system is the defined impulse response function h(t), ie. the response to a Dirac impulse,

$$h(t) = \mathcal{F}^{-1}[H(\omega)] = \frac{1}{2\pi}. \int_{-\infty}^{\infty} H(\omega).e^{j\omega t}.d\omega \tag{4.15}$$

Since all physical systems are causal, the impulse response function must satisfy the condition,

$$h(t) = 0 \text{ for } t < 0 \tag{4.16}$$

It is possible to show that, for the impulse response function determined from a measured transfer function using the Fourier transformation, the condition for causality corresponds to all the poles of the transfer function having an imaginary part which is positive [62].
If at least one pole has a negative imaginary part then,

$$h(t) \neq 0 \quad \text{for } t < 0 \tag{4.17}$$

The property of causality can be used to prescribe a relatively simple test to determine whether or not a system can be characterised by a linear FRF [63]. Any function h(t) can always be broken down into the sum of an even function h,(t) plus an odd function h,(t),

$$h(t) = h_e(t) \; t \; h_o(t) \tag{4.18}$$

For a causal function,

$$h_o(t) = (\text{sgn } t)\ h_e(t)\ , \text{ where sgn } t = \begin{cases} 1\ ;\ t > 0 \\ -1\ ;\ t < 0 \end{cases} \tag{4.19}$$

Hence,

$$\begin{aligned} \mathcal{F}[h_o(t)] &= \text{Re}\{H(\omega)\} = \mathcal{F}[(\text{sgn } t)\ h_e(t)] \\ \mathcal{F}[h_e(t)] &= \text{Im}\{H(\omega)\} = \mathcal{F}[(\text{sgn } t)\ h_o(t)] \end{aligned} \tag{4.20}$$

Consequently, the real part of H(w) can be determined from the imaginary part and vice-versa using the Fourier transform and the sgn-function.

Frequently the relationship which exists between the real and imaginary parts is written [64],

$$\begin{aligned} \mathcal{H}[\text{Re}\{H(\omega)\}] &= \text{Im}\{H(\omega)\} \\ \mathcal{H}[\text{Im}\{H(\omega)\}] &= \text{Re}\{H(\omega)\} \end{aligned} \tag{4.21}$$

where $\mathcal{H}[\ ]$ represents the Hilbert transform, which can be written formally as,

$$\mathcal{H}[H(\omega_0)] = -\frac{1}{j\pi}\int_{-\infty}^{\infty} \frac{H(\omega)}{\omega - \omega_0}.d\omega \tag{4.22}$$

for any individual spectral line $\omega = \omega_0$.

The application of the Hilbert transforms for detecting non-linearity on first order TFs measured using stepped sine excitation is well established [65], and details of the technique are not particularly relevant to this thesis. However, by using the insight that the TFs are corrupted by degenerative terms from the higher order FRFs, it may be possible to understand in greater depth the differences which are observed when the Hilbert transformed function is compared with the measured TF. In figure 4.3, the measured function and its Hilbert transform are shown for the simulated SDOF system described by equation 4.13. Input amplitudes of 0.1, 0.5 and 0.9 were used to demonstrate the Hilbert transform is very sensitive to level of excitation, particularly if bifurcation occurs.

The first order function quantifies the energy which is transferred from the input sinewave to the sinewave component in the output at the same frequency. It can only be used to characterise the system if there is no transferal of energy to other frequencies, in which case the system will be linear and $TF_1(\omega)$ corresponds to $H_1(\omega)$. If the system is non-linear there will be energy transferal and $TF_1(\omega)$ does not characterise the system. The presence of the degenerative terms means that the first order TF of a non-linear system cannot be written in the form of a linear FRF in which all the poles are simple with positive imaginary parts.

As is apparent from equation 4.12, the second approximation to $\mathrm{TF}_1(\omega)$ includes the the degenerative term of, the third order FRF, which incorporates the component $H_1(-\omega)$ = *H;"(w),* the complex conjugate of the first order FRF. Evidently, for causal $h_1(t)$, the presence of $H_1^*(\omega)$ will always generate poles in $\mathrm{TF}_1(\omega)$ which have a negative imaginary part. This is the reason why the Hilbert transform can be used to detect non-linearity from the first order TF. Effectively it is the degenerative terms which are being detected since, via terms which include $H_1^*(\omega)$, they give rise to poles the first order TF which cause the distinction from the first order FRF.

The term $H_1^*(\omega)$ arises also in the particular cases when either $\beta_2 = 0$ or $\beta_3 = 0$, and would also occur if the non-linearity were in the damping function rather than the stiffness. In fact, since all FRFs can be expressed as sums and products of $H_1(\omega)$, it is evident that all degenerative terms must include the component *H,*(w).* For this reason, the Hilbert transform can be used to detect the presence of non-linearity whenever the system can be characterised by a Volterra series. Clearly in this application the strength of the Hilbert transform technique in detecting non-linearity lies in the fact that it is sufficiently sensitive to be able to identify the degenerative terms even when the distortion on the first order TF is not particularly obyious.

By the same argument, the degenerative contributions to each of the higher order TFs will also include a term with *H,*(w).* Since $H_n(\omega, \ldots, \omega)$ can be expressed as sums and products of $H_1(m\omega)$, where m is any positive integer $1 \leq m \leq n$ (as described in chapter 3), all the poles of $H_n(\omega, \ldots, \omega)$ must have positive imaginary parts, assuming hi(t) is causal. Correspondingly the principal diagonals of all the higher order FRFs must be Hilbert transformable and, as a consequence, it must be possible also to apply the Hilbert transform to detect the distortion on each higher order TFs. In figure 4.4 this is demonstrated on the first three TFs from the simulated SDOF system using data measured at the amplitude 0.5.

This is a new application of the Hilbert transform. One general advantage of analysing only the diagonals of the higher order FRFs is that, since each is a function of a single frequency variable, only single dimensional transformations need be used throughout thus simplifing the computations. As already noted, the single dimensional Fourier transform is all that is required when determining the TFs from the measured time signals. Here the detection of distortion on each of the higher order TFs needs only the conventional, single-dimensional Hilbert transform.

Just as multi-dimensional Fourier transforms can be defined, it would be possible to conceive multi-dimensional Hilbert transforms [19]. This extension would be necessary when attempting to detect distortion on a measured estimate of the multi-dimensional FRFs. As described in chapter 2, the Wiener series approach has been developed with this aim. However, since broadband random excitation is used to measure the Wiener kernels, a non-linear system would be linearised and the first order Wiener kernel transform would have the same form as an FRF of a linear system. Consequently, despite the differences which exist between the Wiener and Volterra kernels (described in section 2.3.3), the Hilbert transform would not be able to detect distortion on the first order Wiener kernel transform. Similarly it is likely that there would be no apparent distortion on any of the higher order Wiener kernel transforms, even if they could be measured accurately, and so a multi-dimensional Hilbert transform would be of little use in this application.

Clearly the distortion of the TFs measured using sine excitation can only be interpreted in terms of the degenerative contributions if the system is governed by polynomial functions, since the higher order FRFs are not defined for systems outside this class. However it may well be possible to measure a first order TF from systems for which the first order FRF is not defined, for example structures governed by a bi-linear stiffness function, or when the amplitude of excitation is high so as to cause bifurcation. One principal advantage of the Hilbert transform technique is that, because of its generality, it can be used to detect non-linearity even in these situations. Applying the Hilbert transform method to the task of identifying the form of non-linearity in structures, for example whether the stiffness effect is hardening or softening, relies on the fact that certain common non-linearities tend to cause the Hilbert transformed function to distort in a characteristic fashion. In this context, the interpretation of the degenerative terms may prove useful in understanding these distortion patterns in greater depth, and is suggested as a topic for further research.

Generally, the main disadvantage of the Hilbert transform is that, although it has proved useful in detecting qualitatively whether or not the system is behaving linearly, it provides very little *quantitive* information, and cannot be used to identify a parametric model. On the other hand, higher order TFs provide a direct measure of the non-linearity. By their presence, higher order TFs indicate whether or not a system is linear, and by their form and presence of secondary and tertiary resonances, etc. they can be used to identify which non-linear effects are influencing the various resonances. Most importantly, they are related *quantitively* to the mathematical model. As is the subject of the next section and subsequent chapters, the higher order TFs measured using sine excitation may prove to play an important role in validating practical techniques for identifying mathematical models of non-linear structures.

4.5 Approach to System Identification

In conventional modal analysis, techniques have been developed to curvefit a modal model onto measured TFs [28]. Having proposed practical methods of estimating the higher order FRFs from measurements, it is clear that a similar 'higher order curvefitting' approach could be conceived in an attempt to identify the non-linear parameters in the equations of motion, using the relationships established with harmonic probing in chapter 3. This approach was first proposed by Gifford, who suceeded in curvefitting to the second order Wiener kernel transforms measured from a structure in order to identify second order parameters [20].

The curvefitting procedure makes use of the fact that a FRF of order *n* is related only to lower order FRFs, and parameters of the same and lower orders. (In this context the effect of the higher order FRFs is only to potentially pollute the TFs on which the curvefit is performed, but this is effectively disregarded when formulating the algorithms in terms of the 'ideal' FRFs.) The relationships which have been determined in section 3.3 can be rearranged into a form which is amenable to solution using a least-squares technique [34].

Consider, for example, the relationship derived (equation 3.57) between the first and second order FRFs for a general SDOF system,

$$H_2(\omega_1,\omega_2) = -\{(j\omega_1).(j\omega_2).c_2 + k_2\}.H_1(\omega_1+\omega_2).H_1(\omega_1).H_1(\omega_2)$$

For i different frequency pairs on the second order FRF, the solution matrix can be written in the form,

$$\begin{Bmatrix} \mathrm{Re}[H_2(\omega_{a_1},\omega_{a_2})] \\ \mathrm{Im}[H_2(\omega_{a_1},\omega_{a_2})] \\ \\ \mathrm{Re}[H_2(\omega_{b_1},\omega_{b_2})] \\ \cdot \\ \mathrm{Im}[H_2(\omega_{i_1},\omega_{i_2})] \end{Bmatrix} = \begin{bmatrix} (\omega_{a_1}).(\omega_{a_2}).\mathrm{Re}[H_1(\omega_{a_1}+\omega_{a_2}).H_1(\omega_{a_1}).H_1(\omega_{a_2})] & -\mathrm{Re}[H_1(\omega_{a_1}+\omega_{a_2}).H_1(\omega_{a_1}).H_1(\omega_{a_2})] \\ (\omega_{a_1}).(\omega_{a_2}).\mathrm{Im}[H_1(\omega_{a_1}+\omega_{a_2}).H_1(\omega_{a_1}).H_1(\omega_{a_2})] & -\mathrm{Im}[H_1(\omega_{a_1}+\omega_{a_2}).H_1(\omega_{a_1}).H_1(\omega_{a_2})] \\ \\ (\omega_{b_1}).(\omega_{b_2}).\mathrm{Re}[H_1(\omega_{b_1}+\omega_{b_2}).H_1(\omega_{b_1}).H_1(\omega_{b_2})] & -\mathrm{Re}[H_1(\omega_{b_1}+\omega_{b_2}).H_1(\omega_{b_1}).H_1(\omega_{b_2})] \\ \\ (\omega_{i_1}).(\omega_{i_2}).\mathrm{Im}[H_1(\omega_{i_1}+\omega_{i_2}).H_1(\omega_{i_1}).H_1(\omega_{i_2})] & -\mathrm{Im}[H_1(\omega_{i_1}+\omega_{i_2}).H_1(\omega_{i_1}).H_1(\omega_{i_2})] \end{bmatrix} \cdot \begin{Bmatrix} c_2 \\ k_2 \end{Bmatrix} \tag{4.23}$$

In order to solve this expression for the parameters c_2 and k_2, an estimate of the first order FRF is required at each frequency point ω_1 and ω_2 and also at the sum frequency $(\omega_1+\omega_2)$.

To be able to fit to all points in the principal quadrant of the second order FRF, ie. such that $0 < \omega_1, \omega_2 \leq \omega_l$, it is necessary to measure the first order FRF between 0 and $2\omega_l$. Theoretically, once the second order parameters have been identified, it would be possible to employ similar algorithms to identify the third and higher order parameters. Naturally, the solution becomes more complicated as the order increases. For example, the third order algorithm incorporates the first and second order FRF diagonals, and the second order parameters.

Throughout the procedure the first order FRF plays a central role because it appears in the algorithm of each order. However, one important observation regarding this approach is that the first order parameters need not be estimated during the higher order curvefitting. Moreover, it is not possible to identify either the mass or the first order stiffness or damping parameters using this procedure alone, and some other method would need to be employed in conjunction in order to estimate all parameters in the equation of motion model.

There are two further problems encountered if attempting to identify the non-linear parameters from FRFs estimated using the Wiener series approach. Firstly, the transforms of the Wiener kernels only approximate to the FRFs defined from the Volterra kernels when the level of excitation is small. The estimates of the parameters are likely to be biased if the differences between equivalent Wiener and Volterra kernel transforms are significant. Perhaps more importantly, since there has been very little reported success in being able to measure Wiener kernels above second order from physical structures, it has not been possible to estimate third or higher order parameters using this approach. Clearly to be able to identify non-linear systems in which the governing polynomials have, for example, odd order symmetry, an estimation of at least the third order parameters must be made.

One approach developed initially focussed on simplifying the curvefitting algorithms considerably by considering only the principal diagonals of the FRFs, and applying the curvefitting to the higher order TFs. As an example, to identify the second order parameters of the SDOF system, the curvefitting algorithm of equation 4.23 reduces to the single-dimensional form,

$$\begin{Bmatrix} \mathrm{Re}[H_2(\omega_a,\,\omega_a)] \\ \mathrm{Im}[H_2(\omega_a,\,\omega_a)] \\ \\ \mathrm{Re}[H_2(\omega_b,\omega_b)] \\ \\ \mathrm{Im}[H_2(\omega_i,\,\omega_i)] \end{Bmatrix} = \begin{bmatrix} (\omega_a)^2.\mathrm{Re}[H_1(2\omega_a).H_1(\omega_a)^2] & -\mathrm{Re}[H_1(2\omega_a).H_1(\omega_a)^2] \\ (\omega_a)^2.\mathrm{Im}[H_1(2\omega_a).H_1(\omega_a)^2] & -\mathrm{Im}[H_1(2\omega_a).H_1(\omega_a)^2] \\ \\ (\omega_b)^2.\mathrm{Re}[H_1(2\omega_b).H_1(\omega_b)^2] & -\mathrm{Re}[H_1(2\omega_b).H_1(\omega_b)^2] \\ \vdots & \vdots \\ (\omega_i)^2.\mathrm{Re}[H_1(2\omega_i).H_1(\omega_i)^2] & -\mathrm{Re}[H_1(2\omega_i).H_1(\omega_i)^2] \end{bmatrix} \cdot \begin{Bmatrix} c_2 \\ k_2 \end{Bmatrix} \tag{4.24}$$

This simplification does not impose any additional restrictions on the model since theoretically the parameters can be retrieved using data from any points on the multi-dimensional FRFs. In the multi-dimensional algorithm, pairs of frequency points are selected with regard to reducing the size of the solution matrices and the time for solution [25]. On the other hand, in the single-dimensional case, frequency points along the leading diagonals can be selected around the resonances in a similar way to conventional curvefitting procedures, or data from all frequencies across a bandwidth can be used [47]. The main advantage with the simplification is that algorithms can be formulated up to any order using the pattern established in section 3.2, and since often high quality TFs to (at least) order three can be measured from physical systems, the potential exists for identifying third and higher order terms in the governing polynomials.

As an example, curvefits were performed on the TFs measured from the simulated SDOF system. The second and third order parameters were identified independently for the low amplitude case of X = 0.1 and the medium amplitude X = 0.5, and the results compared.

Parameter	Actual [X = o]	Estimated x = 0.1	Estimated x = 0.5
k_2 (N/m^2)	1.0 x 10^7	1.004 x 10^7	1.068 x 10^7
k_3 (N/m^3)	1.0 x 10^{10}	1.003 x 10^{10}	1.222 x 10^{10}
c_2 (Ns^2/m^2)	0.0	13.05	246.2
c_3 (Ns^3/m^3)	0.0	44.45	-4765.6

(4.25)

Writing the solution algorithm as $\{z\} = [A].\{\phi\}$, the curvefitted functions $\{\hat{z}\}$ can be regenerated from the estimated parameter vector $\{\hat{\phi}\}$ simply by using $\{\hat{z}\} = [A].\{\hat{\phi}\}$.

To test the accuracy of the curvefit, the curvefitted functions can be compared visually with the corresponding TFs. For the lower level data, there was hardly any discernable difference between the functions. However for the higher level, differences can be observed as shown for the second order in figure 4.5, and for the third order in figure 4.6.

The differences between the results for the two amplitudes highlights the main problem with this approach. The higher order TFs only approximate to the diagonals of the FRFs, and any distortion in the TFs will cause errors in the parameter estimates. In this respect, the TFs are no better than the Wiener kernel transforms for providing an estimate of the FRFs, and are often likely to be worse because, in the distorted TFs, the resonant frequency tends to shift away from the true pole of the FRF.

In the SDOF example, the measured TFs for all orders are used in the curvefit algorithm as estimates of the principal diagonals of the FRFs, but their distortion introduces error into the parameter estimation. By performing curvefits on data from a range of simulated systems, it has been discovered that, since each of the TFs have a similar distortion pattern, the errors in the parameters are reduced. However the errors which arise are particularly significant in the MDOF case since, as described below, it is only possible to use an (undistorted) estimate of the first order FRF rather than the measured first order TFs.

Another important fact is that the higher order curvefit must be performed over a reduced frequency range. In this example, all the TFs were measured from zero up to 38.5 Hz. However the second order curvefit must be restricted to half the bandwidth, 0 to 19.2 Hz, because the component $H_1(2\omega)$ is incorporated. Similarly the third order curvefit is restricted to a third of the bandwidth, 0 to 12.8 Hz because of $H_1(3\omega)$. Clearly this would pose a significant problem if only data from a zoomed bandwidth were available. For this reason, measurements of the TFs should always be made from baseband whenever possible if curvefitting is to be applied.

To maintain generality in the curvefitted model, non-linear damping terms have been included, despite the fact that the non-linearity appears only in the stiffness function. If a-priori information regarding the system is known in this manner, it is usually possible to introduce further simplifications into the curvefitting algorithm. For example, using the data at the higher amplitude, and omitting the higher order damping parameters from the solution, the estimates of the stiffness parameters were revised,

$$k_2 = 0.986 \text{ x } 10^7 (\text{N/m}^2); k_3 = 0.897 \text{ x } 10^{10} (\text{N/m}^2) \tag{4.26}$$

The proximity of the curvefitted functions to the measured TFs did not change significantly, indicating that the damping parameters in the model originally identified have only a small contribution to the curvefitted functions.

To this stage, the discussion has centred on the application to SDOF systems. Theoretically, the curvefitting technique could be extended to identify a MDOF model of a structure. However to achieve this, an accurate estimate of the full first order FRF would be required, and in practice this presents a serious problem. The conventional approach to estimating the FRF matrix from a test on a linear structure would be to excite the structure at a single point, measure TFs at several locations to give an estimate of a single column of the matrix, and then, by using modal curvefitting procedures, attempt to synthesise the full FRF matrix.

As described in chapter 1, any residual effects would have considerable influence on the accuracy of this synthesis. Furthermore when the structure is non-linear and sine excitation is used to measure the first order TFs, the distortion which occurs will affect the validity of the modal curvefit. Resulting errors in the first order FRF matrix will give rise to inaccurate second order parameters, an effect which is compounded when proceeding to the higher orders because the inaccurately estimated parameters must then be used in the algorithm.

Figure 4.7 shows the curvefits to the TFs from the simulated 3 DOF system measured from the direct location using the input signal with amplitude X = 1.5. The discrepancies between the original and curvefitted functions become more significant for the higher orders. The problem of restricting the bandwidth of the curvefit is also obvious.

To circumvent the distortion problem, one modification attempted was to measure the linearised TFs using random excitation, in addition to the higher order TFs with a stepped sine test, and then to perform the modal curvefit on these undistorted functions instead of $TF_1(\omega)$. Nevertheless, the distorted higher order TFs must still be used in the algorithm and this again causes inaccuracies in the parameter estimates. In simulation tests, the modification proved to be largely unsuccessful at improving the parameter estimates. Clearly what is required is undistorted estimates of the FRF diagonals for each order, although there appears to be no obvious solution to this problem.

Another general disadvantage with the higher order curvefitting approach to system identification is that mass and first order stiffness and damping parameters, which are always important in the parametric model of weakly non-linear structures, cannot be estimated using these procedures. An alternative technique would need to be used in conjunction, for example using conventional procedures to fit to the measured first order TF, in order to estimate a complete parametric model. Development of parameter identification procedures for structures which are linear is an area of substantial research in itself, and one technique which is currently being investigated is the DPE approach (described in chapter 1.)

One considerable advantage with the DPE approach is that it can be extended to identify parametric models of MDOF structures directly from time data. In the latter stages of the project, as the potential of the DPE technique became increasingly apparent, the direction to system identification became reorientated away from higher order curvefitting and towards direct parameter estimation.

By identifying the model in the time domain with DPE and, using the relationships which have been established in chapter 3, the higher order FRFs can be regenerated. The time domain models can be validated by comparing the principal diagonals of the regenerated FRFs with measured higher order TFs (the subject of chapter 6.) Clearly most of the problems associated with curvefitting in the frequency domain are avoided by performing the identification in the time domain. However some other problems arise, particularly in relation to the practical implementation of the DPE procedure, and these are addressed in chapter 5.

