NONLINEARITY IN STRUCTURAL DYNAMICS

NONLINEARITY IN STRUCTURAL DYNAMICS

Detection, Identification and Modelling

K Worden and G R Tomlinson

University of Sheffield, UK

IoP

Institute of Physics Publishing
Bristol and Philadelphia

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library.

ISBN 0 7503 0356 5

Library of Congress Cataloging-in-Publication Data are available

Commissioning Editor: James Revill
Production Editor: Simon Laurenson
Production Control: Sarah Plenty
Cover Design: Victoria Le Billon
Marketing Executive: Colin Fenton

Published by Institute of Physics Publishing, wholly owned by The Institute of Physics, London

Institute of Physics Publishing, Dirac House, Temple Back, Bristol BS1 6BE, UK

US Office: Institute of Physics Publishing, The Public Ledger Building, Suite 1035, 150 South Independence Mall West, Philadelphia, PA 19106, USA

Typeset in TEX using the IOP Bookmaker Macros
Printed in the UK by MPG Books Ltd, Bodmin

For Heather and Margaret

‘As you set out for Ithaka
hope your road is a long one,
full of adventure, full of discovery.
Laistrygonians, Cyclops,
angry Poseidon—don’t be afraid of them:
You’ll never find things like that in your way
as long as you keep your thoughts raised high,
as long as a rare sensation
touches your body and spirit.
Laistrygonians, Cyclops,
wild Poseidon—you won’t encounter them
unless you bring them along inside your soul,
Unless your soul sets them up in front of you.’

C P Cavafy, ‘Ithaka’

Contents

Preface

Nonlinearity is a frequent visitor to engineering structures which can modify—sometimes catastrophically—the design behaviour of the systems. The best laid plans for a linear system will often go astray due to, amongst other things, clearances and interfacial movements in the fabricated system. There will be situations where this introduces a threat to human life; several illustrations spring to mind. First, an application in civil engineering. Many demountable structures such as grandstands at concerts and sporting events are prone to substantial structural nonlinearity as a result of looseness of joints, this creates both clearances and friction and may invalidate any linear-model-based simulations of the behaviour created by crowd movement. A second case comes from aeronautical structural dynamics; there is currently major concern in the aerospace industry regarding the possibility of limit cycle behaviour in aircraft, i.e. large amplitude coherent nonlinear motions. The implications for fatigue life are serious and it may be that the analysis of such motions is as important as standard flutter clearance calculations. There are numerous examples from the automotive industry; brake squeal is an irritating but non-life-threatening example of an undesirable effect of nonlinearity. Many automobiles have viscoelastic engine mounts which show marked nonlinear behaviour: dependence on amplitude, frequency and preload. The vast majority of engineers—from all flavours of the subject—will encounter nonlinearity at some point in their working lives, and it is therefore desirable that they at least recognize it. It is also desirable that they should understand the possible consequences and be in a position to take remedial action. The object of this book is to provide a background in techniques specific to the field of structural dynamics, although the ramifications of the theory extend beyond the boundaries of this discipline.

Nonlinearity is also of importance for the diagnosis of faults in structures. In many cases, the occurrence of a fault in an initially linear structure will result in nonlinear behaviour. Another signal of the occurrence of damage is the variation with time of the system characteristics.

The distinction between linear and nonlinear systems is important; nonlinear systems can exhibit extremely complex behaviour which linear systems cannot. The most spectacular examples of this occur in the literature relating to chaotic systems [248]; a system excited with a periodic driving force can exhibit an

apparently random response. In contrast, a linear system always responds to a periodic excitation with a periodic signal at the same frequency. At a less exotic level, but no less important for that, the stability theory of linear systems is well understood [207]; this is emphatically not the case for nonlinear systems.

The subject of nonlinear dynamics is extremely broad and an extensive literature exists. This book is inevitably biased towards those areas which the authors are most familiar with and this of course means those areas which the authors and colleagues have conducted research in. This review is therefore as much an expression of personal prejudice and taste as anything else, and the authors would like to sincerely apologise for any inadvertent omissions. This is not to say that there are no *deliberate* omissions; these have good reasons which are explained here.

- There is no real discussion of nonlinear dynamical systems theory, i.e. phase space analysis, bifurcations of systems and vector fields, chaos. This is a subject best described by the more mathematically inclined and the reader should refer to many excellent texts. Good introductions are provided by [79] and [12]. The monograph [125] is already a classic and an overview suited to the Engineer can be found in [248].
- There is no attempt to summarize many of the developments originating in control theory. The geometrical approach to nonlinearity pioneered by Brockett has led to very little concrete progress in mainstream structural dynamics beyond making rigorous some of the techniques adopted lately. The curious reader is directed to the introduction [259] or to the classic monograph [136]. Further, there is no discussion of any of the schemes based on Kalman filtering—again the feeling of the authors is that this is best left to control engineers.
- There is no discussion of some of the recent approaches based on spectral methods. Many of these developments can be traced back to the work of Bendat, who has summarized the background admirably in his own monograph [25] and the recent update [26]. The 'reverse-path' approach typified by [214] can be traced back through the recent literature survey [2]. The same authors, Adams and Allemang, have recently proposed an interesting method based on frequency response function analysis, but it is perhaps a little early to judge [3].
- There is no discussion of *nonlinear normal modes*. Most research in structural dynamics in the past has concentrated on the effect of nonlinearity on the resonant frequencies of systems. Recently, there has been interest in estimating the effect on the modeshapes. The authors here feel that this has been dealt with perfectly adequately in the monograph [257]. There is also a useful recent review article [258].

So, what *is* in this book? The following is a brief outline.

Chapter 1 describes the relevant background in linear structural dynamics. This is needed to understand the rest of the book. As well as describing

the fundamental measured quantities like the impulse response function (IRF) and the frequency response function (FRF) it serves to introduce notation. The backgrounds for both continuous-time systems (those based on differential equations of motion) and discrete-time (those based on difference equations) are given. The chapter begins by concentrating on single-degree-of-freedom (SDOF) linear systems and finally generalizes to those with multiple-degrees-of-freedom (MDOF) with a discussion of modal analysis.

Chapter 2 gives essentially the 'classical' approaches to nonlinearity which have longest been within reach of structural dynamicists. This basically means approaches which can make use of standard dynamic testing equipment like frequency response analysers. Ideas like FRF distortion and coherence are discussed here. The chapter also discusses how nonlinearity can enter the measurement chain and introduces some of the more common types of nonlinearity. Finally, the idea of linearization is introduced. This chapter is not just of historical interest as most of the instrumentation commonly available commercially is still extremely restricted in its ability to deal with nonlinearity.

Chapter 3. Having discussed FRF distortion, this chapter shows how to compute FRFs for nonlinear systems. It describes how each type of nonlinearity produces its own characteristic distortions and how this can lead to qualitative methods of analysis. The chapter also discusses how nonlinear systems do not follow certain behaviour patterns typical of linear systems. It shows how nonlinear systems subject to periodic forcing can respond at harmonics and combination frequencies of the forcing frequencies. The chapter concludes with an analysis of IRF distortion.

Chapter 4 introduces more modern methods of analysis, in particular those which cannot be implemented on conventional instrumentation. The subject of this chapter is the Hilbert transform. This versatile technique can not only detect nonlinearity but also, in certain circumstances, estimate the equations of motion, i.e. solve the *system identification* problem. All the basic theory is given, together with detailed discussion of how to implement the technique.

Chapter 5 continues the discussion of the Hilbert transform from a completely different viewpoint; namely that of complex analysis. Although this chapter does give some extremely interesting results, it places rather more demands on the reader from a mathematical point of view and it can be omitted on first reading. A background in the calculus of residues is needed.

Chapter 6 provides the first discussion of *system identification*, i.e. the vexed question of estimating equations of motion for systems based only on measurements of their inputs and outputs. The particular viewpoint of this chapter is based on discrete-time equations, more specifically the powerful and general NARMAX method. This chapter also provides the most

complete description in this book of the effects of measurement noise and the need for rigorous model validity testing. Finally, the chapter introduces the idea of neural networks and shows how they can be used to identify models of systems.

Chapter 7 balances the discussion of system identification by giving the continuous-time point of view. The approach is not at all general but follows a class of models devised by Masri and Caughey and termed here restoring force surfaces (RFS). The development of MDOF approaches are addressed and a simpler, more powerful, variant of the idea is discussed. The chapter concludes with a discussion of how the system identification problem can be posed in terms of optimization and how this makes available a number of powerful techniques from mathematics.

Chapter 8 shows one approach to generalizing the idea of the FRF from linear systems to nonlinear. The method—based on a type of functional power series—defines an infinite set of impulse response functions or FRFs which can characterize the behaviour of a class of nonlinear systems. The interpretation of the higher-order FRFs is discussed and it is also shown how the approach can give a means of identifying equations of motion of general MDOF systems—essentially a multi-dimensional version of modal analysis.

Chapter 9 is most concerned with practical matters. The object was to describe some simple (and one not-so-simple) laboratory rigs which can be used to illustrate and validate the techniques developed in the earlier chapters.

A substantial set of appendices contain useful material which would otherwise interrupt the flow of the discussion. Amongst other things these discuss: basic probability theory, neural networks and the integration and differentiation of measured time data.

Having discussed the contents, it is important to identify the potential readership. If the reader has leafed through the remaining pages of this book, it is possible that the number of equations has appeared daunting. This is actually rather deceptive. The mathematics required of the reader is little more than a capability of dealing with matrices, vectors, *linear* differential equations and Fourier analysis. Certainly nothing which would not be covered in a degree in a numerate discipline: mathematics, physics or some flavour of engineering. The exceptions to this rule come in chapter 5 and in one section of chapter 8. There, the reader is required to know a little complex analysis, namely how to evaluate integrals using the calculus of residues. These sections can be omitted on a first reading—or omitted altogether for that matter—without losing the thread of the book. This means that the book is accessible to anyone who is in the later stages of a degree in the disciplines previously identified. It is also suitable for study at a beginning postgraduate level and also as a survey of the field of nonlinearity for an expert structural dynamicist.

A book like this does not spring into being without a lot of help from a lot of people. It is a pleasure to thank them. First of all, much of this material is the result of collaboration with various colleagues and friends over the years; (in roughly chronological order) the authors would like to thank: Matthew Simon, Neil Kirk, Ian Kennedy, Ijaz Ahmed, Hugh Goyder, Steve Billings, Steve Gifford, Khalid Mohammad, Mike Reid, Tunde Oyadiji, David Storer, Roy Chng, Jan Wright, Jonathon Cooper, Wieslaw Staszewski, Qian Chen, Nigel King, Mike Hamilton, Steve Cafferty, Paul Holmes, Graeme Manson, Julian Chance, Brian Deacon, Robin Wardle, Sophoclis Patsias and Andreas Kyprianou. In many cases, the authors have shamelessly lifted figures from the PhD theses and publications of these collaborators and they would like to offer thanks for that. A special mention must go to Professor Tuong Vinh who, as a close friend and valued colleague, provided continuous inspiration and guidance to Geof Tomlinson in his early career; without his encouragement, the road may have been a linear one.

In terms of producing the manuscript, the authors are grateful to: Steve Billings, Steve Gifford and particularly Graeme Manson and Heather Worden for their critical readings of portions of the manuscript. Also Julian Chance and (predominantly) Jonny Haywood did a valiant job of translating a mass of disorganized sketches and photocopies into a beautiful sequence of postscript files. The book would certainly not exist in this form without the efforts of these people; nonetheless, any mistakes or omissions which exist are entirely the fault of the authors (who would be grateful if the readers could bring them to their attention).

Thank you for reading this far, the authors sincerely hope that it will be useful and illuminating to carry on further.

K Worden
G R Tomlinson
Sheffield 2000

Chapter 1

Linear systems

This chapter is provided more or less as a reminder of linear system theory. It is not comprehensive and it is mainly intended to set the scene for the later material on nonlinearity. It brings to the attention of the reader the basic properties of linear systems and establishes notation. Parts of the theory which are not commonly covered in elementary textbooks are treated in a little more detail.

Any book on engineering dynamics or mechanical vibrations will serve as reference for the following sections on continuous-time systems, e.g. Thompson [249] or the more modern work by Inman [135]. For the material on discrete-time systems, any recent book on system identification can be consulted, Söderstrom and Stoica [231] is an excellent example.

1.1 Continuous-time models: time domain

How does one begin to model dynamical systems? Starting with the simplest possible system seems to be sensible; it is therefore assumed that the system is a single point particle of mass m moving in one dimension subject to an applied force $x(t)$[1]. The equation of motion for such an object is provided by Newton's second law,

$$\frac{\mathrm{d}}{\mathrm{d}t}(mv) = x(t) \tag{1.1}$$

where v is the velocity of the particle. If the mass m is constant, the equation becomes

$$ma(t) = x(t) \tag{1.2}$$

where $a(t)$ is the acceleration of the particle. If the displacement $y(t)$ of the particle is the variable of interest, this becomes a second-order *differential*

[1] In general, the structures of Engineering significance are continuous: beams, plates, shells and more complicated assemblies. Such systems have partial differential equations of motion dictating the behaviour of an infinite number of degrees-of-freedom (DOF). This book is concerned only with systems with a finite number of DOF as even a small number is sufficient to illustrate fully the complexities of nonlinear systems.

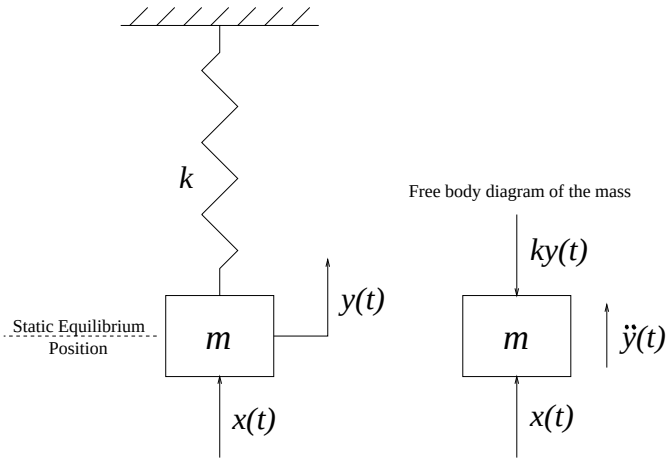


Figure 1.1. SDOF mass–spring system.

equation,

$$m\frac{\mathrm{d}^2 y}{\mathrm{d}t^2} = x(t) \tag{1.3}$$

or

$$m\ddot{y} = x(t) \tag{1.4}$$

in the standard notation where overdots denote differentiation with respect to time. Apart from the obvious restrictions (all real systems have more than one DOF), this equation is unrealistic in that there is no resistance to the motion. Even if $x(t) = 0$, the particle can move with constant velocity. The simplest way of providing resistance to motion is to add an *internal* or *restoring force* $f_{\mathrm{r}}(y)$ which always acts in the opposite direction to the motion, i.e.

$$m\ddot{y} = x(t) - f_{\mathrm{r}}(y). \tag{1.5}$$

The paradigm for this type of equation is a mass on a spring (figure 1.1). The form of the restoring force in this case is given by Hooke's law, for a static displacement y of the mass, the restoring force is given by

$$f_{\mathrm{r}}(y) = ky \tag{1.6}$$

where k is the stiffness constant of the spring. Substituting into the equation of motion gives

$$m\ddot{y} + ky = x(t). \tag{1.7}$$

Note that as the restoring force vanishes when $y = 0$, this will be the static equilibrium position of the motion, i.e. the position of rest when there is no force.

In structural dynamics, it is traditional to use k for the coefficient of y and to refer to it as the *elastic stiffness* or simply *stiffness* of the system.

The solution of (1.7) is elementary and is given in any book on vibrations or differential equations [227]. An interesting special case is where $x(t) = 0$ and one observes the unforced or *free* motion,

$$\ddot{y} + \frac{k}{m} y = 0. \tag{1.8}$$

There is a trivial solution to this equation given by $y(t) = 0$ which results from specifying the initial conditions $y(0) = 0$ and $\dot{y}(0) = 0$. Any point at which the mass can remain without motion for all time is termed an *equilibrium* or *fixed point* for the system. It is clear from the equation that the only equilibrium for this system is the origin $y = 0$, i.e. the static equilibrium position. This is typical of linear systems but need not be the case for nonlinear systems. A more interesting solution results from specifying the initial conditions $y(0) = A$, $\dot{y} = 0$, i.e. the mass is released from rest at $t = 0$ a distance A from the equilibrium. In this case,

$$y(t) = A \cos(\omega_n t). \tag{1.9}$$

This is a periodic oscillation about $y = 0$ with angular frequency $\omega_n = \sqrt{\frac{k}{m}}$ radians per second, frequency $f_n = \frac{1}{2\pi}\sqrt{\frac{k}{m}}$ Hz, and period of oscillation $T_n = 2\pi\sqrt{\frac{m}{k}}$ seconds. Because the frequency is of the free oscillations it is termed the *undamped natural frequency* of the system, hence the subscript n.

The first point to note here is that the oscillations persist without attenuation as $t \to \infty$. This sort of behaviour is forbidden by fundamental thermodynamic constraints, so some modification of the model is necessary in order that free oscillations are not allowed to continue indefinitely. If one thinks in terms of a mass on a spring, two mechanisms become apparent by which energy is dissipated or *damped*. First, unless the motion is taking place in a vacuum, there will be resistance to motion by the ambient fluid (air in this case). Second, energy will be dissipated in the material of the spring. Of these two dissipation processes, only the first is understood to any great extent. Fortunately, experiment shows that it is fairly common. In fact, at low velocities, the fluid offers a resistance proportional to and in opposition to the velocity of the mass. The *damping force* is therefore represented by $f_{\mathrm{d}}(\dot{y}) = -c\dot{y}$ in the model, where c is the *damping constant*. The equation of motion is therefore,

$$m\ddot{y} = x(t) - f_{\mathrm{d}}(\dot{y}) - f_{\mathrm{r}}(y) \tag{1.10}$$

or

$$m\ddot{y} + c\dot{y} + ky = x(t). \tag{1.11}$$

This equation is the equation of motion of a single point mass moving in one dimension, such a system is referred to as single degree-of-freedom (SDOF). If the point mass were allowed to move in three dimensions, the displacement $y(t)$ would be a vector whose components would be specified by three equations

of motion. Such a system is said to have three degrees-of-freedom and would be referred to as a multi-degree-of-freedom (MDOF) system. A MDOF system would also result from considering the motion of an assembly of point particles.

Note that as a differential equation, (1.4) is *linear*. An important consequence of this is the *Principle of Superposition* which can be stated as follows:

> If the response of the system to an arbitrary applied force $x_1(t)$ is $y_1(t)$, and to a second independent input $x_2(t)$ is $y_2(t)$, then the response to the superposition $\alpha x_1(t) + \beta x_2(t)$ (with appropriate initial conditions) is $\alpha y_1(t) + \beta y_2(t)$ for *any* values of the constants α, β.

This is discussed in more detail in chapter 2.

Systems whose equations of motion are differential equations are termed *continuous-time systems* and the evolution of the system from given initial conditions is specified for a continuum of times $t \geq 0$.

Returning now to the equation (1.11), elementary theory shows that the solution for the free motion ($x(t) = 0$) with initial conditions $y(0) = A$, $\dot{y} = 0$ is

$$y_{\mathrm{t}}(t) = A\mathrm{e}^{-\zeta\omega_n t}\cos(\omega_{\mathrm{d}} t) \tag{1.12}$$

where

$$\zeta = \frac{c}{2\sqrt{mk}} \tag{1.13}$$

$$\omega_{\mathrm{d}} = \omega_n(1-\zeta^2)^{\frac{1}{2}} \tag{1.14}$$

and $\omega_n = \sqrt{\frac{k}{m}}$ is the undamped natural frequency. The frequency of free oscillations in this case is $\omega_{\mathrm{d}} \neq \omega_n$ and is termed the *damped natural frequency*; ζ is the *damping ratio*. The main features of this solution can be summarized as follows.

- The damped natural frequency is always less than the undamped natural frequency which it approaches in the limit as $c \to 0$ or equivalently as $\zeta \to 0$.
- If $1 > \zeta > 0$ the oscillations decay exponentially with a certain *time constant* τ_ζ. *This is defined as the time taken for the amplitude to decay from a given value Y, to the value Y/e; where e is the base for natural logarithms.* It follows that $\tau_\zeta = \frac{1}{\zeta\omega_n}$. Because of this, the solution (1.12) is termed the *transient* solution (hence the subscript 't' on the response). If $\zeta < 0$ or, equivalently, $c < 0$ the oscillations grow exponentially (figure 1.3). In order to ensure that the system is stable (in the sense that a bounded input generates a bounded output), ζ and hence c must be positive.
- If $\zeta = 1$, then $\omega_{\mathrm{d}} = 0$ and the system does not oscillate but simply tends monotonically from $y(0) = A$ to zero as $t \to \infty$ (figure 1.4). The system is said to be *critically damped*. The critical value for the damping constant c is easily seen to be $2\sqrt{mk}$.

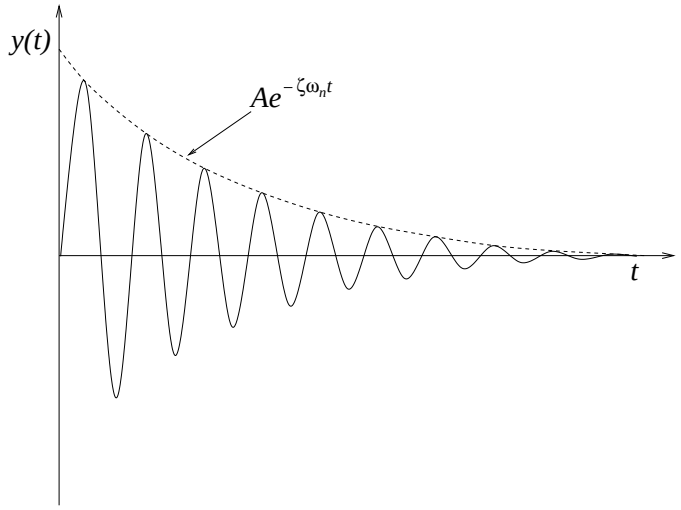


Figure 1.2. Transient motion of a SDOF oscillator with positive damping. The envelope of the response is also shown.

- If $\zeta > 1$, the system is said to be *overdamped* and the situation is similar to critical damping, the system is non-oscillatory but gradually returns to its equilibrium when disturbed. Newland [198] gives an interesting discussion of overdamped systems.

Consideration of the free motion has proved useful in that it has allowed a physical *positivity* constraint on ζ or c to be derived. However, the most interesting and more generally applicable solutions of the equation will be for forced motion. If attention is restricted to deterministic force signals $x(t)$[2], Fourier analysis allows one to express an arbitrary periodic signal as a linear sum of sinusoids of different frequencies. One can then invoke the principle of superposition which allows one to concentrate on the solution where $x(t)$ is a single sinusoid, i.e.

$$m\ddot{y} + c\dot{y} + ky = X\cos(\omega t) \tag{1.15}$$

where $X > 0$ and ω is the *constant* frequency of excitation. Standard differential equation theory [227] asserts that the general solution of (1.15) is given by

$$y(t) = y_{\mathrm{t}}(t) + y_{\mathrm{s}}(t) \tag{1.16}$$

where the *complementary function* (or transient response according to the earlier notation) $y_{\mathrm{t}}(t)$ is the unique solution for the *free* equation of motion and contains arbitrary constants which are fixed by initial conditions. $y_{\mathrm{t}}(t)$ for equation (1.15)

[2] It is assumed that the reader is familiar with the distinction between *deterministic* signals and those which are *random* or *stochastic*. If not, [249] is a good source of reference.

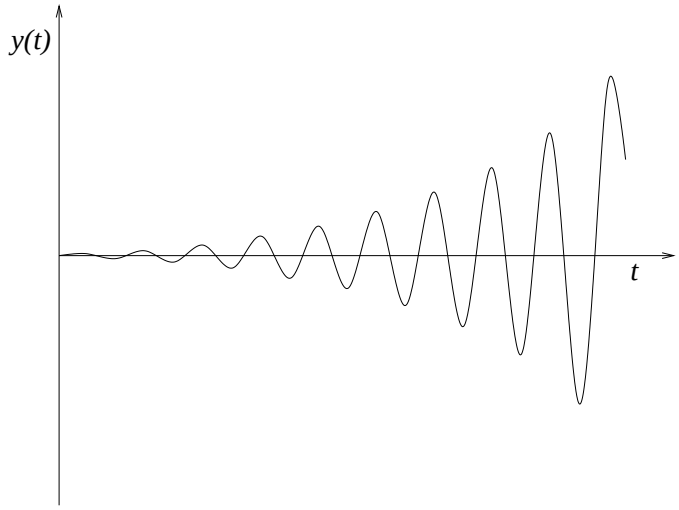


Figure 1.3. Unforced motion of a SDOF oscillator with negative damping. The system displays instability.

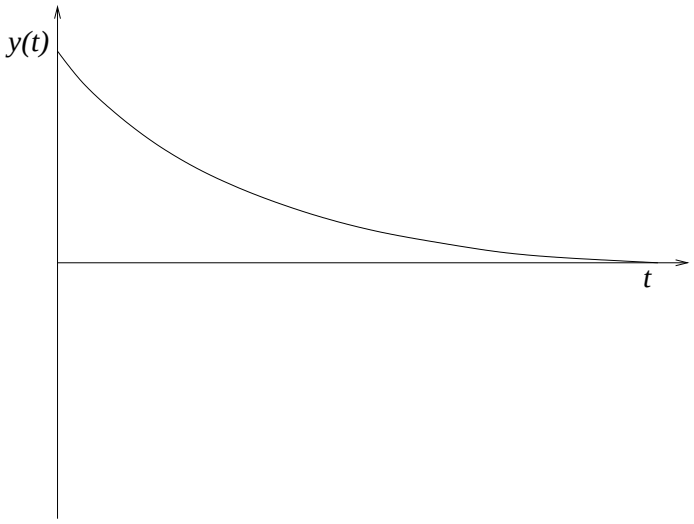


Figure 1.4. Transient motion of a SDOF oscillator with critical damping showing that no oscillations occur.

is therefore given by (1.12). The remaining part of the solution $y_{\mathrm{s}}(t)$, the *particular integral*, is independent of the initial conditions and persists after the transient $y_{\mathrm{t}}(t)$ has decayed away. For this reason $y_{\mathrm{s}}(t)$ is termed the *steady-state*

response of the solution.

For linear systems, the steady-state response to a periodic force is periodic with the same frequency, but not necessarily in phase due to the energy dissipation by the damping term which causes the output to lag the input. In order to find $y_{\rm s}(t)$ for (1.15), one substitutes in the trial solution

$$y_{\rm s}(t) = Y\cos(\omega t - \phi) \tag{1.17}$$

where $Y > 0$ and obtains

$$-m\omega^2 Y\cos(\omega t-\phi)+-c\omega Y\sin(\omega t-\phi)+kY\cos(\omega t-\phi) = X\cos(\omega t). \tag{1.18}$$

A shift of the time variable $t \to t + (\phi/\omega)$ yields the simpler expression,

$$\begin{aligned} -m\omega^2 Y\cos(\omega t) + -c\omega Y\sin(\omega t) + kY\cos(\omega t) &= X\cos(\omega t + \phi) \\ = X\cos(\omega t)\cos\phi - X\sin(\omega t)\sin\phi. \end{aligned} \tag{1.19}$$

Equating coefficients of sin and cos gives

$$-m\omega^2 Y + kY = X\cos\phi \tag{1.20}$$

$$c\omega Y = X\sin\phi. \tag{1.21}$$

Squaring and adding these equations gives

$$\{(-m\omega^2 + k)^2 + c^2\omega^2\}Y^2 = X^2(\cos^2\phi + \sin^2\phi) = X^2 \tag{1.22}$$

so that

$$\frac{Y}{X} = \frac{1}{\sqrt{(-m\omega^2 + k)^2 + c^2\omega^2}}. \tag{1.23}$$

This is the *gain* of the system at frequency ω, i.e. the proportional change in the amplitude of the signal as it passes through the system $x(t) \to y(t)$. Because X and Y are both positive real numbers, so is the gain.

Taking the ratio of equations (1.21) and (1.20) yields

$$\tan\phi = \frac{c\omega}{k - m\omega^2}. \tag{1.24}$$

The *phase* ϕ represents the degree by which the output signal $y(t)$ lags the input $x(t)$ as a consequence of passage through the damped system.

One can now examine how the response characteristics vary as the excitation frequency ω is changed. First, one can rewrite equation (1.23) in terms of the quantities ω_n and ζ as

$$\frac{Y}{X}(\omega) = \frac{1}{m\sqrt{(\omega^2 - \omega_n^2)^2 + 4\zeta^2\omega_n^2\omega^2}}. \tag{1.25}$$

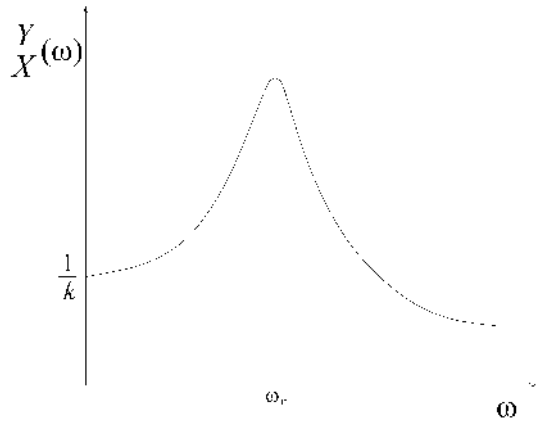


Figure 1.5. SDOF system gain as a function of frequency ω.

This function will clearly be a maximum when

$$(\omega^2 - \omega_n^2)^2 + 4\zeta^2\omega_n^2\omega^2 \tag{1.26}$$

is a minimum, i.e. when

$$\frac{\mathrm{d}}{\mathrm{d}\omega}[(\omega^2 - \omega_n^2)^2 + 4\zeta^2\omega_n^2\omega^2] = 4\omega(\omega^2 - \omega_n^2) + 8\zeta^2\omega_n^2\omega = 0 \tag{1.27}$$

so that

$$\omega^2 = \omega_n^2(1 - 2\zeta^2). \tag{1.28}$$

This frequency corresponds to the *only* extreme value of the gain and is termed the *resonant* or *resonance* frequency of the system and denoted by ω_{r}. Note that for the damped system under study $\omega_{\mathrm{r}} \neq \omega_{\mathrm{d}} \neq \omega_n$. It is easy to show that for an undamped system $\omega_{\mathrm{r}} = \omega_{\mathrm{d}} = \omega_n$ and that the gain of the undamped system is infinite for excitation at the resonant frequency. In general if the excitation is at $\omega = \omega_{\mathrm{r}}$, the system is said to be *at resonance*.

Equation (1.23) shows that $\frac{Y}{X} = \frac{1}{k}$ when $\omega = 0$ and that $\frac{Y}{X} \to 0$ as $\omega \to \infty$. The information accumulated so far is sufficient to define the (qualitative) behaviour of the system gain as a function of the frequency of excitation ω. The resulting graph is plotted in figure 1.5.

The behaviour of the phase $\phi(\omega)$ is now needed in order to completely specify the system response as a function of frequency. Equation (1.24) gives

$$\tan\phi(\omega) = \frac{c\omega}{m(\omega_n^2 - \omega^2)} = \frac{2\zeta\omega_n\omega}{\omega_n^2 - \omega^2}. \tag{1.29}$$

As $\omega \to 0$, $\tan\phi \to 0$ from above, corresponding to $\phi \to 0$. As $\omega \to \infty$, $\tan\phi \to 0$ from below, corresponding to $\phi \to \pi$. At $\omega = \omega_n$ the undamped

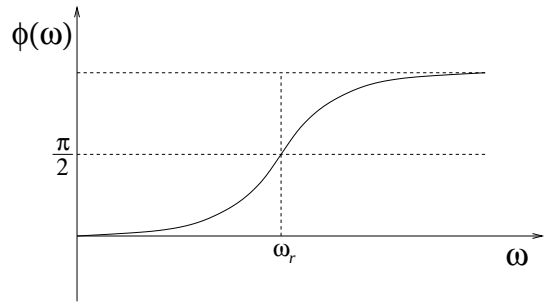


Figure 1.6. SDOF system phase as a function of frequency ω.