Chapter 5

Implementation of Direct Parameter Estimation

5.1 Overview

Throughout this project, emphasis has been placed on developing analytical methods which can be applied in practice within a structural test. Of particular interest has been the identification of parametric models for both linear and non-linear structures, and routines have been implemented and evaluated for this application. Characterising and interpreting the behaviour of structures is the first stage, and conducting a test to identify a mathematical model is necessary before proceeding to attempt to predict the response of the structure.

The primary aim of this chapter is to describe how the Direct Parameter Estimation (DPE) technique has been developed in a practical direction during this project. As discussed in chapter 1, the DPE technique can be applied to structures which are linear, or extended to include non-linear terms. In this chapter, two slightly different methods of measuring representative time data are discussed and evaluated on a structure which is nominally linear. Subsequently, in chapters 6 and 7, results are presented from non-linear structures to illustrate how the extended technique can be used in conjunction with the higher order transfer functions to identify the non-linearity. Also it is appropriate to discuss at this stage some of the features of the computer-controlled, data acquisition system used since this facility has influenced the direction taken during this project towards structural testing and system identification. An account is also included of the approach to experimentation and the equipment used throughout the tests.

5.2 Experimentation and Equipment

5.2.1 Data Acquisition System

Generally, standard experimental procedures for testing structures were used throughout, and only the data acquisition system warrants a detailed description because a particularly versatile system became available during the last year of the project and the latest version of the analysis routines were orientated to make the most appropriate use of this facility.

The main hardware component is a DIFA SCADAS II 'front-end', designed to perform a wide range of signal generation and data acquisition tasks, which can be controlled by a HP-340, 9000 series, workstation computer. A software package produced by LMS was installed on this computer, the primary function of which is to enable the computer to access the front-end directly and to control the generation of a signal using the Digital-to-Analogue Converter (DAC) in the front-end. The DAC output can be used, for example, as an input to an electrodynamic shaker to excite a structure.

By measuring the response of the structure with accelerometers, and feeding these signals through charge amplifiers into the 12 Analogue-to-Digital Converters (ADCs) available in the front-end, each of which have simultaneous sample-and-hold, the response signals from several locations across the structure can be transmitted to the computer. In this way, data can be measured from each location corresponding to the same instant in time. The ADCs in the front-end also have the facility of auto-ranging, amplifying the sampled signals to optimise the 16-bit multiplexer, and enabling time data to be recorded by the computer with high resolution. This data is transferred to the computer via a fast parallel GPIO interface with very high transmission rates, whereas the commands are sent to the front-end using the HPIB serial connection.

Although many alternative data acquisition systems are available which can perform similar tasks, this system is one of the fastest and most accurate, and also has the advantage that the hardware and software are fully integrated, providing a platform on which to code routines for specific tasks within the 'user programming' enviroment of the LMS software. Commands are available for plotting data and performing signal processing operations such as DFTs on blocks of data very quickly. Also, in this application, extensive use was made of the facility for streaming of data to external Fortran programmes, written to perform various specific analysis tasks. Since data processing can be performing immediately following acquisition, 'on-line' analysis is possible, enabling the results to be assessed before the test is completed.

5.2.2 The Linear Beam Rig

One aim of this chapter is to discuss the effectiveness of DPE when applied to a relatively simple, linear, MDOF benchmark structure. A rig consisting of a cantilevered, mild steel beam was made for this purpose, designed to possess three clear, well separated modes lying in the bandwidth 0 to 100 Hz by connecting three mild steel discs of mass 450g, spaced at 0.254 m intervals, along the beam (figure 5.1). One end of the beam was fixed to a rigid bedplate, constraining the vibration motion of the beam to the horizontal plane, whilst a steel coil spring, supported via a rigid bracket to the bedplate, was connected to the free end. The coil spring stiffens the structure and has the effect of raising the first bending mode of the beam from resonance at approximately 6 Hz up to 16.8 Hz. One reason for this modification was that piezoelectric transducers do not perform well at low frequencies, eg. up to 5 Hz, and the DPE procedure is very sensitive to spurious low frequency components in the time signals.

A structure of this construction would be very lightly damped unless an additional form of energy dissipation were included. For example, during a stepped-sine test, transient oscillation must be allowed to die away before measurements are taken at steady-state. To increase the damping, constraining layers were added to both sides, along the length of the beam. A Ling 400 series electrodynamic shaker was connected to the beam via a Kistler type 9311A force transducer, and the response was measured at the locations of each of the three lumped masses using B & K type 4331 piezoelectric accelerometers. Conditioning of the four measured signals was performed using B & K type 2635 charge amplifiers.

The DPE procedure has been found to be very sensitive to any errors in the calibration of the transducers. If the calibration of the force transducer or all of the accelerometers were incorrect, the scaling of all the parameters would be inaccurate. However, if the accelerometer calibrations do not correspond, the effectiveness of the procedure is jeoparised since the response signals are compared directly via the difference terms in the numerical model. Good practice is to use the same type of accelerometer at each location, and to calibrate all at the same time subjected to the same reference excitation. This also ensures other potential factors, such as changes in the ambient temperature, have minimal effect.

Before proceeding to the identification stage, TFs were measured from the beam using broad-band random excitation in order to confirm that the structure can be considered to have only three degrees of freedom. Although modes were detected in the range 100 to 200 Hz, ie. outside the bandwidth of interest, their influence on the bandlimited data should be insignificant. This was checked by performing a conventional curvefit on the TFs over the

first three modes without correcting for stiffness (or mass) residuals. The closeness of the fit for each response point, and in particular the direct point (figure 5.2), indicates that over the bandwidth, the TFs can be adequately represented by a 3 DOF model. The parameters of the curvefit model for each response location are listed in table 5.3.

In the identification the parametric model, the beam was assumed to behave linearly for the excitation levels used. Before proceeding, this was justified by performing a Hilbert transform on the first order TFs measured using stepped-sine excitation. Setting the amplitude of the excitation signal to 0.1V limits the displacement amplitude to approximately 1 mm at the first resonant frequency. The response was measured in terms of acceleration, and the Hilbert Transform across the bandwidth 4 to 96 Hz was computed on the interpolated inertance TFs (including correction for the zoomed frequency range using the 'Ahmed method' [65]), as shown in figure 5.4. The results indicate that the second and third modes of the beam are nominally linear although, since the structure remains rather lightly damped, the resolution is not sufficient to ascertain the linearity around the first mode with confidence. Consequently, the stepped sine test was repeated using 200 increments across the bandwidth 4 to 24 Hz, and the Hilbert Transform computed across the bandwidth 10 to 20 Hz, again using correction. The result, shown in figure 5.5, indicates that there may be a *weak* hardening stiffness non-linearity associated with the first mode, possibly caused by the coil spring which would have the most significant influence around this resonance when the displacement of free end is a maximum, but between 0 and 100 Hz, the structure does behave nominally linearly.

As described in the next section, each of the two methods for measuring representative time data use the same experimental setup. In the first tests, the beam was excited using a broadband random signal bandlimited to 4-96 Hz using an analogue 'reconstruction' filter. The generated signal was limited to 1V RMS, and the setting on the power amplifier adjusted to give a maximum displacement level of approximately 1 mm, a relatively low level when compared with the overall dimensions of the structure. (Throughout the tests, the amplification setting on the power amplifier was maintained, and comparisons between the excitation levels are made by referring to the voltage levels of the signals generated.)
In the second test, this procedure was repeated after a steel disc of approximate diameter 100 mm and thickness 15 mm, trimmed to have a mass of 1.00 kg, was fixed to one side of the beam at location 2. The purpose of this modification was to investigate whether the parameter matrices, and specifically the mass matrix, could be interpreted from a physical viewpoint by repeating the test to identify this added mass. One effect of the modification was to alter the resonant frequencies of each of the three modes of interest, reducing the resonances at 16.8, 31.8 and 78.2 Hz to frequencies of 13.4, 29.1 and 68.7 Hz respectively.

5.2.3 Design of the Excitation Signal

Using the signal generation facility of the front-end, there is a wide range of excitation signals which can be selected. Furthermore it is possible to program the front-end to generate any periodic waveform. However, only the two most commonly used excitation waveforms have been employed in this project, namely sinusoidal and bandlimited random, since they are easy to generate and are the most appropriate for the different applications. The DPE procedure does not rely on any particular excitation waveform, but in this project random excitation was used throughout because with sine excitation a problem known as 'linear dependence' can occur, particularly when the system is linear. To illustrate this, consider the equation of motion of the direct point (cf. equation 1.1),

$$m_r.\ddot{y}_r(t) + c_{rr}.\dot{y}_r(t) + k_{rr}.y_r(t) + \sum_{\substack{p=a \\ p\neq r}}^{N}[c_{rp}.(\dot{y}_r(t) - \dot{y}_p(t)) + k_{rp}.(y_r(t) - y^*(t))] = x_r(t) \quad (5.1)$$

For the linear system, the displacement response to a sinusoidal input, $x(t) = X.\cos(\omega t)$ will also be sinusoidal and can be written $y(t) = Y.\cos(\omega t + \theta)$,
where θ is included to express the phase delay between the signals.
By differentiating, $\dot{y}(t) = -\omega.Y.\sin(\omega t + \theta)$ and $\dot{y}(t) = -\omega^2.Y.\cos(\omega t + \theta)$,
and substituting into the equation of motion (eqn. 5.1),

$$-\omega^2.m_r.Y_r.\cos(\omega t + 6) - \omega.c_{rr}.Y_r.\sin(\omega t + \theta) - \sum_{\substack{p=a \\ p\neq r}}^{N}\omega.c_{rp}.(Y_r - Y_p).\sin(\omega t + \theta)$$
$$+ k_{rr}.Y_r.\cos(\omega t + \theta) + \sum_{\substack{p=a \\ p\neq r}}^{N}k_{rp}.(Y_r - Y_p).\cos(\omega t + \theta)$$
$$= X_r.\cos(\omega t + \theta) \quad (5.2)$$

Since the acceleration and displacement signals are both sinusoidal and in phase,

$$\left[\left(\sum_{p=a}^{N}k_{rp} - \omega^2.m_r\right).Y_r - \sum_{\substack{p=a \\ p\neq r}}^{N}.k_{rp}.Y_p\right].\cos(\omega t + \theta)$$
$$+\left[\omega.\sum_{\substack{p=a \\ p\neq r}}^{N}.c_{rp}.Y_p - \sum_{p=a}^{N}c_{rp}.Y_r\right].\sin(wt + \theta) = X_r.\cos(\omega t + \theta) \quad (5.3)$$

As can be seen, for a single frequency sinusoidal input, there is no way of segregating the displacement term from the acceleration term, and as a consequence it would not be possible to identify the mass parameter independently from the stiffness parameters.

In theory, linear dependence will not arise for non-linear systems because the response to a sinusoidal input is not sinusoidal but includes higher harmonics,

$$y(t) = Y_1.\cos(\omega t + \theta_1) + Y_2.\cos(2\omega t + \theta_2) + Y_3.\cos(3\omega t + \theta_3) + \ldots \quad (5.4)$$

$$\ddot{y}(t) = -\omega^2.Y_1.\cos(\omega t + \theta_1) - 4\omega^2.Y_2.\cos(2\omega t + \theta_2) - 9\omega^2.Y_3.\cos(3\omega t + \theta_3) + \ldots \quad (5.5)$$

Although corresponding terms in the displacement and acceleration signals have the same frequency and phase, they have different amplitudes and theoretically would be segregated. In practice, however, the higher harmonic components are usually relatively low level compared with the fundamental. Consequently sine excitation does not give as good results as random for which linear dependence does not arise. In relation to the restoring force surface, a large area of the phase plane $(y, \dot{y})$ is covered when the excitation is random, and the parameter estimation is improved as a result. As in the example of the electronic circuit (section 1.2.3), sine excitation can give very clear results using the restoring force surface (figure 1.10), incrementing the amplitude or frequency of the excitation to cover the phase plane. However an accurate mass estimate is required which would usually entail employing DPE beforehand. Clearly the problem of linear dependence with sine excitation would be manifest at this stage rather than being avoided.

Apart from selecting the type of excitation waveform, another important decision to be made is with regard to the sampling frequency. Usually this is selected to be an order of magnitude higher than the highest frequency of interest to avoid aliasing, a problem arising from undersampling. Reducing the sampling time interval At ensures that the digitised time signals are smooth. However resolution in the frequency domain is also important, ie. it is also necessary to ensure that Δf, the spacing of the spectral lines, is kept small. Since $\Delta t = \frac{1}{N.\Delta f}$, the only way of maintaining resolution in both the time and frequency domains is to maximise N, the length of the section of sampled time data. For the data acquisition system employed, the maximum value of N is 8192 points per channel.

A third practical consideration which can also influence DPE regards the level of excitation. At low levels, the signal-to-noise ratio decreases and it becomes more difficult to resolve the signals accurately with the ADCs. In addition, if non-linear parameters are to be identified but the excitation is low level, the non-linearity may not be excited sufficiently for accurate estimation of these parameters. Conversely, at very high levels of excitation, the effect of the non-linear behaviour can be exaggerated, particularly if physical limits are approached. Correspondingly, if a linear model is to be identified, a compromise is sought between noise and resolution problems of low level signals, and the tendency to encourage non-linearity at higher levels. (The level of 1V RMS applied to the beam is relatively low.)

5.3 Identifying a Parametric Model

5.3.1 Introduction

To identify a parametric model of a structure, the DPE procedure uses time data representing the force applied at one point, and the response in terms of displacement, velocity and acceleration from several locations on the structure. To obtain all the required signals, the first of the two methods which have been developed and implemented in this project uses the measured force and response time histories and employs numerical integration or differentiation in the time domain. This is the more direct approach. The second, indirect method uses Transfer Functions (TFs), measured from a structure using broadband random excitation, from which representative time data can be generated. This approach has certain advantages when used to identify a linear structure. Furthermore, if the structure is non-linear, the indirect method provides a means of identifying a model of the linearised structure. The aim of this section is to discuss each of these methods while referring to the tests conducted on the linear beam rig shown in figure 5.1.

5.3.2 Method 1

The direct implementation of the DPE procedure uses measured time signals to identify the parametric model. To achieve this, it is necessary obtain a measure of not only the acceleration at each of the locations, which can be measured directly using accelerometers, but also the velocity and displacement. To measure each of these quantities for all locations at the same instant would require a great deal of instrumentation and limit the practical applicability of the approach. Fortunately, having measured only one of the quantities, the other two can be estimated from it using numerical integration or differentiation. It is typical to measure acceleration and to estimate velocity with integration, which in turn can be integrated to give displacement. Alternatively, if it were only possible to measure displacement, as is the case with the analogous electronic circuit system discussed in chapter 1, numerical differentiation to velocity and acceleration could be used. With this approach, the sampled points on each signal correspond to the same instants in time which is important since it has been discovered that the DPE procedure is extremely sensitive to erroneous delays between the signals. However both the numerical calculus procedures have inherent accuracy problems; the accuracy of differentiation falls off at high frequencies, whilst integration can introduce spurious low frequency components.

Differentiation is defined as a limiting process,

$$\frac{dy}{dt} = \underset{St \to 0}{\text{limit}} \frac{\delta y}{St} \quad \frac{\text{Ay}}{At} \tag{5.6}$$

Clearly, the approximation improves as At is reduced, but becomes more sensitive to errors in the measurement of $y(t)$. Since the approximation is the gradient of the straight line between adjacent points in time, any unwanted high frequency corruption of the signals such as measurement noise will be amplified by differentiation. For the applications where differentiation cannot be avoided, a 'centred difference' formulation was implemented which uses data from 4 different instants to determine the differentiated signal at t_i,

$$\dot{y}(t_i) = \frac{1}{6(t_{i+1} - t_{i-1})} . (y(t_{i-2}) - 8y(t_{i-1}) + 8y(t_{i+1}) - y(t_{i+2})) \tag{5.7}$$

$$\ddot{y}(t_i) = \frac{1}{6(t_{i+1} - t_{i-1})} . (\dot{y}(t_{i-2}) - 8\dot{y}(t_{i-1}) + 8\dot{y}(t_{i+1}) - \dot{y}(t_{i+2})) \tag{5.8}$$

Many methods of time domain integration exist, but most have the problem of introducing both spurious low and high frequency components into the data. It is possible to show, however, that the second problem can be avoided by employing the trapezium rule in which data from the previous instant, t_{i-1}, is used to estimate the integrated signal at t_i [39],

$$\dot{y}(t_i) = \dot{y}(t_{i-1}) + \frac{t_i - t_{i-1}}{2} . (\ddot{y}(t_i) + \ddot{y}(t_{i-1})) \tag{5.9}$$

$$y(t_i) = y(t_{i-1}) + \frac{t_i - t_{i-1}}{2} . (\dot{y}(t_i) + \dot{y}(t_{i-1})) \tag{5.10}$$

At each stage, the integration introduces an arbitrary mean level which can be interpreted as the constant of integration, for example a and b,

$$\int \ddot{y}(t).dt = \dot{y}(t) + a \tag{5.11}$$

$$\int \dot{y}(t).dt = y(t) + a.t + b \tag{5.12}$$

The low frequency trends which arise are compounded if the measured acceleration signals contain a zero frequency (d.c.) component or mean level. For example, a zero frequency acceleration would correspond physically to an accelerating drift motion of the structure which would not occur if the structure were connected to ground, as is usually the case. Hence it is possible to justify removing the mean level from the acceleration signal prior to integration.

There are two equivalent ways of removing low-frequency trends from data during post-processing, either by fitting a polynomial trend across the data section, and then removing the trend from the data, or by high-pass filtering.

Removal of the arbitrary mean level from the velocity data resulting from integration can also be justified, since a structure fixed to ground would not tend drift in any direction even with constant velocity. However, a mean displacement level may occur in practice if the structure becomes displaced from the datum which may be arbitrary. In theory, a d.c. component in the displacement can arise also from the mechanism for energy transfer exhibited by a non-linear system. Consequently, distinguishing between a physical zero frequency component in the displacement data and the arbitrary constant of integration is very difficult. Nevertheless, the DPE procedure is extremely sensitive to spurious mean levels in the signals which must be attenuated or removed before using the DPE procedure. Apparently the only option is to bandpass the excitation signal so that the bandwidth includes all resonances of interest, with a high-pass cutoff set some frequency between zero and the first resonance, and then to filter the response signals over a corresponding frequency range.

For a linear structure subject to broadband excitation, there is no mechanism for energy to be transferred between frequencies, and consequently the acceleration, velocity and displacement data can be filtered over the same bandwidth without corrupting the results. The data measured from the beam was filtered in this way. However the response of a non-linear structure may contain components outside the excitation bandwidth due to the non-linearity, and these must be retained if non-linear terms are to be identified.

At this stage of processing the data, if the structure were non-linear, it would be necessary to make an assumption regarding the highest order of non-linearity of interest. For example, to maintain in the data components arising from a cubic term in the governing polynomial, it would be necessary to select the high frequency cutoff to be at least three times the excitation cutoff frequency when numerically filtering the response data. Limiting the polynomial function which can be identified in this way is not particularly restrictive since it would still be a considerable improvement on conventional linear models.

Unfortunately, if the structure is non-linear it is not as easy to justify the selection of the high-pass frequency, ω_0, to be any frequency above zero since an asymmetrical non-linearity can generate output frequency components at frequencies below ω_0 including zero when the excitation is random. Furthermore, it is very difficult to determine whether these low frequency components contribute significantly to the output signals, or if they can be neglected. Research with polyspectra techniques has addressed problems of this nature (section 2.2.4), and observations suggest that the low frequency output components will be significant if there is a structural resonance between zero and w_0, which can be excited by the energy transfer; otherwise filtering out the low frequency components can be justified.

In this implementation, a digital Butterworth algorithm was used to numerically filter the data. This algorithm causes a spurious phase delay in the filtered signal, although the effect can be removed by returning the data backwards through the routine. The phase delay would be manifest as a time delay in the data which could destroy the simultaneity necessary between each of the time signals used in the DPE procedure. Also the filter itself can be considered to act as a dynamical system, and can give rise to transient effects which occur at both ends of the signal after the two passes. For this reason only the central sections of the time signals are used in the DPE procedure.

In the test on the rig, the excitation force and the acceleration responses were measured with a sampling frequency of 2000 Hz, collecting 8192 points simultaneously from each .of the 4 channels. Since accelerometers were found to be unreliable at low frequencies, an operational bandwidth of 4-96 Hz was selected and, following integration, each of the signals were filtered numerically to this frequency range before identifying a linear parametric model. Central sections of the scaled data and corresponding spectra are shown for direct location (1) in figure 5.6. Before proceeding with the identification, a simple check of the data can be made by dividing the response spectra by the force spectrum, and either making a comparison with measured TFs, or simply checking the forms correspond to the familiar receptance, mobility and inertance functions after a single average, as shown in figure 5.7.

Employing the DPE with the central 2048 points of data, which have been scaled appropriately, results in parameter matrices which have physical dimensions,

$$[m] = \begin{bmatrix} 0 & 0.761 & 0 \\ 0 & 0 & 1.004 \\ 0.726 & 0 & 0 \end{bmatrix} \text{(kg)} \; ; \quad [c] = \begin{bmatrix} -4.541 & 9.084 & -2.734 \\ -0.598 & -2.734 & -2.029 \\ 9.560 & -4.541 & -0.598 \end{bmatrix} \text{(N/ms}^{-1}\text{)}$$

$$[k] = 10^3. \begin{bmatrix} 120.6 & -71.4 & 26.5 \\ -71.5 & 68.7 & -35.2 \\ 26.5 & -35.2 & 37.6 \end{bmatrix} \text{(N/m)} \qquad (5.13)$$

5.3.3 Method 2

The second method is less direct because TFs are measured from the structure rather than time signals, and used to generate time data which represent the dynamics. The TFs are measured again using bandlimited random as the excitation signal. Inertance functions,

$$\mathrm{TF}''(\omega) = \frac{Y''(\omega).X^*(\omega)}{X(\omega).X^*(\omega)}$$

are usually measured in practice because accelerometers are used to measure the response.