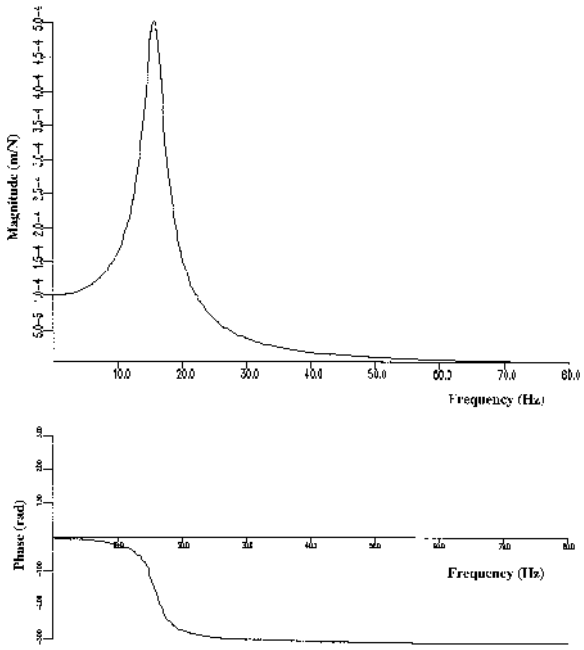


Figure 1.7. Bode plot for system $\ddot{y} + 20\dot{y} + 10^4 y = x(t)$.

natural frequency, $\tan\phi = \infty$ corresponding to $\phi = \frac{\pi}{2}$. This is sufficient to define ϕ (qualitatively) as a function of ω. The plot of $\phi(\omega)$ is given in figure 1.6.

The plots of $\frac{Y}{X}(\omega)$ and $\phi(\omega)$ are usually given together as they specify between them all properties of the system response to a harmonic input. This type of plot is usually called a *Bode plot*. If $\frac{Y}{X}$ and $\phi(\omega)$ are interpreted as the amplitude and phase of a complex function, this is called the *frequency response function* or FRF.

At the risk of a little duplication, an example is given in figure 1.7 for the

Bode plot of an actual SDOF system,

$$\ddot{y} + 20\dot{y} + 10^4 y = x(t). \tag{1.30}$$

(The particular routine used to generate this plot actually shows $-\phi$ in keeping with the conventions of [87].) For this system, the undamped natural frequency is 100 rad s^{-1}, the damped natural frequency is 99.5 rad s^{-1}, the resonance frequency is 99.0 rad s^{-1} and the damping ratio is 0.1 or 10% of critical.

A more direct construction of the system representation in terms of the Bode plot will be given in the following section. Note that the gain and phase in expressions (1.23) and (1.24) are independent of the magnitude of the forcing level X. This means that the FRF is an *invariant* of the amplitude of excitation. In fact, this is only true for linear systems and breakdown in the amplitude invariance of the FRF can be used as a test for nonlinearity as discussed in chapter 2.

1.2 Continuous-time models: frequency domain

The input and output time signals $x(t)$ and $y(t)$ for the SDOF system discussed earlier are well known to have dual *frequency-domain* representations $X(\omega) = \mathcal{F}\{x(t)\}$ and $Y(\omega) = \mathcal{F}\{y(t)\}$ obtained by Fourier transformation where

$$G(\omega) = \mathcal{F}\{g(t)\} = \int_{-\infty}^{+\infty} \mathrm{d}t\, \mathrm{e}^{-\mathrm{i}\omega t} g(t) \tag{1.31}$$

defines the Fourier transform $\mathcal{F}$[3]. The corresponding inverse transform is given by

$$g(t) = \mathcal{F}^{-1}\{G(\omega)\} = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \mathrm{d}\omega\, \mathrm{e}^{-\mathrm{i}\omega t} G(\omega). \tag{1.32}$$

It is natural to ask now if there is a frequency-domain representation of the system itself which maps $X(\omega)$ directly to $Y(\omega)$. The answer to this is yes and the mapping is remarkably simple. Suppose the evolution in time of the signals is specified by equation (1.11); one can take the Fourier transform of both sides of

[3] Throughout this book, the preferred notation for integrals will be

$$\int \mathrm{d}x\, f(x)$$

rather than

$$\int f(x)\, \mathrm{d}x$$

This can be regarded simply as a matter of grammar. The first integral is *the integral with respect to* x *of* $f(x)$, while the second is *the integral of* $f(x)$ *with respect to* x. The meaning is the same in either case; however, the authors feel that the former expression has more formal significance in keeping the integral sign and measure together. It is also arguable that the notation adopted here simplifies some of the manipulations of multiple integrals which will be encountered in later chapters.

the equation, i.e.

$$\int_{-\infty}^{+\infty} \mathrm{d}t\, \mathrm{e}^{-\mathrm{i}\omega t} \left(m\frac{\mathrm{d}^2 y}{\mathrm{d}t^2} + c\frac{\mathrm{d}y}{\mathrm{d}t} + ky \right) = \int_{-\infty}^{+\infty} \mathrm{d}t\, \mathrm{e}^{-\mathrm{i}\omega t} x(t). \tag{1.33}$$

Now, using integration by parts, one has

$$\mathcal{F}\left[\frac{\mathrm{d}^n y}{\mathrm{d}t^n}\right] = (\mathrm{i}\omega)^n Y(\omega) \tag{1.34}$$

and application of this formula to (1.33) yields

$$(-m\omega^2 + \mathrm{i}c\omega + k)Y(\omega) = X(\omega) \tag{1.35}$$

or

$$Y(\omega) = H(\omega)X(\omega) \tag{1.36}$$

where the FRF[4] $H(\omega)$ is defined by

$$H(\omega) = \frac{1}{-m\omega^2 + \mathrm{i}c\omega + k} = \frac{1}{k - m\omega^2 + \mathrm{i}c\omega}. \tag{1.37}$$

So in the frequency domain, mapping input $X(\omega)$ to output is $Y(\omega)$ is simply a matter of multiplying X by a complex function H. All system information is contained in the FRF; all coefficients from the time domain are present and the number and order of the derivatives in (1.4) are encoded in the powers of $\mathrm{i}\omega$ present. It is a simple matter to convince oneself that the relation (1.36) holds in the frequency domain for *any* system whose equation of motion is a linear differential equation although the form of the function $H(\omega)$ will depend on the particular system.

As $H(\omega)$ is a complex function, it has a representation in terms of magnitude $|H(\omega)|$ and phase $\angle H(\omega)$,

$$H(\omega) = |H(\omega)|\mathrm{e}^{\mathrm{i}\angle H(\omega)} \tag{1.38}$$

The $|H(\omega)|$ and $\angle H(\omega)$ so defined correspond exactly to the gain $\frac{Y}{X}(\omega)$ and phase $\phi(\omega)$ defined in the previous section. This result provides a direct interpretation of the FRF $H(\omega)$ in terms of the gain and phase of the response when the system is presented with a harmonic input.

[4] If the Laplace transformation had been used in place of the Fourier transform, equation (1.36) would be unchanged except that it would be in terms of the real Laplace variable s, i.e.

$$Y(s) = H(s)X(s)$$

where

$$H(s) = \frac{1}{ms^2 + cs + k}.$$

In terms of the s-variable, $H(s)$ is referred to as the *transfer function*, the FRF results from making the change of variables $s = \mathrm{i}\omega$.

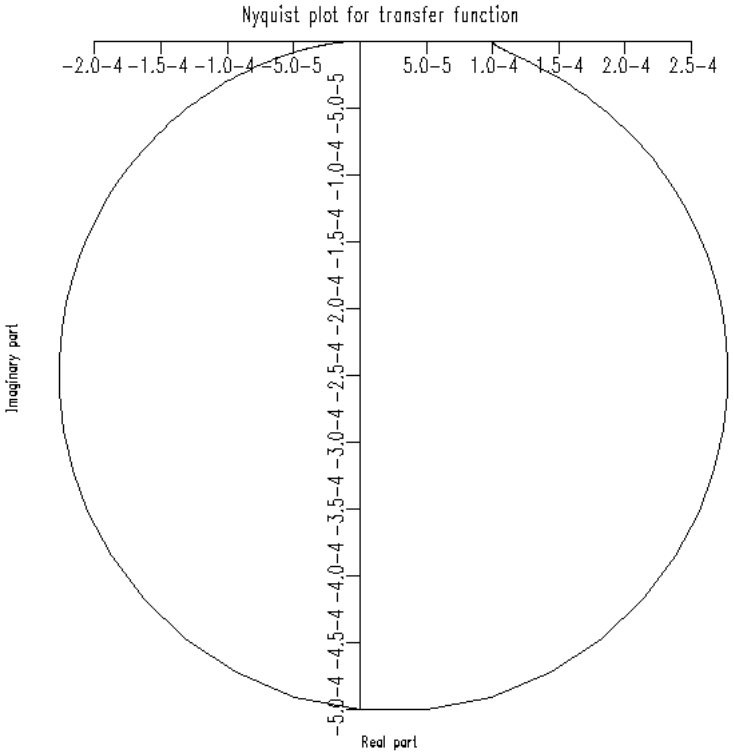


Figure 1.8. Nyquist plot for system $\ddot{y} + 20\dot{y} + 10^4 y = x(t)$—receptance.

It is now clear why the Bode plot defined in the previous section suffices to characterize the system. An alternative means of presenting the information in $H(\omega)$ is the commonly used *Nyquist plot* which describes the locus of $H(\omega)$ in the complex plane or Argand diagram as $w \to \infty$ (or $w \to$ the limit of measurable ω). The Nyquist plot corresponding to the system in (1.30) is given in figure 1.8.

The FRF for the system given in (1.37) for the process $x(t) \to y(t)$. It is called the *receptance* form sometimes denoted $H_{\mathrm{R}}(\omega)$. The FRFs for the processes $x(t) \to \dot{y}(t)$ and $x(t) \to \ddot{y}(t)$ are easily shown to be

$$H_{\mathrm{M}}(\omega) = \frac{\mathrm{i}\omega}{-m\omega^2 + \mathrm{i}c\omega + k} \tag{1.39}$$

and

$$H_{\mathrm{I}}(\omega) = \frac{-\omega^2}{-m\omega^2 + \mathrm{i}c\omega + k}. \tag{1.40}$$

They are respectively referred to as the *mobility* form and *accelerance* or

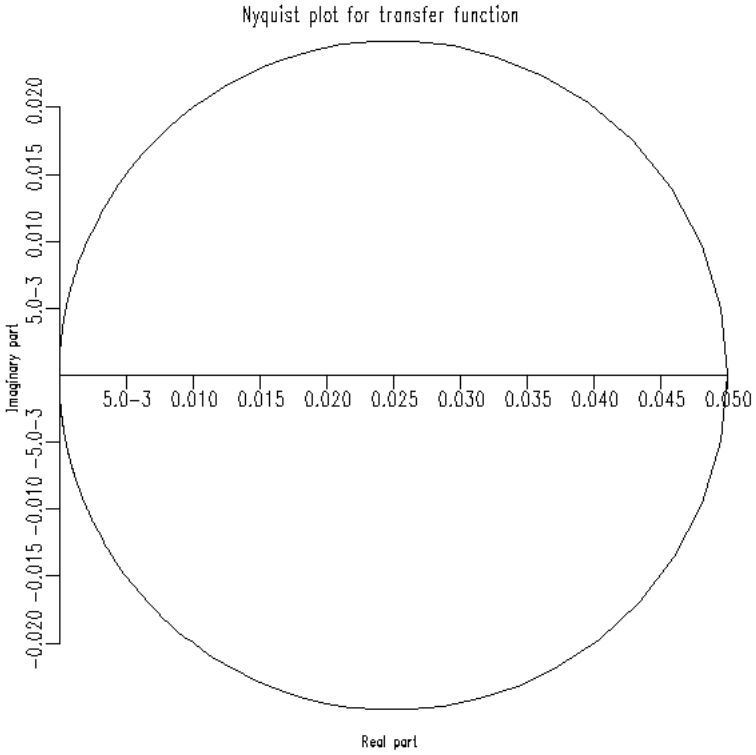


Figure 1.9. Nyquist plot for system $\ddot{y} + 20\dot{y} + 10^4 y = x(t)$—mobility.

accelerance form. The Nyquist plots for these forms of the FRF are given in figures 1.9 and 1.10 for the system in (1.30).

1.3 Impulse response

Given the *general* frequency-domain relationship (1.36) for linear systems, one can now pass back to the time domain and obtain a parallel relationship. One takes the inverse Fourier transform of (1.36), i.e.

$$\frac{1}{2\pi}\int_{-\infty}^{+\infty} \mathrm{d}\omega\, \mathrm{e}^{\mathrm{i}\omega t} Y(\omega) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \mathrm{d}\omega\, \mathrm{e}^{\mathrm{i}\omega t} H(\omega) X(\omega) \qquad (1.41)$$

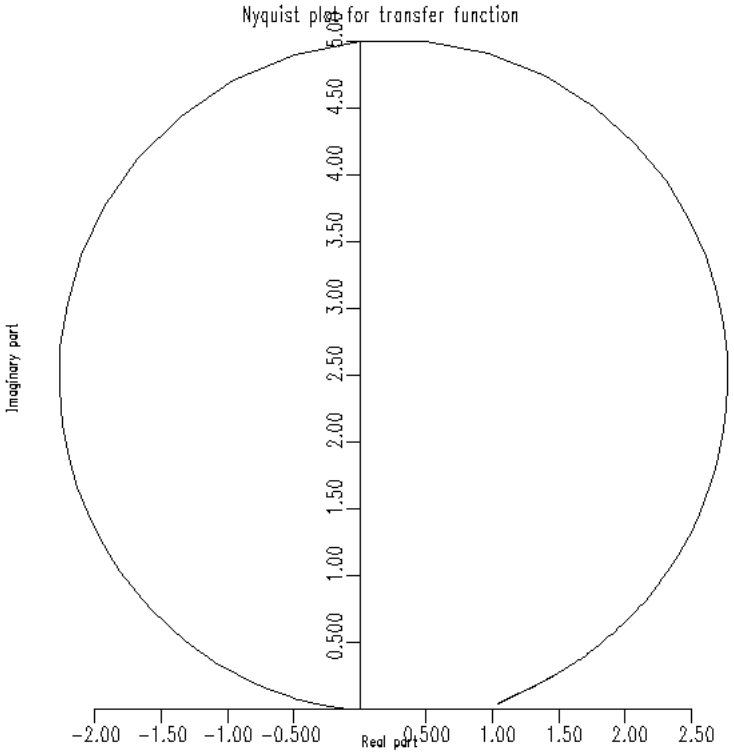


Figure 1.10. Nyquist plot for system $\ddot{y} + 20\dot{y} + 10^4 y = x(t)$—accelerance.

so that

$$\begin{aligned} y(t) &= \frac{1}{2\pi}\int_{-\infty}^{+\infty} \mathrm{d}\omega\, \mathrm{e}^{\mathrm{i}\omega t} H(\omega) X(\omega) \\ &= \frac{1}{2\pi}\int_{-\infty}^{+\infty} \mathrm{d}\omega\, \mathrm{e}^{\mathrm{i}\omega t} H(\omega) \left\{ \int_{-\infty}^{+\infty} \mathrm{d}\tau\, \mathrm{e}^{-\mathrm{i}\omega\tau} x(\tau) \right\}. \end{aligned} \tag{1.42}$$

Interchanging the order of integration gives

$$y(t) = \int_{-\infty}^{+\infty} \mathrm{d}\tau\, x(\tau) \left\{ \frac{1}{2\pi}\int_{-\infty}^{+\infty} \mathrm{d}\omega\, \mathrm{e}^{\mathrm{i}\omega(t-\tau)} H(\omega) \right\} \tag{1.43}$$

and finally

$$y(t) = \int_{-\infty}^{+\infty} \mathrm{d}\tau\, h(t-\tau) x(\tau) \tag{1.44}$$

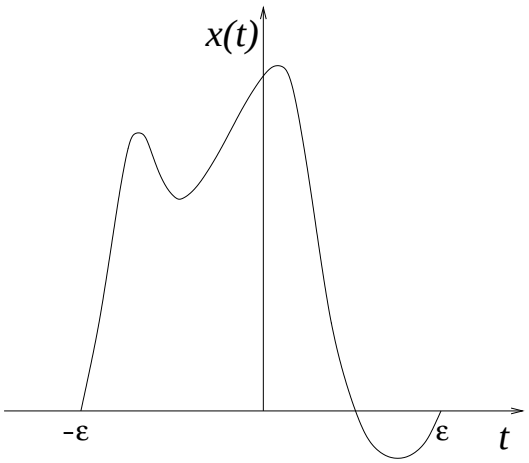


Figure 1.11. Example of a transient excitation whose duration is 2ε.

where the function $h(t)$ is the inverse Fourier transform of $H(\omega)$. If one repeats this argument but takes the inverse transform of $H(\omega)$ before $X(\omega)$ one obtains the alternative expression

$$y(t) = \int_{-\infty}^{+\infty} \mathrm{d}\tau\, h(\tau) x(t-\tau). \tag{1.45}$$

These equations provide another time-domain version of the system's input–output relationship. All system information is encoded in the function $h(t)$. One can now ask if $h(t)$ has a physical interpretation. Again the answer is yes, and the argument proceeds as follows.

Suppose one wishes to know the response of a system to a transient input, i.e. $x(t)$ where $x(t) = 0$ if $|t| > \epsilon$ say (figure 1.11). All the energy is communicated to the system in time 2ϵ after which the system follows the unforced equations of motion. An ideal transient excitation or *impulse* would communicate all energy in an instant. No such *physical* signal exists for obvious reasons. However, there is a mathematical object, the Dirac δ-function $\delta(t)$ [166], which has the properties of an ideal impulse:

infinitesimal duration

$$\delta(t) = 0, \quad t \neq 0 \tag{1.46}$$

finite power

$$\int_{-\infty}^{+\infty} \mathrm{d}t\, |x(t)|^2 = 1. \tag{1.47}$$

The defining relationship for the δ-function is [166]

$$\int_{-\infty}^{+\infty} \mathrm{d}t\, f(t)\delta(t-a) = f(a), \quad \text{for any } f(t). \tag{1.48}$$

Now, according to equation (1.45), the system response to a δ-function input $y_\delta(t)$ is given by

$$y_\delta(t) = \int_{-\infty}^{+\infty} \mathrm{d}\tau\, h(\tau)\delta(t-\tau) \tag{1.49}$$

so applying the relation (1.48) immediately gives

$$y_\delta(t) = h(t) \tag{1.50}$$

which provides the required interpretation of $h(t)$. This is the *impulse response* of the system, i.e. the solution of the equation

$$m\ddot{h}(t) + c\dot{h}(t) + kh(t) = \delta(t). \tag{1.51}$$

It is not an entirely straightforward matter to evaluate $h(t)$ for the general SDOF system, contour integration is needed. Before the rigorous analysis, a more formal argument is provided.

The impulse response is the solution of (1.51) and therefore has the general form

$$y(t) = \mathrm{e}^{-\zeta\omega_n t}[A\cos(\omega_\mathrm{d} t) + B\sin(\omega_\mathrm{d} t)] \tag{1.52}$$

where A and B are fixed by the initial conditions.

The initial displacement $y(0)$ is assumed to be zero and the initial velocity is assumed to follow from the initial momentum coming from the impulsive force $I(t) = \delta(t)$,

$$m\dot{y}(0) = \int \mathrm{d}t\, I(t) = \int \mathrm{d}t\, \delta(t) = 1 \tag{1.53}$$

from (1.48), so it follows that $\dot{y}(0) = 1/m$. Substituting these initial conditions into (1.52) yields $A = 0$ and $B = 1/(m\omega_\mathrm{d})$, and the impulse response is

$$h(t) = \frac{1}{m\omega_\mathrm{d}}\mathrm{e}^{-\zeta\omega_\mathrm{d} t}\sin(\omega_n t) \tag{1.54}$$

for $t > 0$.

The impulse response is therefore a decaying harmonic motion at the damped natural frequency. Note that $h(t)$ is zero before $t = 0$, the time at which the impulse is applied. This is an expression of the principle of causality, i.e. that effect cannot precede cause. In fact, the causality of $h(t)$ will be shown in chapter 5 to follow directly from the fact that $H(\omega)$ has no poles in the lower half of the complex frequency plane. This is generally true for linear dynamical systems and is the starting point for the *Hilbert transform* test of linearity. A further consequence of $h(t)$ vanishing for negative times is that one can change the lower limit of the integral in (1.45) from $-\infty$ to zero with no effect.

Note that this derivation lacks mathematical rigour as the impulsive force is considered to generate the initial condition on velocity, yet they are considered to occur at the same time, in violation of a sensible cause–effect relationship. A

more rigorous approach to evaluating $h(t)$ is simple to formulate but complicated by the need to use the calculus of residues.

According to the definition,

$$h(t) = \mathcal{F}^{-1}\{H(\omega)\} = \frac{1}{2\pi m}\int_{-\infty}^{+\infty} \mathrm{d}\omega \, \frac{\mathrm{e}^{\mathrm{i}\omega t}}{\omega_n^2 - \omega^2 + 2\mathrm{i}\zeta\omega_n\omega}$$
$$= -\frac{1}{2\pi m}\int_{-\infty}^{+\infty} \mathrm{d}\omega \, \frac{\mathrm{e}^{\mathrm{i}\omega t}}{(\omega - \omega_+)(\omega - \omega_-)} \tag{1.55}$$

where $\omega_\pm = \mathrm{i}\zeta\omega_n \pm \omega_\mathrm{d}$ so that $\omega_+ - \omega_- = 2\omega_\mathrm{d}$. Partial fraction expansion of the last expression gives

$$h(t) = \frac{1}{4\pi m\omega_\mathrm{d}}\left\{\int_{-\infty}^{+\infty} \mathrm{d}\omega \, \frac{\mathrm{e}^{\mathrm{i}\omega t}}{(\omega - \omega_-)} - \int_{-\infty}^{+\infty} \mathrm{d}\omega \, \frac{\mathrm{e}^{\mathrm{i}\omega t}}{(\omega - \omega_+)}\right\}. \tag{1.56}$$

The two integrals can be evaluated by contour integration [234],

$$\int_{-\infty}^{+\infty} \mathrm{d}\omega \, \frac{\mathrm{e}^{\mathrm{i}\omega t}}{(\omega - \omega_\pm)} = 2\pi\mathrm{i}\mathrm{e}^{\mathrm{i}\omega_\pm t}\Theta(t) \tag{1.57}$$

where $\Theta(t)$ is the Heaviside function defined by $\Theta(t) = 1$, $t \geq 0$, $\Theta(t) = 0$, $t < 0$, substituting into the last expression for the impulse response gives

$$h(t) = \frac{\mathrm{i}}{2m\omega_\mathrm{d}}(\mathrm{e}^{\mathrm{i}\omega_- t} - \mathrm{e}^{\mathrm{i}\omega_+ t})\Theta(t) \tag{1.58}$$

and substituting for the values of $\omega_\pm$ yields the final result, in agreement with (1.54),

$$h(t) = \frac{1}{m\omega_\mathrm{d}}\mathrm{e}^{-\zeta\omega_\mathrm{d} t}\sin(\omega_n t)\Theta(t). \tag{1.59}$$

Finally, a result which will prove useful later. Suppose that one excites a system with a signal $\mathrm{e}^{\mathrm{i}\omega t}$ (clearly this is physically unrealizable as it is complex), the response is obtained straightforwardly from equation (1.45),

$$y(t) = \int_{-\infty}^{+\infty} \mathrm{d}\tau \, h(\tau)\mathrm{e}^{\mathrm{i}\omega(t-\tau)} \tag{1.60}$$
$$= \mathrm{e}^{\mathrm{i}\omega t}\int_{-\infty}^{+\infty} \mathrm{d}\tau \, h(\tau)\mathrm{e}^{-\mathrm{i}\omega\tau} = H(\omega)\mathrm{e}^{\mathrm{i}\omega t} \tag{1.61}$$

so the system response to the input $\mathrm{e}^{\mathrm{i}\omega t}$ is $H(\omega)\mathrm{e}^{\mathrm{i}\omega t}$. One can regard this result as giving an alternative definition of the FRF.

1.4 Discrete-time models: time domain

The fact that Newton's laws of motion are differential equations leads directly to the continuous-time representation of previously described systems. This

representation defines the motion at *all* times. In reality, most observations of system behaviour—measurements of input and output signals—will be carried out at discrete intervals. The system data are then a discrete set of values $\{x_i, y_i, i = 1, \ldots, N\}$. For modelling purposes one might therefore ask if there exists a model structure which maps the discrete inputs x_i directly to the discrete outputs y_i. Such models do exist and in many cases offer advantages over the continuous-time representation, particularly in the case of nonlinear systems [5].

Consider the general linear SDOF system,

$$m\ddot{y} + c\dot{y} + ky = x(t). \tag{1.62}$$

Suppose that one is only interested in the value of the output at a sequence of regularly spaced times t_i where $t_i = (i-1)\Delta t$ (Δt is called the *sampling interval* and the associated frequency $f_s = \frac{1}{\Delta t}$ is called the *sampling frequency*). At the instant t_i,

$$m\ddot{y}_i + c\dot{y}_i + ky_i = x_i \tag{1.63}$$

where $x_i = x(t_i)$ etc. The derivatives $\dot{y}(t_i)$ and $\ddot{y}(t_i)$ can be approximated by the discrete forms,

$$\dot{y}_i = \dot{y}(t_i) \approx \frac{y(t_i) - y(t_i - \Delta t)}{\Delta t} = \frac{y_i - y_{i-1}}{\Delta t} \tag{1.64}$$

$$\ddot{y}(t_i) \approx \frac{y_{i+1} - 2y_i + y_{i-1}}{\Delta t^2}. \tag{1.65}$$

Substituting these approximations into (1.63) yields, after a little rearrangement,

$$y_i = \left(2 - \frac{c\Delta t}{m} - \frac{k\Delta t^2}{m}\right) y_{i-1} + \left(\frac{c\Delta t}{m} - 1\right) y_{i-2} + \frac{\Delta t^2}{m} x_{i-1} \tag{1.66}$$

or

$$y_i = a_1 y_{i-1} + a_2 y_{i-2} + b_1 x_{i-1} \tag{1.67}$$

where the constants a_1, a_2, b_1 are defined by the previous equation. Equation (1.67) is a *discrete-time* representation of the SDOF system under study[6]. Note that the motion for all discrete times is fixed by the input sequence

[5] i is used throughout as a sampling index *and* the square root of -1, this is not considered to be a likely source of confusion.

[6] The form (1.67) is a consequence of choosing the representations (1.64) and (1.65) for the derivatives. Different discrete-time systems, all approximating to the same continuous-time system, can be obtained by choosing more accurate discrete derivatives. Note that the form (1.67) is still obtained if the *backward difference* (1.64) is replaced by the *forward difference*

$$\dot{y}_i \approx \frac{y_{i+1} - y_i}{\Delta t}$$

or (the more accurate) *centred difference*

$$\dot{y}_i \approx \frac{y_{i+1} - y_{i-1}}{2\Delta t}.$$

Only the coefficients a_1, a_2 and b_1 change.

x_i together with values for y_1 and y_2. The specification of the first two values of the output sequence is directly equivalent to the specification of initial values for $y(t)$ and $\dot{y}(t)$ in the continuous-time case. An obvious advantage of using a discrete model like (1.67) is that it is much simpler to numerically predict the output in comparison with a differential equation. The price one pays is a loss of generality—because the coefficients in (1.67) are functions of the sampling interval Δt, one can only use this model to predict responses with the same spacing in time.

Although arguably less familiar, the theory for the solution of difference equations is no more difficult than the corresponding theory for differential equations. A readable introduction to the relevant techniques is given in chapter 26 of [233].

Consider the free motion for the system in (1.67); this is specified by

$$y_i = a_1 y_{i-1} + a_2 y_{i-2}. \tag{1.68}$$

Substituting a trial solution $y_i = \alpha^i$ with α constant yields

$$\alpha^{i-2}(\alpha^2 - a_1\alpha - a_2) = 0 \tag{1.69}$$

which has non-trivial solutions

$$\alpha_\pm = \frac{a_1}{2} \pm \frac{1}{2}\sqrt{4a_2 + a_1^2}. \tag{1.70}$$

The general solution of (1.68) is, therefore,

$$y_i = A\alpha_+^i + B\alpha_-^i \tag{1.71}$$

where A and B are arbitrary constants which can be fixed in terms of the initial values y_1 and y_2 as follows. According to the previous solution $y_1 = A\alpha_+ + B\alpha_-$ and $y_2 = A\alpha_+^2 + B\alpha_-^2$; these can be regarded as simultaneous equations for A and B, the solution being

$$A = \frac{y_2 - \alpha_- y_1}{\alpha_+(\alpha_+ - \alpha_-)} \tag{1.72}$$

$$B = \frac{\alpha_+ y_1 - y_2}{\alpha_-(\alpha_+ - \alpha_-)}. \tag{1.73}$$

Analysis of the stability of this system is straightforward. If either $|\alpha_+| > 1$ or $|\alpha_-| > 1$ the solution grows exponentially, otherwise the solution decays exponentially. More precisely, if the *magnitudes* of the $alpha$s are greater than one—as they may be complex—the solutions are unstable. In the differential equation case the stability condition was simply $c > 0$. The stability condition in terms of the difference equation parameters is the slightly more complicated expression

$$\left| \frac{a_1}{2} \pm \frac{1}{2}\sqrt{4a_2 + a_1^2} \right| < 1. \tag{1.74}$$

By way of illustration, consider the SDOF system (1.30) again. Equation (1.66) gives the expressions for a_1 and a_2, and if $\Delta t = 0.001$, they are found to be: $a_1 = 1.97$ and $a_2 = -0.98$. The quantities $(a_1 \pm \sqrt{4a_2 + a_1^2})/2$ are found to be 0.985 ± 0.0989i. The magnitudes are both 0.9899 and the stability of the discrete system (1.67) is assured. Note that the stability depends not only on the parameters of the original continuous-time system but also on the sampling interval.

In terms of the original continuous-time parameters m, c and k for this model the stability condition is rather more complex, it is—after substituting (1.66) into (1.74)—

$$\left|\frac{m}{2\Delta} - c - k\Delta t \pm \sqrt{(c + k\Delta)^2 - 4km}\right| < \frac{m}{2\Delta t}. \tag{1.75}$$

Note that each difference equation property parallels a differential equation property. It is this which allows either representation when modelling a system.

As for the differential equation, the principle of superposition holds for linear difference equations so it is sufficient to consider a harmonic excitation $x_i = X\cos(\omega t_i)$ in order to explore the characteristics of the forced equation. As in the continuous-time case, the general solution of the forced equation

$$y_i - a_1 y_{i-1} - a_2 y_{i-2} = X\cos(\omega t_{i-1}) \tag{1.76}$$

will comprise a transient part, specified in equation (1.71), and a steady-state part independent of the initial conditions. In order to find the steady-state solution one can assume that the response will be a harmonic at the forcing frequency; this provides the form of the trial solution

$$y_i = Y\cos(\omega t_i + \phi). \tag{1.77}$$

Substituting this expression into (1.67) and shifting the time $t \to t + \Delta t - \frac{\phi}{\omega}$, yields

$$Y(\cos(\omega t_i + \omega\Delta t) - a_1\cos(\omega t_i) - a_2\cos(\omega t_i - \omega\Delta t)) = X\cos(\omega t_i - \phi). \tag{1.78}$$

Expanding and comparing coefficients for sin and cos in the result yields the two equations

$$Y(-a_1 + (1 - a_2)C) = X\cos\phi \tag{1.79}$$

$$Y(-(1 + a_2)S) = X\sin\phi \tag{1.80}$$

where $C = \cos(\omega\Delta t)$ and $S = \sin(\omega\Delta t)$. It is a now a simple matter to obtain the expressions for the system gain and phase:

$$\frac{Y}{X} = \frac{1}{\sqrt{a_1^2 - 2a_1(1 - a_2)C + (1 - a_2)^2C^2 + (1 + a_2)^2S^2}} \tag{1.81}$$

$$\tan\phi = \frac{(1 + a_2)S}{a_1 + (a_2 - 1)C}. \tag{1.82}$$

One point about these equations is worth noting. The expressions for gain and phase are functions of frequency ω through the variables C and S. However, these variables are periodic with period $\frac{1}{\Delta t} = f_{\mathrm{s}}$. As a consequence, the gain and phase formulae simply repeat indefinitely as $\omega \to \infty$. This means that knowledge of the response functions in the interval $[-\frac{f_{\mathrm{s}}}{2}, \frac{f_{\mathrm{s}}}{2}]$ is sufficient to specify them for all frequencies. An important consequence of this is that a discrete representation of a system can be accurate in the frequency domain *only* on a finite interval. The frequency $\frac{f_{\mathrm{s}}}{2}$ which prescribes this interval is called the *Nyquist frequency*.

1.5 Classification of difference equations

Before moving on to consider the frequency-domain representation for discrete-time models it will be useful to digress slightly in order to discuss the taxonomy of difference equations, particularly as they will feature in later chapters. The techniques and terminology of discrete modelling has evolved over many years in the literature of time-series analysis, much of which may be unfamiliar to engineers seeking to apply these techniques. The aim of this section is simply to describe the basic linear difference equation structures, the classic reference for this material is the work by Box and Jenkins [46].

1.5.1 Auto-regressive (AR) models

As suggested by the name, an auto-regressive model expresses the present output y_i from a system as a linear combination of past outputs, i.e. the variable is regressed on itself. The general expression for such a model is

$$y_i = \sum_{j=1}^{p} a_j y_{i-j} \tag{1.83}$$

and this is termed an AR(p) model.

1.5.2 Moving-average (MA) models

In this case the output is expressed as a linear combination of past inputs. One can think of the output as a weighted average of the inputs over a finite window which moves with time, hence the name. The general form is

$$y_i = \sum_{j=1}^{q} b_j x_{i-j} \tag{1.84}$$

and this is called a MA(q) model.

All linear continuous-time systems have a canonical representation as a moving-average model as a consequence of the input–output relationship:

$$y(t_i) = \int_0^{+\infty} \mathrm{d}\tau \, h(\tau) x(t_i - \tau) \tag{1.85}$$

which can be approximated by the discrete sum

$$y_i = \sum_{j=0}^{\infty} \Delta t h(j\Delta t) x(t_i - j\Delta t). \tag{1.86}$$

As $t_i - j\Delta t = t_{i-j}$, one has

$$y_i = \sum_{j=0}^{\infty} b_j x_{i-j} \tag{1.87}$$

which is an MA(∞) model with $b_j = \Delta t h(j\Delta t)$.

1.5.3 Auto-regressive moving-average (ARMA) models

As the name suggests, these are simply a combination of the two model types discussed previously. The general form is the ARMA(p, q) model,

$$y_i = \sum_{j=1}^{p} a_j y_{i-j} + \sum_{j=1}^{q} b_j x_{i-j} \tag{1.88}$$

which is quite general in the sense that any discretization of a linear differential equation will yield an ARMA model. Equation (1.67) for the discrete version of a SDOF system is an ARMA$(2, 1)$ model.

Note that a given continuous-time system will have in general many discrete-time representations. By virtue of the previous arguments, the linear SDOF system can be modelled using either an MA(∞) or an ARMA$(2, 1)$ structure. The advantage of using the ARMA form is that far fewer past values of the variables need be included to predict with the same accuracy as the MA model.

1.6 Discrete-time models: frequency domain

The aim of this short section is to show a simple construction of the FRF for a discrete-time system. The discussion of the preceding section shows that the ARMA(p, q) structure is sufficiently general in the linear case, i.e. the system of interest is given by (1.88).

Introducing the backward shift operator $\triangle$ defined by its action on the signals $\triangle^k y_i = y_{i-k}$, allows one to rewrite equation (1.88) as

$$y_i = \left(\sum_{j=1}^{p} a_j \triangle^j \right) y_i + \left(\sum_{j=1}^{q} b_j \triangle^j \right) x_i \tag{1.89}$$

or

$$\left(1 - \sum_{j=1}^{p} a_j \triangle^j \right) y_i = \left(\sum_{j=1}^{q} b_j \triangle^j \right) x_i. \tag{1.90}$$

Now one defines the FRF $H(\omega)$ by the means suggested at the end of section 1.3. If the input to the system is $\mathrm{e}^{\mathrm{i}\omega t}$, the output is $H(\omega)\mathrm{e}^{\mathrm{i}\omega t}$. The action of $\triangle$ on the signals is given by

$$\triangle^m x_k = \triangle^m \mathrm{e}^{\mathrm{i}\omega k\Delta t} = \mathrm{e}^{\mathrm{i}\omega(k-m)\Delta t} = \mathrm{e}^{-\mathrm{i}m\omega\Delta t}x_k \tag{1.91}$$

on the input and

$$\begin{aligned}\triangle^m y_k &= \triangle^m H(\omega)x_k = H(\omega)\triangle^m \mathrm{e}^{\mathrm{i}\omega k\Delta t}\\ &= H(\omega)\mathrm{e}^{\mathrm{i}\omega(k-m)\Delta t} = H(\omega)\mathrm{e}^{-\mathrm{i}m\omega\Delta t}x_k\end{aligned} \tag{1.92}$$

on the output. Substituting these results into equation (1.90) yields

$$\left(1-\sum_{j=1}^{p}a_j\mathrm{e}^{-\mathrm{i}j\omega\Delta t}\right)H(\omega)x_i = \left(\sum_{j=1}^{q}b_j\mathrm{e}^{-\mathrm{i}j\omega\Delta t}\right)x_i \tag{1.93}$$

which, on simple rearrangement, gives the required result

$$H(\omega) = \frac{\sum_{j=1}^{q}b_j\mathrm{e}^{-\mathrm{i}j\omega\Delta t}}{(1-\sum_{j=1}^{p}a_j\mathrm{e}^{-\mathrm{i}j\omega\Delta t})}. \tag{1.94}$$

Note that this expression is periodic in ω as discussed at the close of section 1.4.

1.7 Multi-degree-of-freedom (MDOF) systems

The discussion so far has been restricted to the case of a single mass point. This has proved useful in that it has allowed the development of most of the basic theory used in modelling systems. However, the assumption of single degree-of-freedom behaviour for all systems is clearly unrealistic. In general, one will have to account for the motion of several mass points or even a continuum. To see this, consider the transverse vibrations of a simply supported beam (figure 1.12). A basic analysis of the statics of the situation, shows that an applied force F at the centre of the beam produces a displacement y given by

$$F = ky, \quad k = \frac{48EI}{L^3} \tag{1.95}$$

where E is the Young's modulus of the beam material, I is the second moment of area and L is the length of the beam. k is called the *flexural stiffness*.

If it is now assumed that the mass is concentrated at the centre (figure 1.13), by considering the kinetic energy of the beam vibrating with a maximum displacement at the centre, it can be shown that the point mass is equal to half the total mass of the beam $M/2$ [249]. The appropriate equation of motion is

$$\frac{M}{2} + ky = x(t) \tag{1.96}$$

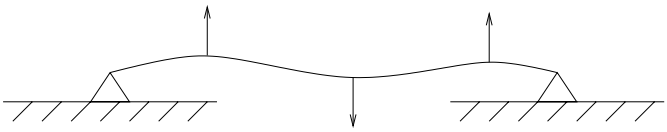


Figure 1.12. A uniform simply supported beam under transverse vibration.

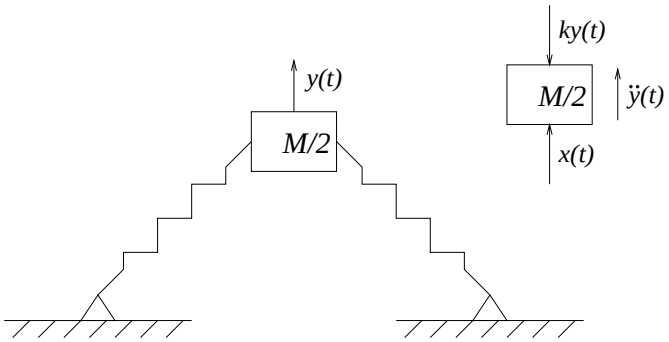


Figure 1.13. Central point mass approximation for the beam of figure 1.12.

for the displacement of the centre point, under a time-dependent excitation $x(t)$. Damping effects are neglected for the present. If $x(t)$ is assumed harmonic, the theory developed in previous sections shows that the response will be harmonic at the same frequency. Unfortunately, as the beam has been replaced by a mass point in this approximation, one cannot obtain any information about the profile of the beam while vibrating. If the free equation of motion is considered, a natural frequency of $\omega_n = \sqrt{\frac{2k}{M}}$ follows. Extrapolation from the static case suggests that the profile of the beam at this frequency will show its maximum displacement in the centre, the displacement of other points will fall monotonically as they approach the ends of the beam. No points except the end points will have zero displacement for all time. This mode of vibration is termed the *fundamental* mode. The word 'mode' has acquired a technical sense here: it refers to the shape of the beam vibrating at its natural frequency.

In order to obtain more information about the profile of the beam, the mass can assumed to be concentrated at two points spaced evenly on the beam (figure 1.14). This time an energy analysis shows that one-third of the beam mass should be concentrated at each point. The equations of motion for this system are

$$\frac{M}{3}\ddot{y}_1 + k^{\mathrm{f}}_{11}y_1 + k^{\mathrm{f}}_{12}(y_1 - y_2) = x_1(t) \tag{1.97}$$

$$\frac{M}{3}\ddot{y}_2 + k^{\mathrm{f}}_{22}y_2 + k^{\mathrm{f}}_{12}(y_2 - y_1) = x_2(t) \tag{1.98}$$

where y_1 and y_2 are the displacement responses. The k^{f}_{ij} are flexural stiffnesses

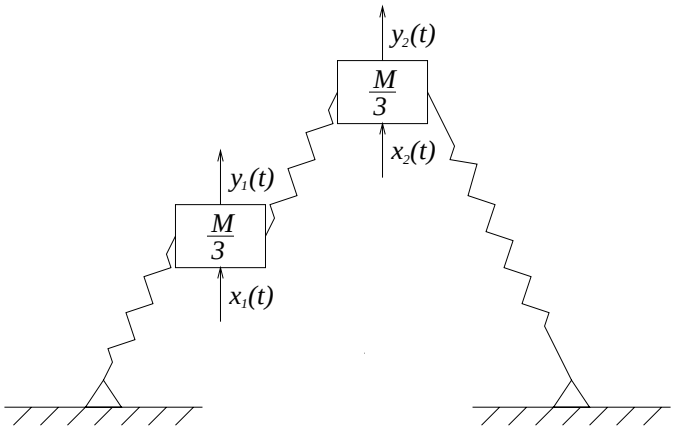


Figure 1.14. Double mass approximation for the beam of figure 1.12 with the masses located at one-third and two-thirds of the length.

evaluated from basic beam theory. Note that the equations of motion are coupled. A little rearrangement gives

$$\frac{M}{3}\ddot{y}_1 + k_{11}y_1 + k_{12}y_2 = x_1(t) \tag{1.99}$$

$$\frac{M}{3}\ddot{y}_2 + k_{21}y_1 + k_{22}y_2 = x_1(t) \tag{1.100}$$

where $k_{11} = k_{11}^{\mathrm{f}} + k_{12}^{\mathrm{f}}$ etc. Note that $k_{12} = k_{21}$; this is an expression of a general principle—that of *reciprocity*. (Again, reciprocity is a property which only holds for linear systems. Violations of reciprocity can be used to indicate the presence of nonlinearity.) These equations can be placed in a compact matrix form

$$[m]\{\ddot{y}\} + [k]\{y\} = \{x\} \tag{1.101}$$

where curly braces denote vectors and square braces denote matrices.

$$[m] = \begin{pmatrix} \frac{M}{3} & 0 \\ 0 & \frac{M}{3} \end{pmatrix}, \quad [k] = \begin{pmatrix} k_{11} & k_{12} \\ k_{21} & k_{22} \end{pmatrix} \tag{1.102}$$

$$\{y\} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, \quad \{x\} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}. \tag{1.103}$$

$[m]$ and $[k]$ are called the mass and stiffness matrices respectively.

In order to find the natural frequencies (it will turn out that there are more than one), consider the unforced equation of motion

$$[m]\{\ddot{y}\} + [k]\{y\} = \{0\}. \tag{1.104}$$

To solve these equations, one can make use of a result of linear algebra theory which asserts that there exists an *orthogonal matrix* $[\psi]$ (i.e. $[\psi]^{\mathrm{T}} = [\psi]^{-1}$ where $^{\mathrm{T}}$ denotes the transpose and $^{-1}$ denotes the inverse), which simultaneously diagonalizes $[m]$ and $[k]$, i.e.

$$[\psi]^{\mathrm{T}}[m][\psi] = [M] = \begin{pmatrix} m_1 & 0 \\ 0 & m_2 \end{pmatrix} \tag{1.105}$$

$$[\psi]^{\mathrm{T}}[k][\psi] = [K] = \begin{pmatrix} k_1 & 0 \\ 0 & k_2 \end{pmatrix}. \tag{1.106}$$

Now, make the linear change of coordinates from $\{y\}$ to $\{z\}$ where $\{y\} = [\psi]\{z\}$, i.e.

$$\begin{array}{rcl} y_1 &=& \psi_{11}z_1 + \psi_{12}z_2 \\ y_2 &=& \psi_{21}z_1 + \psi_{22}z_2. \end{array} \tag{1.107}$$

Equation (1.104) becomes

$$[m][\psi]\{\ddot{z}\} + [k][\psi]\{z\} = \{0\} \tag{1.108}$$

and on premultiplying this expression by $[\psi]^{\mathrm{T}}$, one obtains

$$[M]\{\ddot{z}\} + [K]\{z\} = \{0\} \tag{1.109}$$

which represents the following scalar equations,

$$\begin{array}{rcl} m_1\ddot{z}_1 + k_1z_1 &=& 0 \\ m_2\ddot{z}_2 + k_2z_2 &=& 0 \end{array} \tag{1.110}$$

which represent two uncoupled SDOF systems. The solutions are[7]

$$\begin{array}{rcl} z_1(t) &=& A_1\cos(\omega_1 t) \\ z_2(t) &=& A_2\cos(\omega_2 t). \end{array} \tag{1.111}$$

The *two* undamped natural frequencies are $\omega_{n1} = \sqrt{\frac{k_1}{m_1}}$ and $\omega_{n2} = \sqrt{\frac{k_2}{m_2}}$. Each of the z-coordinates is associated with a distinct frequency and, as will be shown later, a distinct mode of vibration. For this reason the z-coordinates are referred to as *modal coordinates*. The elements of the diagonal mass and stiffness matrices are referred to as the *modal masses* and *modal stiffnesses* respectively.

On transforming back to the physical y-coordinate system using (1.107), one obtains

$$\begin{array}{rcl} y_1 &=& \psi_{11}A_1\cos(\omega_1 t) + \psi_{12}A_2\cos(\omega_2 t) \\ y_2 &=& \psi_{21}A_1\cos(\omega_1 t) + \psi_{22}A_2\cos(\omega_2 t). \end{array} \tag{1.112}$$

[7] These solutions are not general, for example the first should strictly be

$$z_1(t) = A_1\cos(\omega_1 t) + B_1\cos(\omega_1 t).$$

For simplicity, the sine terms are ignored. This can be arranged by setting the initial conditions appropriately.

One observes that both natural frequencies are present in the solution for the physical coordinates.

This solution is unrealistic in that the motion is undamped and therefore persists indefinitely; some damping mechanism is required. The equations of motion of the two-mass system should be modified to give

$$[m]\{\ddot{y}\} + [c]\{\dot{y}\} + [k]\{y\} = \{0\} \tag{1.113}$$

where $[c]$ is called the *damping matrix*. A problem arises now if one tries to repeat this analysis for the damped system. Generally, there is no matrix $[\psi]$ which will simultaneously diagonalize *three* matrices $[m]$, $[c]$ and $[k]$. Consequently, no transformation exists which uncouples the equations of motion. The simplest means of circumnavigating this problem is to assume *proportional* or *Rayleigh damping*. This means

$$[c] = \alpha[m] + \beta[k] \tag{1.114}$$

where α and β are constants. This is a fairly restrictive assumption and in many cases it does not hold. In particular, if the damping is nonlinear, one cannot apply this assumption. However, with this form of damping, one finds that the diagonalizing matrix $[\psi]$ for the undamped motion also suffices for the damped motion. In fact,

$$[\psi]^{\mathrm{T}}[c][\psi] = [C] = \alpha[M] + \beta[K] \tag{1.115}$$

with diagonal entries the *modal dampings*, given by

$$c_i = \alpha m_i + \beta k_i. \tag{1.116}$$

For this type of damping, the equations of motion uncouple as before on transforming to modal coordinates so that

$$\begin{array}{rcrcrcl} m_1\ddot{z}_1 & + & c_1\dot{z}_1 & + & k_1 z_1 & = & 0 \\ m_2\ddot{z}_2 & + & c_2\dot{z}_2 & + & k_2 z_2 & = & 0. \end{array} \tag{1.117}$$

The solutions are

$$\begin{array}{rcl} z_1 & = & A_1 \mathrm{e}^{-\zeta_1\omega_1 t}\cos(\omega_{\mathrm{d}1}t) \\ z_2 & = & A_2 \mathrm{e}^{-\zeta_2\omega_2 t}\cos(\omega_{\mathrm{d}2}t) \end{array} \tag{1.118}$$

where the damped natural frequencies and *modal* damping ratios are specified by

$$\zeta_i = \frac{c_i}{2\sqrt{m_i k_i}}, \quad \omega_{\mathrm{d}i}^2 = \omega_i^2(1-\zeta_i^2). \tag{1.119}$$

On transforming back to the physical coordinates, one obtains

$$\begin{array}{rcrcl} y_1 & = & \psi_{11}A_1\mathrm{e}^{-\zeta_1\omega_1 t}\cos(\omega_{\mathrm{d}1}t) & + & \psi_{12}A_2\mathrm{e}^{-\zeta_2\omega_2 t}\cos(\omega_{\mathrm{d}2}t) \\ y_2 & = & \psi_{21}A_1\mathrm{e}^{-\zeta_1\omega_1 t}\cos(\omega_{\mathrm{d}1}t) & + & \psi_{22}A_2\mathrm{e}^{-\zeta_2\omega_2 t}\cos(\omega_{\mathrm{d}2}t) \end{array} \tag{1.120}$$

and the free motion is a sum of damped harmonics at the damped natural frequencies. Note that the rates of decay are different for each frequency component.

The forced response of the system can be obtained in much the same manner as for the SDOF system. In order to simplify matters slightly, the excitation vector is assumed to have the form,

$$\{x\} = \begin{pmatrix} x_1(t) \\ 0 \end{pmatrix}. \tag{1.121}$$

On transforming the forced equation to modal coordinates, one obtains

$$[M]\{\ddot{z}\} + [C]\{\dot{z}\} + [K]\{z\} = \{p\} = [\psi]^{\mathrm{T}}\{x\} \tag{1.122}$$

where

$$\{p\} = \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} = \begin{pmatrix} \psi_{11}x_1 \\ \psi_{12}x_1 \end{pmatrix} \tag{1.123}$$

so that

$$\begin{array}{rcrcrcl} m_1\ddot{z}_1 & + & c_1\dot{z}_1 & + & k_1 z_1 & = & p_1 \\ m_2\ddot{z}_2 & + & c_2\dot{z}_2 & + & k_2 z_2 & = & p_2. \end{array} \tag{1.124}$$

For a harmonic input $x_1(t)$ these SDOF equations can be solved directly as in section 1.1.

The representation of the system in the frequency domain is obtained by Fourier transforming the equations (1.124). The results are

$$Z_1(\omega) = \frac{\psi_{11}}{-m_1\omega^2 + \mathrm{i}c_1\omega + k_1} X_1(\omega) \tag{1.125}$$

$$Z_2(\omega) = \frac{\psi_{12}}{-m_2\omega^2 + \mathrm{i}c_2\omega + k_2} X_1(\omega) \tag{1.126}$$

and linearity of the Fourier transform implies (from (1.107)),

$$\begin{aligned} Y_1(\omega) &= \psi_{11}Z_1(\omega) + \psi_{12}Z_2(\omega) \\ &= \left\{ \frac{\psi_{11}^2}{-m_1\omega^2 + \mathrm{i}c_1\omega + k_1} + \frac{\psi_{12}^2}{-m_2\omega^2 + \mathrm{i}c_2\omega + k_2} \right\} X_1(\omega) \end{aligned} \tag{1.127}$$

$$\begin{aligned} Y_2(\omega) &= \psi_{21}Z_1(\omega) + \psi_{22}Z_2(\omega) \\ &= \left\{ \frac{\psi_{21}\psi_{11}}{-m_1\omega^2 + \mathrm{i}c_1\omega + k_1} + \frac{\psi_{12}\psi_{22}}{-m_2\omega^2 + \mathrm{i}c_2\omega + k_2} \right\} X_1(\omega). \end{aligned} \tag{1.128}$$

Recalling that $Y(\omega) = H(\omega)X(\omega)$, the overall FRFs for the processes $x_1(t) \to y_1(t)$ and $x_1(t) \to y_2(t)$ are therefore given by

$$H_{11}(\omega) = \frac{Y_1(\omega)}{X_1(\omega)} = \frac{\psi_{11}^2}{-m_1\omega^2 + \mathrm{i}c_1\omega + k_1} + \frac{\psi_{12}^2}{-m_2\omega^2 + \mathrm{i}c_2\omega + k_2} \tag{1.129}$$

$$H_{12}(\omega) = \frac{Y_2(\omega)}{X_1(\omega)} = \frac{\psi_{21}\psi_{11}}{-m_1\omega^2 + \mathrm{i}c_1\omega + k_1} + \frac{\psi_{12}\psi_{22}}{-m_2\omega^2 + \mathrm{i}c_2\omega + k_2}. \tag{1.130}$$

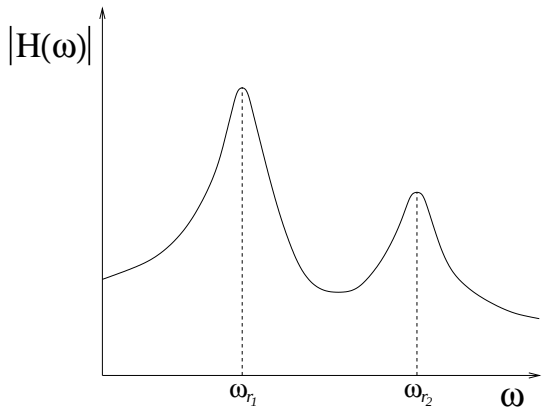


Figure 1.15. Magnitude of the gain of the FRF for an underdamped 2DOF system showing two resonant conditions. The equation of motion is (1.122).

On referring back to the formula for the resonant frequency of a SDOF system, it is clear from these expressions that the Bode plot for each of these expressions will show two peaks or resonances (figure 1.15), at the frequencies

$$\begin{aligned} \omega_{\mathrm{r}1} &= \omega_1\sqrt{1-2\zeta_1^2} \\ \omega_{\mathrm{r}2} &= \omega_2\sqrt{1-2\zeta_2^2}. \end{aligned} \tag{1.131}$$

As an example, the Bode plots and Nyquist plots for the system,

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} \ddot{y}_1 \\ \ddot{y}_2 \end{pmatrix} + 20\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} \dot{y}_1 \\ \dot{y}_2 \end{pmatrix} + 10^4\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ 0 \end{pmatrix} \tag{1.132}$$

are given in figures 1.16–1.19. (Note that there appears to be a discontinuity in the phase of figure 1.18. This is simply a result of the fact that phase possesses a 2π periodicity and phases in excess of π will be continued at $-\pi$.)

It has proved useful to consider a 2DOF system to discuss how natural frequencies etc. generalize to MDOF systems. However, as one might expect, it is possible to deal with linear systems with arbitrary numbers of DOF at the expense of a little more abstraction. This is the subject of the last section.

1.8 Modal analysis

1.8.1 Free, undamped motion

The object of this section is to formalize the arguments given previously for MDOF systems and state them in their full generality. As before, the theory will be provided in stages, starting with the simplest case, i.e. that of an undamped

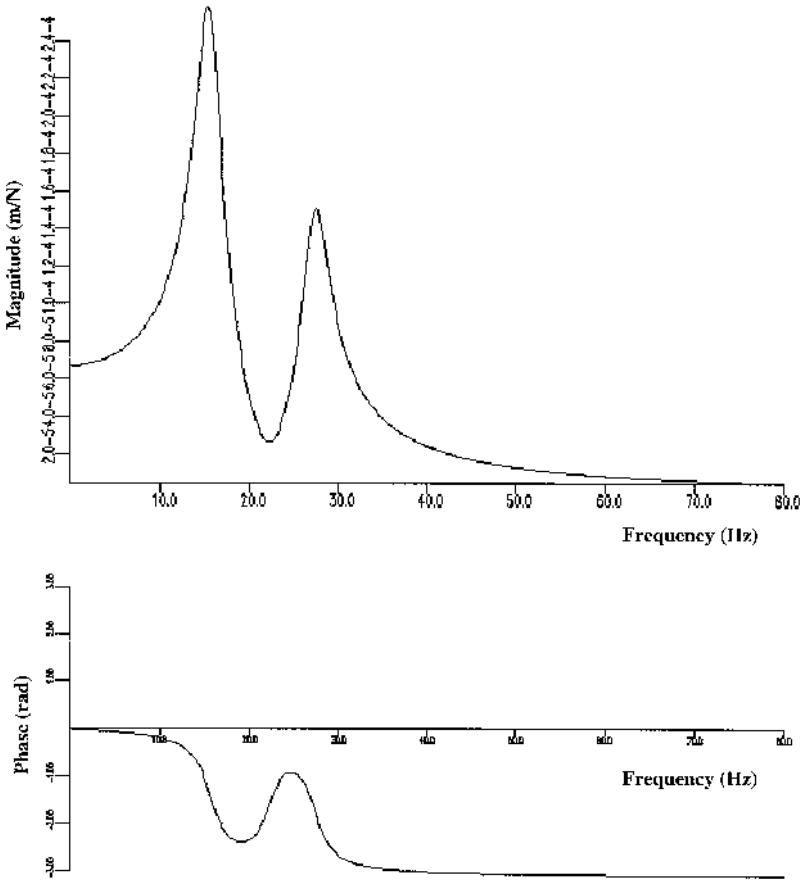


Figure 1.16. H_{11} Bode plot for a 2DOF system.

unforced system. The equation of motion for such a linear system is

$$[m]\{\ddot{y}\} + [k]\{y\} = 0 \tag{1.133}$$

where $\{y\}$ is now an $n \times 1$ column vector and $[m]$ and $[k]$ are $n \times n$ matrices. As always, the excitation is assumed to be harmonic, so the solution is assumed to have the form

$$\{y(t)\} = \{\psi\}\mathrm{e}^{\mathrm{i}\omega t} \tag{1.134}$$

where $\{\psi\}$ is a constant $n \times 1$ vector. This ansatz basically assumes that all points on the structure move in phase with the same frequency. Substituting into (1.133) yields

$$-\omega^2[m]\{\psi\} + [k]\{\psi\} = 0 \tag{1.135}$$

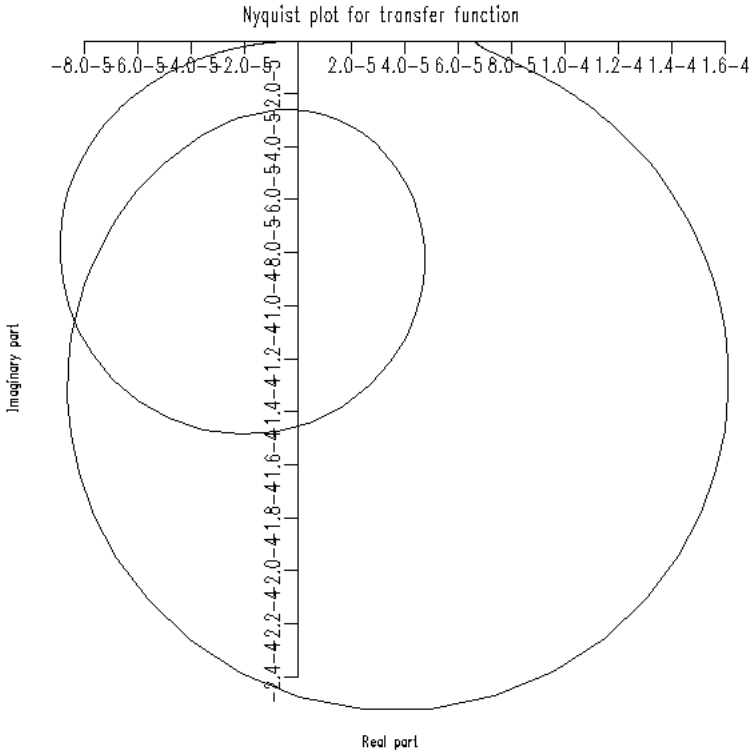


Figure 1.17. H_{12} Bode plot for a 2DOF system.

which is a standard linear eigenvalue problem with n solutions ω_{ni} and $\{\psi_i\}$. These are the undamped natural frequencies and the modeshapes. The interpretation is well known: if the system is excited at a frequency ω_{ni}, all points will move in phase with a profile given by $\{\psi_i\}$.

If it is assumed that $[m]$ is invertible (and this is usually true), it is a simple matter to rewrite equation (1.135) in the more usual form for an eigenvalue problem:

$$[m]^{-1}[k]\{\psi_i\} - \frac{1}{\omega_{ni}^2}\{\psi_i\} = [D]\{\psi_i\} - \lambda_i\{\psi_i\} = 0 \tag{1.136}$$

with a little notation added. Note that the normalization of $\{\psi_i\}$ is arbitrary, i.e. if $\{\psi_i\}$ is a solution of (1.136), then so is $\alpha\{\psi_i\}$ for any real number α. Common normalizations for modeshapes include setting the largest element to unity or setting the length of the vector to unity, i.e. $\{\psi_i\}^{\mathrm{T}}\{\psi_i\} = 1$.

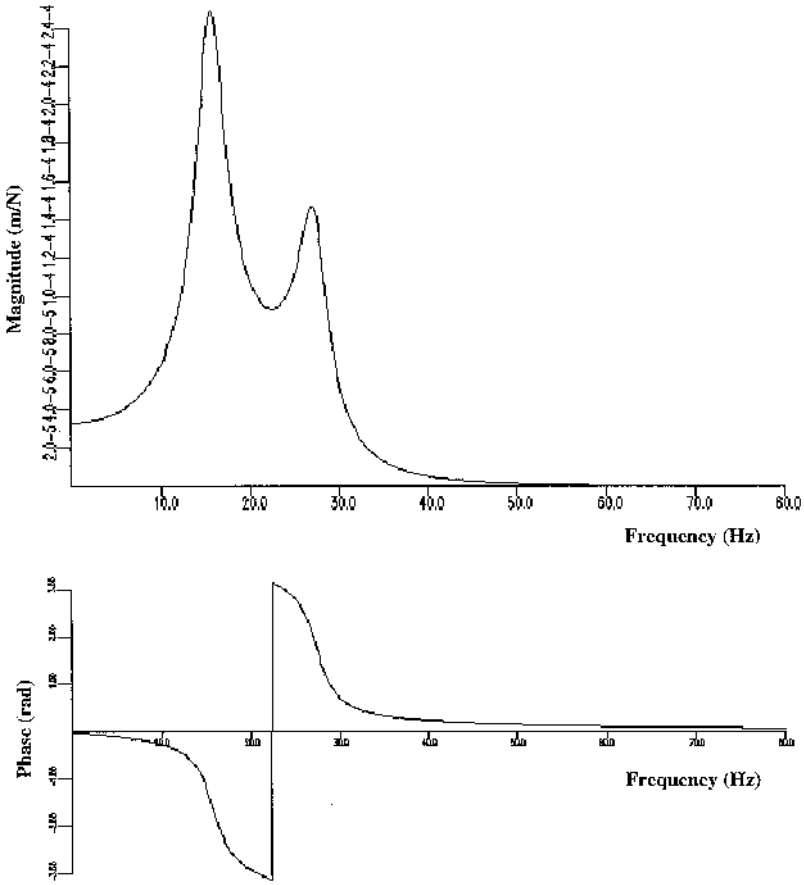


Figure 1.18. H_{11} Nyquist plot for a 2DOF system.

Non-trivial solutions of (1.136) must have $\{\psi_i\} \neq \{0\}$. This forces the *characteristic equation*

$$\det([D] - \lambda_i[1]) = 0 \tag{1.137}$$

which has n solutions for the λ_i as required.