If the structure is non-linear, the TFs measured correspond to the optimum linear model as defined in section 1.3. The original reason for developing this indirect approach was to provide a practical means of identifying the optimum linear model, using the TFs to generate time data to represent the linearised structure. However there are also considerable advantages for using this technique rather than the direct method even when the structure is linear. Firstly, the procedure is quicker because time domain integration / differentiation and filtering can be avoided. Furthermore, it is possible to average the TFs to reduce the effect of uncorrelated, measurement noise, remembering that the process of linearisation requires the TFs be averaged since, in this context, the non-linearity is apparent only as a source of noise to the response signals. The inertance TFs were measured from the three locations on the beam using 50 averages, sufficient to ensure the functions reached 'steady-state'.

Once measured, the TFs effectively replace the physical structure in the subsequent stages of the procedure, and it is essential to make sure before proceeding from the test stage that the TFs measured are of high quality, and have adequate resolution in the frequency domain, particularly around resonances. The aim is to generate time data to be used in the DPE procedure from the TFs by employing the inverse Fourier transform. As described previously, the only way to increase the resolution in both the frequency and time domains is to maximise the number of lines in the Fourier Transform. In this implementation, the frequency domain resolution is 4096 complex data points corresponding to 8192 real points in the time domain. The resolution problem is exactly the same as for method 1, and the selection of the sampling frequency is made with the same considerations. In fact, each step in the indirect method has a corresponding one in the direct method. The principal difference is that the data remain in the frequency domain until the last stage in which the inverse Fourier transform is used to generate the time signals to be used in the DPE.

After measuring either the inertance, mobility or receptance TFs, the other two can be computed using the relations,

$$\text{Inertance} = (j\omega).\text{Mobility} = (j\omega)^2.\text{Receptance}$$

observing that multiplication by (jw) is the frequency domain equivalent of differentiation, and division by $(j\omega)$ corresponds to integration.

In this implementation, the frequency of each spectral line of the measured TF can be determined with accuracy. The problems which are encountered arise from difficulties with measuring high quality TFs. For example, any distortion in the measured inertance TFs at low frequency will be amplified when calculating mobility and receptance. The calculated functions can be displayed and checked to have characteristic forms as suggested previously for method 1 (figure 5.7). Incidentally, the differences which can be observed between the functions with the two methods highlights the extent to which averaging away uncorrelated noise can affect the data.

With frequency domain integration, it is also necessary to stipulate that the zero frequency component of the inertance function is zero to avoid division by $(j0)$. This corresponds to removing the mean level from the acceleration signals in the time domain, and can be justified accordingly. Conversely, high frequency noise on the TFs would be amplified if receptance functions were measured and 'differentiated' to give mobility and inertance, corresponding to the problem of time domain differentiation described previously. Thus similar problems arise with integration and differentiation in the frequency domain as in the time domain, although for this implementation, execution in the frequency domain is far quicker.

Having determined the receptance, mobility and inertance functions for each response location, the next stage is to multiply each by a designed force spectrum in order to determine the spectra of representative displacement, velocity and acceleration signals. In this context, the TFs can be regarded as characterising the linearised structure, and multiplying by the force spectrum corresponds to exciting the optimum linear system. Clearly there are no practical limitations on the selection of the excitation signal since it is computed rather than generated, but it is designed using the same criteria as before. In particular, the bandwidth for the force spectrum must incorporate the resonances of interest but reject, for example, spurious low frequency data.

Hence, this stage of the procedure can also be regarded as corresponding to frequency domain filtering since the force spectrum is also designed to include high-pass and low-pass cutoffs to avoid the problems of spurious low frequency trends and high frequency errors which

may be manifest in the data spectra. Appropriate selection of the cutoffs in the force amplitude spectrum ensures that any corruption in the response spectra, caused by errors in the measured TFs, is not manifest in the time data computed from the spectra.

Sharp cutoffs in the amplitude function, which would give rise to ripple in the time data (known as Gibbs phenomenon), are avoided in the designed force spectra, and a simple cosine taper is used across an arbitrary range at both the low and high frequencies ends of the bandwidth. This is still a relatively sharp cutoff when compared to real analogue or numerical digital filters, and an advantage of frequency domain filtering in this way is that many of the practical problems encountered when optimising the filter characteristics are avoided. In designing the force spectrum, it is also necessary to randomise the phase to prevent the energy in the force signal from being concentrated at the ends of the periodic time section after the inverse transformation.

Finally, the response time data, which represents the linearised system if the structure is non-linear, is computed using the inverse Fourier transform on each spectrum. The resulting data is used in the DPE procedure in exactly the same way as the measured signals in the direct method. The only difference is that because of the linearisation, only first order terms can be selected for identification. It is also important to remember that the original TFs were measured with excitation applied at one specific location and, when solving the identification problem, the computed force signal must be used in the equation of the direct point. Central sections of the representative time data are shown for location 1 in figure 5.8, and were in the DPE procedure to yield the following parameter matrices,

$$[m] = \begin{bmatrix} 0.736 & 0 & 0 \\ 0 & 0.764 & 0 \\ 0 & 0 & 0.989 \end{bmatrix} \text{(kg)} \quad ; \quad [c] = \begin{bmatrix} 9.285 & -3.920 & -1.373 \\ -3.920 & 9.963 & -2.734 \\ -1.373 & -2.734 & -0.340 \end{bmatrix} (\text{N/ms}^{-1})$$

$$[k] = 10^3. \begin{bmatrix} 121.6 & -72.2 & 26.4 \\ -72.2 & 69.8 & -35.4 \\ 26.4 & -35.4 & 38.5 \end{bmatrix} (\text{N/m}) \qquad (5.14)$$

Since the structure is nominally linear, it is possible to compare these results with the equivalent matrices identified using the direct method (equation 5.13). Although there are differences between each of the matrices, there is reasonably good agreement particularly between the mass and stiffness matrices. For example, using both methods the total mass is estimated to be approximately 2.49 kg.

The entire procedure was repeated after the 1 kg mass was added to location 2, with the result,

$$[m] = \begin{bmatrix} 0.800 & 0 & 0 \\ 0 & 1.691 & 0 \\ 0 & 0 & 0.985 \end{bmatrix} \text{(kg)} \quad ; [c] = \begin{bmatrix} 11.300 & -0.206 & 2.567 \\ -0.206 & 3.755 & -5.400 \\ 2.567 & -5.400 & 1.038 \end{bmatrix} \text{(N/ms}^{-1}\text{)}$$

$$[k] = 10^3. \begin{bmatrix} 124.2 & -71.5 & 26.9 \\ -71.5 & 68.1 & -35.4 \\ 26.9 & -35.4 & 36.8 \end{bmatrix} \text{Nm}^{-1} \qquad (5.15)$$

As mentioned previously, the reason for conducting this second test was to investigate whether or not the identified parameters have a physical interpretation. By comparing equation 5.15 with equation 5.14, the identified mass element at location 2 is seen to have increased by 0.93 kg, with the total mass becoming 3.48 kg, an increase of 0.987 kg, which approximates to the 1.00 kg added. Furthermore, the stiffness matrix is largely unaffected, with each element of the matrix differing by less than 2.5% from the previous test. On this basis it would appear that the procedure could be used to identify a physical mass matrix, and begin to detect and quantify modifications to the structure.

Despite the generally close agreement between the mass and stiffness matrices, the identified damping matrices do not correspond; method 1 gives a negative term on the diagonal of the matrix which would correspond to energy being input to the system at location 3 and is clearly wrong. The inaccuracy of the damping matrices arises because, in a lightly damped structure such as this, the restoring forces associated with the damping are small compared with the stiffness and inertia terms, making the identification of damping parameters from time signals very difficult.

5.4 Validating the Identified Model

Immediately following the DPE procedure, one validity test which can performed is to make a comparison of a measured time signal with data regenerated using the estimated parameters. In this implementation, the acceleration signals are used as the basis of the comparison, and at each location the data used in the model are compared with the regenerated acceleration signal $\hat{\ddot{y}}(t)$ (where the carat denotes an estimated quantity). For example, at the direct point (cf. equation 5.1),

$$\hat{\ddot{y}}_r(t) = \frac{1}{\hat{m}_r}.\left(\hat{c}_{rr}.\dot{y}_r(t) + \hat{k}_{rr}.y_r(t) + \sum_{\substack{p=a \\ p\neq r}}^{N}\left[\hat{c}_{rp}.(\dot{y}_r(t)-\dot{y}_p(t)) + \hat{k}_{rp}.(y_r(t)-y_p(t))\right] - x_r(t)\right) \tag{5.16}$$

The error between the signals can be quantified using the standard, Normalised Mean Square Error (NMSE) criterion, expressed as a percentage,

$$\text{NMSE} = \frac{100}{N.(\sigma_{\hat{\ddot{y}}})^2}\sum_{i=i_0}^{i_1}\left(\ddot{y}(t_i)-\hat{\ddot{y}}(t_i)\right)^2 \tag{5.17}$$

where i_0 is the first data point in the interval, and i_1 is the last.

This test measures the success of the least-squares procedure to fit a model to the time data, and can be used as a basis of comparing the results from similar tests. For example, with the beam rig, the NMSE can be used to compare methods 1 and 2,

Identification Met hod	Percentage NMSE		
	$\ddot{y}_1(t)$	$\ddot{y}_2(t)$	$\ddot{y}_3(t)$
1	0.0190	0.0224	0.4468
2	0.0032	0.0142	0.5177
2 (+ mass)	0.0277	0.0239	1.3040

(5.18)

The NMSE indicates that the model does fit the time data very closely, and that the model identified using method 2 gave a closer fit to the acceleration data at locations 1 and 2 than method 1, although is slightly less accurate at location 3. The second method is likely to be more accurate than the first because the TFs have been averaged to reduce the influence of measurement noise. With the added 1 kg mass, the accuracy of the fit using the second method decreased because of the reduced resolution around the first resonance.

Unfortunately, such a test provides little information about the validity of the parameters, or the ability of the model to predict the response to an alternative excitation signal.

5.4 Validating the Identified Model

Immediately following the DPE procedure, one validity test which can performed is to make a comparison of a measured time signal with data regenerated using the estimated parameters. In this implementation, the acceleration signals are used as the basis of the comparison, and at each location the data used in the model are compared with the regenerated acceleration signal $\hat{\ddot{y}}(t)$ (where the carat denotes an estimated quantity). For example, at the direct point (cf. equation 5.1),

$$\hat{\ddot{y}}_r(t) = \frac{1}{\hat{m}_r}.\left(\hat{c}_{rr}.\dot{y}_r(t) + \hat{k}_{rr}.y_r(t) + \sum_{\substack{p=a \\ p\neq r}}^{N}\left[\hat{c}_{rp}.(\dot{y}_r(t) - \dot{y}_p(t)) + \hat{k}_{rp}.(y_r(t) - y_p(t))\right] - x_r(t)\right) \tag{5.16}$$

The error between the signals can be quantified using the standard, Normalised Mean Square Error (NMSE) criterion, expressed as a percentage,

$$\text{NMSE} = \frac{100}{N.(\sigma_{\hat{\ddot{y}}})}\sum_{i=i_0}^{i_1}\left(\ddot{y}(t_i) - \hat{\ddot{y}}(t_i)\right)^2 \tag{5.17}$$

where i_0 is the first data point in the interval, and i_1 is the last.

This test measures the success of the least-squares procedure to fit a model to the time data, and can be used as a basis of comparing the results from similar tests. For example, with the beam rig, the NMSE can be used to compare methods 1 and 2,

Identification Method	Percentage NMSE		
	$\ddot{y}_1(t)$	$\ddot{y}_2(t)$	$\ddot{y}_3(t)$
1	0.0190	0.0224	0.4468
2	0.0032	0.0142	0.5177
2 (+ mass)	0.0277	0.0239	1.3040

(5.18)

The NMSE indicates that the model does fit the time data very closely, and that the model identified using method 2 gave a closer fit to the acceleration data at locations 1 and 2 than method 1, although is slightly less accurate at location 3. The second method is likely to be more accurate than the first because the TFs have been averaged to reduce the influence of measurement noise. With the added 1 kg mass, the accuracy of the fit using the second method decreased because of the reduced resolution around the first resonance.

Unfortunately, such a test provides little information about the validity of the parameters, or the ability of the model to predict the response to an alternative excitation signal.

A more rigorous validity check would be to identify a parametric model from one test, and then to excite the structure with a different signal and measure the accelerations. Using the parameters estimated from the first test, the acceleration signals for the second case could be predicted via numerical simulation, and compared with the measured response.

In this implementation, the same simulation routines incorporated to check the identification algorithms were used in this prediction test procedure, employing the Bulirsch-Stoer method to integrate the differential equations and incorporating Richardson extrapolation to adapt the stepsize hence improving the accuracy [34]. The only modification is that the measured excitation signal replaces the filtered random number generator, although it is important to remember that the simulation generates transients whenever there is a discontinuity in the input, for example at the start, which would be different from the measured signals. For this reason, the complete 8192-point section of the excitation signal is passed through the simulation, but only the central section of 2048 points is used for comparison (figure 5.9).

Again the NMSE criterion can be used to quantify errors between the measured and predicted signals. With the beam rig, the parametric model identified with method 2 (equation 5.14), which is thought to be more accurate, was used to predict the acceleration responses for the excitation signal used in method 1. At locations 1 to 3 the NMSE errors were 2.346%, 4.720% and 13.110% respectively. These errors are far larger than from the previous validity check, but prediction of the response provides a more realistic method of validating the parameters, since effectively the comparison is made between two independent tests.

A third method of validating the model in the frequency domain rather than the time domain was developed in this project. As mentioned previously, models of linear structures are often compared in terms of the eigen-frequencies, determined from the mass and stiffness matrices,

$$\left|[k] - \omega^2 . [m]\right| = 0 \tag{5.19}$$

Usually the resonant frequencies of the structure can be measured without difficulty. If, however, parameter matrices have been identified, this test can be extended to compare instead the TFs measured from the structure with the FRFs regenerated from the parameter matrices (cf. equation 3.9),

$$[H(\omega)] = \left[(j\omega)^2 . [m] + (j\omega) . [c] + [k]\right]^{-1} \tag{5.20}$$

(The complex matrix inversion on the right hand side can be computed accurately using SVD.)

Usually only part of the full FRF matrix is used for this comparison test, the column which corresponds to the same excitation location as the measured TFs. (This does not preclude the identification being used to predict the TFs measured for a different excitation location.)

Clearly, by disregarding the damping matrix, the simple eigen-frequency comparison is made. However, this test also checks for correspondence at all frequencies across the bandwidth of interest, not only the resonances. For example, the mass and stiffness asymptotes of the FRF can be checked since as $\omega \rightarrow 0, [H(\omega)] \rightarrow [k]^{-1}$ and as $\omega \rightarrow \infty, [H(\omega)] \rightarrow [(j\omega)^2.[m]]^{-1}$.

Plotting the measured TF together with the regenerated FRF illustrates clearly any discrepancies. Spurious low frequency trends on data will have significant influence on the stiffness parameters identified, and, since the model would adjust to give agreement between the resonant frequencies, errors in the stiffness estimates would also affect the mass parameters identified. One of the main reasons for including damping terms in the identification procedure is to enable the comparison to be made across the bandwidth, including the resonance regions, in a similar way to the conventional approach of modal curvefitting. Clearly, the damping matrix identified will be particularly sensitive to discrepancies around the resonances.

For the example of the beam, the FRFs were synthesised using the parameters identified with both methods 1 and 2 (equations 5.13 and 5.14), then taking the column corresponding to the direct location to compare with the inertance TFs measured. In figures 5.10 and 5.11, this comparison is shown for these parameter sets which each include damping for the direct point only, although very similar differences occurred for locations 2 and 3.

For the direct method, there is generally close agreement across the bandwidth (figure 5.10), although there is a small discrepancy around the first resonance, the model resonance occurring at 16.2 Hz which is an error of approximately 4%. The discrepancy extends through the low frequency region and causes the small deviation in the second mode. If this were the only test performed on the structure, non-linearity may be considered to account for this slight error since the time domain model has been restricted to first order. However, the Hilbert transform tests indicate that any non-linearity is weak, and the same discrepancy can be observed with the second method for which the structure was effectively linearised (figure 5.11). Although the difference is more noticeable around the second mode, the predicted resonant frequency of 32.3 Hz being in error by +2%, the likelihood is that it is the error in the first mode which is affecting the results and causing the discrepancy in the central frequency region. In the test repeated with the added 1 kg mass, the errors between the model and the data become more significant in the low frequency region (figure 5.12).

In this case, the predicted first resonance moves to 15.6 Hz (-7%) although the discrepancy in the second resonant frequency remains approximately the same (+2%).

Overall, since the discrepancies which occur are associated with the low frequency region of the bandwidth, the indication is that in each case there may be a small frequency domain resolution problem around the first resonance which remains relatively lightly damped even after adding the constraining layers. The lack of resolution in the measured data in this frequency range is particularly evident from the Nyquist plots (eg. figure 5.11). Since adding the mass has the effect of reducing the frequency of this resonance, the resolution decreases further in this region, and the discrepancy between the measured and regenerated functions increases (figure 5.12). The first mode influences most significantly the low frequency components of the data, and inaccuracies in this region are particularly significant in the displacement data, hence affecting the estimates of the stiffness parameters.

As stated previously, the maximum number of spectral lines was selected, and the only route for reducing the frequency increment would be to reduce the sampling frequency, also decreasing the resolution in the time domain. Nevertheless, despite the resolution problem, the measured and regenerated functions do show reasonably close agreement, at least sufficient to proceed in the attempt to interpret the identified parameters from a physical viewpoint.

5.5 Interpreting the Identified Model

Although the checks detailed in the previous section can be used to quantify the agreement between the measured signals and data regenerated from the identified model, there is no indication that the parameters are unique, or that a physical interpretation can be imposed on the model. One aim of parameter estimation may be to identify a model which can be used to predict the response of a structure to a different excitation. Further applications may be to use the estimated model to predict the effect of modifying the structure, or as a means of updating a FE model. Both these applications regard the structure from a physical viewpoint. Consequently it is necessary to ascertain whether or not the identified parameters can be related to the physical components of the structure.

Repeating the test to identify the 1 kg mass was the first stage in interpreting the model from a physical point of view, as described previously. The aim of this section is to use a theoretical approach to construct physical mass and stiffness matrices and to make a direct comparison with the identified matrices.

This structure is, by design, more amenable than most structures for this type of analysis since it behaves linearly, and the concentrated masses make it possible to treat the structure as a 3 DOF lumped parameter system. However for other structures it may be possible to validate the mass and (linear) stiffness matrices using a similar approach.

The mass of beam (1.26 kg/m) within a distance of $\frac{d}{2}(d$ = 0.254 m) is assigned to each location in addition to the concentrated mass elements (450g + 30g accelerometer), to form the theoretical mass matrix,

$$[m] = \begin{bmatrix} 0 & 0.800 & 0 \\ 0.800 & 0 & 0.640 \end{bmatrix} \text{ kg}$$

To determine a representative stiffness matrix, the flexibility of the beam is considered, initially omitting the effect of the coil spring. Denoting α_{ij} as the deflection at point i per unit force applied at point j, the flexibility equation for a cantilevered beam is,

$$\alpha_{ij} = \frac{1}{6EI}\left([l_i - l_j]^3 - 3l_i[l - l_j]^2 + 3l[l - l_i]^2 - [l - l_j]^3\right)$$

where l_i and l_j are the distances from the free end of the beam of points i and j respectively, and $l = 0.762$ m is the free length of the beam. Since each of the points are equally spaced by length d, the stiffness matrix can be written directly as the inverse of the flexibility matrix,

$$[k] = [\alpha]^{-1} = \left[\frac{d^3}{6EI} \cdot \begin{bmatrix} 2 & 5 & 8 \\ 5 & 16 & 28 \\ 8 & 28 & 54 \end{bmatrix}\right]^{-1} = \frac{6EI}{26d^3} \cdot \begin{bmatrix} 80 & -46 & 12 \\ -46 & 44 & -16 \\ 12 & -16 & 7 \end{bmatrix}$$

where Young's Modulus E = 2.00 x 10^{11}N/m^2, and the second moment of area I = 5.42 x 10^{-10}m^4. The effect of the coil spring is included in the stiffness matrix by increasing the stiffness at k_{33} by 11.0 x 10^3 N/m,

$$[k] = 10^3 \cdot \begin{bmatrix} 122.1 & -70.2 & 18.3 \\ -70.2 & 67.2 & -24.4 \\ 18.3 & -24.4 & 21.7 \end{bmatrix} \text{(N/m)} = \frac{6EI}{26d^3} \cdot \begin{bmatrix} 80 & -46 & 12 \\ -46 & 44 & -16 \\ 12 & -16 & 14.2 \end{bmatrix}$$

To first check the validity of these theoretical estimates, the eigenvalue problem (equation 5.19) can be solved, predicting resonances at frequencies 16.83, 31.06 and 74.86 Hz which correspond very closely to the measured frequencies with percentage errors of $+0.2$, -2.2 and -4.4 respectively.

Matrices based on the physical properties of the system are bound to be inaccurate. In particular, attributing sectional masses of the beam to the locations of interest can only be approximate, since a distributed parameter system will always behave differently from the lumped parameter model. In comparison, the mass estimates do correspond well at locations 1 and 2, although there are appreciable differences between the matrices at location 3, the identification giving a relatively high overestimate of this parameter. Perhaps more confidence can be placed in the theoretical stiffness matrix since it is derived in terms of flexibility of, in effect, a distributed parameter beam. Again, the identified stiffness matrix is found to correspond with the estimated matrix very well around locations 1 and 2, but the parameters associated with location 3 are significantly overestimated.

Clearly, an overestimate in both the stiffness and mass matrix can still allow the model to predict the resonant frequencies accurately. Unfortunately if the stiffness matrix is overestimated, the mass matrix will also be overestimated because the DPE attempts to fit to the time signals which are dominated by the components around the resonances, in particular the first.

In this case, the theoretical model cannot be regarded as being perfectly reliable because the approach for developing the mass and stiffness matrices was rather simplistic. Apart from attempting to represent the mass of the beam as a distributed parameter system, modifications to the model could include refining the flexibility analysis to provide a more realistic account of the influence of the coil spring connected at location 3.

To determine whether or not the DPE procedure can reliably identify physical mass and stiffness matrices from structures which do not necessarily resemble the lumped-parameter linear model, the procedure must be applied to more complicated, distributed parameter structures which also may be non-linear. This subject is addressed in the next two chapters.

Chapter 6

Tests on a Non-Linear Beam Rig

6.1 Overview

Whereas the experiment discussed in chapter 5 concentrated on a beam rig which had been designed to be nominally linear and physically resemble a lumped parameter model, in this chapter a more flexible beam without lumped masses is investigated (figure 6.1). Since the beam is flexible, non-linearity can be introduced by exciting the beam so that the deflections are large. From theory, the stiffness restoring force function has been shown to approximate to a symmetrical polynomial with first and third order terms, verified by performing a simple static test. Alternatively, by preloading the beam using a spring in tension, the equilibrium position can be offset from the point of inflexion causing the stiffness function to become asymmetrical and introducing a quadratic term into the governing polynomial.

A principal aim of this chapter is to demonstrate that this relatively simple structure behaves non-linearly. Clear higher order TFs can be measured from the beam for a range of excitation levels, each exhibiting the distortion characteristics of a structure governed by a non-linear stiffness function. Concentrating on the case with preload applied, the two approaches to DPE described in chapter 5 were employed, and by regenerating the first order FRFs, the two methods are compared again, but in the context of a structure which is non-linear. The direct method can also be extended to identify non-linear terms in the model, which can be used to regenerate the principal diagonals of the FRFs for comparison with the measured TFs as discussed in chapter 4. This new approach to the identification of structures has been developed in this project, and details of the implementation are included throughout this chapter.

6.2 The Non-Linear Beam Rig

The rig consists of a flexible mild steel beam with thickness $\frac{1}{8}''$ (3.17 mm), width 1" (25.4 mm) and length (between the supports) of 1.2 m supported at both ends by brackets fixed to a rigid test bed, as shown in figure 6.1. Preload can be introduced by connecting a coil spring in tension at centre span and, in the tests described subsequently, applying preload caused the centre of the beam to deflect transversely by approximately 5 mm. The shaker was connected 0.2 m from one end of the beam via a link sufficiently flexible to avoid additional stiffening and to reduce any problems of shaker-structure interaction, a Ling type V201 electrodynamic shaker was used, the smallest available for the experiment. The transducers and all other instrumentation were retained from the previous test (chapter 5).

Five response locations were selected at symmetrical locations along the beam including the direct point (1) and the mid-point (3). Unfortunately, the accelerometers each have a mass of 30 g which is significant compared to the mass of the beam (0.78 kg), and the dynamics of the structure are unavoidably affected by this mass-loading (although this does not alter the validity of the subsequent discussion.)

Direct measurement of the higher order TFs employs sine excitation and requires the structure be oscillating in steady-state, ie. after each increment in frequency the transients must be allowed to die away before the measurements are taken. The steel beam was very lightly damped and so, to improve energy dissipation and shorten the delay between measurements, constrained layer damping was added using $\frac{1}{8}''$ thick foam tape to adhere a $\frac{1}{32}''$ thick strip of steel, length 20" (508 mm), across the centre-section of the beam, on the opposite side to the tension spring connection. Although the constraining layer tends to increase the stiffness and mass of the beam in addition to the damping, the general discussion is again unaffected.

Using broadband random excitation set to 1V RMS, the beam was tested to determine the resonant frequencies both with and without preload, and the inertance TFs after 10 averages are shown in figure 6.2. Apart from altering the form of the stiffness restoring force function, the effect of the tension spring was to stiffen the entire beam, most obviously affecting the first resonance. Upto 200 Hz, five clear resonances were apparent at frequencies 9.7, 25.4, 56.7, 95.3 and 127.2 Hz, which correspond to 20.4, 32.9, 64.9, 101.6, and 134.4 Hz when the spring was connected and the beam preloaded. In addition, at 86.7 Hz a very lightly damped wave resonance of the tension spring can be observed. In both cases, location 3 is located close to the nodal point of the 2^{nd} and 4^{th} modes. (Below 6.5 Hz, the inertance TFs measured from location 1, the direct point, exhibit an unusual low frequency trend which is

less significant on TFs measured at the other locations. Since the excitation signal was not filtered in this test and the accelerometers used are reliable at frequencies above 2 Hz, the reason for this trend was not clear.) In subsequent sections of this chapter, locations 1,2 and 3 have been chosen to illustrate the results. Since the beam is geometrically symmetrical, with the exception of the link to the shaker, the results at locations 4 and 5 resembled 2 and 1 respectively.

The tests focussed on the first 3 resonances of the preloaded beam, over the bandwidth 0 to 80 Hz, since the stiffness non-linearity has the most influence in the low frequency region. Before proceeding, it is appropriate to mention two tests which were conducted to confirm that the beam behaved non-linearly when excited by a dynamic load.

The first test compared the TFs measured using random excitation, bandlimited between 6 and 80 Hz, after 10 averages at 2 levels (figure 6.3). 1V RMS was considered to be a relatively low level with maximum deflection limited to 1.0 mm, whereas 4V RMS was a relatively high level. Despite the stiffness non-linearity, the frequencies of the resonances did not change noticably as the level was increased because random excitation is effective in linearising the behaviour of the structure. However a significant difference between the coherence functions is evident, and the apparent increase in the noise on the data can be attributed to the effect of the non-linearity on the linearised TF (section 1.3).

In general, sinusoidal excitation is considerably more effective for detecting non-linearity. In the second test, first order inertance TFs were measured using sine excitation at two different amplitudes, 0.1V and 0.4V, stepping in frequency between 2 and 80 Hz as shown in figure 6.4. At each response location, the distortion is significant, and around the first resonance there is a bifurcation even at the lower of the two levels. Whereas the resonances 2 and 3 tend to increase in frequency as the amplitude level increases, usually an indication of a hardening stiffness non-linearity, the first resonance distorts towards the lower frequencies. This effect can be explained by the presence of the second order term in the stiffness function which is dominant when the beam is preloaded, as discussed later in this chapter.

The distortion was also detected using the Hilbert transform applied to the first order TFs measured at the lower level (figure 6.5). The fact that the transforms do not overlay the measured data is sufficient to indicate that the structure is non-linear, but little can be concluded about the form of the non-linearity despite the structure being conceptually very simple, particularly around the first resonance where the bifurcation and lack of resolution conspire to make further interpretation difficult.

6.3 Theoretical Model of the Beam Dynamics

A theoretical development of the dynamics of the beam can be made for the first resonance to show how non-linear restoring forces arise for large deflections, using the Rayleigh energy principle and the Euler-Lagrange equations to relate the kinetic energy of motion to the potential strain energy. Since the aim is to derive the non-linear stiffness function and estimate the first resonant frequency of the structure, energy dissipated due to damping is not considered in the method.

The basis of the procedure is to impose a dynamic deflection shape to represent the oscillation of the structure, and in this example, to correspond with the first resonance of a clamped-clamped beam, the appropriate shape function is expressed as a function of y, the instantaneous deflection at the ***centre*** of the beam during vibration [66],

$$\text{Shape Function, } v = \frac{Y}{2}\left(1-\cos\left(\frac{2\pi u}{l}\right)\right) \tag{6.1}$$

$$\frac{\partial v}{\partial y} = \frac{1}{2}.\left(1-\cos\left(\frac{2\pi u}{l}\right)\right) \tag{6.2}$$

$$\frac{dv}{du} = \frac{Y}{2}.\left(\frac{2\pi}{l}\right)\sin\left(\frac{2\pi u}{l}\right) \tag{6.3}$$

$$\frac{d^2v}{du^2} = \frac{Y}{2}.\left(\frac{2\pi}{l}\right)^2\cos\left(\frac{2\pi u}{l}\right) \tag{6.4}$$

To develop an expression for the equation of motion in terms of the parameters of the beam, one can consider the kinetic energy of a small element of the beam, length δu, to construct the integral for the total kinetic energy across the length 1 of the beam.

Denoting m' as the *mass* per ***unit length,***

$$\text{Kinetic Energy, } T = \frac{m'}{2}\int_0^l\left(\frac{dv}{dt}\right)^2.du \tag{6.5}$$

$$= \frac{m'}{2}\int_0^l\left(\frac{dy}{dt}.\frac{\partial v}{\partial y}\right)^2.du \tag{6.6}$$

$$= \frac{m'.\dot{y}^2}{8}\int_0^l\left(1-\cos\left(\frac{2\pi u}{l}\right)\right)^2.du \tag{6.7}$$

$$T = \frac{3m'l\dot{y}^2}{16} \tag{6.8}$$

During oscillation, two independent strain energy terms can be considered, one arising from the bending of the beam, and the second due to tension. Consequently the total strain energy can be written,

$$\text{Strain Energy, } U = \text{ Bending Strain Energy + Tension Strain Energy} \tag{6.9}$$

$$= \frac{EI}{2}.\int_0^l \left(\frac{d^2v}{du^2}\right)^2 .du + \frac{EAl}{2}.\left(\frac{1}{2l}\int_0^l \left(\frac{dv}{du}\right)^2 .du\right)^2 \tag{6.10}$$

$$= \frac{EI}{2}.\int_0^l \left(\frac{y}{2}\right)^2 \left(\frac{2\pi}{l}\right)^4 \cos^2\left(\frac{2\pi u}{l}\right).du + \frac{EA}{8l}.\left(\int_0^l \left(\frac{y}{2}\right)^2 \left(\frac{2\pi}{l}\right)^2 \sin^2\left(\frac{2\pi u}{l}\right).du\right)^2 \tag{6.11}$$

$$= \frac{EI}{2}.\left(\frac{y}{2}\right)^2 \left(\frac{2\pi}{l}\right)^4 \left[\frac{l}{2}\right] + \frac{EA}{8l}.\left(\frac{y}{2}\right)^4 \left(\frac{2\pi}{l}\right)^4 \left[\frac{l}{2}\right]^2 \tag{6.12}$$

$$U = \frac{EI\pi^4.y^2}{l^3} + \frac{EA\pi^4.y^4}{32l^3} \tag{6.13}$$

If the excitation force is denoted by $x(t)$, the Euler-Lagrange equation relating kinetic and strain energies is,

$$\frac{d}{dt}\left(\frac{dT(\dot{y})}{d\dot{y}}\right) + \frac{dU(y)}{dy} = x(t) \tag{6.14}$$

$$\frac{3m'l}{8}.\ddot{y}(t) + \frac{2EI\pi^4}{l^3}.y(t) + \frac{EA\pi^4}{8l^3}.y^3(t) = x(t) \tag{6.15}$$

which has the familiar differential equation of motion form,

$$m.\ddot{y}(t) + k_1.y(t) + k_3.y^3(t) = x(t) \tag{6.16}$$

where k_1, the first order stiffness, has arisen from the *bending* strain energy term, and k_3, the cubic stiffness, is due to the *tension*.

Calculating the parameters for the beam,

$$\begin{aligned} m &= \frac{3m'l}{8} = 0.29 \text{ kg} \\ k_1 &= \frac{2EI\pi^4}{l^3} = 1432.9 \text{ N/m} \\ k_3 &= \frac{EA\pi^4}{8l^2} = 106.6 \times 10^6 \text{ N/m}^3 \end{aligned} \tag{6.17}$$

the fundamental resonant frequency of the beam, which corresponds to the natural frequency for small deflections, can be estimated,

$$\omega_0 = \sqrt{\frac{k_1}{m}} = 70.3 \text{ rad/s} = 11.2 \text{ Hz} \tag{6.18}$$

The estimated 'modal' mass is $\frac{3}{8}$ x total mass of the beam and, to account for the mass-loading effect of the accelerometers which are evenly spaced along the beam, the modal mass can be considered to increase to 0.35 kg, giving a revised estimate of 10.2 Hz for the first resonance of the structure. The first resonance measured from the beam occurred at 9.7 Hz but, since the beam is MDOF, this is a coupled 'mode', incorporating mass residual effects from the second and higher modes. The coupling tends to cause a slight decrease in the resonant frequency; nevertheless, the theoretical value is in error by less than 5 %.

If a spring of stiffness k_0 is connected at midspan, causing a deflection δ of the midpoint, it is possible to show that the equation of motion becomes,

$$m.\ddot{y} + k_0.(y+\gamma) + k_1.(y+\delta) + k_3.(y+\delta)^3 = x(t) \quad (6.19)$$

$$m.\ddot{y}(t) + (k_0\gamma + k_1\delta + k_3\delta^3) + (ko + k_1 + 3k_3\delta^2).y(t) + 3k_3\delta.y^2(t) + k_3.y^3(t) = x(t) \quad (6.20)$$

where γ is the length the spring extends when preloading the beam.

In this expression, the conditions of equilibrium require that,

$$k_0\gamma + k_1\delta + k_3\delta^3 = 0 \quad (6.21)$$

which simplifies the non-linear equation of motion to,

$$m.\ddot{y}(t) + (k_0 + k_1 + 3k_3\delta^2).y(t) + 3k_3\delta.y^2(t) + k_3.y^3(t) = x(t) \quad (6.22)$$

$$m.\ddot{y}(t) + \beta_1.y(t) + \beta_2.y^2(t) + \beta_3.y^3(t) = x(t) \quad (6.23)$$

Linearity of the tension spring has been assumed, and in an auxiliary test, the spring was found to deflect approximately 15 mm per 1 kg under the conditions of the preload, which corresponds to a stiffness of 650.0 N/m. When setting up the tests, the spring was adjusted such that $\delta \approx 5$ mm, and correspondingly first order stiffness parameter, β_1, can be calculated,

$$\beta_1 = k_0 + k_1 + 3k_3\delta^2 = 650.0 + 1432.9 + 7950.0 = 10032.9 \quad (6.24)$$

This stiffness would result in a fundamental resonant frequency at 26.9 Hz, which compares with the measured first resonance at 20.4 Hz. This comparison indicates that k_3 has been overestimated in the theoretical model and, from equation 6.24, it is immediately evident that the stiffness term is extremely sensitive to errors in δ which can only be measured approximately in practice.

Clearly it was the stiffening non-linear nature of the beam, and not the stiffness of the tension spring itself, which caused the significant rise in in the first resonant frequency which more than doubled when the preload was applied.

For structures as simple as the flexible beam, one effective method of identifying non-linearity in the stiffness restoring force function is to perform a static test, applying a static load to the beam and measuring the deflection. Attaching a cable to the centre of the beam, known masses were added to the free end of the cable passing over a pulley as shown in figure 6.6, and the deflection of the centre of the beam, y, was measured using a dial gauge. Two similar tests were conducted, the second with the tension spring connected, applying the static load in the opposite direction to the preload. For the condition when the two loads are equal, the beam returns through the point of inflexion, and there is a noticeable discontinuity in the load-deflection curve probably caused by the beam tending to buckle (figure 6.7).

From the static test data, the stiffness parameters can be identified directly by fitting polynomial functions to the load-deflection curves,

1. With no preload, the function $f(y) = \alpha_1 y + \alpha_3 y^3$ was fitted (figure 6.8),

$$\begin{aligned} \alpha_1 &= 0.174 \text{ kg/mm} = 1710.0 \text{ N/m} \\ \alpha_3 &= 6.46 \times 10^{-3} \text{ kg/mm}^3 = 63.4 \times 10^6 \text{ N/m}^3 \end{aligned} \tag{6.25}$$

   Referring to the equation of motion (eqn. 6.16), α_1 corresponds to k_1 and α_3 to k_3. When compared with equation 6.17, k_1 was underestimated in the theoretical model, but k_3 was overestimated. Incidentally, when a second order term α_2 was included in the polynomial fit, α_1 remained unchanged, α_3 increased by only 0.5 %, and $\alpha_2 = 0.11 \times 10^{-3}$ kg/mm^2, which is very small and verifies that the polynomial is approximately symmetrical, accurately represented by first order plus cubic terms.

2. With preload, the function f(y) = $\beta_1 y + \beta_2 y^2 + \beta_3 y^3$ was fitted over the region above the buckling discontinuity (figure 6.9),

$$\begin{aligned} \beta_1 &= 0.882 \text{ kg/mm} = 8650.0 \text{ N/m} \\ \beta_2 &= 0.123 \text{ kg/mm}^2 = 1200.0 \times 10^3 \text{ N/m}^2 \\ \beta_3 &= 6.14 \times 10^{-3} \text{ kg/mm}^3 = 60.2 \times 10^6 \text{ N/m}^3 \end{aligned} \tag{6.26}$$

   Referring to the equation of motion (eqn. 6.22), each of these three coefficients can be used to estimate k_3. From β_1, and using the value of k_1 estimated in the previous test (eqn. 6.25), k_3 is estimated to be 83.9 x 10^6 N/m^3.
   This compares with 80.1 x 10^6 and 60.2 x 10^6 N/m^3 from β_2 and β_3 respectively.

The polynomial functions correspond well to the measured data in both cases.

6.4 Higher Order Transfer Functions

6.4.1 Introduction

In section 6.2, two conventional dynamic tests for detecting non-linearity were applied to the beam. A third possible test would be to excite the beam, again with a single sinusoid, and measure for example the displacement response at a frequency close to the fundamental resonance. This test was performed on the preloaded beam and figure 6.10 shows sections of response data measured from the direct location using 10 Hz sine excitation at the lower amplitude of 0.1V. From the corresponding frequency spectra, it is evident that the higher harmonic content of the signals is considerable. Furthermore, the relative size of the third harmonic to the second harmonic provides evidence of the relatively strong cubic term in the polynomial, as predicted by theory and observed here and in the static tests.

As mentioned previously, the existence of harmonic terms in itself indicates that the structure is non-linear. However, by quantatively relating each of the higher harmonic response components to the sinusoidal force, a series of Transfer Functions (TFs) can be measured to characterise the energy transfer taking place. As described in chapter 4, these higher order TFs provide a great deal of information regarding the influence of the non-linearity present across the bandwidth of interest.

The aim of this section is to describe how the TFs defined by equation 4.6 (or 4.9) can be measured from a structure in practice, including details of the implementation using the computer-controlled data acquisition system (introduced in section 5.2.1). Initially, however, it is appropriate to illustrate the higher order TFs determined by simulation, corresponding to the parameters identified in the previous section.

6.4.2 Simulating the Higher Order TFs

In order to illustrate the distortion in the first and higher order TFs measured using sine excitation, particularly around the first resonance of the beam, two non-linear SDOF systems were simulated using the parameters identified from the static tests. In both systems, the simulated stepped sine test was repeated, incrementing the amplitude of the sinewave input X to broadly coincide with the range of the static test. Also a small viscous damping term (c = 3 Ns/m) was incorporated to dissipate transients in the simulation.

System 1: Without preload (cf. equation 6.25),

$$m.\ddot{y}(t) + c.\dot{y}(t) + \alpha_1.y(t) + \alpha_3.y^3(t) = x(t)$$

in which, m = 0.35 kg, α_1 = **1.7 x** 10^3 N/m, α_3 = 6.5 x $10^7\,N/m^3$.
The linear viscous damping term corresponds to the modal coefficient, $\zeta = \frac{c}{2\sqrt{\alpha_1.m}} \approx 6.2\%$,

The first and third order inertance TFs are shown in figure 6.11 for X=0.2, 0.4, 0.6.

The second order TF is not shown since, as explained in chapter 4, a system which is governed by symmetrical polynomial functions has null even order FRFs. Therefore the second order FRF is null and, since there can be no degenerative contributions from higher, even order FRFs, the second order TF is also null.

The fundamental resonance at 11 Hz is clearly evident on both the first and third order functions, and for increasing excitation amplitude, the distortion of the peak is toward higher frequency, characteristic of the hardening stiffness type non-linearity, with bifurcation occurring when X exceeds approximately 0.5 N. The third order TF also shows the tertiary resonance clearly, but there is no secondary resonance because the governing polynomial has odd order symmetry.

System 2: With preload (cf. equation 6.26),

$$m.\ddot{y}(t) + c.\dot{y}(t) + \beta_1.y(t) + \beta_2.y^2(t) + \beta_3.y^3(t) = x(t)$$

in which, β_1 = 8.6 x 10^3 N/m, β_2 = 1.2 x $10^6\,N/m^2$, β_3 = 6.0 x $10^7\,N/m^3$.
In this system, the damping term corresponds to the coefficient, $\zeta = \frac{c}{2\sqrt{\beta_1.m}} \approx 2.7\%$.