This apparently flexible system of equations turns out to have rather constrained solutions for the modeshapes. The reason is that $[m]$ and $[k]$ can almost always be assumed to be symmetric. This is a consequence of the property of reciprocity mentioned earlier.

Suppose that ω_{ni}^2 and ω_{nj}^2 are *distinct* eigenvalues of (1.136), then

$$\begin{aligned} -\omega_{ni}^2[m]\{\psi_i\} &= [k]\{\psi_i\} \\ -\omega_{nj}^2[m]\{\psi_j\} &= [k]\{\psi_j\}. \end{aligned} \tag{1.138}$$

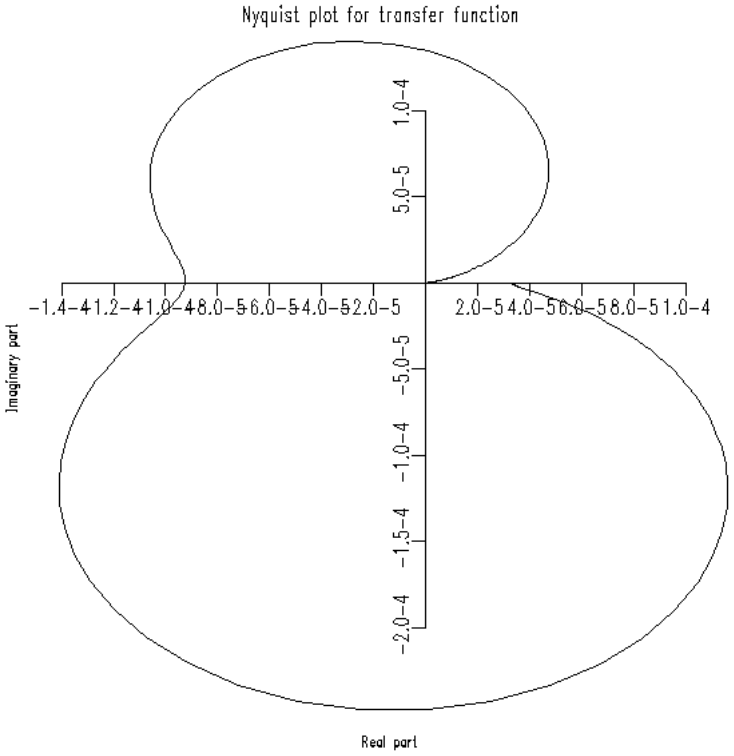


Figure 1.19. H_{12} Nyquist plot for a 2DOF system. (Note that the Real and Imaginary axes do not have equal scales.)

Now, premultiplying the first of these expressions by $\{\psi_j\}^{\mathrm{T}}$ and the second by $\{\psi_i\}^{\mathrm{T}}$ gives

$$\begin{aligned} -\omega_{ni}^2\{\psi_j\}^{\mathrm{T}}[m]\{\psi_i\} &= \{\psi_j\}^{\mathrm{T}}[k]\{\psi_i\} \\ -\omega_{nj}^2\{\psi_i\}^{\mathrm{T}}[m]\{\psi_j\} &= \{\psi_i\}^{\mathrm{T}}[k]\{\psi_j\} \end{aligned} \tag{1.139}$$

and as $[m]$ and $[k]$ are symmetric, it follows that

$$\begin{aligned} (\{\psi_j\}^{\mathrm{T}}[m]\{\psi_i\})^{\mathrm{T}} &= \{\psi_i\}^{\mathrm{T}}[m]\{\psi_j\} \\ (\{\psi_j\}^{\mathrm{T}}[k]\{\psi_i\})^{\mathrm{T}} &= \{\psi_i\}^{\mathrm{T}}[k]\{\psi_j\} \end{aligned} \tag{1.140}$$

so transposing the first expression in (1.139) and subtracting from the second expression yields

$$(\omega_{ni}^2 - \omega_{nj}^2)\{\psi_i\}^{\mathrm{T}}[m]\{\psi_j\} = 0 \tag{1.141}$$

and as $\omega_{ni} \neq \omega_{nj}$, it follows that

$$\{\psi_i\}^{\mathrm{T}}[m]\{\psi_j\} = 0 \tag{1.142}$$

and from (1.139) it follows that

$$\{\psi_i\}^{\mathrm{T}}[k]\{\psi_j\} = 0. \tag{1.143}$$

So the modeshapes belonging to distinct eigenvalues are *orthogonal with respect to the mass and stiffness matrices*. This is referred to as *weighted orthogonality*. The situation where the eigenvalues are not distinct is a little more complicated and will not be discussed here, the reader can refer to [87]. Note that unless the mass or stiffness matrix is the unit, the eigenvectors or modeshapes are *not* orthogonal in the usual sense, i.e. $\{\psi_i\}^{\mathrm{T}}\{\psi_j\} \neq 0$. Assuming n distinct eigenvalues, one can form the *modal matrix* $[\Psi]$ by taking an array of the modeshapes

$$[\Psi] = \{\{\psi_1\}, \{\psi_2\}, \ldots, \{\psi_n\}\}. \tag{1.144}$$

Consider the matrix

$$[M] = [\Psi]^{\mathrm{T}}[m][\Psi]. \tag{1.145}$$

A little algebra shows that the elements are

$$M_{ij} = \{\psi_i\}^{\mathrm{T}}[m]\{\psi_j\} \tag{1.146}$$

and these are zero if $i \neq j$ by the weighted orthogonality (1.142). This means that $[M]$ is diagonal. The diagonal elements $m_1, m_2, \ldots, m_n$ are referred to as the *generalized masses* or modal masses as discussed in the previous section. By a similar argument, the matrix

$$[K] = [\Psi]^{\mathrm{T}}[k][\Psi] \tag{1.147}$$

is diagonal with elements $k_1, k_2, \ldots, k_n$ which are termed the *generalized* or modal stiffnesses. The implications for the equations of motion (1.133) are important. Consider the change of coordinates

$$[\Psi]\{u\} = \{y\} \tag{1.148}$$

equation (1.133) becomes

$$[m][\Psi]\{\ddot{u}\} + [k][\Psi]\{u\} = 0 \tag{1.149}$$

and premultiplying by $[\Psi]^{\mathrm{T}}$ gives

$$[\Psi]^{\mathrm{T}}[m][\Psi]\{\ddot{u}\} + [\Psi]^{\mathrm{T}}[k][\Psi]\{u\} = 0 \tag{1.150}$$

or

$$[M]\{\ddot{u}\} + [K]\{u\} = 0 \tag{1.151}$$

by virtue of equations (1.145) and (1.147). The system has been decoupled into n SDOF equations of motion of the form

$$m_i\ddot{u}_i + k_i u_i = 0, \quad i = 1, \ldots, n \tag{1.152}$$

and it follows, by premultiplying the first equation of (1.138) by $\{\psi_i\}$, that

$$\omega_{ni}^2 = \frac{k_i}{m_i} \tag{1.153}$$

and (1.152) becomes

$$\ddot{u}_i + \omega_{ni}^2 u_i = 0 \tag{1.154}$$

the equation of an undamped SDOF oscillator with undamped natural frequency ω_{ni}. The coordinates u_i are termed *generalized*, *modal* or *normal* coordinates. Now, following the SDOF theory developed in the course of this chapter, the solution of (1.154) is simply

$$u_i = U_i \cos(\omega_{ni} t) \tag{1.155}$$

and in the original physical coordinates, the response can contain components at *all* natural frequencies,

$$y_i = \sum_{j=1}^{n} \Psi_{ij} U_j \cos(\omega_{nj} t) \tag{1.156}$$

Before passing to the damped case, it is worthwhile to return to the question of normalization. Different normalizations lead to different modal masses and stiffness; however, they are always constrained to satisfy $k_i/m_i = \omega_{ni}^2$. A common approach is to use *mass normalization* as follows. Suppose a modal matrix $[\Psi]$ is specified such that the modal mass matrix is $[M]$; if one defines $[\Phi]$ by

$$[\Phi] = [\Psi][M]^{-\frac{1}{2}} \tag{1.157}$$

it follows that

$$\begin{aligned} [\Phi]^{\mathrm{T}}[m][\Phi] &= [1] \\ [\Phi]^{\mathrm{T}}[k][\Phi] &= [\Lambda]^2 \end{aligned} \tag{1.158}$$

where

$$[\Lambda] = \mathrm{diag}(\omega_{n1}, \omega_{n2}, \ldots, \omega_{nn}) \tag{1.159}$$

and this representation is unique. Equation (1.157) amounts to choosing

$$\{\phi_i\} = \frac{1}{\sqrt{m_i}}\{\psi_i\}. \tag{1.160}$$

1.8.2 Free, damped motion

It is a simple matter to generalize (1.133) to the damped case, the relevant equation is

$$[m]\{\ddot{y}\} + [c]\{\dot{y}\} + [k]\{y\} = 0 \tag{1.161}$$

with $[c]$ termed the (viscous) *damping matrix*. (In many cases, it will be desirable to consider structural damping, the reader is referred to [87].) The desired result is to decouple the equations (1.160) into SDOF oscillators in much the same way as for the damped case. Unfortunately, this is generally impossible as observed in the last section. While it is (almost) always possible to find a matrix $[\Psi]$ which diagonalizes two matrices ($[m]$ and $[k]$), this is not the case for three ($[m]$, $[c]$ and $[k]$). Rather than give up, the usual recourse is to assume Rayleigh or proportional damping as in (1.114)[8]. In this case,

$$[\Psi]^{\mathrm{T}}[c][\Psi] = [C] = \mathrm{diag}(c_1, \ldots, c_n) \tag{1.162}$$

with

$$c_i = \alpha m_i + \beta k_i. \tag{1.163}$$

With this assumption, the modal matrix decouples the system (1.160) into n SDOF systems in much the same way as for the undamped case, the relevant equations are (after the transformation (1.148)),

$$m_i \ddot{u}_i + c_i \dot{u}_i + k_i u_i = 0, \quad i = 1, \ldots, n \tag{1.164}$$

and these have solutions

$$u_i = A_i \mathrm{e}^{-\zeta_i \omega_{ni} t} \sin(\omega_{\mathrm{d}i} t - \theta_i) \tag{1.165}$$

where A_i and θ_i are fixed by the initial conditions and

$$\zeta_i = \frac{c_i}{2\sqrt{m_i k_i}} \tag{1.166}$$

is the ith modal damping ratio and

$$\omega_{\mathrm{d}i}^2 = \omega_{ni}^2 (1 - \zeta_i^2) \tag{1.167}$$

is the ith damped natural frequency. Transforming back to physical coordinates using (1.148) yields

$$y_i = \sum_{j=1}^{n} \Psi_{ij} A_j \mathrm{e}^{-\zeta_i \omega_{ni} t} \sin(\omega_{\mathrm{d}i} t - \theta_i). \tag{1.168}$$

[8] One can do slightly better than traditional proportional damping. It is known that if a matrix $[\Psi]$ diagonalizes $[m]$, then it also diagonalizes $f([m])$ where f is a restricted class of matrix functions. (f must have a Laurent expansion of the form

$$f([m]) = \ldots a_{-1}[m]^{-1} + a_0[1] + a_1[m] + a_2[m]^2 \ldots$$

functions like $\det[m]$ are not allowed for obvious reasons.) Similarly, if $[\Psi]$ diagonalizes $[k]$, it will also diagonalize $g([k])$ if g belongs to the same class as f. In principle, one can choose any damping matrix

$$[c] = f([m]) + g([m])$$

and $[\psi]$ will diagonalize it, i.e.

$$[\Psi]^{\mathrm{T}}[c][\Psi] = \mathrm{diag}(f(m_1) + g(k_1), \ldots, f(m_n) + g(k_n)).$$

Having said this, this freedom is never used and the most common choice of damping prescription is proportional.

1.8.3 Forced, damped motion

The general forced linear MDOF system is

$$[m]\{\ddot{y}\} + [c]\{\dot{y}\} + [k]\{y\} = \{x(t)\} \tag{1.169}$$

where $\{x(t)\}$ is an $n \times 1$ vector of time-dependent excitations. As in the free, damped case, one can change to modal coordinates, the result is

$$[M]\{\ddot{u}\} + [C]\{\dot{u}\} + [K]\{u\} = [\Psi]^{\mathrm{T}}\{x(t)\} = \{p\} \tag{1.170}$$

which serves to define $\{p\}$, the vector of *generalized forces*. As before (under the assumption of proportional damping), the equations decouple into n SDOF systems,

$$m_i\ddot{u}_i + c_i\dot{u}_i + k_iu_i = p_i, \quad i = 1, \ldots, n \tag{1.171}$$

and all of the analysis relevant to SDOF systems developed previously applies.

It is instructive to develop the theory in the frequency domain. Suppose the excitations p_i are broadband random, it is sensible to think in terms of FRFs. The ith modal FRF (i.e. the FRF associated with the process $p_i \longrightarrow u_i$) is

$$G_i(\omega) = \frac{S_{u_ip_i}(\omega)}{S_{u_iu_i}(\omega)} = \frac{1}{-m_i\omega^2 + \mathrm{i}c_i\omega + k_i}. \tag{1.172}$$

In order to allow a simple derivation of the FRFs in physical coordinates, it will be advisable to abandon rigour[9] and make the formal definition,

$$\{Y(\omega)\} = [H(\omega)]\{X(\omega)\} \tag{1.173}$$

of $[H(\omega)]$, the FRF matrix. According to (1.172), the corresponding relation in modal coordinates is

$$\{U(\omega)\} = [G(\omega)]\{P(\omega)\} \tag{1.174}$$

with $[G(\omega)] = \mathrm{diag}(G_1(\omega), \ldots, G_n(\omega))$ diagonal. Substituting for $\{U\}$ and $\{P\}$ in the last expression gives

$$[\Psi]^{-1}\{Y(\omega)\} = [G(\omega)][\Psi]^{\mathrm{T}}\{X(\omega)\} \tag{1.175}$$

or

$$\{Y(\omega)\} = [\Psi][G(\omega)][\Psi]^{\mathrm{T}}\{X(\omega)\} \tag{1.176}$$

which identifies

$$\{H(\omega)\} = [\Psi][G(\omega)][\Psi]^{\mathrm{T}}. \tag{1.177}$$

[9] Strictly speaking, it is not allowed to Fourier transform random signals $x(t)$, $y(t)$ as they do not satisfy the Dirichlet condition. The reader may rest assured that a more principled analysis using correlation functions yields the same results as those given here.

In terms of the individual elements of $[H]$, (1.177) yields

$$H_{ij}(\omega) = \sum_{l=1}^{n}\sum_{k=1}^{n}\psi_{il}[G(\omega)_{lk}\psi_{kj}^{\mathrm{T}} = \sum_{k=1}^{n}\psi_{ik}G_k(\omega)\psi_{jk} \tag{1.178}$$

and finally

$$H_{ij}(\omega) = \sum_{k=1}^{n}\frac{\psi_{ik}\psi_{jk}}{-m_i\omega^2 + \mathrm{i}c_i\omega + k_i} \tag{1.179}$$

or

$$H_{ij}(\omega) = \sum_{k=1}^{n}\frac{{}_kA_{ij}}{-(\omega^2 - \omega_{nk}^2) + 2\mathrm{i}\zeta_k\omega_{nk}\omega} \tag{1.180}$$

where

$${}_kA_{ij} = \frac{\psi_{ik}\psi_{jk}}{m_k} = \phi_{ik}\phi_{jk} \tag{1.181}$$

are the *residues* or modal constants.

It follows from these equations that the FRF for any process $x_i \longrightarrow y_j$ of a MDOF linear system is the sum of n SDOF FRFs, one for each natural frequency. It is straightforward to show that each individual mode has a resonant frequency,

$$\omega_{\mathrm{r}i} = \omega_{ni}\sqrt{1 - 2\zeta_i^2}. \tag{1.182}$$

Taking the inverse Fourier transform of the expression (1.180) gives the general form of the impulse response for a MDOF system

$$h_{ij}(t) = \sum_{k=1}^{n}\frac{{}_kA_{ij}}{\omega_{\mathrm{d}k}}\mathrm{e}^{-\zeta_k\omega_k t}\cos(\omega_{\mathrm{d}k}t - \theta_k) \tag{1.183}$$

and the response of a general MDOF system to a transient is a sum of decaying harmonics with individual decay rates and frequencies.

A final remark is required about the proportionality assumption for the damping. For a little more effort than that expended here, one can obtain the system FRFs for an arbitrarily damped linear system [87]. The only change in the final form (1.181) is that the constants ${}_kA_{ij}$ become complex.

All these expressions are given in receptance form; parallel mobility and accelerance forms exist and are obtained by multiplying the receptance form by $\mathrm{i}\omega$ and $-\omega^2$ respectively.

There are well-established signal-processing techniques which allow one to experimentally determine the FRFs of a system. It is found for linear structural systems that the representation as a sum of resonances given in (1.181) is remarkably accurate. An example of a MDOF FRF is given in figure 1.20. After obtaining an experimental curve for some $H(\omega)$ the data can be curve-fitted to the form in equation (1.181) and the best-fit values for the parameters m_i, c_i, k_i,

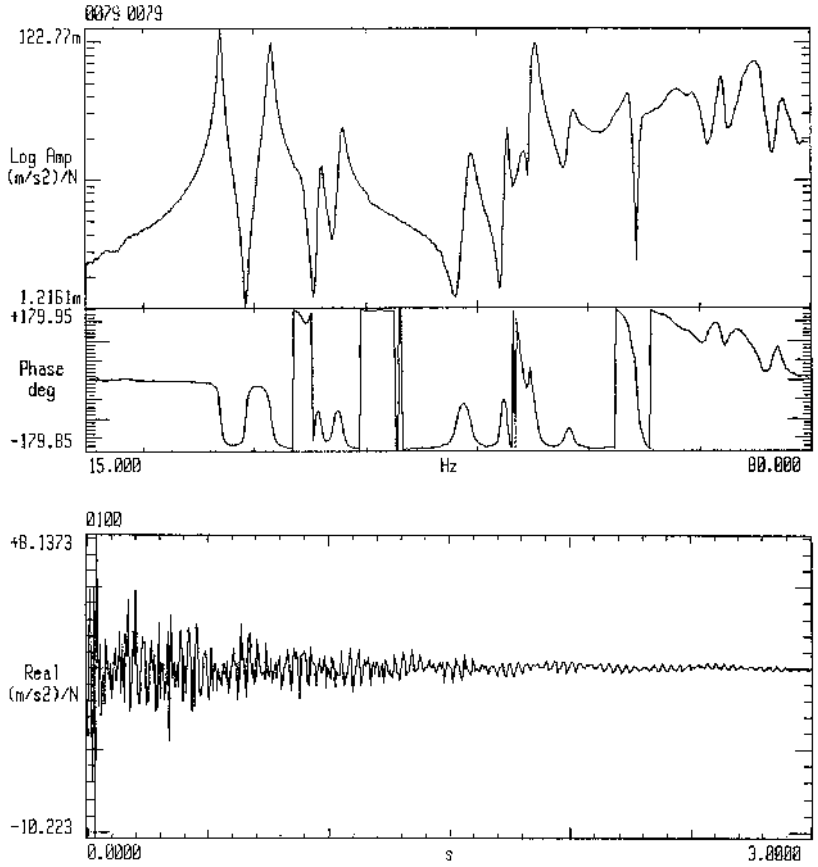


Figure 1.20. FRF and impulse response for multi-mode system.

$i = 1, \ldots, N$ can be obtained. The resulting model is called a *modal model* of the system.

This discussion should convince the reader of the effectiveness of modal analysis for the description of linear systems. The technique is an essential part of the structural dynamicist's repertoire and has no real rivals for the analysis of linear structures. Unfortunately, the qualifier *linear* is significant. Modal analysis is a linear theory *par excellence* and relies critically on the principle of superposition. This is a serious limitation in a world where nonlinearity is increasingly recognized to have a significant effect on the dynamical behaviour of systems and structures.

In the general case, the effect of nonlinearity on modal analysis is rather destructive. All the system invariants taken for granted for a linear system—resonant frequencies, damping ratios, modeshapes, frequency response functions

(FRFs)—become dependent on the level of the excitation applied during the test. As the philosophy of modal analysis is to characterize systems in terms of these 'invariants', the best outcome from a test will be a model of a *linearization* of the system, characteristic of the forcing level. Such a model is clearly incapable of predictions at other levels and is of limited use. Other properties of linear systems like reciprocity are also lost for general nonlinear systems.

The other fundamental concept behind modal analysis is that of *decoupling* or dimension reduction. As seen earlier, the change from physical (measured by the transducers) coordinates to normal or modal coordinates converts a linear n-degree-of-freedom system to n independent SDOF systems. This decoupling property is lost for generic nonlinear systems.

In the face of such a breakdown in the technique, the structural dynamicist—who still needs to model the structure—is faced with essentially three possibilities:

(1) Retain the philosophy and basic theory of modal analysis but learn how to characterize nonlinear systems in terms of the particular ways in which amplitude invariance is lost.
(2) Retain the philosophy of modal analysis but extend the theory to encompass objects which *are* amplitude invariants of nonlinear systems.
(3) Discard the philosophy and seek theories which address the nonlinearity directly.

The aim of the current book is to illustrate examples of each course of action.

Chapter 2

From linear to nonlinear

2.1 Introduction

It is probable that all practical engineering structures are nonlinear to some extent, the nonlinearity being caused by one, or a combination of, several factors such as structural joints in which looseness or friction characteristics are present, boundary conditions which impose variable stiffness constraints, materials that are amplitude dependent or components such as shock absorbers, vibration isolators, bearings, linkages or actuators whose dynamics are input dependent. There is no unique approach to dealing with the problem of nonlinearity either analytically or experimentally and thus we must be prepared to experiment with several approaches in order to ascertain whether the structure can be classified as linear or nonlinear. It would be particularly helpful if the techniques employed in modal testing could be used to test nonlinear structures and it is certainly essential that some form of test for linearity is carried out at the beginning of any dynamic test as the majority of analysis procedures currently available are based on linearity. If this principle is violated, errors may be introduced by the data analysis. Thus the first step is to consider simple procedures that can be employed to establish if the structure or component under test is linear. In the following it is assumed that the structure is time invariant and stable.

2.2 Symptoms of nonlinearity

As stated at the end of the last chapter, many of the properties which hold for linear structures or systems break down for nonlinear. This section discusses some of the more important ones.

2.2.1 Definition of linearity—the principle of superposition

The principle of superposition discussed briefly in the first chapter is more than a property of linear systems; in mathematical terms it actually *defines* what is linear

and what is not.

The principle of superposition can be applied statically or dynamically and simply states that the total response of a linear structure to a set of simultaneous inputs can be broken down into several experiments where each input is applied individually and the output to each of these separate inputs can be summed to give the total response.

This can be stated precisely as follows. If a system in an initial condition $S_1 = \{y_1(0), \dot{y}_1(0)\}$ responds to an input $x_1(t)$ with an output $y_1(t)$ and in a separate test an input $x_2(t)$ to the system initially in state $S_2 = \{y_2(0), \dot{y}_2(0)\}$ produces an output $y_2(t)$ then superposition holds if and only if the input $\alpha x_1(t) + \beta x_2(t)$ to the system in initial state $S_3 = \{\alpha y_1(0) + \beta y_2(0), \alpha \dot{y}_1(0) + \beta \dot{y}_2(0)\}$ results in the output $\alpha y_1(t) + \beta y_2(t)$ for all constants α, β, and all pairs of inputs $x_1(t), x_2(t)$.

Despite its fundamental nature, the principle offers limited prospects as a test of linearity. The reason being that in order to establish linearity beyond doubt, an infinity of tests is required spanning all α, β, $x_1(t)$ and $x_2(t)$. This is clearly impossible. However, to show nonlinearity without doubt, only one set of $\alpha, \beta, x_1(t), x_2(t)$ which violate superposition are needed. In general practice it may be more or less straightforward to establish such a set.

Figure 2.1 shows an example of the static application of the principle of superposition to a uniform beam rigidly clamped at both ends subject to static loading at its centre. It can be seen that superposition holds to a high degree of approximation when the static deflections are small, i.e. less than the thickness of the beam; however, as the applied load is increased, producing deflections greater than the beam thickness, the principle of superposition is violated since the applied loads $F_1 + F_2$ do not result in the sum of the deflections $y_1 + y_2$. What is observed is a stiffness nonlinearity called a *hardening stiffness* which occurs because the boundary conditions restrict the axial straining of the middle surface (the neutral axis) of the beam as the lateral amplitude is increased. It is seen that the rate of increase of the deflection begins to reduce as the load continues to increase. The symmetry of the situation dictates that if the applied load direction is reversed, the deflection characteristic will follow the same pattern resulting in an *odd* nonlinear stiffness characteristic as shown in figure 2.2. (The defining property of an odd function is that $F(-y) = -F(y)$.)

If the beam were pre-loaded, the static equilibrium point would not be centred at $(0, 0)$ as in figure 2.2 and the resulting force-deflection characteristic would become a general function lacking symmetry as shown in figure 2.3.

This is a common example of a stiffness nonlinearity, occurring whenever clamped beams or plates are subjected to flexural displacements which can be considered large, i.e. well in excess of their thickness. The static analysis is fairly straightforward and will be given here; a discussion of the dynamic case is postponed until chapter 9.

Consider an encastré beam (a beam with fully clamped boundary conditions) under a centrally applied static load (figure 2.4). The deflection shape, with

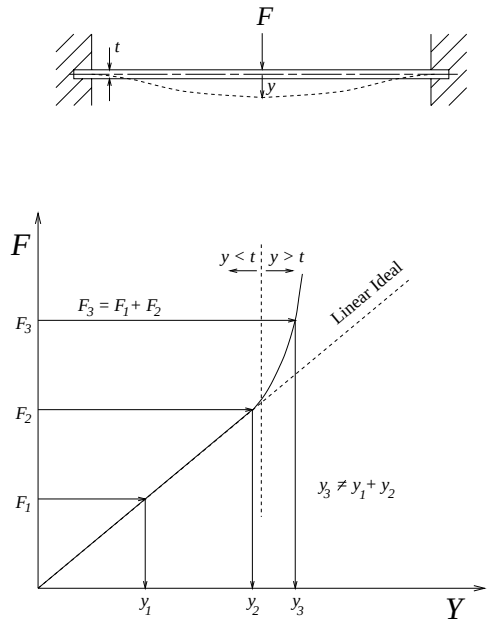


Figure 2.1. Example of the static application of the principle of superposition to a uniform clamped–clamped beam showing that for static deflections in excess of the beam thickness a 'hardening' stiffness is induced which violates the principle.

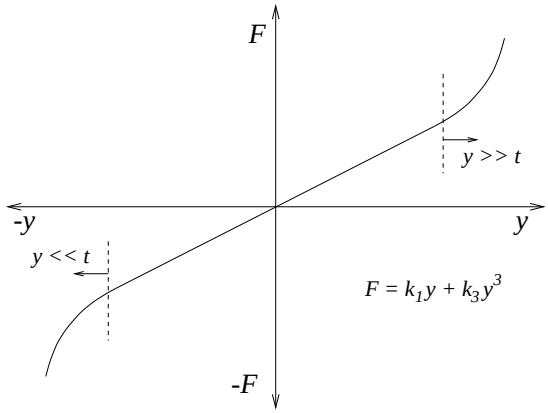


Figure 2.2. The effect of reversing the applied load on the beam of figure 2.1: a symmetric 'hardening' stiffness nonlinearity.

the coordinates located at the mid-point of the beam, can be assumed to be a polynomial which satisfies all the boundary conditions and the eigenvalue

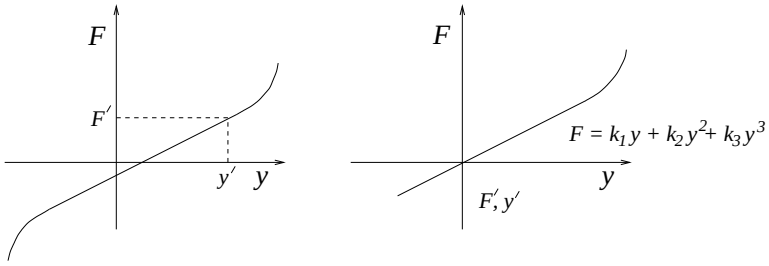


Figure 2.3. The result of pre-loading the beam in figure 2.1 is a general cubic form for the stiffness, lacking the symmetry of figure 2.2.

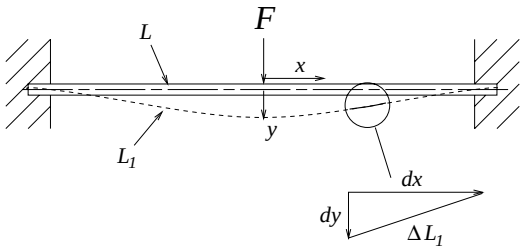


Figure 2.4. An encastré (clamped–clamped) beam under a centrally applied static load resulting in a change of length from L to L_1. The elemental length represents the axial extension.

problem, i.e. an admissible function

$$y(x) = Y\left(1 - \frac{ax^2}{\left[\frac{L}{2}\right]^2} + \frac{bx^4}{\left[\frac{L}{2}\right]^4} - \frac{cx^6}{\left[\frac{L}{2}\right]^6} + \cdots\right). \tag{2.1}$$

Using this assumed shape and by deriving the axial and flexural strain energies, an expression for the lateral stiffness at the centre of the beam can be found. If only the first three terms in the series are used with the appropriate values for the constants, the expression for the deflection is

$$y(x) = Y\left(1 - 2.15\frac{x^2}{\left[\frac{L}{2}\right]^2} + 1.30\frac{x^4}{\left[\frac{L}{2}\right]^4} - 0.15\frac{x^6}{\left[\frac{L}{2}\right]^6}\cdots\right) \tag{2.2}$$

and the flexural strain energy V_{F} is found from

$$V_{\mathrm{F}} = \int_{-L/2}^{L/2} \mathrm{d}x\, \frac{M^2}{2EI} = \frac{EI}{2}\int_{-L/2}^{L/2} \mathrm{d}x\, \left(\frac{\mathrm{d}^2 y}{\mathrm{d}x^2}\right)^2 \tag{2.3}$$

to be

$$= \frac{EIY^2}{\left[\frac{L}{2}\right]^{12}} \int_0^{L/2} \mathrm{d}x \left(-4.3 \left[\frac{L}{2}\right]^4 + 15.6 \left[\frac{L}{2}\right]^2 x^2 - 4.5x^4 \right)^2 \tag{2.4}$$

so finally

$$V_{\mathrm{F}} = 98.9EI\frac{Y^2}{L^3}. \tag{2.5}$$

The strain energy due to the in-plane axial load is found from the expression governing the axial extension,

$$\Delta L_1 = (\mathrm{d}x^2 + \mathrm{d}y^2)^{\frac{1}{2}} = \mathrm{d}x \left[1 + \left(\frac{\mathrm{d}y}{\mathrm{d}x} \right)^2 \right]^{\frac{1}{2}} \tag{2.6}$$

i.e.

$$\Delta L_1 = \mathrm{d}x \left[1 + \frac{1}{2} \left(\frac{\mathrm{d}y}{\mathrm{d}x} \right)^2 - \frac{1}{8} \left(\frac{\mathrm{d}y}{\mathrm{d}x} \right)^4 + \cdots \right] \approx \mathrm{d}x \left[1 + \frac{1}{2} \left(\frac{\mathrm{d}y}{\mathrm{d}x} \right)^2 \right]. \tag{2.7}$$

Therefore,

$$L_1 = \int_{-L/2}^{L/2} \mathrm{d}x \left[1 + \frac{1}{2} \left(\frac{\mathrm{d}y}{\mathrm{d}x} \right)^2 \right] = L + \frac{1}{2} \int_{-L/2}^{L/2} \mathrm{d}x \left(\frac{\mathrm{d}y}{\mathrm{d}x} \right)^2 \tag{2.8}$$

and ΔL, the change in axial length of the beam, is given by

$$L_1 - L = \frac{1}{2} \int_{-L/2}^{L/2} \mathrm{d}x \left(\frac{\mathrm{d}y}{\mathrm{d}x} \right)^2. \tag{2.9}$$

Substituting for $y(x)$ from equation (2.1) gives

$$\Delta L = 2.44 \frac{Y^2}{L}. \tag{2.10}$$

Thus, the axial strain energy is

$$V_{\mathrm{A}} = \frac{1}{2} \frac{EA}{L} (\Delta L)^2 = 2.98EA \frac{Y^4}{L^3}. \tag{2.11}$$

From Lagrange's equations, the stiffness terms are given by

$$\frac{\partial}{\partial Y}(V_{\mathrm{F}} + V_{\mathrm{A}}) = 197.8 \frac{EIY}{L^3} + 11.92 \frac{EAY^3}{L^3} \tag{2.12}$$

i.e. the linear elastic stiffness term is $k_1 = 197.8EI/L^3$ and the nonlinear hardening-stiffness term is $k_3 = 11.92EAY^2/L^3$. (Note that the linear elastic

stiffness term k_1 should be $192EI/L^3$ from simple bending theory. The small error is due to limiting the assumed deflection polynomial to only three terms.)