The first, second and third order inertance TFs for system 2 are shown in figure **6.12** for X=1.0, 2.0, 3.0.

The general effect of increasing the stiffness is to move the fundamental resonance to 27 Hz, and each TF exhibits primary, secondary and teriary resonances accordingly. However, as the amplitude is increased, the primary resonant peak distorts toward lower frequency in each function which indicates that, over this range of amplitudes, the quadratic term introduced into the function gives rise to a softening stiffness effect. From the measured load-deflection curve (figure 6.9), it is evident that when the beam moves towards the equilibrium position (ie. the original state of the beam without preload), the tension in the beam decreases and the stiffness function tends to soften. In each TF, bifurcation occurs as X exceeds approximately 1.8 N.

6.4.3 Measuring the Higher Order TFs

As mentioned in chapter 3, the simulation routines were written to emulate the measurement procedure, generating a high resolution sinewave, inputting this signal into the system, and measuring periods of the steady-state response signals. In practice, it is possible to perform these tasks with the DIFA SCADAS 'front-end' employed. Also, using this equipment, data acquisition can be synchronised with the generation of the sinewave excitation signal so that the Fourier transforms, used to calculate the spectra and then the higher order TFs, operate on sections of the measured signals containing whole periods of the waveform. In so doing, the problems of leakage are circumvented.

An accurate sinewave can be generated easily using the DAC in the front-end using the LMS software. Firstly, the spectrum of the sinewave is created in a frequency domain data block, consisting of a single 'spike' at one frequency line. By performing the inverse Fourier transform, a section of time data containing whole sinewaves is formed. This implementation requires that all block lengths be a power of 2 (ie. 2^i, where i is an integer) to a maximum of 8192 real lines (time domain) or 4096 complex lines (frequency domain). For example, a 4096 complex block is initialised, with all real and imaginary parts set to zero; then the real value of line 1 (line 0 is the d.c. line) is set to a value representing the voltage amplitude of the sinewave to be generated. Using the inverse Fourier transform block operation, 8192 points of time data are created which correspond to one period of a sinewave.

The sinewave time data block is then loaded from the computer into the buffer memory of the front-end to be repeatedly and continuously output from the DAC at a controlled rate. Since this signal is used to drive the shaker via a power amplifier, the DAC output rate determines the frequency at which the structure is excited. For this application, an excitation signal generated within the' front-end has the important advantage over an externally generated signal in that, because the front-end is also used to acquire the data, leakage is avoided by matching the ADC and DAC rates, easily achieved in practice. Consequently, the harmonic components of each response signal can be resolved very accurately. However, if there were a mismatch between the rates of signal generation and acquisition, each interval of time data measured would include fractions of periods, and the significance of the leakage problem would become proportionally worse for increasing harmonic orders.

In practice, the hardware imposes a maximum on the ADC and DAC rates, and hence the frequencies of sine excitation which can be achieved. Furthermore, since the front-end employs multiplexing, the maximum ADC sampling frequency is also governed by the number of acquisition channels selected. The multiplexer has a sampling time of 3 microseconds

per channel plus an operational 'overhead' of 6 microseconds when more than one channel is selected. For example, if N is the total number of response locations of interest, $N+1$ channels are required (the excitation force is also measured), and the maximum ADC rate is limited to $\frac{10^6}{3.(N+1)+6}$ Hz. Correspondingly, if measurements from 3 locations were taken, the maximum ADC or DAC rate would be limited to 55.56 kHz. Consequently, a sinewave resolved by the maximum 8192 lines per period would be limited to a frequency of $55,560/8192$ = 6.78 Hz. Clearly, to increase the upper limit on the generated sinewave frequency, it is necessary to compromise by reducing the resolution from 8192 points progressively by factors of 2. Down to the minimum of 16 lines/period, a smooth sinewave can still be produced since the front-end incorporates smoothing filters on the DAC outputs (and additional reconstruction filters could be used.) However the resolution of acquisition is also affected, and the realistic lower limit is 32 lines per fundamental period if up to third order harmonics are to be estimated from the spectra.

After each frequency increment, the transient vibration of the structure is left to die away simply by introducing a delay between signal generation and acquisition stages, before the sections of time data are measured and transferred to the computer. The TFs are constructed incremently, each different frequency of excitation producing a line on the TFs of each order. Then the generated sinewave is stepped up (or down) to the next frequency, and the measurements repeated. Clearly, the better the accuracy with which each measured time signal can be resolved, the higher the order of TF which can be estimated.

The front-end has optional analogue filters on each of the DAC channels and ADC channels. The DAC filters act to smooth the generated signals formed digitally with zero order hold, whereas the ADC filters reduce the effect of high frequency experimental noise on the measured signals. However care is needed to ensure the relevant higher harmonics are retained in the response signals. If the structure is excited at a frequency f Hz, and the TFs up to n^{th} order are of interest, the ADC filters must be set to a low-pass frequency exceeding $n.f$ Hz to retain the harmonics.

In addition, the roll-off characteristic, inherent in the analogue filters, is very important since the effect introduced is manifest both as a distortion in the gain and a shift in phase, the errors becoming more significant as the filter cut-off frequency is approached. When only the conventional, first order TF is of interest, usually it is sufficient to set filter low-pass cutoffs to be the same for the response and excitation channels. In this way, as the fundamental spectral component of the output is divided by the input component at the same frequency,

the roll-off effects tend to cancel each other. However in the higher order TFs, response components at higher harmonic frequencies are compared with the fundamental component of the excitation signal, and consequently the low-pass frequency of each filter must be selected to ensure the roll-off effects are negligable across the bandwidth of interest. The ADC filter setting of 1 kHz was selected since TFs up to order three are of particular interest, and each structure was tested across a bandwidth not exceeding 250 Hz.

All the frequency domain analysis procedures, such as the Hilbert transform routine, assume *equal* frequency spacing between each spectral line of the measured TFs, and that the TFs of all orders from all response locations coincide on an equally-spaced frequency 'grid'. However, sinewaves can seldom be generated by the DAC at exactly the specified frequency since the operation is digital, and the DAC rate is governed by the front-end 'clock' frequency. Considerable differences can arise between the desired and actual frequencies, particularly at high frequencies, and this effectively limits the minimum frequency increment which can be achieved. To overcome the problem, an auxiliary routine was incorporated to interpolate each of the measured TFs onto a designed frequency grid succeeding the testing stage. As mentioned previously, the measurement instrumentation, and particularly the accelerometers, can be inaccurate at low frequencies. Although the analysis procedures may also require the TFs to be baseband, a structure is often tested over a zoomed bandwidth with a stepped sine procedure. Consequently the 'grid' interpolation routine also has the option of extrapolating zoomed TFs onto a baseband frequency range.

One of the principal objectives of this chapter is to demonstrate that higher order TFs can be measured in practice from structures. Figures 6.13 to 6.15 show first, second and third order inertance TFs measured from locations 1, 2 and 3 using sine excitation at the lower excitation amplitude, 0.1V. At each frequency increment, a 10 second delay was allowed for the transients to die away, and then 50 consecutive periods of time data were averaged before the spectra computed. The high quality of the TFs, even away from resonance where the response is small, provides motivation for developing this approach, particularly since it can be regarded as a straightforward extension to conventional testing procedures. Higher order TFs are measured at the same instant as the conventional (first order) TF with hardly any additional effort.

When using this method it is imperative to ensure that the input to the system does approximate closely to a sinewave. In particular, any harmonic distortion on the excitation signal, as measured by the force transducer, would jeoparise the method. Higher harmonic components of the output due to non-linear interactions with the fundamental input component would coincide with contributions from the linear interactions with the input harmonics.

In forming the higher order TFs, the linear interaction contributions would be uncorrelated with the fundamental input component and be manifest as noise. Theoretically this effect can be reduced by averaging, although in practice it is best to improve the experimentation to minimise harmonic components in the measured force signal. (The problem is similar to the one encountered in measuring second order Wiener kernels with random excitation where the first order components dominate the response and prevent the second order kernel being determined by correlation, as described in section 2.3.3.)

The higher harmonic input components may be generated by the structure interacting with the shaker, and because the structure is non-linear, the force transducer picks up the harmonics in the response. Usually, however, this situation can be improved by matching the shaker with the structure to reduce interaction. In this implementation, the generated signal has constant voltage amplitude and consequently the structure is excited with nominally the same level of forcing for all frequencies, although shaker-structure interaction may result in some drop-out particularly around the resonances, as shown in figure 6.16.

Often non-linear structures are governed by, for example, stiffness functions which may only approximate to the polynomial functions assumed in the mathematical models, particularly if there are clearances or other discontinuous non-linear effects. Whereas close to resonance bifurcation may occur, away from resonance the non-linearity will not be excited strongly as the structure tends to linearity, particularly when close to an anti-resonance. Nevertheless, the stepped sine method of measuring the TFs using constant input amplitude assumes the non-linear function approximates to a continuous polynomial for all responses levels, relying on the higher harmonic response components being resolvable at all frequencies across the bandwidth.

Although the primary resonances are usually discernable on each of the higher order TFs, the secondary and tertiary resonances may coincide with frequencies at which the response is relatively low level, making the harmonic terms indistinguishable from the background noise level. Consequently, the TFs become noisy in these regions rather than having the characteristic form predicted by theory. Future developments of this stepped sine technique could attempt to circumvent this problem by controlling the signal generation so that the structure responds at nominally the same level for each frequency, as described in section 4.3. However, the problem is not paricularly evident on the TFs measured from the preloaded beam (figures 6.13 to 6.15) since the non-linear stiffness function does approximate well to a continuous polynomial as seen from the static test results (figures 6.8 and 6.9).

The measured TFs clearly demonstrate the energy transfer pattern of non-linear systems, described in chapter 4 with relation to simulated systems. The bifurcation of the first resonance is evident in the Nyquist plot of the TF of each order (eg. figure 6.13), and the primary and secondary resonances are obvious on each second order TF. Similarly the primary and tertiary resonances are very clear on each third order function although, for the three fundamental resonances in the range, the secondary resonances do not appear. This is particularly evident from the functions in figure 6.15 which were measured from the midpoint of the beam (location 3) corresponding to the node of the second beam bending mode. Consequently only the first and third fundamental resonances (at approximately 20 and 65 Hz) are apparent on the TFs measured from this location. The second order function has secondary resonances at approximately 10 and 32 Hz, and the tertiary resonances on the third order function occur at approximately 7 and 22 Hz. Clearly the third order TF lacks secondary resonances, and the small peaks at approximately 34 and 45 Hz correspond to the tertiary resonances of the fourth and fifth fundamental resonances (cf. figure 6.2).

In order to illustrate the non-linearity of the beam without preload, the stepped sine test was repeated after the tension spring had been removed. The first, second and third order TFs measured from the direct location are shown in figure 6.17. Theory predicts that, if the second order term of the polynomial were zero, the second order FRF would be null. Despite this the governing non-linear function is never exactly symmetrical in practice and a small quadratic term arises in the stiffness function even when the beam is not preloaded. Correspondingly, clear second order TFs were measured which show primary resonances at approximately 10, 25 and 57 Hz and secondary resonances at 5, 13 and 29 Hz. The third order functions also show the primary resonances of the three fundamental resonances, whereas the tertiary resonances at approximately 3, 9, 19, 32 and 42 Hz correspond to the first five fundamental resonances. Again the secondary resonances are not apparent on the third order TFs because the quadratic term in the governing stiffness function was small.

Finally, to demonstrate that the higher order TFs exhibit significant distortion at higher levels of excitation, the stepped sine test was repeated at the amplitude 0.4V, and a comparison made between the TFs at the two levels (figure 6.18). The distortion patterns can be seen to correspond in each of the TFs as described previously. In particular the second order term caused the magnitude of each TF around the first resonance to distort toward lower frequency, and bifurcation occurred around the first primary resonances at both levels of excitation. Similar effects were observed in the simulation study of the previous section. These results also help to establish that the TFs do approximate to 'the diagonals of the FRFs of the structure which are defined to be independent of the level of excitation.

6.5 Identifying a Parametric Model

6.5.1 Introduction

The DPE procedure was described in detail in chapter 1, and in chapter 5 the technique was used to identify a first order model from a linear structure. One of the aims of this section is to apply the two methods of measuring representative time data to the preloaded beam, a structure which has been identified as being non-linear. Initially an attempt is made to determine a model with first order stiffness functions, before proceeding to extend the DPE procedure to identify non-linear stiffness terms. Chapter 4 concluded by arguing that higher order TFs measured using the stepped sine procedure may have an application in validating the terms in the identified non-linear model, and this subject is also addressed with reference to results from the non-linear beam rig.

6.5.2 Identifying a First Order Model

Both approaches to DPE presented in section 5.3 were applied to the preloaded beam, attempting to identify a model for the structure which included first order terms only. In each test, the broadband random excitation signal was set at the lower level of 1V RMS, using an operational bandwidth of 6-80 Hz to encompass the 3 resonances of interest whilst avoiding the region of the low frequency trend observed on the inertance TF measured from the direct location.

Firstly acceleration was measured from each of the 5 locations, collecting 8192 of time data points from each of the 6 channels using a sampling frequency of 1000 Hz. Following integration and filtering of the data, the DPE procedure was applied, restricting the solution to first order terms only. The second test used the same experimental setup but inertance TFs were measured using random excitation from each location and then representative time data was calculated by specifying a force spectrum with constant magnitude across the bandwidth 6 to 80 Hz.

Using the models identified from each of the two methods, the first order FRF matrix was regenerated, and a comparison was made with measured first order inertance TFs, as shown in figures 6.19 and 6.20. Using method 2, the results correspond closely with the measured data, and the only noticeable discrepancy occurs in the low frequency region where the measurements appear to be spurious particularly at the direct point, as already noted.

The results using method **1** also agree reasonably with the measurements although less accurately than with method 2. This difference is also reflected in the NMSE criterion used to compare acceleration signal regenerated from the identified model and the data used in the estimation,

Identification Method	Percentage NMSE					
	$x(t)$	$\ddot{y}_1(t)$	$\ddot{y}_2(t)$	$\ddot{y}_3(t)$	$\ddot{y}_4(t)$	$\ddot{y}_5(t)$
1	0.6354	0.4207	0.1591	9.0797	0.9197	0.8638
2	0.5683	0.2961	0.0111	0.0352	0.2507	0.1796

(6.27)

Consistently the model identified using the indirect method agrees more closely with the measurements, **particularly** at location 3, the centre point of the beam. In chapter 5, the results indicated that for a linear system more accurate results can be obtained using the indirect method since the time data is generated from **TFs** which have been averaged, and hence measurement noise affect the solution minimally. However, in this example of a non-linear structure, an alternative account for the differences is proposed.

Using the data generated from the measured **TFs,** the model identified via the indirect method represents the linearised structure, and consequently the appropriate mathematical model includes only first order parameters. For this reason, the first order model corresponds well with the measured, linearised **TFs** even if the structure is non-linear. However, for a non-linear structure, identifying a first order model directly from measured time data is only appropriate when the excitation level is sufficiently low that the structure tends to behave linearly. Otherwise, the parametric model should be extended to include non-linear terms.

The lower level of excitation was used in the tests to promote linearity, although the relatively poor correspondence between the measured TF and the FRF regenerated from the parameters identified using method 1 would indicate that the structure **still** behaves appreciably non-linearly at this level. Certainly, when evaluating the direct method in this way with the preloaded beam, the DPE procedure proved to be very sensitive to the level of excitation applied, whereas for method 2 it was not. The model identified with the indirect approach is expected to be relatively independent of the level of excitation, since **usually** the **TFs** measured from a weakly non-linear system using broadband excitation hardly change as the level of excitation is increased (provided a sufficient number of averages have been taken to ensure the non-linearity is linearised). More extensive studies with strongly non-linear systems, which may include clearances for example, would be required to substantiate these conclusions as being generally applicable to all structures.

6.5.3 Identifying a Higher Order Model

All the tests which have been outlined in this chapter have indicated that the structure has a stiffness restoring force function which is not linear, and that significant non-linear behaviour is apparent even at the lower levels of excitation. To further promote the non-linear behaviour, a test was performed on the preloaded beam using random excitation at the higher level, 4V RMS, with the aim of identifying a parametric model including stiffness terms up to third order using the extended DPE procedure described in section 1.2.2.

Apart from the increasing the level of the generated signal, the experimental setup was maintained from the previous tests, again bandlimiting the excitation to the range 6 to 80 Hz. The acceleration responses were measured from the five locations, and then were integrated and filtered over the bandwidth 6 to 300 Hz, to retain all components in the response signals resulting from the non-linear energy transfer mechanism. Applying the extended DPE procedure, the mass, damping and first order stiffness matrices were identified, in conjuction with the second and third order stiffness parameters,

$$[m] = \begin{bmatrix} 0.234 & 0 & 0 & 0 & 0 \\ 0 & 0.233 & 0 & 0 & 0 \\ 0 & 0 & 0.114 & 0 & 0 \\ 0 & 0 & 0 & 0.211 & 0 \\ 0 & 0 & 0 & 0 & 0.264 \end{bmatrix} \quad (\mathrm{kg})$$

$$[c] = \begin{bmatrix} 19.29 & 23.56 & 28.64 & -22.53 & -26.20 \\ 5.26 & 17.65 & 9.11 & 9.94 & 15.60 \\ -2.48 & -2.26 & 17.81 & 2.81 & 9.18 \\ -22.80 & -52.30 & -72.00 & 37.70 & -65.50 \\ 61.47 & 14.32 & 257.2 & 94.39 & 215.7 \end{bmatrix} \quad (\mathrm{N/ms^{-1}}) \qquad (6.28)$$

$$[k_1] = 10^3 \times \begin{bmatrix} 63.64 & 11.51 & 7.52 & -32.16 & -7.38 \\ 11.51 & 42.72 & 18.55 & -26.58 & -31.73 \\ 7.52 & 18.55 & 0.64 & -12.95 & -15.60 \\ -32.16 & -26.58 & 12.95 & 3.60 & 17.01 \\ -7.38 & -31.73 & -15.60 & 17.01 & 34.56 \end{bmatrix} \quad (\mathrm{N/m})$$

Locations		Parameters	
		k_2 (N/m^2) $\times 10^6$	k_3 (N/m^3) $\times 10^9$
1	ground	1.090	-4.140
1	2	0.916	0.001
1	3	0.667	-8.929
1	4	0.003	-2.398
1	5	0.306	-5.652
2	ground	0.168	0.885
2	3	0.313	0.364
2	4	0.789	0.828
2	5	0.803	-4.421
3	ground	0.743	0.124
3	4	0.100	-0.870
3	5	0.107	21.36
4	ground	0.730	1.251
4	5	0.532	15.05
5	ground	0.552	-6.139

Table 6.21 Second and third order stiffness parameters identified using the direct method from the preloaded beam.

The first order inertance FRF matrix was regenerated using the parameters listed in equation 6.28, and for the direct location a comparison was made with the inertance TF measured at the higher level, as shown in figure 6.22.

Direct interpretation of the non-linear parameters listed in table 6.21, is difficult partly because the non-linearity of the beam is distributed as opposed to being 'localised' to parts of the structure. (For example, non-Iinearities are often localised around the points of connection of sub-assemblies in a built-up structure, as mentioned in the Introduction.)

One method of checking at least the order of magnitude of the non-linear terms is to regenerate the diagonals of the higher order FRFs from the identified parameters using the expressions developed in section 3.2, and then to compare the second and third order functions with their corresponding TFs to verify the second and third order stiffness parameters respectively. For the direct location, these comparisons are also shown in figure 6.22.

In addition, figures 6.23 and 6.24 illustrate the multi-dimensional second and third order FRFs respectively which also have been regenerated from the parametric model

The comparison test provides strong indication that the non-linear parameters are reasonably accurate. In particular the second order functions show close agreement across the frequency range. Since the regeneration of the second order FRF entails synthesising the full first order FRF matrix, differences can be partly accounted for, arising from the discrepancies between the first order functions. Moreover, the third order functions show increasing error is due to inaccuracies in the second order parameters which are used in the regeneration. Nevertheless the functions correspond satisfactorily to conclude that this comparison technique is useful to validate the higher order model.

Clearly, even if exact higher order parameters were identified, differences between the functions would still be observed because the higher order TFs, measured using stepped sine excitation of nominally constant amplitude, are always subject to distortion as described in chapter 4. Realistically, maintaining the excitation amplitude constant at a level high enough to resolve high quality TFs across the entire bandwidth is likely to cause significant distortion around the resonances, and further research would be well directed at avoiding this problem, perhaps by regulating the excitation amplitude as suggested previously.

A further question which needs to be addressed is the uniqueness of the identified model, in particular the mass and the first order stiffness parameters. Although the first order FRFs show good agreement with the measured data, the validation tests used do not guarantee the identified parameters are physical. For example, although the lumped mass matrix is approximately symmetrical (corresponding to the physics of the beam), the mass elements are evidently overestimated because the total estimated mass is 1.06 kg, in error by over 35% with respect to the actual mass of the beam. As mentioned in section 5.5, errors in the mass matrix will cause corresponding inaccuracies in the first order stiffness parameters, which in turn will affect the validity of the higher order parameters identified. Thus, before employing the DPE technique in earnest to identify non-linear models of structures, it is first necessary to ensure the mass matrix is accurate, a subject which is addressed in the next chapter.

Chapter 7

A Structure with Localised Non-Linearity

7.1 Introduction

With the pre-loaded beam investigated in chapter 6, the non-linear function can be controlled to some extent by adjusting the tension spring, although the source of the non-linearity cannot be considered to be localised in the structure. Consequently, to comply with the specifications of the project proposal, a 'structural model' was designed, attempting to localise the predominant source of non-linearity. The intention was that this model act as a benchmark MDOF non-linear structure on which to apply and validate the procedures developed for identifying mathematical models which incorporate non-linear terms.

The design of the structural model was based on the non-linearity which often arises with the connection of a store to a pylon attached to an aircraft wing. A unit was designed to represent the pylon-store sub-assembly with a symmetrical, hardening stiffness restoring force function, but only very light damping. This sub-assembly unit is a SDOF system and is therefore amenable to identification by a range of procedures, including static load-deflection testing. Also the restoring force surface method has proved to be extremely effective for non-linear SDOF systems of this type, and can be used to estimate directly the parameters of the stiffness function once the mass term has been estimated. Furthermore, when the unit is mounted on a cantilevered beam intended to represent the aircraft wing (shown in figure 7.1), the oscillation motion of the unit couples with the vibration of the beam, and the combined assembly is a MDOF structure in which the non-linearity is localised.

7.2 Design of the Structural Model

The model of the complete assembly of the structural model is shown in figure 7.1, and TF and coherence function measured from the direct location (1) with and without the non-linear unit attached, are shown in figure 7.2. Details of the design and the tests performed on this structure are discussed subsequently. Beforehand, however, it is appropriate to mention a different idea suggested by Gifford [25] for incorporating a controllable, localised non-linearity into an otherwise linear structure.

The method, which was evaluated in this project using the linear 3DOF beam described in chapter 5, was to connect a second, auxiliary shaker to the structure, in addition to the drive shaker, to apply an independently controlled excitation at a single location. In this way, it is possible to generate a non-linear force to act at the connection point by measuring the response at this location and using the signal as the input to an electronic circuit designed to generate a specific non-linear function. For example, if the displacement response y(t) were measured (using the B & K type 2635 charge amplifier units to also integrate the signals from the accelerometers), the circuit could be designed to generate the function, $z(t) = \alpha_1.y + \alpha_2.y^2 + \alpha_3.y^3$. Then the output from the circuit z(t) is fed back via a power amplifier and used as the input to the auxiliary shaker connected at the same point from which the response y(t) was measured, as shown in figure 7.3. Effectively this would cause an additional non-linear stiffness term to be introduced into the governing equations of motion at the single location selected.

This concept is attractive because, assuming the original structure were nominally linear, the non-linearity would be effectively localised and could be controlled by regulating the amplifiers $\alpha_1, \alpha_2, \alpha_3$ in the controlling circuit. Morever the method could be applied equally easily to the damping restoring force by using the measured velocity response $\dot{y}(t)$ instead of the displacement signal y(t). Also it would appear that the non-linear function would have the form of a polynomial which is assumed in the mathematical models, and is therefore desirable for a benchmark structure. Unfortunately this is where one considerable practical problem is encountered.

For the governing function to be polynomial, it is essential that there be no appreciable time delay between the measurement of the response and the application of the non-linear force by the auxiliary shaker. Unfortunately, by feeding back the response signal, the dynamics of the auxiliary shaker become involved and, as a consequence, a time delay is introduced between the voltage input to the shaker coil and the displacement output of the armature.

Electrodynamic shakers can exhibit complicated dynamics [67], despite being approximately SDOF, and usually have a resonance in the range 5 to 50 Hz depending on the mass of the armature and the effective stiffness generated by the electromagnetic field and the armature spider support. In this application, an important fact is that the resonant frequency of the auxiliary shaker occurs within the bandwidth of interest with respect to the structure, and the phase shift associated with this resonance destroys the simultaneity required to generate a non-linear restoring force which is polynomial. For example, the small Gearing and Watson type V4/II shaker was used as the auxiliary in the evaluation tests performed. From the transfer function measured (figure 7.4) relating the measured acceleration of the armature to the random voltage input (for the constant current setting of the Gearing and Watson type SS30 power amplifier), this shaker was found to have resonance at approximately 47 Hz.

The auxiliary shaker was connected at location 2 on the linear beam rig (described in section 5.5.2) to provide the feedback forcing, and the beam was driven at location 1 by a 47 Hz sinewave using the larger Ling 400 series shaker as before. The displacement response measured at location 2, $y_2(t)$, was input to the electronic circuit and both the output of the circuit, $z(t)$, and the force applied by the auxiliary shaker, $x_2(t)$, were measured as shown in figure 7.5. The considerable harmonic content in the signal z(t), which practically matches the spectrum of $x_2(t)$ (plotted on *linear* scales), indicates that at this level of forcing, the non-linear circuit was introducing a significant non-linear effect.

Figure 7.6 illustrates the transfer characteristics of the non-linear circuit alone $(z(t)$ vs. $y_2(t))$, and also of the complete feedback section $(x_2(t)$ vs. $y_2(t))$ which incorporates the power amplifier and the auxiliary shaker in addition to the circuit. Although it is clear that the circuit generates a good polynomial function for both random and sine excitation, the delay introduced by the auxiliary shaker was very evident as the complete feedback section formed a hysteresis loop. Since the 47 Hz sine excitation coincides with the resonance of the auxiliary shaker, the hysteresis loop approached maximum width. Furthermore, with the O-100 Hz random excitation, the function departed completely from the polynomial form.

Methods of compensating for the dynamics of the auxiliary shaker were investigated briefly before proceeding to look for an alternative design. Unfortunately compensating for the phase delay across a relatively wide band of frequencies, and in particular starting from zero frequency, is evidently very difficult [68]. P01e cancellation/placement control strategies are only able to effectively shift a resonance up or down in frequency, but not remove a time delay altogether. Consequently, the feedback approach for creating a localised non-linearity was abandoned, and an alternative mechanical solution adopted.

Throughout this thesis, examples of structures governed by non-linear stiffness functions have been considered. In some physical systems, however, the dynamic behaviour is governed predominantly by the damping forces which may also be markedly non-linear. A common example is the automotive shock absorber [38]. As mentioned above, the feedback approach originally seemed ideal for generating a damping non-linearity, which would have been useful for validating the testing and identification procedures. However, no feasible mechanical designs for introducing well-behaved non-linearity into the damping function were arrived at. Instead, it was decided to base the structural model on the non-linear stiffness which arises in the attachment of a store via a pylon to an aircraft wing due to the clearances at the connections.

The unit which models the pylon-store sub-assembly and is detailed in figure 7.7, was originally designed using the same general principle as the flexible beam described in chapter 6 to introduce a hardening stiffness term into the governing function. In this case the scale was reduced, with the flexible beam taking the form of a leaf spring approximately 25 mm long, clamped at the centre, and acting to resist rotational motion of a pivoted mass element. In the context of the aircraft structure, the pivot-spring represents the pylon which supports the store (mass element) on the wing (the cantilevered beam in figure 7.1). To amplify the hardening effect for smaller amplitude oscillations of the mass element, the leaf spring was supported on a curved surface. As the spring deflects, contact is made with the surface, shortening the effective length of the spring and thus increasing the stiffness smoothly and non-linearly with respect to the angle of rotation, which is also kept small in this way.

Before arriving at the final arrangement, several adjustments were necessary, particularly to the thickness of the leaf spring which had to be flexible whilst retaining elasticity, ie. without permanent deformation. In the design it was also important that friction between each contacting surface be minimised in order to ensure that the stiffness restoring force dominates over the dissipative forces. Moreover, any 'stiction' in the unit would cause severe problems. Consequently ball bearing races were used for each pivot, and very small roller bearings were needed at the point of contact to the leaf spring.

By designing the unit to be an independent, SDOF system in this way, the stiffness function can be identified relatively easily using either a static test or the restoring force surface method, as described in the next section. On the other hand, by connecting the unit to a cantilevered beam modelling the aircraft wing (figure 7.1), the intention was that a MDOF structure with a localised non-linearity be created (section 7.4).

The cantilevered beam consisted of a 4 mm thick mild steel plate of width 0.25 m and length 1 m, although after one end of the beam was bolted to a rigid pillar mounted the test bed, a length of approximately 0.9 m remained free. To increase the damping, a constraining layer was added by bonding a sheet of 1 mm thick polystyrene foam over the top surface, then bonding a thin 0.5 mm steel plate on top of the foam.

A preliminary analysis of the dynamics of the beam was made by constructing a simple Finite Element (FE) model using only beam elements with the appropriate material properties. FE was used as opposed to direct calculation since a relatively large number of nodes and elements can be used, removing restrictions on the number of degrees of freedom. Consequently, the model resembles the physical system more closely than a lumped parameter model. Furthermore, to select the locations of the accelerometers for the subsequent tests, reference was made to the master degrees of freedom determined from the FE model by specifying a frequency range and the modes of interest.

The FE model was constructed using 19 simple beam elements, each of length 50 mm, constraining the motion node 1 in all 6 directions to represent the fixed end condition, and specifying master DOFs for the vertical motion of nodes 6 and 9, respectively the location of the connection to the shaker, and the position selected for attaching the non-linear unit. Two automatic masters were also specified, determined from the first processing stage to be at nodes 14 and 18 in the vertical direction. Hence, to measure the response in each of the tests described subsequently, accelerometers are placed at the locations 1 to 4 with distances of 0.25, 0.4, 0.65 and 0.85 m from the fixture to correspond with nodes 6, 9, 14 and 18 respectively. Performing the eigenvalue analysis, the FE model gave the first four resonant frequencies to be 4.5, 27.5, 70.3 and 141.4 Hz, which correspond reasonably well with values 4.7, 26.4, 68.8 and 133.0 Hz, determined from the original transfer function test on the beam (figure 7.2), considering the FE model was kept very simple without subsequent adjustment.

One conclusion which can be drawn from the correlation between the measured and the predicted resonances is that, since the FE model is constructed from the physical dimensions and properties of the structure, a numerical model can be formulated to represent the dynamics of the beam. Although this is a fundamental observation, the ultimate objective of performing system identification is to determine a model from experimental data which not only predicts the resonant frequencies and regenerates the measured TFs accurately, but also can be related to the physics of the structure in a similar way to the FE model. A further application may be to use the experimentally identified parameters to update the FE model, which can then be used to predict the response of the structure to a different excitation.

7.3 Identification of the Sub-Assembly Unit

7.3.1 Static Test

The sub-assembly unit was tested statically, initially to check that the device worked as designed, particularly the polynomial nature of the stiffness function, and subsequently to identify the stiffness parameters.

As shown in figure 7.8, the unit was inverted for the static test, fixing one end to a rigid test bed, and applying the load horizontally to the free, store end by adding known masses via a cable over a pulley wheel. (Care was taken to ensure the cable attachment did not restrict the rotational movement of the store around the pivot.) Since the angles of rotation were small, transverse deflections of the store could be measured with reasonable accuracy using a dial gauge. The static load was added in 100 gram increments to a total of 5 kg, causing a maximum displacement of approximately 6 mm from equilibrium position, to correspond with the deflection limit imposed in the dynamic tests described next. Consequently, the rotation of the store around its pivot was limited to maximum of 4". The test was repeated in the reverse direction by turning the unit around, and then the data plotted to give the load-deflection curve shown in figure 7.9.

To identify the parameters of the stiffness function, polynomials were fitted to the measured curve:

1. Linear + Cubic terms, f(y) = $\alpha_1 y + \alpha_3 y^3$,

$$\begin{aligned} \alpha_1 &= 0.3252 \text{ kg/mm} = 3191.0 \text{ N/m} \\ \alpha_3 &= 0.0116 \text{ kg/mm}^3 = 114.0 \times 10^6 \text{ N/m}^3 \end{aligned} \tag{7.1}$$

2. Linear + Quadratic + Cubic terms, f(y) = $\beta_1 y + \beta_2 y^2 + \beta_3 y^3$,

$$\begin{aligned} \beta_1 &= 0.3171 \text{ (kg/mm)} = 3110.0 \text{ (N/m)} \\ \beta_2 &= -0.0047 \text{ (kg/mm}^2\text{)} = -45.0 \times 10^3 \text{ (N/m}^2\text{)} \\ \beta_3 &= 0.0120 \text{ (kg/mm}^3\text{)} = 117.0 \times 10^6 \text{ (N/m}^3\text{)} \end{aligned} \tag{7.2}$$

By introducing the second order term, $\beta_2 y^2$, into the restoring force function, the first order parameter decreased by only 2.5 % and the third order increased by 2.8 %. Furthermore, the accuracy of the curvefit was not significantly improved, and the small discrepancies between the measured data and the curvefits are attributable to the inaccuracies of the test rather than behaviour of the unit itself. The conclusion is that the appropriate polynomial for the restoring force function includes first and third order terms only, and the quadratic term is relatively insignificant in the stiffness function.

7.3.2 Transmissibility test

Following the static test, an unsuccessful attempt was made to measure the higher order TFs of the unit directly by applying excitation from a shaker to the store through a narrow, flexible link connected via the force transducer to a nylon collar around the store, as shown in figure 7.10. The reason for including the collar and the flexible link was to avoid introducing any additional stiffening effect on the rotational movement of the store around the pivot. Nevertheless from the tests conducted, it was evident that the connection arrangement did increase the apparent stiffness of the unit, the fundamental resonant frequency increasing from approximately 10 to 14 Hz. In addition, the need for the collar gave rise to problems of stiction/friction, and the accelerometer and force transducer also increased the effective mass of the store significantly.

To avoid all these problems, an alternative arrangement for the subsequent dynamic tests on the unit was used; as illustrated in figure **7.11.** The unit was mounted on a rigid bracket fixed to the armature of an electrodynamic shaker positioned so that the centre of gravity was coincident with the axis of the shaker armature to reduce the torque loading on the shaker coil. Importantly, the need for including a force link between the shaker and the unit was avoided by representing the dynamics of the unit in terms of transmissibility, hence reducing the influence of the peripheral apparatus on the data.

Relating the accelerations at the pylon, $\ddot{y}_p(t)$, and on the store, $\ddot{y}_s(t)$, to the voltage input to the shaker power amplifier, v(t), the linearised TFs of the combined system (non-linear unit plus electrodynamic shaker) were measured using broadband random excitation, as shown in figure 7.12. The combined system exhibits two clear resonances over the frequency range, as the oscillation of the unit couples with the motion of the shaker. The first resonance at 8.8 Hz is close to the resonant frequency of the unit, whereas the second resonance at 20 Hz corresponds to the resonant frequency of the shaker.

However from these measured transfer functions it was possible to estimate the linearised transmissibility function, G(w), (also shown in figure 7.12) by observing,

$$G(\omega) = \frac{Y_s''(\omega)}{Y_p''(\omega)} = \frac{Y_s''(\omega)/X(\omega)}{Y_p''(\omega)/X(\omega)} \tag{7.3}$$

Clearly, by taking the ratio of the TFs, the oscillation of the unit effectively becomes isolated. Correspondingly, over the frequency range of interest, G(w) exhibited a single resonance at 9.2 Hz which did not alter as the level of the random excitation was changed from 1V to 4V RMS.

For this simple system, an expression for G(w) can be formulated directly from the equation of motion of the store,

$$m_s.\ddot{y}_s(t) + c_{sp}(\dot{y}_s(t) - \dot{y}_p(t)) + k_{sp}(y_s(t) - y_p(t)) = 0 \tag{7.4}$$

in which m_s denotes the mass of the store, and k_{sp} the linearised stiffness of the leaf spring between the pylon and store. The energy dissipation is representated by the viscous damping term with coefficient c_{sp}.

Since a formulation for the linearised transmissibility function is of interest, the method of harmonic probing can be applied as if the system were linear. Using the phasor notation introduced in section 1.1.3,

$$\begin{aligned} y_p(t) &= Y_p(\omega),\ \dot{y}_p(t) = (j\omega).Y_p(\omega),\ \ddot{y}_p(t) = (j\omega)^2.Y_p(\omega) \\ y_s(t) &= Y_s(\omega),\ \dot{y}_s(t) = (j\omega).Y_s(\omega),\ \ddot{y}_s(t) = (j\omega)^2.Y_s(\omega) \end{aligned} \tag{7.5}$$

Substituting these expressions into the equation of motion (eqn. 7.4),

$$\left((j\omega)^2.m_s + (j\omega).c_{sp} + k_{sp}\right).Y_s(\omega) = ((j\omega).c_{sp} + k_{sp}).Y_p(\omega) \tag{7.6}$$

Hence,

$$G(\omega) = \frac{Y_s''(\omega)}{Y_p''(\omega)} = \frac{Y_s(\omega)}{Y_p(\omega)} = \frac{(j\omega).c_{sp} + k_{sp}}{(j\omega)^2.m_s + (j\omega).c_{sp} + k_{sp}} \tag{7.7}$$

Theoretically, the linearised transmissibility function has same form as the conventional receptance function for this relatively simple SDOF system (cf. equation 1.13). This can be confirmed in practice by comparing G(w), shown in figure 7.12, with the first order receptance FRF of the unit, shown in figure 7.15, which was generated from the model identified using the restoring force surface method, as described in the next section.

However, the magnitude of the two functions are different and so too are the low and high frequency asymptotes. For example, as $\omega \rightarrow 0$, G(w) $\rightarrow$ 1.

As can be seen from figure 7.12, the linearised transmissibility function does tend towards unity magnitude and zero phase at low frequencies, although below approximately 5 Hz the acceleration signals became very small, and so the calculated transmissibility function is noisy in this range.

7.3.3 Restoring Force Surface Method

The original reason for developing the transmissibility test was to provide a means of isolating the dynamics of the unit in order that the restoring force method, which has proved to be very useful in applications of this type, could be employed to perform an accurate identification of the non-linearity in the unit.

In measuring the conventional TF, random excitation linearises the system. However, the same excitation signal can be used to identify the non-linear parameters if the time response is measured rather than the TF, hence avoiding the linearisation. In this application, the equation of motion of the store (eqn. 7.4) must be reformulated to express the linear plus cubic form of the non-linear stiffness function,

$$m_s.\ddot{y}_s(t) + c_{sp}(\dot{y}_s(t) - \dot{y}_p(t)) + {}_{sp}k_1.(y_s(t) - y_p(t)) + {}_{sp}k_3.(y_s(t) - y_p(t))^3 = 0 \qquad (7.8)$$

in which ${}_{sp}k_1$ and ${}_{sp}k_3$ are respectively the first and third order stiffness parameters of the pylon-store connection. The restoring force method can be used to identify directly the mass-normalised stiffness and damping parameters in the equation,

$$\frac{c_{sp}}{m_s}.\dot{\delta}_{ps}(t) + \frac{{}_{sp}k_1}{m_s}.\delta_{ps}(t) + \frac{{}_{sp}k_3}{m_s}.\delta_{ps}(t)^3 = \ddot{y}_s(t) \qquad (7.9)$$

where $\delta_{ps} = y_p(t) - y_s(t)$. Importantly, in this alternative formulation of the restoring force method, the response signals ($\ddot{y}(t), \dot{y}(t)$, y(t)) measured at the pylon and at the store are required but *not* the excitation force. Correspondingly, when constructing the force surface, it is necessary to plot the acceleration at the store, $\ddot{y}_s(t)$, against the *differences* between the velocity and displacement signals, $\dot{\delta}_{ps}$ and δ_{ps} respectively.

Using broadband random excitation bandlimited to between 6 Hz and 14 Hz, and set at the higher level of 4V RMS to promote non-linear behaviour, the acceleration responses were measured at the two locations, integrated and filtered, and the restoring force surface formed as shown in figure 7.13 together with the stiffness cross-section. The approximation of the stiffness function to a straight line over the central region again highlights the need for applying a very high level of excitation when using random excitation with the restoring force surface method. For the level of excitation used, the response tended to be concentrated in the low-level region where the first order stiffness term dominates. Correspondingly, along the stiffness cross-section, excursions to the higher relative displacements are infrequent. On the basis of these results alone, an incorrect conclusion might have been that the system behaves linearly for this range of displacements. However, by repeating the restoring force test but using sine excitation instead, it became clear that this was not the case.

As mentioned in section 1.2, sine excitation is particularly effective at sustaining the oscillation at a higher level, promoting non-linearity to become more clearly apparent in the restoring force surface, and hence better for identifying a non-linear parametric model. The second restoring force test was performed using sinusoidal excitation of amplitude $0.4\mathrm{V}$, incrementing the frequency by 0.5 Hz over the range 6 Hz to 14 Hz to pass through the resonance. At each increment, the acceleration signals were measured with a sampling frequency of 560 Hz, and following integration, the response signals were filtered to the range 2.5 to 150 Hz in order to retain higher harmonic components up to third order in the response.

As can be seen from the force surface shown in figure 7.14, the displacement range of the sine test was nominally the same as with the random excitation (cf. figure 7.13) but the surface is of considerably higher quality and, most importantly, the stiffness trajectory clearly illustrates the non-linearity in the restoring force function. The fact that this trajectory resembles the measured load-deflection curve (figure 7.9) confirms that the dynamics of non-linear unit are governed by a stiffness function which is non-linear and approximates very closely to a linear plus cubic polynomial.

Fitting the polynomial function f(y) = $a_1 y + a_3 y^3$ to the stiffness cross-section, the *mass*-normalised first and third order parameters were identified,

$$a_1 = 2735.5\ \mathrm{Nm^{-1}/kg}, a_3 = 80.20\ \mathbf{x}\ 10^6\ \mathrm{Nm^{-3}/kg} \tag{7.10}$$

The stiffness parameters can be estimated from a_1 and us only if an, estimate of m_s, the mass of the store, is known. Alternatively, by comparison with the results from the static test (equation 7.1), two estimates of the mass parameter m_s can be made,

$$\frac{\alpha_1}{a_1} = \frac{3191.0}{2735.5} = 1.17\ \mathrm{kg}$$

$$\frac{\alpha_3}{a_3} = \frac{114.0\ \mathrm{x}\ 10^6}{80.2\ \mathrm{x}\ 10^6} = 1.42\ \mathrm{kg} \tag{7.11}$$

(Since the first order components tend to dominate the response, the first estimate of the mass is likely to be more reliable.)