In practise, because it is not possible to fully implement the principle of superposition, i.e. spanning all the possibilities of inputs, simpler procedures are employed. Since best practice in dynamic testing should always include some check of linearity, it is important that easy-to-use procedures for detecting nonlinearity are available. The most commonly used procedures are based on harmonic distortion, homogeneity and reciprocity.

2.2.2 Harmonic distortion

Harmonic or waveform distortion is one of the clearest indicators of the presence of nonlinearity. It is a straightforward consequence of the principle of superposition. If the excitation to a linear system is a monoharmonic signal, i.e. a sine or cosine wave of frequency ω, the response will be monoharmonic at the same frequency (after any transients have died out). The proof is elementary [1] and proceeds as follows.

Suppose $x(t) = \sin(\omega t)$ is the input to a linear system. First of all, it is observed that $x(t) \longrightarrow y(t)$ implies that $\dot{x}(t) \longrightarrow \dot{y}(t)$ and $\ddot{x}(t) \longrightarrow \ddot{y}(t)$. This is because superposition demands that

$$\frac{x(t+\Delta t) - x(t)}{\Delta t} \longrightarrow \frac{y(t+\Delta t) - y(t)}{\Delta t} \tag{2.13}$$

and $\dot{x}(t) \longrightarrow \dot{y}(t)$ follows in the limit as $\Delta t \longrightarrow 0$. (Note that there is also an implicit assumption of *time invariance* here, namely that $x(t) \longrightarrow y(t)$ implies $x(t+\tau) \longrightarrow y(t+\tau)$ for any τ.) Again, by superposition,

$$x_1(t) + \omega^2 x_2(t) \longrightarrow y_1(t) + \omega^2 y_2(t) \tag{2.14}$$

so taking $x_1(t) = \ddot{x}(t)$ and $x_2(t) = x(t)$ gives

$$\ddot{x}(t) + \omega^2 x(t) \longrightarrow \ddot{y}(t) + \omega^2 y(t). \tag{2.15}$$

Now, as $x(t) = \sin(\omega t)$,

$$\ddot{x}(t) + \omega^2 x(t) = 0. \tag{2.16}$$

In the steady state, a zero input to a linear system results in a zero output. It therefore follows from (2.15) that

$$\ddot{y}(t) + \omega^2 y(t) = 0 \tag{2.17}$$

and the general solution of this differential equation is

$$y(t) = A\sin(\omega t - \phi) \tag{2.18}$$

[1] The authors learnt this proof from Dr Hugh Goyder.

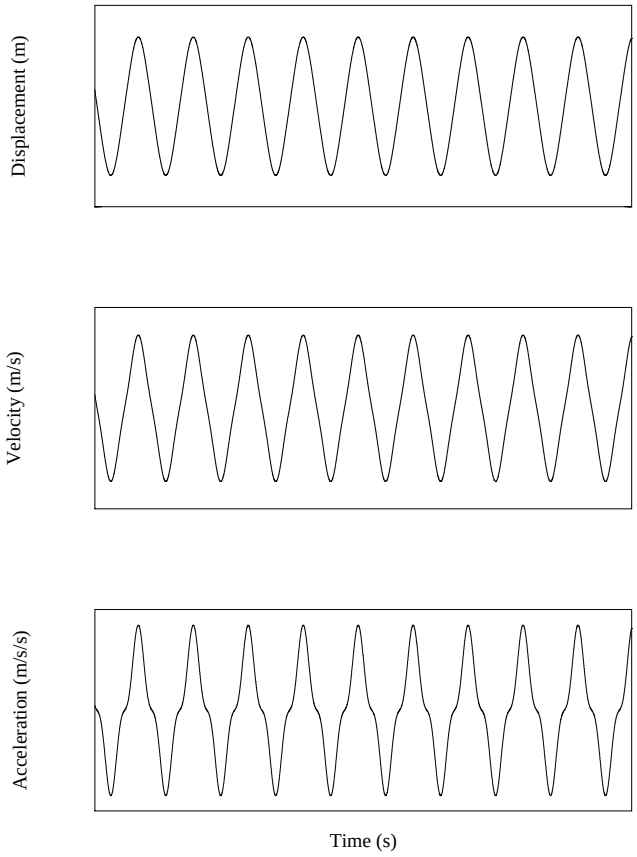


Figure 2.5. Response signals from a nonlinear system showing clear distortion only on the acceleration signal.

and this establishes the result. This proof is rather interesting as it only uses the fact that $x(t)$ satisfies a homogeneous linear differential equation to prove the result. The implication is that any such function will not suffer distortion in passing through a linear system.

It is not a corollary of this result that a sine-wave input to a nonlinear system will not generally produce a sine-wave output; however, this is usually the case and this is the basis of a simple and powerful test for nonlinearity as sine waves are simple signals to generate in practice. The form of the distortion will be discussed in chapter 3, it will be revealed that the change in form is due to the appearance of higher harmonics in the response such as $\sin(3\omega t)$, $\sin(5\omega)$ etc.

Distortion can be easily detected on an oscilloscope by observing the input and output time response signals. Figures 2.5 and 2.6 show examples of

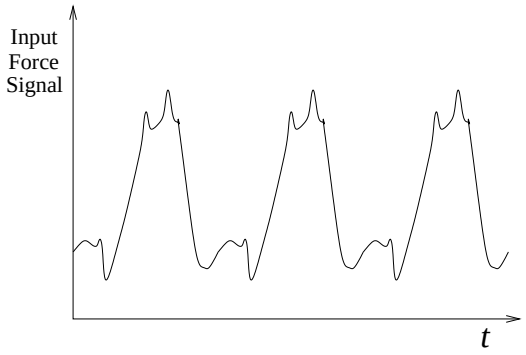


Figure 2.6. Distortion on the input force signal arising from vibration exciter misalignment (the severe distortion is due to the exciter coil rubbing against the magnet).

harmonic waveform distortion where a sinusoidal excitation signal is warped due to nonlinearity.

In figure 2.5 the output response from a nonlinear system is shown in terms of the displacement, velocity and acceleration. The reason that the acceleration is more distorted compared with the corresponding velocity and displacement is easily explained. Let $x(t) = \sin(\omega t)$ be the input to the nonlinear system. As previously stated, the output will generally (at least for weak nonlinear systems [2]) be represented as a Fourier series composed of harmonics written as

$$y(t) = A_1 \sin(\omega t + \phi_1) + A_2 \sin(2\omega t + \phi_2) + A_3 \sin(3\omega t + \phi_3) + \cdots \quad (2.19)$$

and the corresponding acceleration is

$$\begin{aligned} y(t) = & -\omega^2 A_1 \sin(\omega t + \phi_1) - 4\omega^2 B_2 \sin(2\omega t + \phi_2) \\ & - 9\omega^2 B_3 \sin(3\omega t + \phi_2) - \cdots . \end{aligned} \quad (2.20)$$

Thus the nth output acceleration term is weighted by the factor n^2 compared to the fundamental.

In figure 2.6 the signal represents the output of a force transducer during a modal test. The distortion is due to shaker misalignment resulting in friction between the armature of the shaker and the internal magnet—a nonlinearity.

If non-sinusoidal waveforms are used, such as band-limited random signals, waveform distortion is generally impossible to detect and additional procedures are required such as the coherence function described in section 2.5.2.

[2] There are a number of opinions as to what constitutes *weak* nonlinearity. What it means here is simply that the system does not undergo transition to chaos or show subharmonic generation.

2.2.3 Homogeneity and FRF distortion

This represents a restricted form of the principle of superposition. It is undoubtedly the most common method in use for detecting the presence of nonlinearity in dynamic testing. Homogeneity is said to hold if $x(t) \longrightarrow y(t)$ implies $\alpha x(t) \longrightarrow \alpha y(t)$ for all α. In essence, homogeneity is an indicator of the system's insensitivity to the magnitude of the input signal. For example, if an input $\alpha x_1(t)$ always produces an output $\alpha y_1(t)$, the ratio of output to input is independent of the constant α. The most striking consequence of this is in the frequency domain. First, note that $\alpha x(t) \longrightarrow \alpha y(t)$ implies $\alpha X(\omega) \longrightarrow \alpha Y(\omega)$. This means that if $x(t) \longrightarrow \alpha x(t)$,

$$H(\omega) = \frac{Y(\omega)}{X(\omega)} \longrightarrow \frac{\alpha Y(\omega)}{\alpha X(\omega)} = H(\omega) \tag{2.21}$$

and the FRF is invariant under changes of α or effectively of the level of excitation.

Because of this, the homogeneity test is usually applied in dynamic testing to FRFs where the input levels are usually mapped over a range encompassing typical operating levels. If the FRFs for different levels overlay, linearity is assumed to hold. This is not infallible as there are some systems which are nonlinear which nonetheless show homogeneity; the bilinear system discussed in the next chapter is an example. The reason for this is that homogeneity is a weaker condition than superposition.

An example of the application of a homogeneity test is shown in figure 2.7. In this case band-limited random excitation has been used but, in principle, any type of excitation signal may be employed. Although a visual check is often sufficient to see if there are significant differences between FRFs, other metrics can be used such as a measure of the mean-square error between the FRFs. The exact form of the distortion in the FRF depends on the type of the nonlinearity, some common types of FRF distortion produced by varying the level of excitation are discussed in the following section.

One possible problem with the homogeneity test is caused by force 'drop-out'. Drop-out is a common phenomenon which occurs when forced vibration tests are carried out during dynamic testing. As its description implies, this is a reduction in the magnitude of the input force spectrum measured by the force transducer and occurs in the vicinity of the resonant frequencies of the structure under test. It is a result of the interaction between an electrodynamic exciter and the structure [251]. A typical experimental force drop-out characteristic is shown in figure 2.8.

If homogeneity is being used as a detection method for nonlinearity, force drop-out can create misleading results. This is because the test for homogeneity assumes that the input is persistently exciting, i.e. exercises the system equally across the whole excitation bandwidth, whereas the effect of force drop-out is to effectively notch-filter the input at the resonant frequency. This results in less

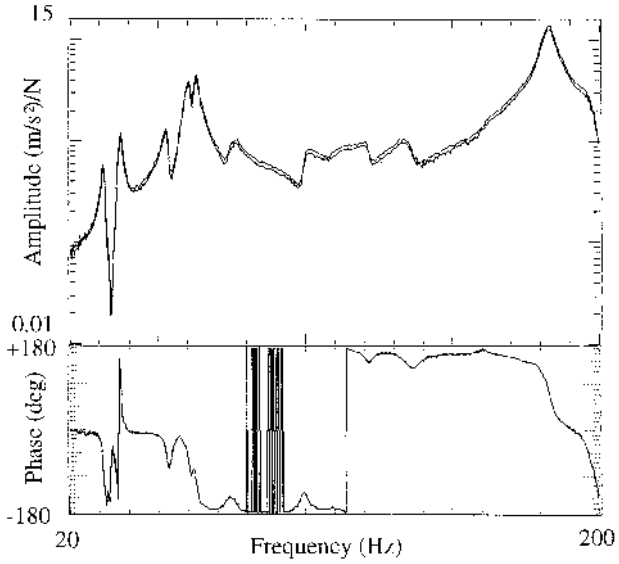


Figure 2.7. Application of a homogeneity test on a real structure. The close agreement of the results is an indicator that the structure is linear within the excitation bounds used.

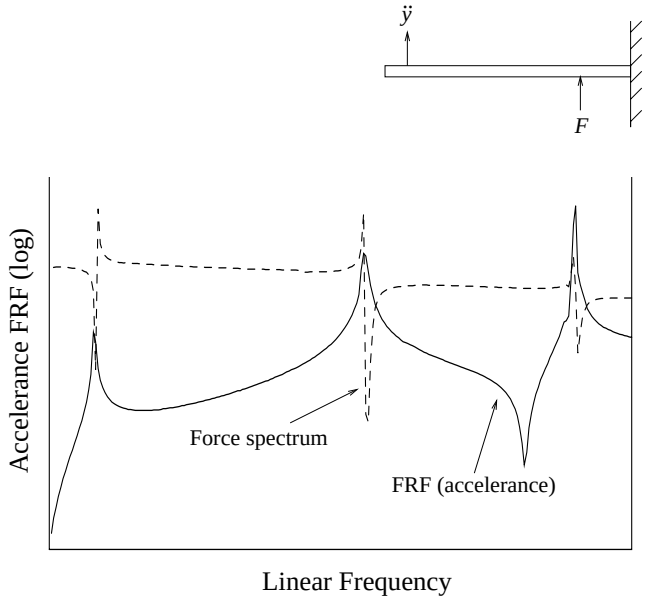


Figure 2.8. A typical force 'drop-out' characteristic overlayed on the FRF of a cantilever beam. Note the correspondence between the force spectrum minima and the FRF maxima.

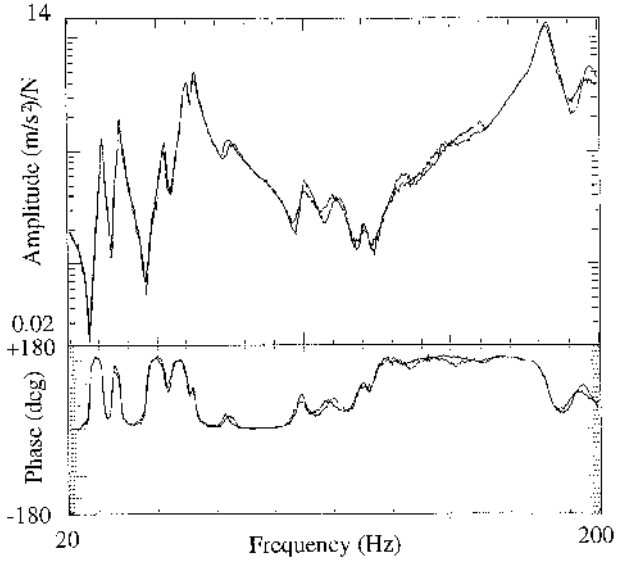


Figure 2.9. Application of a reciprocity test on a real structure. The close agreement of the results is an indicator that the structure is linear within the test bounds.

force communicated to the structure near resonance and the response may be linearized. If a control system is employed to maintain a constant excitation force spectrum, nonlinearity can easily be detected using homogeneity.

2.2.4 Reciprocity

Reciprocity is another important property which, if violated, can be used to detect the presence of nonlinearity. For linearity to hold reciprocity is a necessary but not a sufficient condition since some symmetrical nonlinear systems may exhibit reciprocity but will not satisfy the principle of superposition. Reciprocity holds if an output y_{B} at a point B due to an input x_{A} at a point A, gives a ratio $y_{\mathrm{B}}/x_{\mathrm{A}}$ numerically equal to that when the input and output points are reversed giving $y_{\mathrm{A}}/x_{\mathrm{B}}$. It follows that if this condition holds, the FRFs for the processes $x_{\mathrm{A}} \longrightarrow y_{\mathrm{B}}$ and $x_{\mathrm{B}} \longrightarrow y_{\mathrm{A}}$ are equal. This is the basis of the experimental test.

Figure 2.9 shows the results of a reciprocity test on a structure using band-limited random excitation and the FRFs between two different points, A and B. As in the homogeneity test, the difference is usually assessed by eye.

When employing reciprocity it is important to note that all the response parameters must be the same, e.g. displacements or accelerations and all the inputs must be forces. If reciprocity holds, then by definition the stiffness matrix of a structure will be symmetric as will the FRF matrix.

2.3 Common types of nonlinearity

The most common types of nonlinearity encountered in dynamic testing are those due to polynomial stiffness and damping, clearances, impacts, friction and saturation effects. As one would expect, these nonlinearities are usually amplitude, velocity and frequency dependent. However, it is usual to simplify and idealize these in order that they can be incorporated into analysis, simulation and prediction capabilities. Consider an SDOF oscillator with nonlinear damping and stiffness terms:

$$m\ddot{y} + f_{\mathrm{d}}(\dot{y}) + f_{\mathrm{s}}(y) = x(t). \tag{2.22}$$

Figure 2.10 summarizes the most common types of nonlinearity in terms of their idealized force against displacement or force against velocity characteristics.

Some examples of the effects of several of the nonlinearities shown in figure 2.10 on the vibration characteristics of an isolated mode of vibration (in this case considered as an SDOF) in the FRF subject to sinusoidal excitation can be seen in figure 2.11. Here, the frequency response characteristics are shown in terms of the Argand plane in the Nyquist plot) and the modulus of the receptance FRF. Distortions are clearly seen which, if not recognized and understood, may produce errors in the parameters which are extracted from these FRFs if curve-fitting is used. A detailed discussion of the origin of these distortions is postponed until chapter 3, only brief observations will be made here. If a structure incorporates actuators, bearings, linkages or elastomeric elements, these can act as localized nonlinearities whose characteristics may be represented by one or more of those shown in figure 2.10.

It is instructive to consider each nonlinearity briefly in turn.

2.3.1 Cubic stiffness

In this case, the force displacement characteristic has the form,

$$f_{\mathrm{s}}(y) = ky + k_3 y^3 \tag{2.23}$$

and k_3 may be positive or negative. If $k_3 > 0$, one can see that at high levels of excitation the restoring force will be greater than that expected from the linear term alone. The extent of this excess will increase as the forcing level increases and for this reason such systems are referred to as having a *hardening* characteristic. Examples of such systems are clamped plates and beams as discussed earlier. If $k_3 < 0$, the effective stiffness decreases as the level of excitation increases and such systems are referred to as *softening*. Note that softening cubic systems are unphysical in the sense that the restoring force changes sign at a certain distance from equilibrium and begins to drive the system to infinity. Systems with such characteristics are always found to have higher-order polynomial terms in the stiffness with positive coefficients which dominate at high levels and restore stability. Systems which appear to show softening cubic behaviour over limited ranges include buckling beams plates.

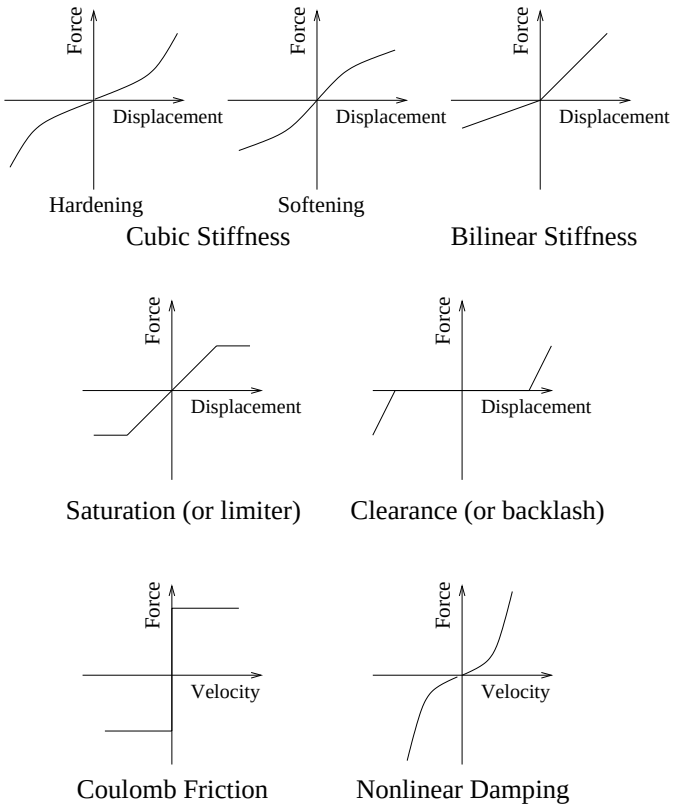


Figure 2.10. Idealized forms of simple structural nonlinearities.

The equation of motion of the SDOF oscillator with linear damping and stiffness (2.23) is called *Duffing's equation* [80],

$$m\ddot{y} + c\dot{y} + ky + k_3 y^3 = x(t) \tag{2.24}$$

and this is the single most-studied equation in nonlinear science and engineering. The reason for its ubiquity is that it is the simplest nonlinear oscillator which possesses the odd symmetry which is characteristic of many physical systems. Despite its simple structure, it is capable of showing almost all of the interesting behaviours characteristic of general nonlinear systems. This equation will re-occur many times in the following chapters.

The FRF distortion characteristic of these systems is shown in figures 2.11(*b*) and (*c*). The most important point is that the resonant frequency shifts up for the hardening system as the level of excitation is raised, this is consistent with the

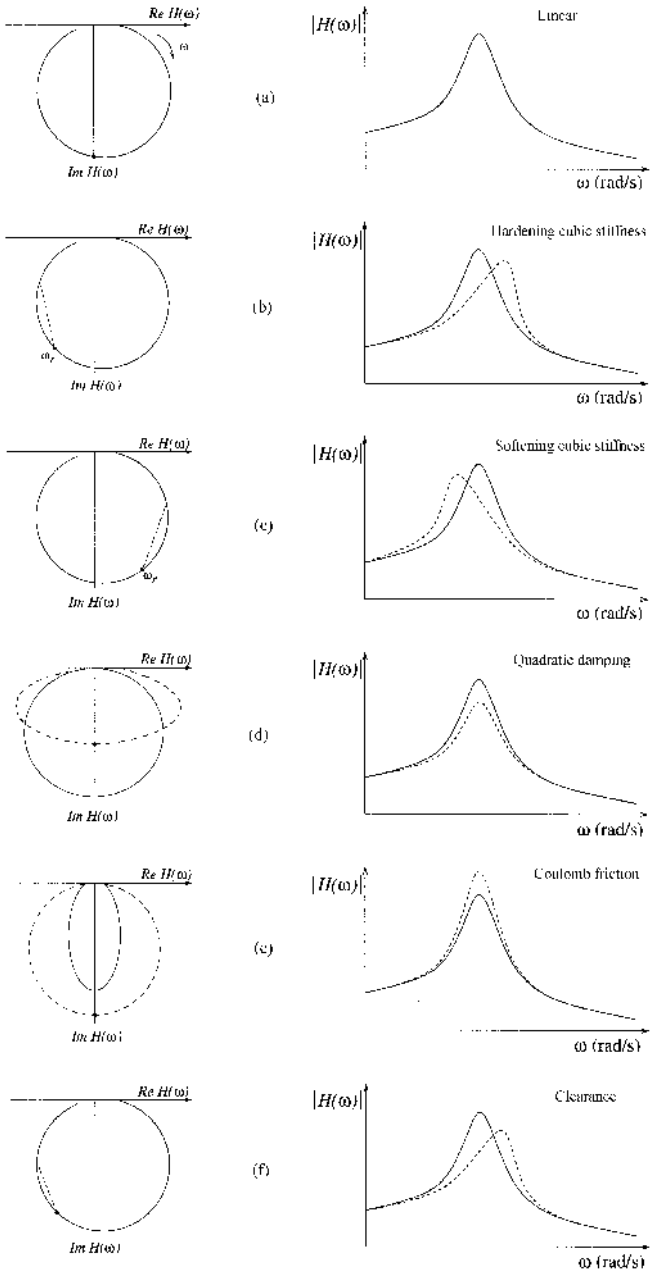


Figure 2.11. SDOF system Nyquist and FRF (Bode) plot distortions for five types of nonlinear element excited with a constant amplitude sinusoidal force; —— low level, – – – high level.

increase in effective stiffness. As one might expect, the resonant frequency for the softening system shifts down.

2.3.2 Bilinear stiffness or damping

In this case, the stiffness characteristic has the form,

$$f_{\mathrm{s}}(y) = \begin{cases} k_1 y, & y > 0 \\ k_2 y, & y < 0 \end{cases} \tag{2.25}$$

with a similar definition for bilinear damping. The most extreme example of a bilinear system is the impact oscillator for which $k_1 = 0$ and $k_2 = \infty$; this corresponds to a ball bouncing against a hard wall. Such systems can display extremely complex behaviour indeed (see chapter 15 of [248]). One system which approximates to a bilinear damping system is the standard automotive damper or shock absorber which is designed to have different damping constants in compression and rebound. Such systems are discussed in detail in chapters 7 and 9.

Figure 2.11 does not show the FRF distortion characteristic of this system because it is one of the rare nonlinear systems which display homogeneity. (This last remark is only true if the position of the change in stiffness is at the origin, if it is offset by any degree, the system will fail to show homogeneity if the level of excitation is taken sufficiently high.)

2.3.3 Piecewise linear stiffness

The form of the stiffness function in this case is

$$f_{\mathrm{s}}(y) = \begin{cases} k_2 y + (k_1 - k_2)d, & y > d \\ k_1 y, & |y| < d \\ k_2 y - (k_1 - k_2)d, & y < -d. \end{cases} \tag{2.26}$$

Two of the nonlinearities in figure 2.10 are special cases of this form. The saturation or limiter nonlinearity has $k_2 = 0$ and the clearance or backlash nonlinearity has $k_1 = 0$.

In aircraft ground vibration tests, nonlinearities of this type can arise from assemblies such as pylon–store–wing assemblies or pre-loading bearing locations. Figure 2.12 shows typical results from tests on an aircraft tail-fin where the resonant frequency of the first two modes reduces as the input force level is increased and then asymptotes to a constant value. Such results are typical of pre-loaded backlash or clearance nonlinearities.

Typical FRF distortion is shown in figure 2.11(*f*) for a hardening piecewise linear characteristic ($k_2 > k_1$).

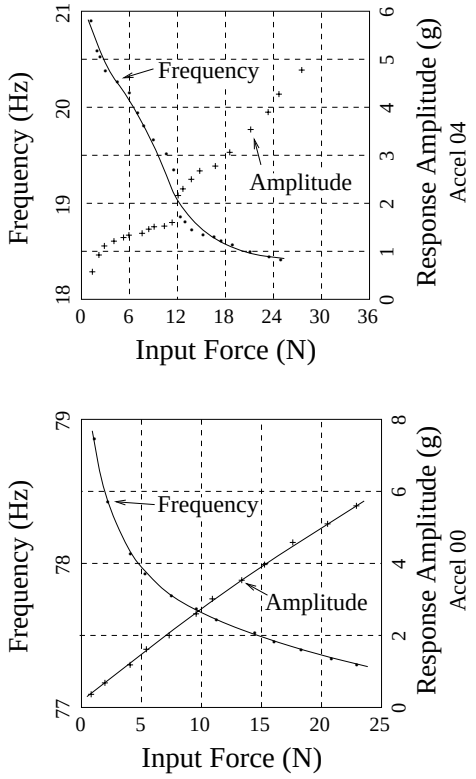


Figure 2.12. Results from ground vibration tests on the tail-fin of an aircraft showing significant variation in the resonant frequency with increasing excitation level. This was traced to clearances in the mounting brackets.

2.3.4 Nonlinear damping

The most common form of polynomial damping is quadratic:

$$f_{\mathrm{d}}(\dot{y}) = c_2 \dot{y}|\dot{y}| \tag{2.27}$$

(where the absolute value term is to ensure that the force is always opposed to the velocity). This type of damping occurs when fluid flows through an orifice or around a slender member. The former situation is common in automotive dampers and hydromounts, the latter occurs in the fluid loading of offshore structures. The fundamental equation of fluid loading is Morison's equation [192],

$$F(t) = c_1 \dot{u}(t) + c_2 u(t)|u(t)| \tag{2.28}$$

where F is the force on the member and u is the velocity of the flow. This system will be considered in some detail in later chapters.

The effect of increasing excitation level is to increase the effective damping as shown in figure 2.11(*d*).

2.3.5 Coulomb friction

This type of damping has characteristic,

$$f_{\mathrm{d}}(\dot{y}) = c_{\mathrm{F}}\,\mathrm{sgn}(\dot{y}) \tag{2.29}$$

as shown in figure 2.10. This type of nonlinearity is common in any situation with interfacial motion. It is particularly prevalent in demountable structures such as grandstands. The conditions of constant assembly and disassembly are suitable for creating interfaces which allow motion. In this sort of structure friction will often occur in tandem with clearance nonlinearities. It is unusual here in the sense that it is most evident at low levels of excitation, where in extreme cases, stick–slip motion can occur. At higher levels of excitation, the friction 'breaks out' and the system will behave nominally linearly. The characteristic FRF distortion (figure 2.11(*e*)) is the reverse of the quadratic damping case, with the higher damping at low excitation.

2.4 Nonlinearity in the measurement chain

It is not uncommon for nonlinearity to be unintentionally introduced in the test programme through insufficient checks on the test set-up and/or the instrumentation used. There are several common sources of nonlinearity whose effects can be minimized at the outset of a test programme and consideration should be given to simple visual and acoustic inspection procedures (listening for rattles etc) before the full test commences.

The principal sources of nonlinearity arising from insufficient care in the test set-up are:

- misalignment
- exciter problems
- looseness
- pre-loads
- cable rattle
- overloads/offset loads
- temperature effects
- impedance mismatching
- poor transducer mounting

Most of these problems are detectable in the sense that they nearly all cause waveform distortion of some form or other. Unless one observes the actual input and output signals periodically during testing it is *impossible* to know whether or not any problems are occurring. Although tests frequently involve the

measurement of FRFs or spectra it is strongly recommended that a visual check is maintained of the individual drive/excitation and response voltage signals. This can be done very simply by the use of an oscilloscope.

In modal testing it is usual to use a force transducer (or transducers in the case of multi-point testing) as the reference input signal. Under such circumstances it is strongly recommended that this signal is continuously (or at least periodically) monitored on an oscilloscope. This is particularly important as harmonic distortion of the force excitation signal is not uncommon, often due to shaker misalignment or 'force drop-out' at resonance. Distortion can create errors in the measured FRF which may not be immediately apparent and it is very important to ensure that the force input signal is not distorted.

Usually in dynamic testing one may have the choice of observing the waveform in terms of displacement, velocity or acceleration. For a linear system in which no distortion of the signal occurs it makes little difference which variable is used. However, when nonlinearity is present this generally results in harmonic distortion. As discussed earlier in this chapter, under sinusoidal excitation, harmonic distortion is much easier to observe when acceleration is measured. Thus it is recommended that during testing with a sine wave, a simple test of the quality of the output waveform is to observe it on an oscilloscope in terms of the acceleration response. Any distortion or noise present will be more easily visible.

Due to their nature, waveform distortion in random signals is more difficult to observe using an oscilloscope than with a sine-wave input. However, it is still recommended that such signals are observed on an oscilloscope during testing since the effect of extreme nonlinearities such as clipping of the waveforms can easily be seen.

The first two problems previously itemized will be discussed in a little more detail.

2.4.1 Misalignment

This problem often occurs when electrodynamic exciters are used to excite structures in modal testing. If an exciter is connected directly to a structure then the motion of the structure can impose bending moments and side loads on the exciter armature and coil assembly resulting in misalignment, i.e. the coil rubbing against the internal magnet of the exciter. Misalignment can be detected by using a force transducer between the exciter and the test structure, the output of which should be observed on an oscilloscope. If a sine wave is injected into the structure, misalignment will produce a distorted force signal which, if severe, may appear as shown in figure 2.6. If neglected, this can create significant damage to the vibration exciter coil, resulting in a reduction in the quality of the FRFs and eventual failure of the exciter. To minimize this effect it is recommended that a 'stinger' or 'drive-rod' is used between the exciter and the test structure described in [87].

2.4.2 Vibration exciter problems

Force drop-out was briefly mentioned in section 2.2.3. When electrodynamic vibration exciters are employed to excite structures, the actual force that is applied is the reaction force between the exciter and the structure under test. The magnitude and phase of the reaction force depends upon the characteristics of the structure and the exciter. It is frequently (but mistakenly) thought that if a force transducer is located between the exciter and the structure then one can forget about the exciter, i.e. it is outside the measurement chain. In fact, the quality of the actual force applied to the structure, namely the reaction force, is very dependent upon the relationship between the exciter and the structure under test.

Detailed theory shows that, in order to apply a constant-magnitude force to a structure as the frequency is varied, it would be necessary to use an exciter whose armature mass and spider stiffness are negligible. This can only be achieved using special exciters such as non-contact electromagnetic devices or electrodynamic exciters based on magnets which are aligned with lightweight armatures that are connected to the structure, there then being no spider stiffness involved.

When a sine wave is used as the excitation signal and the force transducer signal is observed on an oscilloscope, within the resonance region the waveform may appear harmonically distorted and very small in magnitude. This is particularly evident when testing lightly damped structures. The harmonic distortion in the force signal is due to the fact that at resonance the force supplied by the exciter has merely to overcome the structural damping. If this is small (as is often the case), the voltage level representing the force signal becomes very small in relation to the magnitude of the nonlinear harmonics present in the exciter. These nonlinearities are created when the structure and hence armature of the exciter undergoes large amplitudes of vibration (at resonance) and begins to move into the non-uniform flux field in the exciter. This non-uniform flux field produces strong second harmonics of the excitation frequency which distorts the fundamental force signal.

2.5 Two classical means of indicating nonlinearity

It is perhaps facetious to use the term ‘classical’ here as the two techniques discussed are certainly very recent in historical terms. The reason for the terminology is that they were both devised early in the development of modal testing, many years before most of the techniques discussed in this book were developed. This is not to say that their time is past—coherence, in particular, is arguably the simplest test for nonlinearity available via mass-produced instrumentation.

2.5.1 Use of FRF inspections—Nyquist plot distortions

FRFs can be visually inspected for the characteristic distortions which are indicative of nonlinearity. In particular, the resonant regions of the FRFs will be the most sensitive. In order to examine these regions in detail, the use of the the Nyquist plot (i.e. imaginary versus real part of the FRF) is commonly used. (If anti-resonances are present, they can also prove very sensitive to nonlinearity.)

The FRF is a complex quantity, i.e. it has both magnitude and phase, both of which can be affected by nonlinearity. In some cases it is found that the magnitude of the FRF is the most sensitive to the nonlinearity and in other cases it is the phase. Although inspecting the FRF in terms of the gain and phase characteristics separately embodies all the information, combining these into one plot, namely the Nyquist plot, offers the quickest and most effective way of inspecting the FRF for distortions.

The type of distortion which is introduced in the Nyquist plot depends upon the type of nonlinearity present in the structure and on the excitation used, as discussed elsewhere in this chapter. However, a simple rule to follow is that if the FRF characteristics in the Nyquist plane differ significantly from a circular or near-circular locus in the vicinity of the resonances then nonlinearity is a suspect. Examples of common forms of Nyquist plot distortion as a result of structural nonlinearity, obtained from numerical simulation using sinusoidal excitation, are shown in figure 2.11. It is interesting to note that in the case of the non-dissipative nonlinearities under low levels of excitation, e.g. the polynomial and piecewise nonlinear responses, the Nyquist plot appears as a circular locus. However, by inspecting the $\Delta\omega$ spacings (proportional to the change in phase) it is possible to detect a phase distortion. When the input excitation level is increased to the point at which the effect of the nonlinearity becomes severe enough to create the 'jump' phenomenon (discussed in more detail in the next chapter), the Nyquist plot clearly shows this.

In the case of dissipative nonlinearities and also friction, the distortion in the Nyquist plot is easily detected with appropriate excitation levels via the unique characteristic shapes appearing which have been referred to as the 'apples and pears' of FRFs.

An example of nonlinearity from an attached element is shown in figure 2.13 where a dynamic test was carried out on a cantilever beam structure which had a hydraulic, passive, actuator connected between the beam and ground. Under low-level sinusoidal excitation the friction in the actuator seals dominates the response producing a distorted 'pear-shaped' FRF as shown in figure 2.13.

When the excitation level was increased by a factor of three (from a $2N$ to a $6N$ peak), the FRF distortion changed to an oval shape. These changes in the FRF can be attributed to the nonlinearity changing from a friction characteristic at low input excitation levels to a nonlinear velocity-dependent characteristic such as a quadratic damping effect.

It is relatively straightforward to demonstrate that such distortions occur

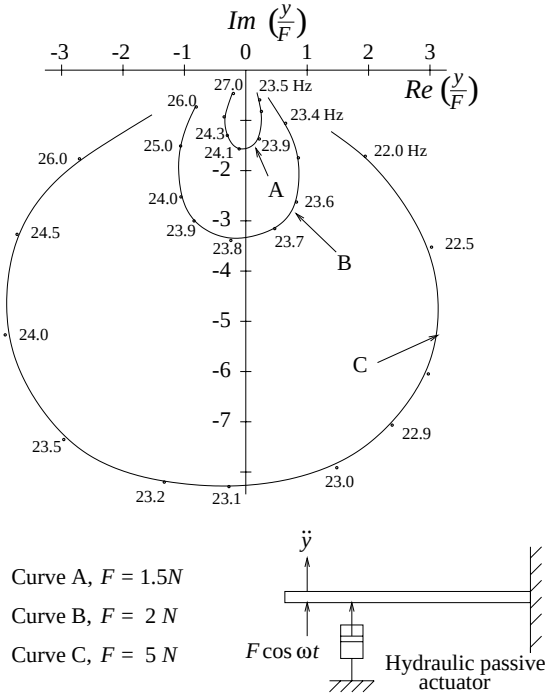


Figure 2.13. Nyquist plot distortions arising from a combination of seal friction nonlinearity in the passive hydraulic actuator at low excitation levels and a velocity-squared nonlinearity at higher excitation levels.

in the Argand plane when nonlinearity is present. Anticipating the theme of the next chapter a little, consider the case of a simple oscillator, with structural damping constant δ and Coulomb friction of magnitude c_{F}, given by the equation of motion,

$$m\ddot{y} + k(1 + \mathrm{i}\delta)y + c_{\mathrm{F}}\,\mathrm{sgn}(\dot{y}) = P\mathrm{e}^{\mathrm{i}\omega t}. \tag{2.30}$$

By using the method of harmonic balance (see chapter 3) the Coulomb friction function can be represented by an equivalent structural damping constant h^*, where

$$h^* = \frac{4c_{\mathrm{F}}}{\pi}|Y| \tag{2.31}$$

where Y is the peak displacement. Thus equation (2.30) can be written as

$$m\ddot{y} + k(1 + \mathrm{i}\delta^*)y = P\mathrm{e}^{\mathrm{i}\omega t} \tag{2.32}$$

with

$$\delta^* = \delta + \frac{4c_{\mathrm{F}}}{\pi}|Y|. \tag{2.33}$$

The solution to equation (2.32) can be written as

$$y(t) = Y\mathrm{e}^{\mathrm{i}\omega t} \quad \text{with } Y = |Y|\mathrm{e}^{\mathrm{i}\phi} \tag{2.34}$$

i.e.

$$|Y| = \left(\frac{P}{k}\right)\left[(1-\beta^2)^2 + \delta^{*2}\right]^{\frac{1}{2}}, \quad \tan\phi = \frac{\delta^*}{(1-\beta^2)} \tag{2.35}$$

where $\beta = \omega/\omega_{\mathrm{n}}$. Substituting (2.33) in (2.35) gives the magnitude of the response as

$$|Y| = \frac{-\delta r + \left[\left(\frac{P}{k}\right)\{(1-\beta^2)^2 + \delta^2\} - r^2(1-\beta^2)^2\right]^{\frac{1}{2}}}{(1-\beta^2)^2 + \delta^2} \tag{2.36}$$

and the phase as

$$\phi = \tan^{-1}\frac{\left[\delta + \frac{r}{|Y|}\right]}{(1-\beta^2)} \tag{2.37}$$

where $r = 4c_{\mathrm{F}}/\pi k$.

A solution for $|Y|$ is only possible when $r < P/k$. If this condition is violated, stick–slip motion occurs and the solution is invalid. When the vector response is plotted in the Argand plane the loci change from a circular response for $r = 0$, i.e. a linear system, to a distorted, pear-shaped response as r increases. In the case of viscously damped systems, the substitution $\delta = 2\zeta$ can generally be made without incurring any significant differences in the predicted results.

2.5.2 Coherence function

The coherence function is a spectrum and is usually used with random or impulse excitation. It can provide a quick visual inspection of the quality of an FRF and, in many cases, is a rapid indicator of the presence of nonlinearity in specific frequency bands or resonance regions. It is arguably the most often-used test of nonlinearity, by virtue of the fact that almost all commercial spectrum analysers allow its calculation.

Before discussing nonlinearity, the coherence function will be derived for *linear* systems subject to measurement noise on the output (figure 2.14). Such

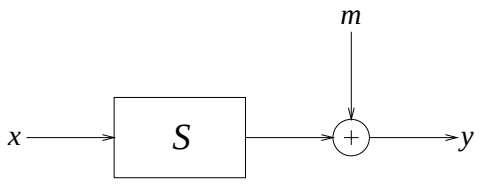


Figure 2.14. Block diagram of a linear system with noise on the output signal.

systems have time-domain equations of motion,

$$y(t) = S[x(t)] + m(t) \tag{2.38}$$

where $m(t)$ is the measurement noise. In the frequency domain,

$$Y(\omega) = H(\omega)X(\omega) + M(\omega). \tag{2.39}$$

Multiplying this equation by its complex conjugate yields

$$Y\overline{Y} = HX\overline{HX} + HX\overline{M} + \overline{HX}M + M\overline{M} \tag{2.40}$$

and taking expectations gives[3]

$$S_{yy}(\omega) = |H(\omega)|^2 S_{xx}(\omega) + H(\omega)S_{xm}(\omega) + \overline{H}(\omega)S_{mx}(\omega) + S_{mm}(\omega). \tag{2.41}$$

Now, if x and m are uncorrelated signals (unpredictable from each other), then $S_{wx}(\omega) = S_{xw}(\omega) = 0$ and equation (2.41) reduces to

$$S_{yy}(\omega) = |H(\omega)|^2 S_{xx}(\omega) + S_{mm}(\omega) \tag{2.42}$$

and a simple rearrangement gives

$$\frac{|H(\omega)|^2 S_{xx}(\omega)}{S_{yy}(\omega)} = 1 - \frac{S_{mm}(\omega)}{S_{yy}(\omega)}. \tag{2.43}$$

The quantity on the right-hand side is the fraction of the output power, which can be linearly correlated with the input. It is called the *coherence* function and denoted $\gamma^2(\omega)$. Now, as $\gamma^2(\omega)$ and $S_{mm}(\omega)/S_{yy}(\omega)$ are both positive quantities, it follows that

$$0 \leq \gamma^2 \leq 1 \tag{2.44}$$

with $\gamma^2 = 1$ only if $S_{mm}(\omega) = 0$, i.e. if there is no measurement noise. The coherence function therefore detects if there is noise in the output. In fact, it will be shown later that $\gamma^2 < 1$ if there is noise anywhere in the measurement chain. If the coherence is plotted as a function of ω, any departures from unity will be readily identifiable. The coherence is usually expressed as

$$\gamma^2(\omega) = \frac{|S_{yx}(\omega)|^2}{S_{yy}(\omega)S_{xx}(\omega)}. \tag{2.45}$$

Note that all these quantities are easily computed by commercial spectrum analysers designed to estimate $H(\omega)$; this is why coherence facilities are so readily available in standard instrumentation.

[3] It is assumed that the reader is familiar with the standard definitions of auto-spectra and cross-spectra, e.g.

$$S_{yx}(\omega) = E[Y\overline{X}].$$

The coherence function also detects nonlinearity as previously promised. The relationship between input and output spectra for nonlinear systems will be shown in later chapters to have the form (for many systems)

$$Y(\omega) = H(\omega)X(\omega) + F[X(\omega)] \tag{2.46}$$

where F is a rather complicated function, dependent on the nonlinearity. Multiplying by $\overline{Y}$ and taking expectations gives

$$S_{yy}(\omega) = |H(\omega)|^2 S_{xx}(\omega) + H(\omega)S_{xf}(\omega) + \overline{H}(\omega)S_{fx}(\omega) + S_{ff}(\omega) \tag{2.47}$$

where this time the cross-spectra S_{fx} and S_{xf} will not necessarily vanish; in terms of the coherence,

$$\gamma^2(\omega) = 1 - 2\,\mathrm{Re}\left(H(\omega)\frac{S_{xf}(\omega)}{S_{yy}(\omega)}\right) - \frac{S_{ff}(\omega)}{S_{yy}(\omega)} \tag{2.48}$$

and the coherence will generally only be unity if $f = 0$, i.e. the system is linear. The test is not infallible as unit coherence will also be observed for a nonlinear system which satisfies

$$2\,\mathrm{Re}\,H(\omega)S_{xf}(\omega) = -S_{ff}(\omega) \tag{2.49}$$

However, this is very unlikely.

Consider the Duffing oscillator of equation (2.24). If the level of excitation is low, the response y will be small and y^3 will be negligible in comparison. In this regime, the system will behave as a linear system and the coherence function for input and output will be unity (figure 2.15). As the excitation is increased, the nonlinear terms will begin to play a part and the coherence will drop (figure 2.16). This type of situation will occur for all polynomial nonlinearities. However, if one considers Coulomb friction, the opposite occurs. At high excitation, the friction breaks out and a nominally linear response will be obtained and hence unit coherence.

Note that the coherence is *only* meaningful if averages are taken. For a one-shot measurement, a value of unity will always occur, i.e.

$$\gamma^2 = \frac{Y\overline{X}X\overline{Y}}{Y\overline{Y}X\overline{X}} = 1. \tag{2.50}$$

Finally, it is important to stress again that in order to use the coherence function for detecting nonlinearity it is necessary to realize that a reduction in the level of coherency can be caused by a range of problems, such as noise on the output and/or input signals which may in turn be due to incorrect gain settings on amplifiers. Such obvious causes should be checked before structural nonlinearity is suspected.

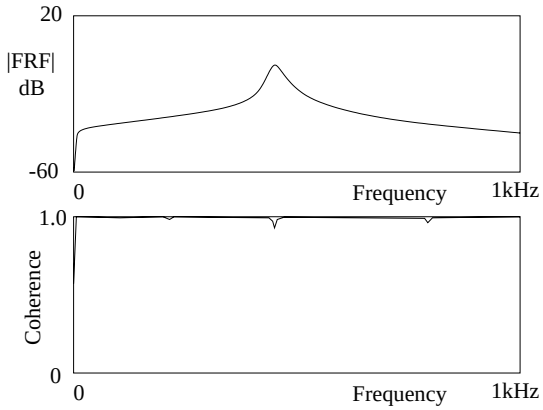


Figure 2.15. FRF gain and coherence plots for Duffing oscillator system given by equation (2.24) subject to low-level random excitation showing almost ideal unit coherence.

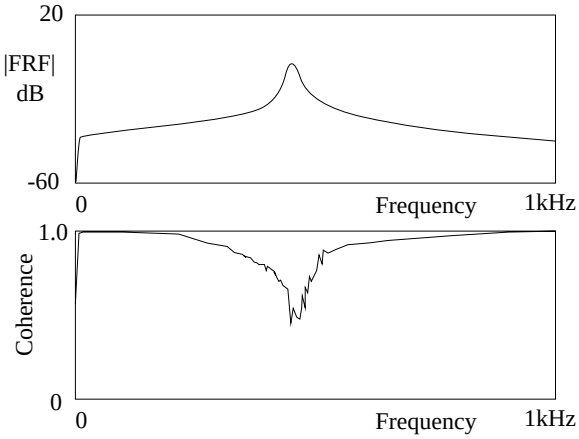


Figure 2.16. The effect of increasing the excitation level for the Duffing oscillator of figure 2.15, the coherence drops well below unity in the resonant region.

2.6 Use of different types of excitation

Nonlinear systems and structures respond in different ways to different types of input excitation. This is an important observation in terms of detecting the presence of nonlinearity or characterizing or quantifying it, some excitations will be superior to others. In order to fully discuss this, it will be useful to consider a concrete example of a nonlinear system. The one chosen is the Duffing oscillator

(with fairly arbitrary choices of parameter here),

$$\ddot{y} + 0.377\dot{y} + 39.489y + 0.4y^3 = x(t). \tag{2.51}$$

The excitation, $x(t)$ will be chosen to represent four common types used in dynamic testing namely steady-state sine, impact, rapid sine sweep (chirp) and random excitation.

2.6.1 Steady-state sine excitation

It is well known that the use of sinusoidal excitation usually produces the most vivid effects from nonlinear systems. For example, a system governed by a polynomial stiffness function can exhibit strong nonlinear effects in the FRF such as bifurcations (the jump phenomenon) where the magnitude of the FRF can suddenly reduce or increase. With stepped sinusoidal excitation, all the input energy is concentrated at the frequency of excitation and it is relatively simple, via integration, to eliminate noise and harmonics in the response signal (a standard feature on commercial frequency response function analysers).

As such, the signal-to-noise ratio is very good compared with random or transient excitation methods, an important requirement in all dynamic testing scenarios, and the result is a well-defined FRF with distortions arising from nonlinearity being very clear, particularly when a constant magnitude force excitation is used.

It should be remembered that one of the drawbacks of using stepped sine excitation methods is that they are slow compared with transient or random input excitation methods. This is because at each stepped frequency increment, time is required for the response to attain a steady-state condition (typically 1–2 s) before the FRF at that frequency is determined. However, this is usually a secondary factor compared with the importance of obtaining high-quality FRFs.

Consider figure 2.17(*a*). This FRF was obtained using steady-state sinusoidal excitation. At each frequency step a force was applied consisting of a constant amplitude sinewave. The displacement response was allowed to reach a steady-state condition and the amplitude and phase at the excitation frequency in the response were determined. The modulus of the ratio of the amplitude to the force at each frequency increment constitutes the modulus of the FRF (see chapter 1) shown in figure 2.17(*a*). The same (constant) amplitude of force was chosen for each frequency and this amplitude was selected so that the displacement of the system would be similar for all the excitation methods studied here. The FRF was obtained by stepping the frequency of excitation from 0.4 to 1.6 Hz (curve a–b–c–d) and then down from 1.6 Hz (curve d–c–e–a). As previously discussed, the distortion of the FRF from the usual linear form is considerable. The discontinuity observable in the curve will be discussed in considerable detail in chapter 3.

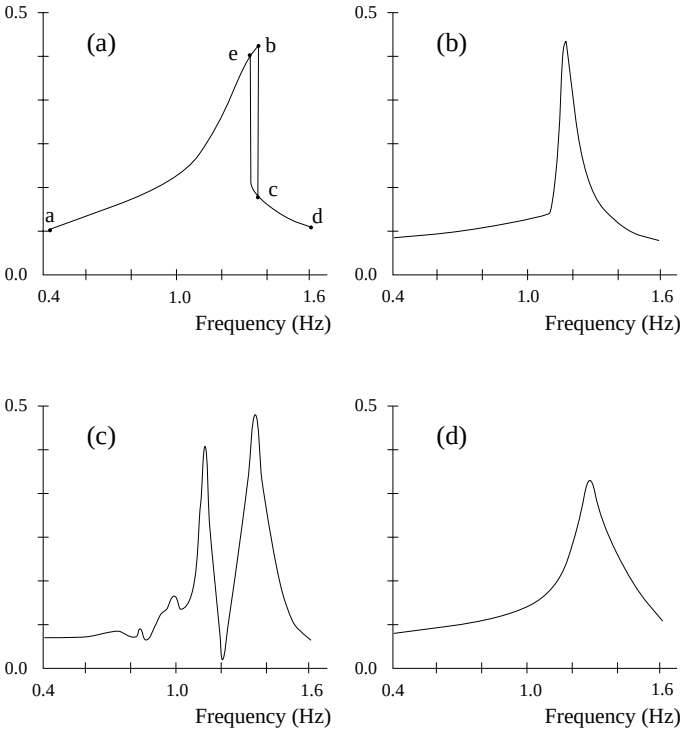


Figure 2.17. Measurement of the FRF of a single degree-of-freedom nonlinear oscillator with polynomial stiffness subject to different types of oscillation signals: (*a*) sinusoidal input; (*b*) pulse input; (*c*) rapid sweep (chirp) input; (*d*) random input.

2.6.2 Impact excitation

The most well-known excitation method for measuring FRFs is the impact method. Its popularity lies in its simplicity and speed. Impact testing produces responses with high crest factors (ratio of the peak to the rms value). This property can assist in nonlinearity being excited and hence observed in the FRFs and their corresponding coherence functions, usually producing distortions in the FRFs opposite to those obtained from sinusoidal excitation. The use of impact testing methods however, suffers from the same problems as those of random excitation, namely that the input is a broad spectrum and the energy associated with an individual frequency is small, thus it is much more difficult to excite structural nonlinearity. Impact is a form of transient excitation.

The FRF in figure 2.17(*b*) was obtained by applying the force as a very short impact (a pulse). In practice pulses or impacts of the type chosen are often obtained by using an instrumented hammer to excite the structure. This makes the method extremely attractive for *in situ* testing. The FRF is obtained

by dividing the Fourier transform of the response by the Fourier transform of the force. Averaging is usually carried out and this means that a coherence function can be estimated. The pulse used here was selected so that the maximum value of the response in the time domain was similar to the resonant amplitude from the sine-wave test of the last section. The results in figure 2.17(*b*) confirm the earlier remarks in that a completely different FRF is obtained to that using sine excitation.

2.6.3 Chirp excitation

A second form of transient excitation commonly used for measuring FRFs is chirp excitation. This form of excitation can be effective in detecting nonlinearity and combines the attraction of being relatively fast with an equal level of input power across a defined frequency range. Chirp excitation can be linear or nonlinear where the nonlinear chirp signal can be designed to have a specific input power spectrum that can vary within a given frequency range [265]. The simplest form of chirp has a linear sweep characteristic so the signal takes the form

$$x(t) = X \sin(\alpha t + \beta t^2) \tag{2.52}$$

where α and β are chosen to give appropriate start and end frequencies. At any given time, the instantaneous frequency of the signal is

$$\omega(t) = \frac{\mathrm{d}}{\mathrm{d}t}(\alpha t + \beta t^2) = \alpha + 2\beta t. \tag{2.53}$$

As one might imagine, the response of a nonlinear system to such a comparatively complex input may be quite complicated. The FRF in figure 2.17(*c*) was obtained using a force consisting of a frequency sweep between 0 and 2 Hz in 50 s. (This sweep is rapid compared with the decay time of the structure.) The FRF was once again determined from the ratio of the Fourier transforms. The excitation level was selected so that the maximum displacement in the time-domain was the same as before. The 'split' response in figure 2.17(*c*) is due to the presence of the nonlinearity.

2.6.4 Random excitation

The FRF of a nonlinear structure obtained from random (usually band-limited) excitation often appears undistorted due to the randomness of the amplitude and phase of the excitation signal creating a 'linearized' or 'averaged' FRF.

Due to this linearization, the only way in which random excitation can assist in detecting nonlinearity is for several tests to be carried out at different rms levels of the input excitation (auto-spectrum of the input) and the resulting FRFs overlaid to test for homogeneity. A word of warning here. Since the total power in the input spectrum is spread over the band-limited frequency range used, the ability to excite nonlinearities is significantly reduced compared with sinusoidal

excitation. In fact, experience has shown that it is often difficult to drive structures into their nonlinear regimes with random excitation unless narrower-band signals are used. This effect is also compounded by the fact that if an electrodynamic exciter is being used to generate the FRFs in an open-loop configuration (no feedback control for the force input) the force spectrum will suffer from force drop-out in the resonant regions. This makes it even more difficult to drive a structure into its nonlinear regimes and the measured FRFs corresponding to different input spectrum levels may not show a marked difference. However, the speed at which FRFs can be measured with random excitation and the combined use of the coherence function makes random excitation a useful tool in many practical situations for detecting nonlinearity.

Note that pseudo-random excitation is not recommended for use in nonlinearity detection via FRF measurements. Pseudo-random excitation is periodic and contains harmonically related discrete frequency components. These discrete components can be converted (via the nonlinearity) into frequencies which coincide with the harmonics in the input frequency. These will not average out due to their periodic nature and hence the coherence function may appear acceptable (close to unity) even though the FRF looks very 'noisy'.

The FRF in figure 2.17(*d*) was obtained by using a random force and determining spectral density functions associated with the force and response. These were then used to estimate the FRF using

$$H(\omega) = \frac{S_{yx}(\omega)}{S_{xx}(\omega)}. \tag{2.54}$$

2.6.5 Conclusions

These examples have been chosen to demonstrate how different answers can be obtained from the same nonlinear model when the input excitation is changed. It is interesting to note that the only FRF which one would recognize as 'linear' in terms of its shape is the one shown in figure 2.17(*d*), due to a random excitation input. This is because random excitation introduces a form of 'linearization' as discussed in later chapters. As opposed to linear systems, the importance of the type of excitation employed in numerical simulation or practical testing of nonlinear systems has been demonstrated. Many of the detection and parameter extraction methods for nonlinear systems, described later in this book, are dependent upon the type of input used and will only provide reliable answers under the correct excitation conditions.

2.7 FRF estimators

In the section on coherence, a linear system subject to measurement noise on the output was studied. It was shown that the coherence dips below unity if such noise is present. This is unfortunately not the only consequence of noise. The object of

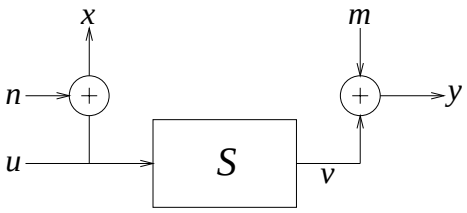


Figure 2.18. Block diagram of a linear system with input and output measurement noise.

the current section is to show that noise also leads to erroneous or *biased* estimates of the FRF when random excitation is used via equation (2.54).

This time a general system will be assumed which has noise on both input and output (figure 2.18). The (unknown) clean input is denoted $u(t)$ and after the addition of (unknown) noise $n(t)$, gives the measured input $x(t)$. Similarly, the unknown clean output $v(t)$ is corrupted by noise $m(t)$ to give the measured output $y(t)$. It is assumed that $m(t)$, $n(t)$ and $x(t)$ are pairwise uncorrelated. The basic equations in the frequency domain are

$$X(\omega) = U(\omega) + N(\omega) \tag{2.55}$$

and

$$Y(\omega) = H(\omega)U(\omega) + M(\omega). \tag{2.56}$$

Multiplying (2.55) by $\overline{X}$ and taking expectations gives

$$S_{xx}(\omega) = S_{uu}(\omega) + S_{nn}(\omega). \tag{2.57}$$

Multiplying (2.56) by $\overline{X}$ and taking expectations gives

$$S_{yx}(\omega) = H(\omega)S_{uu}(\omega) \tag{2.58}$$

as $S_{mx}(\omega) = 0$. Taking the ratio of (2.58) and (2.57) yields

$$\frac{S_{yx}(\omega)}{S_{xx}(\omega)} = \frac{H(\omega)S_{uu}(\omega)}{S_{uu}(\omega) + S_{nn}(\omega)} = \frac{H(\omega)}{\left(1 + \frac{S_{nn}(\omega)}{S_{uu}(\omega)}\right)}. \tag{2.59}$$

This means that the estimator S_{yx}/S_{xx}—denoted $H_1(\omega)$—is only equal to the correct FRF $H(\omega)$ if there is no noise on the input ($S_{nn} = 0$). Further, as $S_{nn}/S_{uu} > 0$, the estimator is always an underestimate, i.e. $H_1(\omega) < H(\omega)$ if input noise is present. Note that the estimator is completely insensitive to noise on the output.

Now, multiply (2.56) by $\overline{Y}$ and take expectations, the result is

$$S_{yy}(\omega) = |H(\omega)|^2 S_{uu}(\omega) + S_{mm}(\omega). \tag{2.60}$$

Multiplying (2.55) by $\overline{Y}$ and averaging yields

$$S_{xy}(\omega) = \overline{H}(\omega)S_{uu}(\omega) \tag{2.61}$$

and taking the ratio of (2.60) and (2.61) gives

$$\frac{S_{yy}(\omega)}{S_{xy}(\omega)} = H(\omega)\left(1 + \frac{S_{mm}(\omega)}{S_{uu}(\omega)}\right) \tag{2.62}$$

and this means that the estimator S_{yy}/S_{xy}—denoted by $H_2(\omega)$—is only equal to $H(\omega)$ if there is no noise on the output ($S_{mm} = 0$). Also, as $S_{mm}/S_{uu} > 0$, the estimator is always an overestimate, i.e. $H_2(\omega) > H(\omega)$ if output noise is present. The estimator is insensitive to noise on the input.

So if there is noise on the input only, one should always use H_2: if there is noise only on the output, one should use H_1. If there is noise on both signals a compromise is clearly needed. In fact, as H_1 is an underestimate and H_2 is an overestimate, the sensible estimator would be somewhere in between. As one can always interpolate between two numbers by taking the mean, a new estimator H_3 can be defined by taking the geometric mean of H_1 and H_2,

$$H_3(\omega) = \sqrt{H_1(\omega)H_2(\omega)} = H(\omega)\sqrt{\frac{S_{mm}(\omega) + S_{uu}(\omega)}{S_{nn}(\omega) + S_{uu}(\omega)}} \tag{2.63}$$

and this is the estimator of choice if both input and output are corrupted.

Note that a byproduct of this analysis is a general expression for the coherence,

$$\gamma^2(\omega) = \frac{|S_{yx}(\omega)|^2}{S_{yy}(\omega)S_{xx}(\omega)} = \frac{1}{\left(1 + \frac{S_{mm}(\omega)}{S_{vv}(\omega)}\right)\left(1 + \frac{S_{nn}(\omega)}{S_{vv}(\omega)}\right)} \tag{2.64}$$

from which it follows that $\gamma^2 < 1$ if either input or output noise is present. It also follows from (2.64), (2.62) and (2.59) that $\gamma^2 = H_1/H_2$ or

$$H_2(\omega) = \frac{H_1(\omega)}{\gamma^2(\omega)} \tag{2.65}$$

so the three quantities are not independent.

As the effect of nonlinearity on the FRF is different to that of input noise or output noise acting alone, one might suspect that H_3 is the best estimator for use with nonlinear systems. In fact it is shown in [232] that H_3 is the best estimator for nonlinear systems in the sense that, of the three estimators, given an input density S_{xx}, H_3 gives the best estimate of S_{yy} via $S_{yy} = |H|^2 S_{xx}$. This is a useful property if the object of estimating the FRF is to produce an effective linearized model by curve-fitting.

2.8 Equivalent linearization

As observed in the last chapter, modal analysis is an extremely powerful theory of linear systems. It is so effective in that restricted area that one might be tempted to apply the procedures of modal analysis directly to nonlinear systems without modification. In this situation, the curve-fitting algorithms used will associate a linear system with each FRF—in some sense the linear system which explains it best. In the case of a SDOF system, one might find the equivalent linear FRF

$$H_{\mathrm{eq}}(\omega) = \frac{1}{-m_{\mathrm{eq}}\omega^2 + \mathrm{i}c_{\mathrm{eq}}\omega + k_{\mathrm{eq}}} \tag{2.66}$$

which approximates most closely that of the nonlinear system. In the time domain this implies a best linear model of the form

$$m_{\mathrm{eq}}\ddot{y} + c_{\mathrm{eq}}\dot{y} + k_{\mathrm{eq}}y = x(t) \tag{2.67}$$

and such a model is called a *linearization*. As the nonlinear system FRF will usually change its shape as the level of excitation is changed, any linearization is only valid for a given excitation level. Also, because the form of the FRF is a function of the *type* of excitation as discussed in section 2.6, different forcing types of nominally the same amplitude will require different linearizations. These are clear limitations.

In the next chapter, linearizations based on FRFs from harmonic forcing will be derived. In this section, linearizations based on random excitation will be discussed. These are arguably more fundamental because, as discussed in section 2.6, random excitation is the only excitation which generates nonlinear systems FRFs which *look* like linear system FRFs.

2.8.1 Theory

The basic theory presented here does not proceed via the FRFs, one operates directly on the equations of motion. The technique—*equivalent* or more accurately *statistical linearization*—dates back to the fundamental work of Caughey [54]. The following discussion is limited to SDOF systems; however, this is not a fundamental restriction of the method[4].