By assuming the governing stiffness function is symmetrical, the second order order FRF will be null for this system, as discussed in chapter 3. However, as shown in figure 7.15, it is possible to regenerate the principal diagonal of the third order FRF, in addition to the first order receptance function, directly from the identified parametric model of the non-linear unit,

$$1.2\ \ddot{y}(t) + 5.\dot{y}(t) + 3200.y(t) + (1.1\ \mathrm{x}\ 10^7).y^3(t) = x(t) \tag{7.12}$$

The coefficient of (linear) viscous damping was selected arbitrarily to give $\zeta \approx 4\%$.

7.4 Identification of Wing-Pylon-Store Assembly

7.4.1 Introduction

The original purpose of constructing the model of the wing with detachable pylon-store sub-assembly was to investigate the effectiveness of the identification routines for detecting not only the presence but also the location of non-linearity in a structure by imposing a physical interpretation on the parameters identified using the DPE method.

The complete assembly included the non-linear unit attached firmly to the beam 0.4 m from the fixture, as shown in figure 7.1. The response signals along the beam were measured using standard B & K type 4331 accelerometers at the four locations specified in section 7.2 at distances of 0.25, 0.4, 0.65 and 0.85 m from the fixture. Also a smaller, B & K type 4393, accelerometer was employed to measure the response at the store (location 5). Although this accelerometer is less sensitive, it was considered necessary to use the lightest piezoelectric transducer available in order to minimise the mass loading on the store.

7.4.2 Detection of Non-Linearity

Before embarking on the parameter identification stage, tests were conducted in an attempt to establish that the complete assembly of the unit attached to the beam is non-linear. Although the pylon-store sub-assembly unit is known to be non-linear, its influence on the behaviour of the beam may be prove to be negligable.

In figure 7.2 the coherence functions, measured at location 1 on the beam without the unit attached, showed degradation in the low frequency region (below 5 Hz) where the accelerometers were inaccurate, and around the first anti-resonance where the resolution was poor. For the complete assembly, the resonance of the non-linear unit oscillation occurred at approximately 10 Hz which almost coincided with the first anti-resonance. There was also further degradation in the coherence function in this range.

Although, as discussed in 2.4, the additional degradation may be consistent with the unit introducing non-linearity into the structure, there are other factors which can affect the coherence function, and certainly it is not possible to state conclusively that the complete assembly was behaving appreciably non-linearly on the basis of these results alone. On the other hand, substantial evidence of non-linearity would be provided if clear higher order TFs could be measured with the stepped sine testing procedure described in section 6.4.3.

Also the Hilbert transform procedure may be performed on the first order TF to confirm that non-linearity has been detected, and by repeating the test for a different excitation level, distortion patterns in the TFs could be interpreted.

Higher order TFs were measured from locations 1 to 5 using sine excitation at the higher amplitude of 0.4V with 50 averages, as shown in figures 7.16 to 7.20. From each location, the first order TFs are very clear, although the resonance of the store is only just detectable at the locations along the beam (1 to 4), despite the store element oscillating with relatively high amplitude (eg. 3 mm) when the drive excitation reached 10 Hz. Naturally the oscillation resonance is very apparent at location 5 where the first and second beam resonances are also clear. It is also evident from the Nyquist plot for location 5 that bifurcation occurred in the oscillation resonance of the store element.

The second order TFs are also reasonably clear for locations 1 to 4, although for location 5 the functions are relatively noisy. The presence of strong secondary and primary resonances indicates that both the first and second resonances of the beam were non-linear to some extent. The primary resonance of the store oscillation at 10 Hz is very clear on each of the second order TFs, but the secondary resonance is more difficult to discern around 5 Hz partly because of the proximity to the first beam resonance.

For each of the locations along the beam, the second and third order TFs become less well defined at frequencies away from the resonances as the response signals become smaller and approach the noise level. Nevertheless each primary resonance can be distinguished on all the third order TFs, and the tertiary peak corresponding to the second beam resonance is particularly obvious, for example at locations 1 and 2.

Although overall the higher order TFs do exhibit the energy transfer patterns characteristic of a structure governed by polynomial type non-linearity, as described in chapter 4, it must be admitted that the results were rather disappointing, especially from location 5 where it was expected the non-linearity of the store would be particularly apparent. Certainly it is still not possible to conclude from these results that the unit had a significant influence on the behaviour of the assembly across the frequency range 2 to 28 Hz. Indeed, rather than the non-linearity being caused by the oscillation of the store coupling with the vibration of the beam, the trends exhibited by the TFs are consistent with the cantilevered beam being very weakly non-linear at the relatively high level of excitation used.

The stepped sine test was repeated at the lower amplitude of 0.1V in order to compare the measured first order TFs for the two amplitudes. Figure 7.21 shows the result from the direct location, which was typical of each of the locations along the beam, and figure 7.22 illustrates location 5. At both these locations, there was very little difference between the two levels around the first resonance, whereas the store oscillation resonance did exhibit a slight distortion toward higher frequency for increasing amplitude caused by the hardening stiffness in the unit between the pylon and store elements. Also there was significant decrease in the frequency of the second beam resonance falling from 21.8 to 21.2 Hz, indicating a weak, quadratic type, stiffness non-linearity in the structure.

These general observations were confirmed by the Hilbert transform tests conducted on the first order functions at the higher amplitude from locations 1 and 5, shown in figures 7.23 and 7.24 respectively. Certainly any non-linearity associated with the vibration of the beam was only very weak, because only minimal distortion is apparent in the transformed function with respect to the measured function. The oscillation of the store element appears to have had very little influence in terms of introducing non-linearity into the vibration of the beam, except at the response location on the store (5) where again the bifurcation near 10 Hz was detected.

One conclusion which can be drawn from these results is that the higher order TFs provide an extremely sensitive method for detecting non-linearity. However, despite the TFs quantifying non-linear energy transfer, it is difficult to determine the relative strength of the non-linearity present using only the TFs, mainly because each order has different dimensions and their magnitudes cannot be compared directly. It was evident from the coherence function (figure 7.2), the minimal distortion around the second resonance (figure 7.21) and from the Hilbert transform test (figure 7.23) that the beam was only very weakly non-linear. Nevertheless the measured second order order TFs exhibited clear primary and secondary resonances, particularly associated with the second beam resonance (eg. figure 7.16).

The relative sensitivity of the TFs can be viewed as an advantage in the context of making a comparison between the TFs and the principal diagonals of the FRFs regenerated from identified non-linear models. On the other hand, since higher order TFs are poor at quantifying the strength of the non-linearity in the structure, it is clear that the conventional technique of the Hilbert transform does still have an important role to play, supporting the evidence of non-linearity provided by higher order TFs measured in a stepped sine test.

7.4.3 Parameter Identification

Direct parameter estimation was performed on the complete assembly in an attempt to identify a mathematical model which could be interpreted from a physical viewpoint, and to determine whether extending the model to include non-linear terms causes a significant improvement. A total of four similar cases were considered in order to assess the consistency of the parameters identified, particularly the mass and first order stiffness matrices. In each test, the bandwidth of interest was selected as 2 to 28 Hz in order to encompass the first and second resonances of the beam and the oscillation resonance of the store element.

In the first case, the indirect method (method 2, re. section 5.3.3) was used, generating representative time data from inertance functions measured using 1V RMS random excitation, a relatively low level to encourage linear behaviour. Moreover, the oscillation of the store is hardly evident on the TFs measured along the beam, as shown in figure 7.2 for the direct location. Consequently the DPE routine was executed specifying a first order model and incorporating only the responses from the locations along the beam, omitting location 5 in an attempt to identify the linearised model of the structure with respect to the beam. Having performed the usual data processing procedures, the following parameter matrices were identified,

$$[m] = \begin{bmatrix} 9.390 & 0 & 0 & 0 \\ 0 & 0 & 3.977 & 0 \\ 0 & 1.122 & 0 & 1.093 \end{bmatrix} \text{ (kg)}$$

$$[c] = 10^3. \begin{bmatrix} 1.29 & -1.61 & 1.47 & -0.66 \\ -0.58 & 0.20 & -0.63 & 0.49 \\ 0.45 & -0.59 & 53.4 & -0.22 \end{bmatrix} \text{ (N/ms}^{-1}\text{)} \qquad (7.13)$$

$$[k] = 10^3. \begin{bmatrix} -10.99 & 324.81 & -428.83 & 161.80 \\ 324.81 & -324.48 & 278.49 & -108.82 \\ -428.83 & 278.49 & -66.60 & 3.46 \\ 161.80 & -108.82 & 3.46 & 16.80 \end{bmatrix} \text{ (N/m)}$$

The comparison between the regenerated inertance FRF and the measured TF is shown in figure 7.25 for locations 1 and 4, with excellent agreement across the frequency range.

In the second case, the direct route to DPE was taken (method 1), measuring the acceleration response at each location along the beam, and then numerically integrating the signals before identifying the parameters. On this occasion the higher level of 4V RMS random excitation was used to encourage non-linear behaviour, but in order to draw a comparison with the previous results, location 5 was again omitted from the model, and the identification was restricted to first order,

$$[m] = \begin{vmatrix} 11.68 & 0 & 0 & 0 \\ 0 & -0.23 & 0 & 0 \\ 0 & 0 & 5.59 & 0 \\ 0 & 0 & 0 & 0.90 \end{vmatrix} \text{(kg)}$$

$$[c] = 10^3. \begin{bmatrix} -1.88 & 2.25 & -2.06 & 0.79 \\ -0.19 & 0.25 & -0.24 & 0.10 \\ 0.02 & 0.08 & -0.08 & 0.07 \\ 0.10 & -0.19 & 0.19 & 0.00 \end{bmatrix} (\text{N/ms}^{-1}) \qquad (7.14)$$

$$[k] = 10^3. \begin{bmatrix} 210.20 & 53.68 & -142.01 & 40.53 \\ 53.68 & -47.79 & 23.49 & -6.36 \\ -142.01 & 23.49 & 105.58 & -53.62 \\ 40.53 & -6.36 & -53.62 & 33.00 \end{bmatrix} (\text{N/m})$$

There is a slight increase in the difference between FRFs regenerated from the parameter estimates and the TFs measured using 4V RMS random excitation (figure 7.26) when compared with the previous model (cf. figure 7.25), although the correspondence is still very good. The deterioration in the fit can be attributed partly to the weakly non-linear effects detected but mainly to the fact that the the representative data used in the previous model was derived from functions which were averaged to reduce the influence of measurement noise on the identified parameters. However, comparison of the matrices listed in equation 7.14 with 7.13 does reveal a considerable difference between the two models, and clearly the presence of relatively weak non-linearity cannot be entirely responsible for this difference.

Furthermore it is obvious that the parameters identified cannot be interpreted as being physical, since the total mass estimated in both cases exceeds 16 kg, far greater than the actual mass of the beam, and the stiffness matrices and damping matrices each have negative terms along the leading diagonal. The apparent lack of a physical interpretation of these models is very important, and the topic is addressed again later in this section.

In the third case, the response signals measured and used in case 2 were used again, but on this occasion the model was extended to include also location 5, the response measured at the store element,

$$[m] = \begin{bmatrix} 12.05 & 0 & 0 & 0 & 0 \\ 0 & -0.25 & 0 & 0 & 0 \\ 0 & 0 & 5.37 & 0 & 0 \\ & & & & \\ 0 & 0 & 0 & 0.920 & 1.25\ 0 \end{bmatrix} \text{(kg)}$$

$$[c] = 10^3. \begin{vmatrix} -2.10 & 2.48 & -2.25 & 0.86 & 0.04 \\ -0.22 & 0.28 & -0.28 & 0.12 & 0.01 \\ 0.07 & 0.01 & -0.02 & 0.04 & 0.00 \\ 0.08 & -0.16 & 0.16 & -0.08 & 0.01 \\ 0.43 & -0.52 & 0.49 & -0.20 & 0.01 \end{vmatrix} \text{(N/ms}^{-1}\text{)} \qquad (7.15)$$

$$[k] = 10^3. \begin{vmatrix} 207.74 & 56.45 & -138.22 & 37.86 & 2.00 \\ 56.45 & -49.18 & 24.35 & -6.45 & 0.03 \\ -138.22 & 24.35 & 100.06 & -51.05 & 0.20 \\ 37.86 & -6.45 & -51.05 & 31.99 & -1.03 \\ 2.00 & 0.03 & 0.20 & -1.03 & 5.68 \end{vmatrix} \text{(N/m)}$$

The purpose identifying two very similar models (cases 2 and 3) was to determine whether or not including an extra degree of freedom would affect the parametric model identified for the beam (locations 1 to 4), particularly since the non-linearity in the stiffness between the pylon and store is most apparent at location 5.

Comparing equations 7.15 with 7.14, it is clear that there is a change in the mass and stiffness parameters associated with locations 1 to 4, although relatively small. As illustrated in figure 7.27, the correspondence between the regenerated FRF and measured TF is almost identical with the previous model at locations 1 and 4 (cf. figure 7.26). Also shown in figure 7.27 is the comparison for the fifth location, where very significant differences between the two functions can be observed, particularly apparent from the Nyquist plot.

The conclusion drawn is that the parametric model was not influenced significantly by the inclusion of data from location 5 because the response at this location was low level when compared to the vibration of the beam measured at locations 1 to 4 over the frequency range selected.

Finally, in case 4, DPE was applied in the same way as in case 3 except that the model was extended to include also second and third order stiffness terms,

$$[m] = \begin{vmatrix} 11.52 & 0 & 0 & 0 & 0 \\ 0 & -0.25 & 0 & 0 & 0 \\ 0 & 0 & 5.88 & 0 & 0 \\ 0 & 0 & 0 & 0.88 & 0 \\ 0 & 0 & 0 & 0 & 1.67 \end{vmatrix} ; \quad [c] = 10^3. \begin{vmatrix} -2.02 & 2.42 & -2.22 & 0.85 & 0.03 \\ -0.16 & 0.20 & -0.19 & 0.08 & 0.00 \\ 0.01 & 0.07 & -0.05 & 0.05 & 0.00 \\ -0.09 & 0.04 & -0.03 & -0.01 & 0.00 \\ 0.97 & -1.17 & 1.08 & -0.44 & 0.01 \end{vmatrix}$$

$$[k_1] = 10^3. \begin{vmatrix} 221.23 & 47.51 & -142.3 & 42.52 & -1.94 \\ 47.51 & -40.76 & 18.73 & -4.43 & 0.06 \\ -142.86 & 18.73 & 113.78 & -57.55 & -0.24 \\ 42.52 & -4.43 & -57.55 & 35.16 & -1.75 \\ -1.94 & 0.06 & -0.24 & -1.75 & 8.76 \end{vmatrix} \tag{7.16}$$

I Locations		Parameters	
		k_2 (N/m^2) $\times 10^6$	k_3 (N/m^3) $\times 10^9$
1	1	-8.451	18.58
1	2	-3.373	-43.49
1	3	1.334	16.94
1	4	-1.632	2.80
1	5	-2.895	-0.43
2	2	1.418	2.81
2	3	-9.354	-7.66
2	4	2.253	12.30
2	5	1.080	2.03
3	3	-1.591	1.75
3	4	-4.427	-78.53
3	5	-1.524	-3.66
4	4	0.698	-1.70
4	5	0.310	-0.77
5	5	1.536	-12.16

Table 7.29 Second and third order stiffness parameters identified using the direct method from the complete assembly.

As can be seen from figure 7.28, the comparison between the first order FRF and the measured TF does not significantly improve on the previous case (cf. figure 7.27) as a result of extending the model to be non-linear, although differences can be observed at location 5.

These four applications of the DPE procedure serve to highlight two very important points.

Firstly, it was intended that the combined structure of the non-linear unit attached to the cantilevered beam would act as a benchmark structure in which the non-linearity was localised. The results indicated, however, that the non-linear stiffness between the pylon and store did not significantly affect the response signals at the locations along the beam. Developments to the existing rig are required to increase the influence of this localised non-linearity on the dynamics of the complete assembly. One solution may be simply to increase the mass of the store element to reduce the resonant frequency of the oscillation so that significant coupling occurs with the first resonance of the beam.

The second point regards the lack of physical interpretation of the parameter matrices identified. One conclusion drawn from the previous tests on the non-linear beam (chapter 6) was that the physical structure more closely resembles a distributed system rather than the lumped parameter model. Clearly this is also true for the structure considered in this section. Even at the lower level of excitation, the linearised model identified (equation 7.13) departs from the physical model, and it is evident that one of the sources of error is the deficiency of the mathematical model at representing a distributed parameter structure.

However an appropriate way to proceed is revealed from these tests, namely to find a means of fixing either the first order stiffness parameters or, perhaps more practically, the lumped mass elements in order to force the mathematical model to correspond more closely with the physical structure. By reducing the freedom of the DPE solution in this way, the identified stiffness parameters would more likely have a physical interpretation if the mass elements were either selected with regard the distribution of mass in the structure, or alternatively from a FE model.

The suggestion of incorporating data from a previously constructed FE model within the DPE procedure, together with measured response signals, also has implications regarding the task of FE model updating on the basis of test results, since it may be possible to compare directly the identified stiffness matrix with the corresponding FE matrix. Furthermore, the facility of matrix reduction in FE procedures may prove useful in specifying the appropriate response locations to be incorporated in the DPE to identify a lumped parameter model, before expanding back to the physical distributed parameter model. Using a pre-determined mass matrix in the DPE procedure would have the important advantage that parameters need not be passed down to scale subsequent scale equations, making the solution quicker. In this way there would be no need for optimising the selection of the scaling parameters, and errors would no longer be compounded since each equation would be solved independently.

Chapter 8

Summary, Conclusions and Recommendations for Further Work

8.1 Summary

Before listing the conclusions which have been drawn from this research, it is worthwhile to summarise the main points which have arisen from the discussions presented in each chapter.

Chapter 1 addressed the problem of distinguishing between structures which can be represented by a linear model, and those which are non-linear. Conventional non-linearity detection techniques include testing for homogeneity of the system, examining the resonant peaks of measured TFs for amplitude dependent distortion, or using the Hilbert transform (which is effectively a sensitive tool for performing the same task.) Alternatively, random excitation may be employed to measure linearised functions from a non-linear structure. However, looking along the single-dimensional line of conventional TFs reveals very little information regarding the characteristic behaviour of non-linear systems, and considerable advantages can be gained by broadening the perspective.

For SDOF structures, one very powerful method of both detecting non-linearity and identifying a parametric model is provided by the restoring force surface method. This technique can be reformulated for MDOF structures, referred to as direct parameter estimation (DPE) in this text. The non-linear electronic circuit system which was discussed throughout chapter 1 appeared to be non-linear, but the results were inconclusive because the observed effects were consistent a high pass filter being connected in series with a SDOF system.

One of the principal aims of the project was to relate the non-linear differential equations of motion to the Volterra series representation of a non-linear system, since this alternative model provides considerable insight into the characteristics of the response behaviour in both the time and frequency domains. This was the main theme of chapter 2. Unfortunately it is evidently impossible to measure directly the homogeneous kernels which define the Volterra series. Certainly very few results have been published in which the Volterra kernels have been identified for a physical structure. The Wiener series was originally formulated from the Volterra series and defined to have useful properties from the viewpoint of being able to identify the kernels from the response of a system. Nevertheless it is very difficult to measure accurately the Wiener kernels using the two established techniques of cross-correlation and polyspectral analysis. These methods tend to be time consuming and remain impractical for estimating Volterra kernels of a structure of order higher than two.

In chapter 3, relationships were established between the differential equation of motion model of a non-linear structure, and the higher order FRFs which are defined as the multi-dimensional Fourier transforms of the Volterra kernels. The expressions derived in this text make it possible to regenerate the FRFs directly from a parametric model, for example as identified using DPE, and this was demonstrated for both single- and multi-degree of freedom systems. These multi-dimensional FRFs are important for understanding and representing the behaviour of non-linear systems in the frequency domain because they characterise the mechanism for energy transfer which is exhibited by many physical structures.

Chapter 4 followed the discussion of the FRFs by focussing on an alternative to the Wiener series for measuring higher order FRFs from a non-linear structure. Numerical simulations based on the non-linear equation of motion model were used to illustrate the discussion. Higher order TFs can be defined which quantify the transfer energy to the higher harmonic frequencies when the system is excited by a sinewave, and it has been shown that the TFs defined for the equation of motion model approximate to the principal diagonals of the higher order FRFs. Theoretically, for systems governed by polynomial functions, this approximation improves as the amplitude of the excitation is reduced. For practical situations in which the excitation waveform can be controlled, the measurement of the higher order TFs provides a relatively straightforward method for estimating the higher order FRFs up to third order and beyond and, by using this procedure, estimates of second and third order FRFs have been obtained from physical structures.

The main conclusion drawn from chapter 5, which was concerned with the practicalities of implementing the DPE to identify a structure, was that it is possible to determine a parametric model which can be interpreted from a physical viewpoint when the structure resembles

the lumped parameter model. A nominally linear structure was investigated which consisted of a cantilevered beam with three lumped mass elements connected at equal spacings. The parameter matrices determined by DPE were able to regenerate an FRF matrix which agreed very closely with the TFs measured using random excitation. For this structure, the identified model could be validated by several techniques including comparison of the mass and stiffness matrices with a theoretical model constructed from a physical standpoint. Also a method was presented for determining a linearised model of a non-linear structure.