Given a general SDOF nonlinear system,

$$m\ddot{y} + f(y, \dot{y}) = x(t) \tag{2.68}$$

one seeks an equivalent linear system of the form (2.67). As the excitation is random, an apparently sensible strategy would be to minimize the average difference between the nonlinear force and the linear system (it will be assumed

[4] The following analysis makes rather extensive use of basic probability theory, the reader who is unfamiliar with this can consult appendix A.

that the apparent mass is unchanged, i.e. $m_{\rm eq} = m$), i.e. find the $c_{\rm eq}$ and $k_{\rm eq}$ which minimize

$$J_1(y, c_{\rm eq}, k_{\rm eq}) = E[f(y, \dot{y}) - c_{\rm eq}\dot{y} - k_{\rm eq}y]. \tag{2.69}$$

In fact this is not sensible as the differences will generally be a mixture of negative and positive and could still average to zero for a wildly inappropriate system. The correct strategy is to minimize the expectation of the squared differences, i.e.

$$J_2(y, c_{\rm eq}, k_{\rm eq}) = E[(f(y, \dot{y}) - c_{\rm eq}\dot{y} - k_{\rm eq}y)^2] \tag{2.70}$$

or

$$\begin{aligned} J_2(y, c_{\rm eq}, k_{\rm eq}) = E[&(f(y, \dot{y})^2 + c_{\rm eq}^2\dot{y}^2 + k_{\rm eq}^2y^2 - 2f(y, \dot{y})c_{\rm eq}\dot{y} \\ &- 2f(y, \dot{y})k_{\rm eq}y + 2c_{\rm eq}k_{\rm eq}y\dot{y}]. \end{aligned} \tag{2.71}$$

Now, using elementary calculus, the values of $c_{\rm eq}$ and $k_{\rm eq}$ which minimize (2.71) are those which satisfy the equations

$$\frac{\partial J_2}{\partial c_{\rm eq}} = \frac{\partial J_2}{\partial k_{\rm eq}} = 0. \tag{2.72}$$

The first of these yields

$$E[c_{\rm eq}\dot{y}^2 - \dot{y}f(y, \dot{y}) + k_{\rm eq}y\dot{y}] = c_{\rm eq}E[\dot{y}^2] - E[\dot{y}f(y, \dot{y})] + k_{\rm eq}E[y\dot{y}] = 0 \tag{2.73}$$

and the second

$$E[k_{\rm eq}y^2 - yf(y, \dot{y}) + c_{\rm eq}y\dot{y}] = k_{\rm eq}E[y^2] - E[yf(y, \dot{y})] + c_{\rm eq}E[y\dot{y}] = 0 \tag{2.74}$$

after using the linearity of the expectation operator. Now, it is a basic theorem of stochastic processes $E[y\dot{y}] = 0$ for a wide range of processes[5]. With this assumption, (2.73) and that (2.74) become

$$c_{\rm eq} = \frac{E[\dot{y}f(y, \dot{y})]}{E[\dot{y}^2]} \tag{2.75}$$

and

$$k_{\rm eq} = \frac{E[yf(y, \dot{y})]}{E[y^2]} \tag{2.76}$$

[5] The proof is elementary and depends on the processes being *stationary*, i.e. that the statistical moments of $x(t)$, mean, variance etc do not vary with time. With this assumption

$$\frac{{\rm d}\sigma_y^2}{{\rm d}t} = 0 = \frac{\rm d}{{\rm d}t}E[y^2] = E\left[\frac{{\rm d}y^2}{{\rm d}t}\right] = 2E[y\dot{y}].$$

and all that remains is to evaluate the expectations. Unfortunately this turns out to be non-trivial. The expectation of a function of random variables like $f(y,\dot{y})$ is given by

$$E[f(y,\dot{y})] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{d}y\,\mathrm{d}\dot{y}\, p(y,\dot{y}) f(y,\dot{y}) \tag{2.77}$$

where $p(y,\dot{y})$ is the probability density function (PDF) for the processes y and $\dot{y}$. The problem is that as the PDF of the response is not known for general nonlinear systems, estimating it presents formidable problems of its own. The solution to this problem is to approximate $p(y,\dot{y})$ by $p_{\mathrm{eq}}(y,\dot{y})$—the PDF of the equivalent linear system (2.67); this still requires a little thought. The fact that comes to the rescue is a basic theorem of random vibrations of *linear* systems [76], namely: if the excitation to a linear system is a zero-mean Gaussian signal, then so is the response. To say that $x(t)$ is Gaussian zero-mean is to say that it has the PDF

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma_x}\exp\left(-\frac{x^2}{2\sigma_x^2}\right) \tag{2.78}$$

where σ_x^2 is the variance of the process $x(t)$.

The theorem states that the PDFs of the responses are Gaussian also, so

$$p_{\mathrm{eq}}(y_{\mathrm{eq}}) = \frac{1}{\sqrt{2\pi}\sigma_{y_{\mathrm{eq}}}}\exp\left(-\frac{y_{\mathrm{eq}}^2}{2\sigma_{y_{\mathrm{eq}}}^2}\right) \tag{2.79}$$

and

$$p_{\mathrm{eq}}(\dot{y}_{\mathrm{eq}}) = \frac{1}{\sqrt{2\pi}\sigma_{\dot{y}_{\mathrm{eq}}}}\exp\left(-\frac{\dot{y}_{\mathrm{eq}}^2}{2\sigma_{\dot{y}_{\mathrm{eq}}}^2}\right) \tag{2.80}$$

so the joint PDF is

$$p_{\mathrm{eq}}(y_{\mathrm{eq}},\dot{y}_{\mathrm{eq}}) = p_{\mathrm{eq}}(y_{\mathrm{eq}})p_{\mathrm{eq}}(\dot{y}_{\mathrm{eq}}) = \frac{1}{\sqrt{2\pi}\sigma_{y_{\mathrm{eq}}}\sigma_{\dot{y}_{\mathrm{eq}}}}\exp\left(-\frac{y_{\mathrm{eq}}^2}{2\sigma_{y_{\mathrm{eq}}}^2}-\frac{\dot{y}_{\mathrm{eq}}^2}{2\sigma_{\dot{y}_{\mathrm{eq}}}^2}\right). \tag{2.81}$$

In order to make use of these results it will be assumed from now on that $x(t)$ is zero-mean Gaussian.

Matters can be simplified further by assuming that the nonlinearity is separable, i.e. the equation of motion takes the form

$$m\ddot{y} + c\dot{y} + ky + \phi(\dot{y}) + \psi(y) = x(t) \tag{2.82}$$

in this case, $f(y,\dot{y}) = c\dot{y} + ky + \phi(\dot{y}) + \psi(y)$.

Equation (2.75) becomes

$$c_{\mathrm{eq}} = \frac{E[\dot{y}(c\dot{y} + ky + \phi(\dot{y}) + \psi(y))]}{E[\dot{y}^2]} \tag{2.83}$$

or, using the linearity of E,

$$c_{\mathrm{eq}} = \frac{cE[\dot{y}^2] + kE[\dot{y}y] + E[\dot{y}\phi(\dot{y})] + E[\dot{y}\psi(y))]}{E[\dot{y}^2]} \tag{2.84}$$

which reduces to

$$c_{\mathrm{eq}} = c + \frac{E[\dot{y}\phi(\dot{y})] + E[\dot{y}\psi(y))]}{E[\dot{y}^2]} \tag{2.85}$$

and a similar analysis based on (2.75) gives

$$k_{\mathrm{eq}} = k + \frac{E[y\phi(\dot{y})] + E[y\psi(y))]}{E[y^2]}. \tag{2.86}$$

Now, consider the term $E[y\phi(\dot{y})]$ in (2.86). This is given by

$$E[y\phi(\dot{y})] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{d}y\,\mathrm{d}\dot{y}\, p_{\mathrm{eq}}(y,\dot{y})y\phi(\dot{y}) \tag{2.87}$$

and because the PDF factors, i.e. $p_{\mathrm{eq}}(y_{\mathrm{eq}},\dot{y}_{\mathrm{eq}}) = p_{\mathrm{eq}}(y_{\mathrm{eq}})p_{\mathrm{eq}}(\dot{y}_{\mathrm{eq}})$, so does the integral, hence,

$$E[y\phi(\dot{y})] = \left(\int_{-\infty}^{\infty} \mathrm{d}y\, p_{\mathrm{eq}}(y)y\right)\left(\int_{-\infty}^{\infty} \mathrm{d}\dot{y}\, p_{\mathrm{eq}}(\dot{y})\phi(\dot{y})\right) = E[y]E[\phi(\dot{y})] \tag{2.88}$$

but the response is zero-mean Gaussian and therefore $E[y] = 0$. It follows that $E[y\phi(\dot{y})] = 0$ and therefore (2.86) becomes

$$k_{\mathrm{eq}} = k + \frac{E[y\psi(y))]}{E[y^2]} \tag{2.89}$$

and a similar analysis for (2.85) yields

$$c_{\mathrm{eq}} = c + \frac{E[\dot{y}\phi(\dot{y})]}{E[\dot{y}^2]}. \tag{2.90}$$

Now, assuming that the expectations are taken with respect to the linear system PDFs ((2.79) and (2.80)), equation (2.90) becomes

$$c_{\mathrm{eq}} = c + \frac{1}{\sqrt{2\pi}\sigma_{\dot{y}_{\mathrm{eq}}}^3}\int_{-\infty}^{\infty} \mathrm{d}\dot{y}\,\dot{y}\phi(\dot{y})\exp\left(-\frac{\dot{y}^2}{2\sigma_{\dot{y}_{\mathrm{eq}}}^2}\right) \tag{2.91}$$

and (2.89) becomes

$$k_{\mathrm{eq}} = k + \frac{1}{\sqrt{2\pi}\sigma_{\dot{y}_{\mathrm{eq}}}^3}\int_{-\infty}^{\infty} \mathrm{d}y\, y\psi(y)\exp\left(-\frac{y^2}{2\sigma_{\dot{y}_{\mathrm{eq}}}^2}\right) \tag{2.92}$$

which are the final forms required. Although it may now appear that the problem has been reduced to the evaluation of integrals, unfortunately things are not quite that simple. It remains to estimate the variances in the integrals. Now standard theory (see [198]) gives

$$\sigma_{y_{\mathrm{eq}}}^2 = \int_{-\infty}^{\infty} \mathrm{d}\omega \, |H_{\mathrm{eq}}(\omega)|^2 S_{xx}(\omega) = \int_{-\infty}^{\infty} \mathrm{d}\omega \, \frac{S_{xx}(\omega)}{(k_{\mathrm{eq}} - m\omega^2)^2 + c_{\mathrm{eq}}^2\omega^2} \tag{2.93}$$

and

$$\sigma_{\dot{y}_{\mathrm{eq}}}^2 = \int_{-\infty}^{\infty} \mathrm{d}\omega \, \frac{\omega^2 S_{xx}(\omega)}{(k_{\mathrm{eq}} - m\omega^2)^2 + c_{\mathrm{eq}}^2\omega^2} \tag{2.94}$$

and here lies the problem. Equation (2.92) expresses k_{eq} in terms of the variance $\sigma_{y_{\mathrm{eq}}}^2$ and (2.93) expresses $\sigma_{y_{\mathrm{eq}}}^2$ in terms of k_{eq}. The result is a rather nasty pair of coupled nonlinear algebraic equations which must be solved for k_{eq}. The same is true of c_{eq}. In order to see how progress can be made, it is useful to consider a concrete example.

2.8.2 Application to Duffing's equation

The equation of interest is (2.24), so

$$\psi(y) = k_3 y^3 \tag{2.95}$$

and the expression for the effective stiffness, from (2.92) is

$$k_{\mathrm{eq}} = k + \frac{k_3}{\sqrt{2\pi}\sigma_{y_{\mathrm{eq}}}^3} \int_{-\infty}^{\infty} \mathrm{d}y \, y^4 \exp\left(-\frac{y^2}{2\sigma_{y_{\mathrm{eq}}}^2}\right). \tag{2.96}$$

In order to obtain a tractable expression for the variance from (2.93) it will be assumed that $x(t)$ is a *white* zero-mean Gaussian signal, i.e. $S_{xx}(\omega) = P$ a constant. It is a standard result then that [198]

$$\sigma_{y_{\mathrm{eq}}}^2 = P \int_{-\infty}^{\infty} \mathrm{d}\omega \, \frac{1}{(k_{\mathrm{eq}} - m\omega^2)^2 + c_{\mathrm{eq}}^2\omega^2} = \frac{\pi P}{c k_{\mathrm{eq}}}. \tag{2.97}$$

This gives

$$k_{\mathrm{eq}} = k + \frac{k_3}{\sqrt{2\pi}\left(\frac{\pi P}{c k_{\mathrm{eq}}}\right)^{\frac{3}{2}}} \int_{-\infty}^{\infty} \mathrm{d}y \, y^4 \exp\left(-\frac{c k_{\mathrm{eq}} y^2}{2\pi P}\right). \tag{2.98}$$

Now, making use of the result[6],

$$\int_{-\infty}^{\infty} \mathrm{d}y\, y^4 \exp(-ay^2) = \frac{3\pi^{\frac{1}{2}}}{4a^{\frac{5}{2}}} \tag{2.99}$$

gives

$$k_{\mathrm{eq}} = k + \frac{3\pi k_3 P}{ck_{\mathrm{eq}}} \tag{2.100}$$

and the required k_{eq} satisfies the quadratic equation

$$ck_{\mathrm{eq}}^2 - ckk_{\mathrm{eq}} - 3\pi k_3 P = 0. \tag{2.101}$$

The desired root is (after a little algebra)

$$k_{\mathrm{eq}} = \frac{k}{2} + \frac{k}{2}\sqrt{1 + \frac{12\pi k_3 P}{ck^2}} \tag{2.102}$$

which shows the expected behaviour, i.e. k_{eq} increases if P or k_3 increase. If $k_3 P$ is small, the binomial approximation gives

$$k_{\mathrm{eq}} = k + \frac{3k_3 P}{ck} + \mathrm{O}(k_3^2 P^2). \tag{2.103}$$

[6] Integrals of the type

$$\int_{-\infty}^{\infty} \mathrm{d}y\, y^n \exp(-ay^2)$$

occur fairly often in the equivalent linearization of polynomial nonlinearities. Fortunately, they are fairly straightforward to evaluate. The following trick is used: it is well known that

$$I = \int_{-\infty}^{\infty} \mathrm{d}y\, \exp(-ay^2) = \frac{\pi^{\frac{1}{2}}}{a^{\frac{1}{2}}}.$$

Differentiating with respect to the parameter a yields

$$\frac{\mathrm{d}I}{\mathrm{d}a} = -\int_{-\infty}^{\infty} \mathrm{d}y\, y^2 \exp(-ay^2) = -\frac{\pi^{\frac{1}{2}}}{2a^{\frac{3}{2}}}$$

and differentiating again, gives the result in (2.99)

$$\frac{\mathrm{d}^2 I}{\mathrm{d}a^2} = \int_{-\infty}^{\infty} \mathrm{d}y\, y^4 \exp(-ay^2) = \frac{3\pi^{\frac{1}{2}}}{4a^{\frac{5}{2}}}.$$

Continuing this operation will give results for all integrals with n even. If n is odd, the sequence is started with

$$I = \int_{-\infty}^{\infty} \mathrm{d}y\, y \exp(-ay^2)$$

but this is the integral of an odd function from $-\infty$ to ∞ and it therefore vanishes. This means the integrals for all odd n vanish.

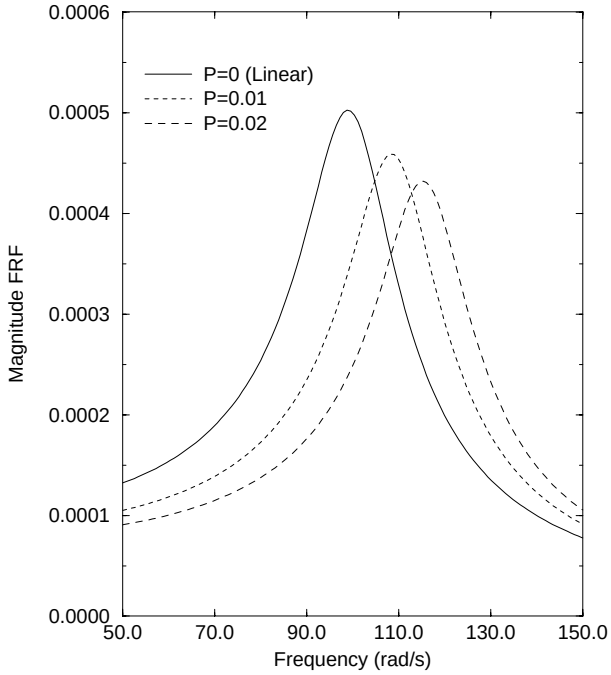


Figure 2.19. Linearized FRF of a Duffing oscillator for different levels of excitation.

To illustrate (2.102), the parameters $m = 1$, $c = 20$, $k = 10^4$ and $k_3 = 5 \times 10^9$ were chosen for the Duffing oscillator. Figure 2.19 shows the linear FRF with $k_{\rm eq}$ given by (2.102) with P = 0, 0.01 and 0.02. The values of $k_{\rm eq}$ found are respectively 10 000.0, 11 968.6 and 13 492.5, giving natural frequencies of $\omega_{\rm n} = 100.0$, 109.4 and 116.2.

In order to validate this result, the linearized FRF for $P = 0.02$ is compared to the FRF estimated from the full nonlinear system in figure 2.20. The agreement is good, the underestimate of the FRF from the simulation is probably due to the fact that the H_1 estimator was used (see section 2.7).

2.8.3 Experimental approach

The problem with using (2.75) and (2.76) as the basis for an experimental method is that they require one to know what $f(y, \dot{y})$ is. In practice it will be useful to extract a linear model without knowing the details of the nonlinearity. Hagedorn and Wallaschek [127, 262] have developed an effective experimental procedure for doing precisely this.

Suppose the linear system (2.67) (with $m_{\rm eq} = m$) is assumed for the

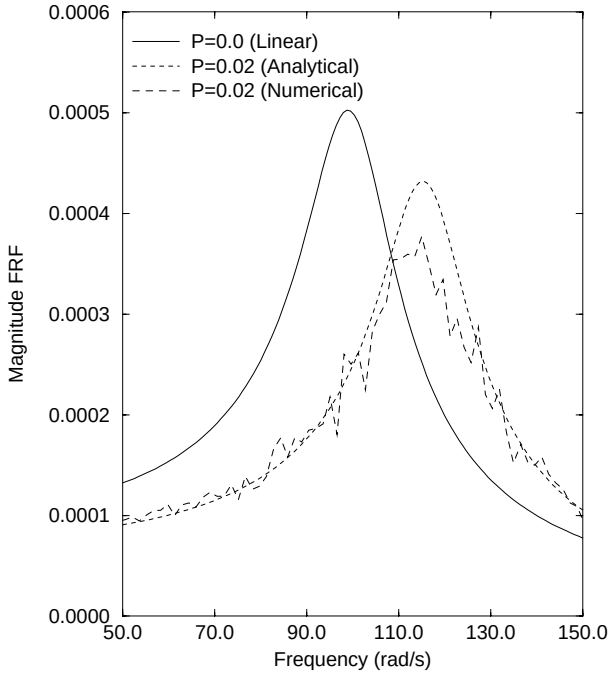


Figure 2.20. Comparison between the nonlinear system FRF and the theoretical FRF for the linearized system.

experimental system. Multiplying (2.67) by $\dot{y}$ and taking expectations yields

$$mE[\dot{y}\ddot{y}] + c_{\mathrm{eq}}E[\dot{y}^2] + k_{\mathrm{eq}}E[\dot{y}y] = E[x\dot{y}]. \tag{2.104}$$

Stationarity implies that $E[y\dot{y}] = E[\dot{y}\ddot{y}] = 0$, so

$$c_{\mathrm{eq}} = \frac{E[x\dot{y}]}{E[\dot{y}^2]}. \tag{2.105}$$

(All processes are assumed zero-mean, the modification if they are not is fairly trivial.) Similarly, multiply (2.67) by y and take expectations

$$mE[y\ddot{y}] + c_{\mathrm{eq}}E[y\dot{y}] + k_{\mathrm{eq}}E[y^2] = E[xy]. \tag{2.106}$$

Now using stationarity and $E[y\ddot{y}] = -E[\dot{y}^2]$ which follows from

$$\frac{\mathrm{d}}{\mathrm{d}t}E[y\dot{y}] = 0 = E[\dot{y}^2] + E[y\ddot{y}] \tag{2.107}$$

yields

$$k_{\mathrm{eq}} = \frac{E[xy] + E[\dot{y}^2]}{E[y^2]} \tag{2.108}$$

and it follows that the equivalent stiffnesses and dampings can be obtained experimentally if the signals $x(t)$, $y(t)$ and $\dot{y}(t)$ are measured. In fact, the experimental approach to linearization is superior in the sense that the equivalent damping and stiffness are unbiased. The theoretical procedure yields biased values simply because the statistics of the linearized process are used in the calculation in place of the true statistics of the nonlinear process.

This analysis concludes the chapter, rather neatly reversing the title by going from nonlinear to linear.

Chapter 3

FRFs of nonlinear systems

3.1 Introduction

In the field of structural dynamics, probably the most widely-used method of visualizing the input–output properties of a system is to construct the frequency response function or FRF. So ubiquitous is the technique that it is usually the first step in any vibration test and almost all commercially available spectrum analysers provide FRF functionality. The FRF summarizes most of the information necessary to specify the dynamics of a structure: resonances, anti-resonances, modal density and phase are directly visible. If FRFs are available for a number of response points, the system modeshapes can also be constructed. In addition, the FRF can rapidly provide an indication of whether a system is linear or nonlinear; one simply constructs the FRFs for a number of different excitation levels and searches for changes in the frequency or magnitude of the resonant peaks. Alternatively, in numerical simulations, the FRFs are invaluable for benchmarking algorithms, structural modification studies and updating numerical models.

This chapter describes how FRFs are defined and constructed for nonlinear systems. The interpretation of the FRFs is discussed and it is shown that they provide a representation of the system as it is linearized about a particular operating point. FRF distortions are used to provide information about nonlinearity.

3.2 Harmonic balance

The purpose of applied mathematics is to describe and elucidate experiment. Theoretical analysis should yield information in a form which is readily comparable with observation. The method of harmonic balance conforms to this principle beautifully as a means of approximating the FRFs of nonlinear systems. Recall the definition of an FRF for a linear system from chapter 1. If a signal

$X\sin(\omega t)$ is input to a system and results in a response $Y\sin(\omega t+\phi)$, the FRF is

$$H(\omega) = \left|\frac{Y}{X}(\omega)\right| \mathrm{e}^{\mathrm{i}\phi(\omega)}. \tag{3.1}$$

This quantity is very straightforward to obtain experimentally. Over a range of frequencies $[\omega_{\min}, \omega_{\max}]$ at a fixed frequency increment $\Delta\omega$, sinusoids $X\sin(\omega t)$ are injected sequentially into the system of interest. At each frequency, the time histories of the input and response signals are recorded after transients have died out, and Fourier transformed. The ratio of the (complex) response spectrum to the input spectrum yields the FRF value at the frequency of interest. In the case of a linear system, the response to a sinusoid is always a sinusoid at the same frequency and the FRF in equation (3.1) summarizes the input/output process in its entirety, and does not depend on the amplitude of excitation X. In such a situation, the FRF will be referred to as *pure*.

In the case of a nonlinear system, it will be shown that sinusoidal forcing results in response components at frequencies other than the excitation frequency. In particular, the distribution of energy amongst these frequencies depends on the level of excitation X, so the measurement process described earlier will also lead to a quantity which depends on X. However, because the process is simple, it is often carried out experimentally in an unadulterated fashion for nonlinear systems. The FRF resulting from such a test will be referred to as *composite*[1], and denoted by $\Lambda_{\mathrm{s}}(\omega)$ (the subscript s referring to sine excitation). $\Lambda_{\mathrm{s}}(\omega)$ is often called a *describing function*, particularly in the literature relating to control engineering [259]. The form of the composite FRF also depends on the type of excitation used as discussed in the last chapter. If white noise of constant power spectral density P is used and the FRF is obtained by taking the ratio of the cross- and auto-spectral densities,

$$\Lambda_{\mathrm{r}}(\omega, P) = \frac{S_{yx}(\omega)}{S_{xx}(\omega)} = \frac{S_{yx}(\omega)}{P}. \tag{3.2}$$

The function $\Lambda_{\mathrm{r}}(\omega, P)$ is distinct from the $\Lambda_{\mathrm{s}}(\omega, X)$ obtained from a stepped-sine test. However, for linear systems the forms (3.1) and (3.2) coincide. In all the following discussions, the Λ subscripts will be suppressed when the excitation type is clear from the context.

The analytical analogue of the stepped-sine test is the *method of harmonic balance*. It is only one of a number of basic techniques for approximating the response of nonlinear systems. However, it is presented here in some detail as it provides arguably the neatest means of deriving the FRF.

The system considered here is the most commonly referenced nonlinear system, Duffing's equation,

$$m\ddot{y} + c\dot{y} + ky + k_2y^2 + k_3y^3 = x(t) \tag{3.3}$$

[1] For reasons which will become clear when the Volterra series is discussed in chapter 8.

which represents a low-order Taylor approximation to systems with a more general stiffness nonlinearity,

$$m\ddot{y} + c\dot{y} + ky + f_s(y) = x(t) \tag{3.4}$$

where $f_s(y)$ is an odd function, i.e. $f_s(y) = -f_s(-y)$ with the restoring force always directed towards the origin and with magnitude independent of the sign of the displacement. For such a system, the low-order approximation (3.3) will have $k_2 = 0$.

The Duffing equations with $k_2 = 0$ will be referred to throughout as a *symmetric Duffing*[2] *oscillator.*. If $k_2 \neq 0$, the system (3.3) will be called *asymmetric*. As discussed in the previous chapter, the Duffing oscillator is widely regarded as a benchtest for any method of analysis or system identification and as such will appear regularly throughout this book.

Harmonic balance mimics the spectrum analyser in simply assuming that the response to a sinusoidal excitation is a sinusoid at the same frequency. A trial solution $y = Y\sin(\omega t)$ is substituted in the equation of motion; in the case of the symmetric Duffing oscillator,

$$m\ddot{y} + c\dot{y} + ky + k_3y^3 = X\sin(\omega t - \phi). \tag{3.5}$$

(To simplify matters, k_2 has been zeroed, and the phase has been transferred onto the input to allow Y to be taken as real.) The substitution yields

$$\begin{aligned} -m\omega^2 Y\sin(\omega t) + c\omega Y\cos(\omega t) + kY\sin(\omega t) + k_3Y^3\sin^3(\omega t) \\ = X\sin(\omega t - \phi) \end{aligned} \tag{3.6}$$

and after a little elementary trigonometry this becomes

$$\begin{aligned} -m\omega^2 Y\sin(\omega t) + c\omega Y\cos(\omega t) + kY\sin(\omega t) \\ + k_3Y^3\{\tfrac{3}{4}\sin(\omega t) - \tfrac{1}{4}\sin(3\omega t)\} \\ = X\sin(\omega t)\cos\phi - X\cos(\omega t)\sin\phi. \end{aligned} \tag{3.7}$$

Equating the coefficients of $\sin(\omega t)$ and $\cos(\omega t)$ (the *fundamental* components) yields the equations

$$(-m\omega^2 Y + kY + \tfrac{3}{4}k_3Y^3) = X\cos\phi \tag{3.8}$$

$$c\omega Y = -X\sin\phi. \tag{3.9}$$

Squaring and adding these equations yields

$$X^2 = Y^2[\{-m\omega^2 + k + \tfrac{3}{4}k_3Y^2\}^2 + c^2\omega^2] \tag{3.10}$$

which gives an expression for the gain or modulus of the system,

$$\left|\frac{Y}{X}\right| = \frac{1}{[\{-m\omega^2 + k + \frac{3}{4}k_3Y^2\}^2 + c^2\omega^2]^{\frac{1}{2}}}. \tag{3.11}$$

[2] Strictly speaking, this should be an anti-symmetric oscillator.

The phase is obtained from the ratio of (3.8) and (3.9).

$$\phi = \tan^{-1} \frac{-c\omega}{-m\omega^2 + k + \frac{3}{4}k_3 Y^2}. \tag{3.12}$$

These can be combined into the complex composite FRF,

$$\Lambda(\omega) = \frac{1}{k + \frac{3}{4}k_3 Y^2 - m\omega^2 + \mathrm{i}c\omega}. \tag{3.13}$$

One can regard this as the FRF of a linearized system,

$$m\ddot{y} + c\dot{y} + k_{\mathrm{eq}} y = X \sin(\omega t - \phi) \tag{3.14}$$

where the *effective* or *equivalent stiffness* is amplitude dependent,

$$k_{\mathrm{eq}} = k + \tfrac{3}{4}k_3 Y^2. \tag{3.15}$$

Now, at a fixed level of excitation, the FRF has a natural frequency

$$\omega_n = \sqrt{\frac{k + \frac{3}{4}k_3 Y^2}{m}} \tag{3.16}$$

which depends on Y and hence, indirectly on X. If $k_3 > 0$, the natural frequency increases with X; such a system is referred to as *hardening*. If $k_3 < 0$ the system is *softening*; the natural frequency decreases with increasing X. Note that the expression (3.16) is in terms of Y rather than X, this leads to a sublety which has so far been ignored. Although the apparent resonant frequency changes with X in the manner previously described, the *form* of the FRF is not that of a linear system. For given X and ω, the displacement response Y is obtained by solving the cubic equation (3.10). (This expression is essentially cubic in Y as one can disregard negative amplitude solutions.) As complex roots occur in conjugate pairs, (3.10) will either have one or three real solutions—the complex solutions are disregarded as unphysical. At low levels of excitation, the FRF is a barely distorted version of that for the underlying linear system as the k term will dominate for $Y \ll 1$. A unique response amplitude (a single real root of (3.10)) is obtained for all ω. As X increases, the FRF becomes more distorted, i.e. departs from the linear form, but a unique response is still obtained for all ω. This continues until X reaches a critical value X_{crit} where the FRF has a vertical tangent. Beyond this point a range of ω values, $[\omega_{\mathrm{low}}, \omega_{\mathrm{high}}]$, is obtained over which there are three real solutions for the response. This is an example of a *bifurcation point* of the parameter X; although X varies continuously, the number and stability types of the solutions changes abruptly. As the test or simulation steps past the point ω_{low}, two new responses become possible and persist until ω_{high} is reached and two solutions disappear. The plot of the response looks like figure 3.1. In the interval $[\omega_{\mathrm{low}}, \omega_{\mathrm{high}}]$, the solutions $Y^{(1)}$, $Y^{(2)}$ and $Y^{(3)}$ are possible with $Y^{(1)} > Y^{(2)} > Y^{(3)}$. However,

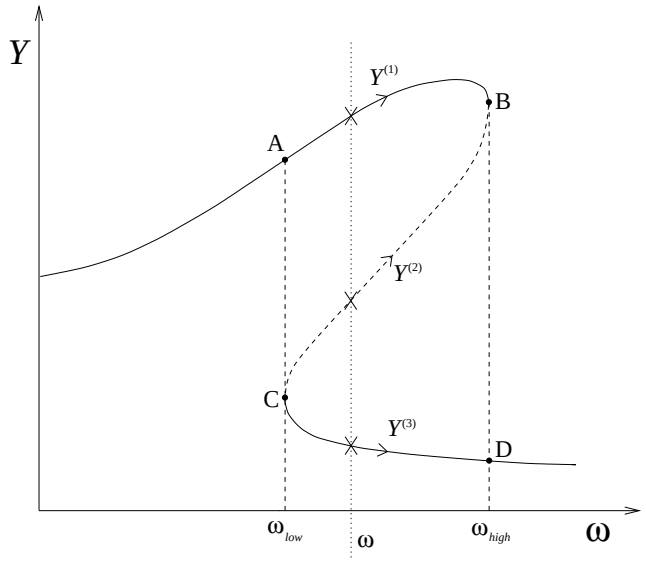


Figure 3.1. Displacement response of a hardening Duffing oscillator for a stepped-sine input. The bifurcation points are clearly seen at B and C.