The main topic of chapter 6 was the application of the stepped sine technique of measuring the higher order TFs from a structure, describing a series of tests performed on a non-linear beam rig. The non-linear behaviour of the beam was shown to be due principally to the tension restoring force which, as a cubic function of the displacement, gives rise to a hardening stiffness effect. This can be altered by preloading the beam. Since the non-linear function is polynomial in form, it was possible to measure very high quality first, second and third order TFs (the latter for the first time). Ultimately the discussion centred on the problem of validating the first order and non-linear parameters, concluding that the regeneration of the principal diagonals of the FRFs from the identified parameters may provide a means of validating the model by direct comparison with the measured higher order TFs. The development of this comparison test is recommended as a topic for future work.

In chapter 7, a structure was described which had been designed and built in this project with the intention of applying the identification procedures to the task of locating the source of non-linearity in a structure, the ultimate aim of the research. The non-linear unit proved to be a well-behaved SDOF system governed by a linear plus cubic stiffness function, and an excellent quality restoring force surface was measured in terms of transmissibility using sine excitation. Using the polynomials fitted to the cross-sections of the force surface in conjunction with the load-deflection curve measured in a static test, a non-linear differential equation of motion model was identified for this system. By attaching the unit to a cantilevered beam, it was intended that the complete assembly act as a multi-degree of freedom structure with localised non-linearity. Unfortunately, there was poor coupling between the bending modes of the beam and the oscillation of the unit, and only very weak non-linearity could be detected at the response locations along the beam. For subsequent tests, it is suggested the location of the unit on the beam be altered, or the mass of the store element be increased, in order to improve the influence of the unit on the dynamics of the beam.

Although further modifications are required, a design for a benchmark mechanical non-linear system has been established on which it will be possible in future to validate procedures for detecting non-linearity and identifying parametric models.

8.2 Conclusions

1. Conventionally differential equation of motion models have been formulated to represent the dynamics of single and multi-degree of freedom linear structures. However, when the non-linearity can be represented by a polynomial type stiffness or damping restoring force function, it is possible to extend the mathematical models to represent multi-degree of freedom non-linear structures.

2. Multi-dimensional Frequency Response Functions (FRFs) can be defined from the Volterra series representation of a non-linear system and related to the parameters of the non-linear differential equation of motion model. Although strictly restricted to weakly non-linear systems governed by polynomial type functions, the higher order FRFs are particularly useful for representing the dynamics structures in the frequency domain since each is defined to be a property of the system, and therefore independent of the type or level of the excitation. By regenerating the FRFs for a parametric model, it is possible to interpret the energy transfer mechanism which is exhibited by many physical structures.

3. Higher order TFs can be defined for a non-linear system which is subjected to sinusoidal excitation. Theoretically, for systems governed by continuous polynomial type non-linear functions, the higher order TFs tend to the principal diagonals of the multi-dimensional FRFs as the amplitude of excitation is decreased towards zero. By performing a stepped sine test, measurement of the higher order TFs presents a practical method of estimating the principal diagonals of the FRFs, and hence of detecting the presence of non-linearity in a structure.

4. In contrast to the FRFs, the higher order TFs are dependent on the type and level of excitation. Deviation of the applied excitation from a sinewave would cause the TFs to be corrupted. In practice some distortion of the sinewave always occurs when the excitation applied via an electrodynamic shaker to a structure. Furthermore, since it is often impractical to use low levels of excitation in a test, the TFs usually exhibit distortion which may be significant when compared to the corresponding FRF diagonal. The degree of distortion depends on the amplitude of the force signal and can be explained in terms of a non-linear energy transfer mechanism.

5. Further problems which may arise in the measurement of higher order TFs include the spurious phase delay effect introduced by bandpass filters on the measured force and response signals, since the higher order TFs correlate the fundamental component of

the excitation with higher harmonic components in the response spectra. High quality first, second and third order functions have been measured from small scale structures in the laboratory, although, to validate the technique as a generally reliable method of detecting non-linearity, it is essential that further developments to the test procedure be directed at large scale non-linear structures such as aircraft.

6. The conventional technique of comparing the FRFs regenerated from an identified mathematical model with measured TFs can be extended to non-linear structures using the principal diagonals of the higher order FRFs and the higher order TFs measured during a stepped sine test. In this way, the higher order TFs can be employed to validate the parametric models, although for this application it would be advantageous to improve the test procedure to reduce the amplitude dependent distortion around the resonances in the TFs.

7. The parameters of the non-linear equations of motion can be identified directly by using time signals representing the force excitation and the acceleration, velocity and displacement response signals at the locations of interest. When the structure physically resembles a lumped parameter model, it may be possible to interpret the parameters as being physical. However the model often departs from being physical for distributed parameter systems, and the inaccuracies which arise prevent the progression on to the interpretion of the non-linear parameters.

8.3 Recommendations for Further Work

Measuring Higher Order TFs

Although the distortion in the higher order TFs measured in a stepped sine test can provide an indication of the strength of the non-linearity in the system, it presents a problem when making the comparison with the (undistorted) FRFs regenerated from the identified model. In chapters 4 and 6 it was suggested that a possible improvement to the testing procedure would be to control the stepped sine test in such a way as to promote non-linear behaviour away from resonances but reduce distortion at resonance. To achieve this, it may be possible to control the excitation to cause the amplitude of either the force applied or the displacement response to remain nominally constant. Certainly further work should be directed at improving the stepped sine procedure for estimating the principal diagonals of the FRFs with accuracy.

Simulating Distributed Parameter Systems

One **requirement of the original project proposal was the development of simulation routines.** The mathematical model used for this task was the same as that employed by DPE, and the routines proved useful for correcting routines during development, as well as investigating the distortions in the TFs measured using sine excitation. However, it has been concluded that the lumped parameter model is generally deficient. at representing distributed parameter structures, and to be able to investigate in depth the differences which occur between the parametric models identified using DPE, and the theoretical models constructed from a physical standpoint, DPE should be applied in future to numerical simulations of distributed parameter systems. Certain FE packages such as PAFEC include routines whereby in the dynamic time domain solution stage localised or distributed non-linearities can be incorporated into a model known to be otherwise linear. In this way, time data can be simulated for a system with a high number of degrees of freedom (hence moving away from the lumped parameter model) whilst avoiding experimental problems, providing the basis for a rigorous investigation of the limitations DPE for identifying physical models.

Investigating the Accuracy of DPE

When the DPE procedure was introduced in chapter 1, comments were made regarding the problem of compounding errors in the parameter estimates by the passing of inaccurate scaling parameters. Although it is unlikely that the errors in the parameters are entirely responsible for the identified matrices departing from the physical model, clearly the solution procedure can be improved. Instead of involving time consuming optimisation procedures in the selection of scaling parameters, an alternative would be to impose, for example, a mass matrix on the solution so that each equation could be solved independently. This would have the effect of restricting the freedom of the solution and, by specifying a mass matrix which is has been determined from a physical standpoint, the identified stiffness and damping matrices would be forced towards the physical model. As mentioned at the end of chapter 7, FE modelling may provide one route to the determination of a physical mass matrix, and also may reveal a different perspective onto the subject of updating FE models.

Validating the Parametric Models

An important part of the identification process is the validation of the models. For the structures discussed in this thesis, it is relatively easy ***to*** determine whether or not the parametric models are physical by assessing the total of the mass estimates for instance. However for many structures, the only way to validate the models may be to predict the response of a structure to a different excitation and to compare the predicted with the

measured response. Although this test is rigorous and reflects the primary objective for identifying a parametric model, relatively small errors in the parameter estimates can cause large discrepancies between the signals compared. Clearly it would be worthwhile to develop alternative checks to assist in the validation and interpretation of the parameters identified.

Interpreting the Parametric Models

The ultimate objective proposed at the outset of the project was the identification of the location of non-linearity in a structure. The route taken was to develop methods for identifying mathematical models, with the intention of interpreting the non-linear parameters to indicate the source of non-linearity. Clearly it is first necessary to develop reliable techniques for identifying linear structures. The MDOF equation of motion models are very general when compared to conventional linear models, despite restricting the non-linear functions to have polynomial form. Unfortunately, by allowing each link to have non-linear stiffness and damping functions, there may be too much freedom in the solution. Consequently, it may prove necessary to restrict non-linearity on the basis of a-priori information or, by iterating the solution procedure, progressively omit terms which are deemed insignificant.

Investigating Alternatives to DPE

The parameter estimation techniques presented in this thesis represent only a fraction of the research which is currently being undertaken to identify mathematical models from physical systems. One alternative technique which appears very promising is the identification of difference models of linear or non-linear systems from input and output time data. It has been shown that the difference models can be related to the higher order FRFs in a similar way to the equations of motion. Future developments may make it possible to combine this technique with the procedures established in this project. For example, the principal FRFs regenerated from a NARMAX model of a structure could be curvefitted to identify a lumped parameter model by using the algorithms developed in chapter 3.

Investigating Strongly Non-Linear Systems

This project has focussed on weakly non-linear structures, drawing comparison with linear systems. Although it is known that the Volterra series is only applicable to weak non-linearities, nothing has been published regarding for example the restrictions of using higher order FRFs to describe bifurcation when the amplitude of excitation exceeds a certain level, or the application to strongly non-linear systems which become chaotic or exhibit parametric excitations. Although these phenomena have been addressed using other analysis techniques such as perturbation methods, no relationship with the Volterra series has been derived.

Bibliography

[1] Volterra V. *Theory of Functionals and of Integral Equations.* Dover Publications, New York, 1959.

[2] Frechet M. On continuous functionals. *Annals Scientifiques de l'Ecole Normale Superieure,* 27:193 to 216, 1910.

[3] Wiener N. Nonlinear *Problems in Random Theory.* John Wiley, New York, 1958.

[4] Barrett J F. The use of functionals in the analysis of non-linear physical systems. *Journal of Electronics and Control,* 15:567 to 615, 1963.

[5] Bedrossian E and Rice S 0. The output properties of volterra systems. *Proceedings of IEEE,* 591:688 to *707,* 1971.

[6] Lee Y W and Schetzen M. Measurement of the wiener kernels of a non-linear system by cross-correlation. *International Journal of Control,* 2(3):237 to *254,* 1965.

[7] Schetzen M. *The Volterra and Wiener Theories of Non-Linear Systems.* John Wiley, New York, 1980.

[8] Marmarelis P Z and Ken-Ika Naka. White noise analysis of a neuron chain: An application of the wiener theory. Science, 175:1276 to 1278, 1972.

[9] Marmarelis V Z, Masri S F, and Caughey T K. Analytical and experimental studies of the modelling of non-linear systems. Nuclear *Engineering and Design,* 55:59 to *68,* 1979.

[10] Marmarelis V Z and Marmarelis P Z. *Analysis of Physiological Systems: The White Noise Approach.* Plenum Press, New York, 1978.

[11] Hong J Y, Kim Y C, and Powers E J. On modelling the non-linear relationship between fluctuations with non-linear transfer functions. Proceedings of IEEE, 68(8):1026 and 1027, 1980.

[12] Bendat J S. *Non-Linear System Analysis and Identification.* John Wiley, 1990.

[13] Sandberg I W. On the properties of some systems that distort signals. *Bell Systems Technical Journal,* 43(1):91 to 112, 1964.

[14] Sandberg I W. The mathematical foundations of associated expansions for mildly non-linear systems. *IEEE Transactions on Circuits and Systems,* 30(7):441 to 455, 1983.

[15] Billings S A, Tsang K M, and Tomlinson G R. *Application of the NARMAX method to Non-Linear Frequency Response Estimation.* Proceedings of 6^{th} International Modal Analysis Conference, Orlando, USA, 1988.

[16] Peyton Jones J C and Billings S A. Recursive algorithm for computing the frequency response of a class of non-linear differential equation models. *International Journal of Control,* 50(4):1925 to 1940, 1989.

[17] Peyton Jones J C and Billings S A. Interpretation of non-linear frequency response functions. *International Journal of Control,* 52(2):319 to *346,* 1990.

[18] Vinh T, Chouychai T, Liu H, and Djouder M. *Second Order Transfer Function: Computation and Physical Interpretation.* Proceedings of 5^{th} International Modal Analysis Conference, London, UK, 1987.

[19] Chouychai T. *Dy*namic *Behaviour of Non-Linear Structures (Linearisation, Multidimensional Hilbert Transforms and the Volterra Series) : PhD Thesis.* ISMCM Saint-Ouen, Paris, France, 1986.

[20] Gifford S J and Tomlinson G R. Recent advances in the application of functional series to non-linear structures. *Journal of Sound and Vibration,* 135(2):289 to 317, 1989.

[21] Gifford S J and Tomlinson G R. *Understanding Multi-Degree of Freedom Non-Linear Systems via Higher Order Frequency Response Functions.* Proceedings of 7^{th} International Modal Analysis Conference, Las Vegas, USA, 1989.

[22] Mohammad K S, Worden K, and Tomlinson G R. *Direct Parameter Estimation for Linear and Non-Linear Structures.* Journal of Sound and Vibration, [To appear], 1991.

[23] Spina J F and Weiner D D. Sinusoidal *Analysis and Modelling of Weakly Non-Linear Circuits.* Van Nostrand Reinhold, 1980.

[24] Boyd S, Tang Y S, and Chua L 0. Measuring volterra kernels. *IEEE Transactions on Circuits and Systems,* 30(8):571 to *577,* 1983.

[25] Gifford S J. *Volterra Series Analysis of Non-Linear Structures : PhD Thesis.* Heriot-Watt University, Edinburgh, UK, 1989.

[26] Worden K. *Parametric and Non-Parametric Identification of Non-Linearity in Structural Dynamics :PhD Thesis.* Heriot-Watt University, Edinburgh ,UK, 1989.

[27] Goyder H G. *Measurement of Frequency Response Functions from Non-Linear Systems Using Random Excitation.* Short Course on the Dynamics of Linear and Non-Linear Structures, Heriot-Watt University, Edinburgh, UK, 1989.

[28] Ewins D J. *Modal Testing: Theory and Practice.* Research Studies Press, Letchworth, UK, 1984.

[29] Goyder H G. Methods and application of structural modelling from measured frequency response data. *Journal of Sound and Vibration,* 68(2):209 to *230,* 1980.

[30] Cooper J E. *Modal Parameter Identification using Time Domain Methods : PhD Thesis.* Queen Mary College, London, UK, 1988.

[31] Fritzen C. Frequency domain direct parameter estimation for modal analysis. *ASME Journal of Vibration, Acoustics, Stress and Reliability in* Design, 108:9 to 16, 1986.

[32] Lembregts F, Leuridan J, and Brussel H V. *Frequency Domain Direct Parameter Estimation for Modal Analysis.* Proceedings of 13^{th} International Conference on Modal Analysis, Katholieke University, Leuven, Belgium, 1988.

[33] Billings S A. Identification of non-linear systems : A survey. *IEE* Proceedings, 127 (D):272 to 285, 1980.

[34] Press W H, Flannery B P, Teukolsky S A, and Vetterling W T. Numerical *Recipes : The Art of Scientific Computing.* Cambridge University Press, 1986.

[35] Mohammad K S. *Identification of the Characteristics of Non-Linear Structures : PhD Thesis.* Heriot-Watt University, Edinburgh, UK, 1989.

[36] Liang Y C and Cooper J E. *Advances in Physical Parameter Estimation of Linear and Non-Linear Structures.* Proceedings of 4th International Conference on Recent Advances in Structural Dynamics, Southampton, UK, 1991.

[37] Masri S F, Miller R K, Saud A F, and Caughey T K. Identifcation of non-linear vibrating structures: Formulation and applications. *Journal of Applied Mechanics,* 54:918 to 929, 1987.

[38] Surace C, Worden K, and Tomlinson G R. On *the Non-Linear Characteristics of Automotive Shock Absorbers.* Proceedings of 10th International Modal Analysis Conference, San Diego, USA, 1992.

[39] Worden K. Data processing and experimental design for the restoring force surface method. *Journal of Mechanical Systems and Signal Processing,* 4(4):295 to 319, 1990.

[40] Al-Hadid M A. Id*enti fiation of* Non-Linear *Dynamic Systems Using the Force State Mapping Technique : PhD Thesis.* Queen Mary College, London, UK, 1989.

[41] Al-Hadid M A and Wright J R. *A Method for the Estimation of Mass and Modal Mass in the Identification of Non-Linear Single- and Multi-Degree of Freedom Systems using the Force-State Mapping Approach.* Journal of Mechanical Systems and Signal Processing, [To appear], 1991.

[42] Worden K and Tomlinson G R. *Identification of Linear/Non-Linear Restoring Force Surfaces in Single and Multi-Mode Systems.* Proceedings of 3rd International Conference on Recent Advances in Structural Dynamics, Southampton, UK, 1988.

[43] Thompson J M and Stewart H B. *Non-Linear Dynamics and Chaos.* John Wiley, 1986.

[44] Hung G, Stark L, and Eykhoff P. On the interpretation of kernels. *Annals of Biomedical Engineering,* 5:130 to 143, 1977.

[45] Bracewell R. *The Fourier Transform and Its Applications.* McGraw-Hill, 1965.

[46] Frachebourg A and Kuster F. Practical application of the volterra series: Modal analysis of higher order. *Annals of the CIRP,* 37(1):347 to *350,* 1988.

[47] Storer D M and Tomlinson G R. *Higher Order Frequency Response Functions and their Relation to Practical Structures.* Proceedings of 16th International Seminar on Modal Analysis, Florence, Italy, 1991.

[48] Tsang K M. *Spectral Analysis for Non-Linear Systems : PhD Thesis.* University of Sheffield, Sheffield ,UK, 1988.

[49] Bendat J S and Piersol A C. *Random Data: Analysis and Measurement Procedures.* John Wiley, 1971.

[50] Powers E J and Miksad R W. *Polyspectral Measurement and Analysis of Non-Linear Wave Interactions.* ASME Symposium on Non-Linear Wave Interactions in Fluids, Boston, USA, 1987.

[51] Kim K I, Powers E J, Ritz C P, Miksad R W, and Fischer F J. Modelling of the non-linear drift oscillations of moored vessels subject to non-gaussian random sea-wave excitation. *IEEE Journal of Oceanic Engineering,* 12(4):568 to 575, 1987.

[52] Raghuveer M R and Nikias C L. Bispectrum estimation: A parametric approach. *IEEE Transactions on Acoustics, Speech and Signal Processing,* 33(4):1213 to 1229, 1985.

[53] Chang J H, Stearman R 0, Choi D, and Powers E J. *Identification of Aeroelastic Phenomenon Employing Bispectml Analysis Techniques.* Proceedings of 3rd International Modal Analysis Conference, Orlando, USA, 1985.

[54] Nam S W, Kim S B, and Powers E J. *Utilization of Digital Polyspectral Analysis to Estimate Transfer Functions of Cubically* Non-Linear *Systems.* International Conference on Acoustics, Speech and Signal Processing, Glasgow, UK, 1989.

[55] Dalle Molle J W and Hinich M J. *Trispectral Analysis.* The University of Texas at Austin, Applied Research Laboratories, Austin, Texas, 1989.

[56] Schetzen M. Measurement of the kernels of a system of finite order. *International Journal of Control,* 1:251 to 263, 1980.

[57] Hamming R W. *Digital Filters.* Prentice Hall, 1983.

[58] Liu H. *Multi-Dimensional Signal Processing for Non-Linear Structural Dynamics : PhD Thesis.* ISMCM Saint-Ouen, Paris, France, 1989.

[59] Kim K I and Powers E J. A digital method of modelling quadratically non-linear systems with a general random input. *IEEE Transactions on Acoustics,Speech and Signal Processing,* 36:1758 to 1769, 1988.

[60] **Bert C W and Stricklin J D. Comparative evaluation of six different numerical integration methods for non-linear dynamic systems.** *Journal of Sound and* **Vibration,** 127(2):221 **to 229, 1988.**

[61] **Simon M and Tomlinson G R. Application of the hilbert transform in the modal analysis of linear and non-linear systems.** *Journal of Sound and Vibration,* 90(2):275 **to** 282, 1984.

[62] Worden **K and Tomlinson G R.** *The High Frequency Behaviour of Frequency Response Functions and their Hilbert Transforms.* **Proceedings of** 9^{th} **International Modal Analysis Conference, Florida, USA, 1990.**

[63] **Goyder H G. S***ome Theory and Applications of the Relationship between the Real and Imaginary Parts of a Frequency Response Function Provided by the Hilbert Transform.* 2^{nd} **International Conference on Recent Advances in Structural Dynamics, Southampton, UK, 1984.**

[64] **Tomlinson G R and Ahmed I.** *Hilbert Transform Procedures for Detecting and Quantifying Non-Linearity in Modal Testing.* **Proceedings of** 8^{th} **National Congress of** AIMETA, **Turin, Italy, 1986.**

[65] **Ahmed I.** *Develop ments in Hilbert Transform Procedures with Applications to Linear and Non-Linear Structures : PhD Thesis.* **University of Manchester, Manchester, UK, 1987.**

[66] **Blevins R D.** *Formulas for Natural Frequencies and Modeshapes.* **Van Nonstrand Reinhold, 1979.**

[67] **Wright J R.** *Notes on Ineraction between the Electrodynamic Exciter and Structure in Vibration Testing and its Influence on the Applied Force.* **Short Course on Signal Processing and Modal Testing, University of Manchester, UK, 1991.**

[68] **Zhang Q, Shelley S, and Brown D L.** *Compensation for Actuator Dynamic Effects using Model Following Adaptive Control.* **Proceedings of** 6^{th} **International Modal Analysis Conference, Orlando, USA, 1988.**