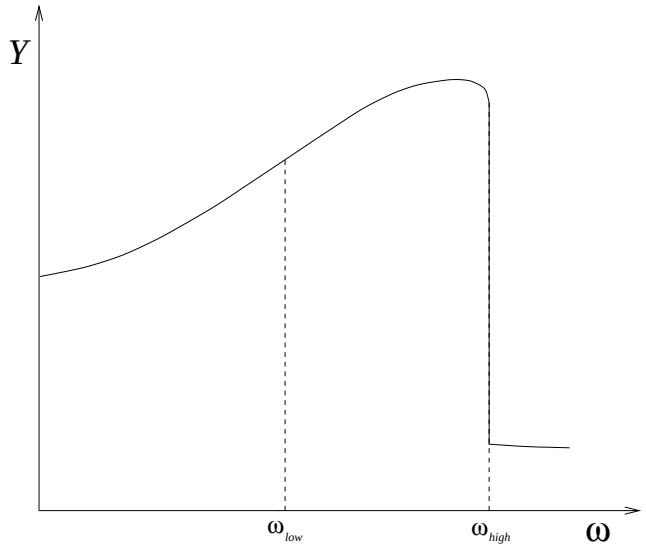


Figure 3.2. Displacement response for hardening Duffing oscillator as the excitation steps up from a low to a high frequency.

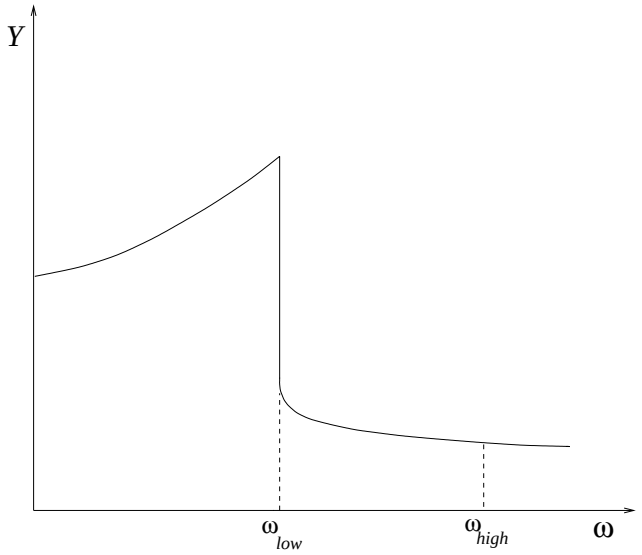


Figure 3.3. Displacement response for hardening Duffing oscillator as the excitation steps down from a high to a low frequency.

it can be shown that the solution $Y^{(2)}$ is unstable and will therefore never be observed in practice.

The corresponding experimental situation occurs in a stepped-sine or sine-dwell test. Consider an upward sweep. A unique response exists up to $\omega = \omega_{\mathrm{low}}$. However, beyond this point, the response stays on branch $Y^{(1)}$ essentially by continuity. This persists until, at frequency ω_{high}, $Y^{(1)}$ ceases to exist and the only solution is $Y^{(3)}$, a jump to this solution occurs giving a discontinuity in the FRF. Beyond ω_{high} the solution stays on the continuation of $Y^{(3)}$ which is the unique solution in this range. The type of FRF obtained from such a test is shown in figure 3.2.

The downward sweep is very similar. When $\omega > \omega_{\mathrm{high}}$, a unique response is obtained. In the multi-valued region, branch $Y^{(3)}$ is obtained by continuity and this persists until ω_{low} when it ceases to exist and the response jumps to $Y^{(1)}$ and thereafter remains on the continuation of that branch (figure 3.3).

If $k_3 > 0$, the resonance peak moves to higher frequencies and the jumps occur on the right-hand side of the peak as described earlier. If $k_3 < 0$, the jumps occur on the left of the peak and the resonance shifts downward in frequency. These discontinuities are frequently observed in experimental FRFs when high levels of excitation are used.

As expected, discontinuities also occur in the phase ϕ, which has the multi-valued form shown in figure 3.4(*a*). The profiles of the phase for upward and downward sweeps are given in figures 3.4(*b*) and (*c*).

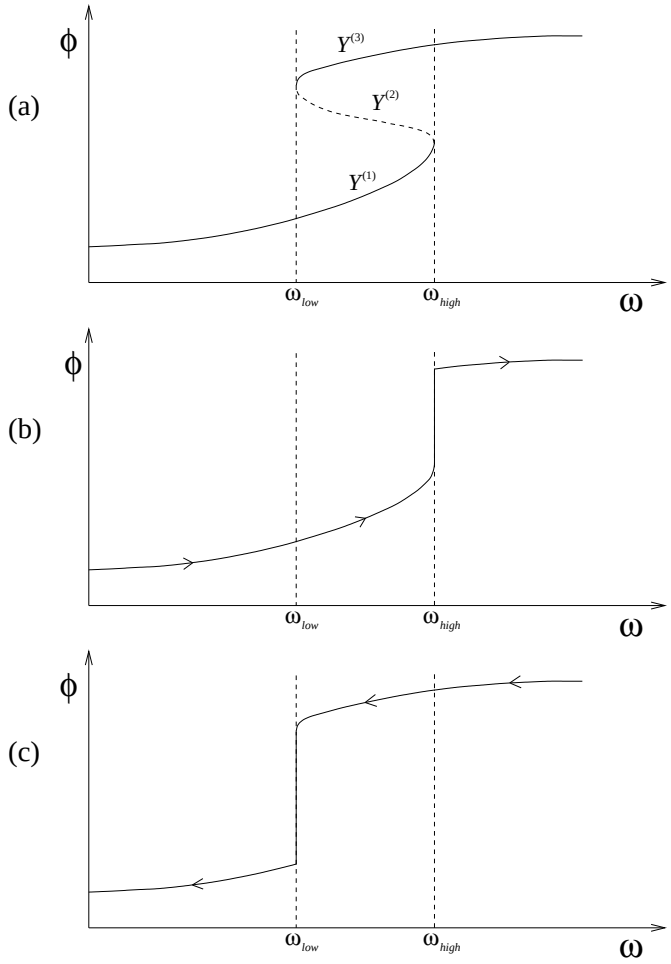


Figure 3.4. Phase characteristics of stepped-sine FRF of hardening Duffing oscillator as shown in figures 3.1–3.3.

It is a straightforward matter to calculate the position of the discontinuities; however, as it would cause a digression here, it is discussed in appendix B.

Before continuing with the approximation of FRFs within the harmonic balance method it is important to recognize that nonlinear systems *do not* respond to a monoharmonic signal with a monoharmonic at the same frequency. The next two sections discuss how departures from this condition arise.

3.3 Harmonic generation in nonlinear systems

The more observant readers will have noticed that the harmonic balance described in section 3.2 is not the whole story. Equation (3.6) is *not* solved by equating coefficients of the fundamental components; a term $\frac{1}{4}k_3Y^2\sin(3\omega t)$ is not balanced. Setting it equal to zero leads to the conclusion that k_3 or Y is zero, which is clearly unsatisfactory. The reason is that $y(t) = Y\sin(\omega t)$ is an unnacceptable solution to equation (3.3). Things are much more complicated for nonlinear systems. An immediate fix is to add a term proportional to $\sin(3\omega t)$ to the trial solution yielding

$$y(t) = Y_1\sin(\omega t + \phi_1) + Y_3\sin(3\omega t + \phi_3) \tag{3.17}$$

(with the phases explicitly represented). This is substituted in the phase-adjusted version of (3.5)

$$m\ddot{y} + c\dot{y} + ky + k_3y^3 = X\sin(\omega t) \tag{3.18}$$

and projecting out the coefficients of $\sin(\omega t)$, $\cos(\omega t)$, $\sin(3\omega t)$ and $\cos(3\omega t)$ leads to the system of equations

$$\begin{aligned}
&-m\omega^2Y_1\cos\phi_1 - c\omega Y_1\sin\phi_1 + kY_1\cos\phi_1 \\
&\quad + \tfrac{3}{4}k_3Y_1^3\cos\phi_1 + \tfrac{3}{2}k_3Y_1Y_3^2\cos\phi_1 - \tfrac{3}{4}k_3Y_1^2y_3\cos\phi_3\cos 2\phi_1 = X
\end{aligned} \tag{3.19}$$

$$\begin{aligned}
&-m\omega^2Y_1\sin\phi_1 - c\omega Y_1\cos\phi_1 + kY_1\sin\phi_1 \\
&\quad + \tfrac{3}{4}k_3Y_1^3\sin\phi_1 + \tfrac{3}{2}k_3Y_1Y_3^2\sin\phi_1 - \tfrac{3}{4}k_3Y_1^2y_3\sin\phi_3\cos 2\phi_1 = 0
\end{aligned} \tag{3.20}$$

$$\begin{aligned}
&-9m\omega^2Y_3\cos\phi_3 - 3c\omega Y_3\sin\phi_3 + kY_3\cos\phi_3 - \tfrac{1}{4}k_3Y_1^3\cos^3\phi_1 \\
&\quad + \tfrac{3}{4}k_3Y_3^3\cos\phi_3 - \tfrac{3}{4}k_3Y_1^3\cos\phi_1\sin^2\phi_1 + \tfrac{3}{2}k_3Y_1^2Y_3\cos\phi_3 = 0
\end{aligned} \tag{3.21}$$

$$\begin{aligned}
&-9m\omega^2Y_3\sin\phi_3 + 3c\omega Y_3\cos\phi_3 + kY_3\sin\phi_3 + \tfrac{1}{4}k_3Y_1^3\sin^3\phi_1 \\
&\quad + \tfrac{3}{4}k_3Y_3^3\sin\phi_3 - \tfrac{3}{4}k_3Y_1^3\cos^2\phi_1\sin\phi_1 + \tfrac{3}{2}k_3Y_1^2Y_3\sin\phi_3 = 0.
\end{aligned} \tag{3.22}$$

Solving this system of equations gives a better approximation to the FRF. However, the cubic term generates terms with $\sin^3(\omega t)$, $\sin^2(\omega t)\sin(3\omega t)$, $\sin(\omega t)\sin^2(3\omega t)$ and $\sin^3(3\omega t)$ which decompose to give harmonics at $5\omega t$, $7\omega t$ and $9\omega t$. Equating coefficients up to third-order leaves these components uncancelled. In order to deal with them properly, a trial solution of the form

$$\begin{aligned}
y(t) = {} & Y_1\sin(\omega t + \phi_1) + Y_3\sin(3\omega t + \phi_3) + Y_5\sin(5\omega t + \phi_5) \\
& + Y_7\sin(7\omega t + \phi_7) + Y_9\sin(9\omega t + \phi_9)
\end{aligned} \tag{3.23}$$

is required, but this in turn will generate higher-order harmonics and one is led to the conclusion that the only way to obtain consistency is to include *all* odd

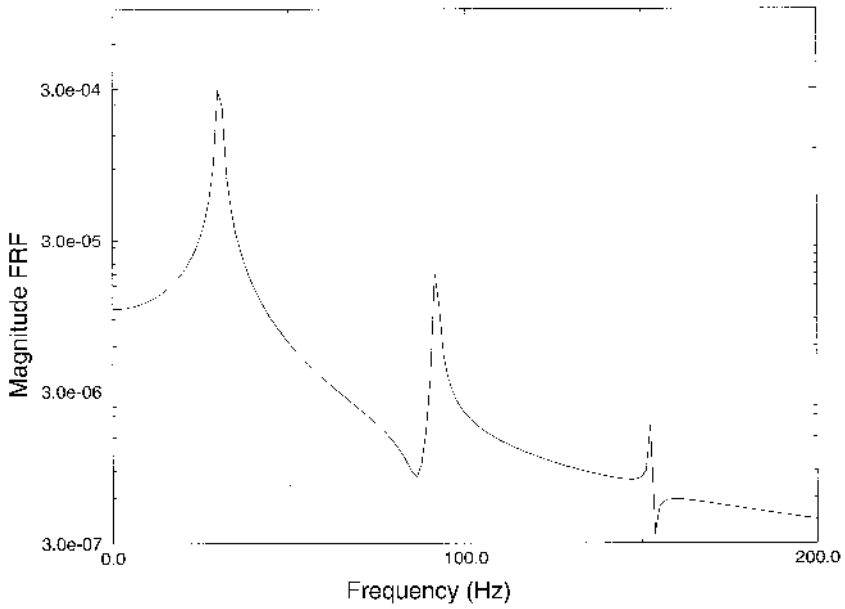


Figure 3.5. Pattern of the harmonics in the response of the hardening Duffing oscillator for a fixed-frequency input.

harmonics in the trial solution, so

$$y(t) = \sum_{i=1}^{\infty} Y_{2i+1} \sin([2i+1]\omega t + \phi_{2i+1}) \tag{3.24}$$

is the necessary expression. This explains the appearance of harmonic components in nonlinear systems as described in chapter 2. The fact that only odd harmonics are present is a consequence of the stiffness function $ky + k_3 y^3$, being odd. If the function were even or generic, all harmonics would be present; consider the system

$$m\ddot{y} + c\dot{y} + ky + k_2 y^2 = X \sin(\omega t - \phi) \tag{3.25}$$

and assume a sinusoidal trial solution $y(t) = Y \sin(\omega t)$. Substituting this in (3.22) generates a term $Y^2 \sin^2(\omega t)$ which decomposes to give $\frac{1}{2}Y^2 - \frac{1}{2}Y^2 \cos(2\omega t)$, so d.c., i.e. a constant (zero frequency) term, and the second harmonic appear. This requires an amendment to the trial solution as before, so $y(t) = Y_0 + Y_1 \sin(\omega t) + Y_2 \sin(2\omega t)$ (neglecting phases). It is clear that iterating this procedure will ultimately generate all harmonics and also a d.c. term.

Figure 3.5 shows the pattern of harmonics in the response of the system

$$\ddot{y} + 20\dot{y} + 10^4 y + 5 \times 10^9 y^3 = 4 \sin(30t). \tag{3.26}$$

(Note the log scale.)

The relative size of the harmonics can be determined analytically by probing the equation of motion with an appropriately high-order trial solution. This results in a horrendous set of coupled nonlinear equations. A much more direct route to the information will be available when the Volterra series is covered in later chapters.

3.4 Sum and difference frequencies

It has been shown earlier that nonlinear systems can respond at multiples of the forcing frequency if the excitation is a pure sinusoid. The situation becomes more complex if the excitation is not a pure tone. Consider equation (3.3) (with $k_3 = 0$ for simplicity) if the forcing function is a sum of two sinusoids or a *two-tone* signal

$$x(t) = X_1 \sin(\omega_1 t) + X_2 \sin(\omega_2 t) \tag{3.27}$$

then the trial solution must at least have the form

$$y(t) = Y_1 \sin(\omega_1 t) + Y_2 \sin(\omega_2 t) \tag{3.28}$$

with Y_1 and Y_2 complex to encode phase. The nonlinear stiffness gives a term

$$\begin{aligned} &k_2(Y_1 \sin(\omega_1 t) + Y_2 \sin(\omega_2 t))^2 \\ &\quad = k_2(Y_1^2 \sin^2(\omega_1 t) + 2Y_1 Y_2 \sin(\omega_1 t)\sin(\omega_2 t) + Y_2^2 \sin(\omega_2 t)) \end{aligned} \tag{3.29}$$

which can be decomposed into harmonics using elementary trigonometry, the result is

$$\begin{aligned} &k_2(\tfrac{1}{2}Y_1^2(1 - \cos(2\omega_1 t) + Y_1 Y_2 \cos([\omega_1 - \omega_2]t) - Y_1 Y_2 \cos([\omega_1 + \omega_2]t) \\ &\quad + \tfrac{1}{2}Y_2^2(1 - \cos(2\omega_2 t)). \end{aligned} \tag{3.30}$$

This means that balancing the coefficients of sines and cosines in equation (3.3) requires a trial solution

$$\begin{aligned} y(t) = Y_0 &+ Y_1 \sin(\omega_1 t) + Y_2 \sin(\omega_2 t) + Y_{11}^{+} \sin(2\omega_1 t) + Y_{22}^{+} \sin(2\omega_2 t) \\ &+ Y_{12}^{+} \cos([\omega_1 + \omega_2]t) + Y_{12}^{-} \cos([\omega_1 - \omega_2]t) \end{aligned} \tag{3.31}$$

where $Y_{ij}^{\pm}$ is simply the component of the response at the frequency $\omega_i \pm \omega_j$.

If this is substituted into (3.3), one again begins a sequence of iterations, which ultimately results in a trial solution containing all frequencies

$$\pm p\omega_1 \pm q\omega_2 \tag{3.32}$$

with p and q integers. If this exercise is repeated for the symmetric Duffing oscillator ($k_2 = 0$), the same result is obtained except that p and q are only allowed to sum to odd values. To lowest nonlinear order, this means that the frequencies $3\omega_1$, $2\omega_1 \pm \omega_2$, $\omega_1 \pm 2\omega_2$ and $3\omega_2$ will be present.

The FRF cannot encode information about sum and difference frequencies, it only makes sense for single-input single-tone systems. In later chapters, the Volterra series will allow generalizations of the FRF which describe the response of multi-tone multi-input systems.

This theory provides the first instance of a nonlinear system violating the principle of superposition. If excitations $X_1 \sin(\omega_1 t)$ and $X_2 \sin(\omega_2 t)$ are presented to the asymmetric Duffing oscillator separately, each case results only in multiples of the relevant frequency in the response. If the excitations are presented together, the new response contains novel frequencies of the form (3.32); novel anyway as long as ω_1 is not an integer multiple of ω_2.

3.5 Harmonic balance revisited

The analysis given in section 3.2 is not very systematic. Fortunately, there is a simple formula for the effective stiffness, given the form of the nonlinear restoring force. Consider the equation of motion,

$$m\ddot{y} + c\dot{y} + f_{\mathrm{s}}(y) = x(t). \tag{3.33}$$

What is needed is a means to obtain

$$f_{\mathrm{s}}(y) \simeq k_{\mathrm{eq}} y \tag{3.34}$$

for a given operating condition. If the excitation is a phase-shifted sinusoid, $X \sin(\omega t - \phi)$, substituting the harmonic balance trial solution $Y \sin(\omega t)$ yields the nonlinear form $f_{\mathrm{s}}(Y \sin(\omega t))$. This function can be expanded as a Fourier series:

$$f_{\mathrm{s}}(Y \sin(\omega t)) = a_0 + \sum_{n=1}^{\infty} a_n \cos(n\omega t) + \sum_{n=1}^{\infty} b_n \sin(n\omega t) \tag{3.35}$$

and this is a finite sum if f_{s} is a polynomial. For the purposes of harmonic balance, the only important parts of this expansion are the fundamental terms. Elementary Fourier analysis applies and

$$a_0 = \frac{1}{2\pi} \int_0^{2\pi} \mathrm{d}(\omega t)\, f_{\mathrm{s}}(Y \sin(\omega t)) \tag{3.36}$$

$$a_1 = \frac{1}{\pi} \int_0^{2\pi} \mathrm{d}(\omega t)\, f_{\mathrm{s}}(Y \sin(\omega t)) \cos(\omega t) \tag{3.37}$$

$$b_1 = \frac{1}{\pi} \int_0^{2\pi} \mathrm{d}(\omega t)\, f_{\mathrm{s}}(Y \sin(\omega t)) \sin(\omega t) \tag{3.38}$$

or, in a more convenient notation,

$$a_0 = \frac{1}{2\pi} \int_0^{2\pi} \mathrm{d}\theta\, f_{\mathrm{s}}(Y \sin\theta) \tag{3.39}$$

$$a_1 = \frac{1}{\pi}\int_0^{2\pi} \mathrm{d}\theta\, f_{\mathrm{s}}(Y\sin\theta)\cos\theta \tag{3.40}$$

$$b_1 = \frac{1}{\pi}\int_0^{2\pi} \mathrm{d}\theta\, f_{\mathrm{s}}(Y\sin\theta)\sin\theta. \tag{3.41}$$

It is immediately obvious from (3.39), that the response will always contain a d.c. component if the stiffness function has an even component. In fact if the stiffness function is purely odd, i.e. $f_{\mathrm{s}}(-y) = -f_{\mathrm{s}}(y)$, then $a_0 = a_1 = 0$ follows straightforwardly. Now, considering terms up to the fundamental in this case, equation (3.34) becomes

$$f_{\mathrm{s}}(Y\sin(\omega t)) \simeq b_1\sin(\omega t) = k_{\mathrm{eq}}Y\sin(\omega t) \tag{3.42}$$

which gives

$$k_{\mathrm{eq}} = \frac{b_1}{Y} = \frac{1}{\pi Y}\int_0^{2\pi} \mathrm{d}\theta\, f_{\mathrm{s}}(Y\sin\theta)\sin\theta \tag{3.43}$$

so the FRF takes the form

$$\Lambda(\omega) = \frac{1}{k_{\mathrm{eq}} - m\omega^2 + \mathrm{i}c\omega} \tag{3.44}$$

(combining both amplitude and phase). It is straightforward to check (3.43) and (3.44) for the case of a symmetric Duffing oscillator. The stiffness function is $f_{\mathrm{s}}(y) = ky + k_3y^3$, so substituting in (3.43) yields

$$k_{\mathrm{eq}} = \frac{k}{\pi Y}\int_0^{2\pi} \mathrm{d}\theta\, \sin\theta\sin\theta + \frac{k_3}{\pi Y}\int_0^{2\pi} \mathrm{d}\theta\, Y^3\sin^3\theta\sin\theta. \tag{3.45}$$

The first integral trivially gives the linear part k; the contribution from the nonlinear stiffness is

$$\frac{k_3}{\pi Y}\int_0^{2\pi} \mathrm{d}\theta\, Y^3\sin^4\theta = \frac{k_3Y^2}{\pi}\int_0^{2\pi} \mathrm{d}\theta\, \frac{1}{8}[3-4\cos 2\theta+\cos 4\theta] = \frac{3}{4}k_3Y^2 \tag{3.46}$$

so

$$k_{\mathrm{eq}} = k + \tfrac{3}{4}k_3Y^2 \tag{3.47}$$

in agreement with (3.15).

As described previously, this represents a naive replacement of the nonlinear system with a linear system (3.14). This begs the question: What is the significance of the linear system. This is quite simple to answer and fortunately the answer agrees with intuition.

A measure of how well the linear system represents the nonlinear system is given by the error function

$$E = \lim_{T\to\infty}\frac{1}{T}\int_0^{\mathrm{T}} \mathrm{d}t\,(y(t) - y_{\mathrm{lin}}(t))^2. \tag{3.48}$$

A system which minimizes E is called an *optimal quasi-linearization*. It can be shown [259], that a linear system minimizes E if and only if

$$\phi_{xy}(\tau) = \phi_{xy_{\mathrm{lin}}}(\tau) \tag{3.49}$$

where ϕ is the cross-correlation function

$$\phi_{pq}(\tau) = \lim_{T\to\infty} \frac{1}{T}\int_0^{\mathrm{T}} \mathrm{d}t\, p(t)q(t+\tau). \tag{3.50}$$

(This is quite a remarkable result, no higher-order statistics are needed.)

It is straightforwardly verified that (3.49) is satisfied by the system with harmonic balance relations (3.40) and (3.41), for the particular reference signal used[3]. It suffices to show that if

$$f(t) = a_0 + \sum_{n=1}^{\infty} a_n \cos(n\omega t) + \sum_{n=1}^{\infty} b_n \sin(n\omega t) \tag{3.51}$$

and

$$f_{\mathrm{lin}}(t) = a_1 \cos(\omega t) + b_1 \sin(\omega t) \tag{3.52}$$

then

$$\phi_{xf}(\tau) = \phi_{xf_{\mathrm{lin}}}(\tau) \tag{3.53}$$

with $x(t) = X\sin(\omega t + \phi)$. This means that the linear system predicted by harmonic balance is an optimal quasi-linearization.

The physical content of equation (3.43) is easy to extract. It simply represents the average value of the restoring force over one cycle of excitation, divided by the value of displacement. This gives a mean value of the stiffness experienced by the system over a cycle. For this reason, harmonic balance, to this level of approximation, is sometimes referred to as an *averaging method*. Use of such methods dates back to the work of Krylov and Boguliubov in the first half of the 20th century. So strongly is this approach associated with these pioneers that it is sometimes referred to as the method of Krylov and Boguliubov [155].

3.6 Nonlinear damping

The formulae presented for harmonic balance so far have been restricted to the case of nonlinear stiffness. The method in principle has no restrictions on the form of the nonlinearity and it is a simple matter to extend the theory to nonlinear damping. Consider the system

$$m\ddot{y} + f_{\mathrm{d}}(\dot{y}) + ky = X\sin(\omega t - \phi). \tag{3.54}$$

[3] Note that linearizations exist for *all* types of reference signal, there is no restriction to harmonic signals.

Choosing a trial output $y(t) = Y\sin(\omega t)$ yields a nonlinear function

$$f_{\mathrm{d}}(\omega Y\cos(\omega t)). \tag{3.55}$$

Now, truncating the Fourier expansion at the fundamental as before gives

$$f_{\mathrm{d}}(\omega Y\cos(\omega t)) \simeq a_0 + a_1\cos(\omega t) + b_1\sin(\omega t) \tag{3.56}$$

and further, restricting f_{d} to be an odd function yields, $a_0 = b_1 = 0$ and

$$a_1 = \frac{1}{\pi}\int_0^{2\pi} \mathrm{d}\theta\, f_{\mathrm{d}}(\omega Y\sin\theta)\cos\theta \tag{3.57}$$

Defining the equivalent damping from

$$f_{\mathrm{d}}(\dot{y}) \simeq c_{\mathrm{eq}}\dot{y} \tag{3.58}$$

so

$$f_{\mathrm{d}}(\omega Y\cos(\omega t)) \simeq c_{\mathrm{eq}}\omega Y\cos(\omega t) = a_1\cos(\omega t) \tag{3.59}$$

gives finally

$$c_{\mathrm{eq}} = \frac{a_1}{\omega Y} = \frac{1}{\pi\omega Y}\int_0^{2\pi} \mathrm{d}\theta\, f_{\mathrm{d}}(\omega Y\sin\theta)\cos\theta \tag{3.60}$$

with a corresponding FRF

$$\Lambda(\omega) = \frac{1}{k - m\omega^2 + \mathrm{i}c_{\mathrm{eq}}\omega}. \tag{3.61}$$

An interesting physical example of nonlinear damping is given by

$$f_{\mathrm{d}}(\dot{y}) = c_2\dot{y}|\dot{y}| \tag{3.62}$$

which corresponds to the drag force experienced by bodies moving at high velocities in viscous fluids. The equivalent damping is given by

$$c_{\mathrm{eq}} = \frac{c_2}{\pi\omega Y}\int_0^{2\pi} \mathrm{d}\theta\,\omega Y\cos\theta|\omega Y\cos\theta|\cos\theta = \frac{c_2\omega Y}{\pi}\int_0^{2\pi} \mathrm{d}\theta\,\cos^2\theta|\cos\theta| \tag{3.63}$$

and it is necessary to split the integral to account for the $|\ |$ function, so

$$\begin{aligned} c_{\mathrm{eq}} &= \frac{2c_2\omega Y}{\pi}\int_0^{\frac{\pi}{2}} \mathrm{d}\theta\,\cos^3\theta - \frac{c_2\omega Y}{\pi}\int_{\frac{\pi}{2}}^{\frac{3\pi}{2}} \mathrm{d}\theta\,\cos^3\theta \\ &= \frac{c_2\omega Y}{2\pi}\int_0^{\frac{\pi}{2}} \mathrm{d}\theta\,(\cos 3\theta + 3\cos\theta) - \frac{c_2\omega Y}{4\pi}\int_{\frac{\pi}{2}}^{\frac{3\pi}{2}} \mathrm{d}\theta\,(\cos 3\theta + 3\cos\theta). \end{aligned} \tag{3.64}$$

After a little manipulation, this becomes

$$c_{\mathrm{eq}} = \frac{8c_2\omega Y}{\pi} \tag{3.65}$$

so the FRF for a simple oscillator with this damping is

$$\Lambda(\omega) = \frac{1}{k - m\omega^2 + \mathrm{i}\frac{8c_2\omega Y}{\pi}\omega} \tag{3.66}$$

which appears to be the FRF of an undamped linear system

$$\Lambda(\omega) = \frac{1}{k - m_{\mathrm{eq}}\omega^2} \tag{3.67}$$

with complex mass

$$m_{\mathrm{eq}} = m + \mathrm{i}\frac{8c_2 Y}{\pi}. \tag{3.68}$$

This is an interesting phenomenon and a similar effect is exploited in the definition of *hysteretic damping*. Damping always manifests itself as the imaginary part of the FRF denominator. Depending on the frequency dependence of the term, it can sometimes be absorbed in a redefinition of one of the other parameters. If the damping has no dependence on frequency, a complex stiffness can be defined $k^* = k(i + \mathrm{i}\eta)$ (where η is called the *loss factor*). This is hysteretic damping and it will be discussed in more detail in chapter 5. Polymers and viscoelastic materials have damping with quite complicated frequency dependence [98].

The analysis of systems with mixed nonlinear damping and stiffness presents no new difficulties. In fact in the case where the nonlinearity is *additively separable*, i.e.

$$m\ddot{y} + f_{\mathrm{d}}(\dot{y}) + f_{\mathrm{s}}(y) = X\sin(\omega t - \phi) \tag{3.69}$$

equations (3.43) and (3.60) still apply and the FRF is

$$\Lambda(\omega) = \frac{1}{k_{\mathrm{eq}} - m\omega^2 + \mathrm{i}c_{\mathrm{eq}}\omega}. \tag{3.70}$$

3.7 Two systems of particular interest

In this section, two systems are studied whose analysis by harmonic balance presents interesting subtleties.

3.7.1 Quadratic stiffness

Consider the system specified by the equation of motion

$$m\ddot{y} + c\dot{y} + ky + k_2 y^2 = X\sin(\omega t - \phi). \tag{3.71}$$

If one naively follows the harmonic balance procedure in this case and substitutes the trial solution $y(t) = Y\sin(\omega t)$, one obtains

$$-m\omega^2 Y\sin(\omega t) + c\omega Y\cos(\omega t) + kY\sin(\omega t) + \tfrac{1}{2}k_2Y_1^2 - \tfrac{1}{2}k_2Y_1^2\cos(2\omega t) = X\sin(\omega t - \phi) \tag{3.72}$$

and equating the coefficients of the fundamentals leads to the FRF of the underlying linear system[4]. The problem here is that the trial solution not only requires a higher-harmonic component, it needs a lower-order part—a d.c. term. If the trial solution $y(t) = Y_0 + Y_1\sin(\omega t)$ is adopted, one obtains, after substitution,

$$\begin{aligned} -m\omega^2 Y_1\sin(\omega t) + c\omega Y_1\cos(\omega t) + kY_0 + kY_1\sin(\omega t) \\ + k_2Y_0^2 + 2k_2Y_0Y_1\sin(\omega t) + \tfrac{1}{2}k_2Y_1^2 - \tfrac{1}{2}k_2Y_1^2\cos(2\omega t) \\ = X\sin(\omega t - \phi). \end{aligned} \tag{3.73}$$

Equating coefficients of $\sin$ and $\cos$ yields the FRF

$$\Lambda(\omega) = \frac{1}{k + 2k_2Y_0 - m\omega^2 + \mathrm{i}c\omega} \tag{3.74}$$

so the effective natural frequency is

$$\omega_n = \sqrt{\frac{k + 2k_2Y_0}{m}} \tag{3.75}$$

and a little more effort is needed in order to interpret this.

Consider the potential energy function $V(y)$, corresponding to the stiffness $f_s(y) = ky + k_2y^2$. As the restoring force is given by

$$f_s = \frac{\partial V}{\partial y} \tag{3.76}$$

then

$$V(y) = -\int \mathrm{d}y\, f_s(y) = \tfrac{1}{2}ky^2 + \tfrac{1}{3}k_2y^3. \tag{3.77}$$

Now, if $k_2 > 0$, a function is obtained like that in figure 3.6. Note that if the forcing places the system beyond point A on this curve, the system falls into an infinitely deep potential well, i.e. escapes to $-\infty$. For this reason, the system must be considered unstable except at low amplitudes where the linear term dominates and always returns the system to the stable equilibrium at B. In any case, if the motion remains bounded, less energy is required to maintain negative displacements, so the mean operating point $Y_0 < 0$. This means the product $k_2Y_0 < 0$. Alternatively, if $k_2 < 0$, a potential curve as in figure 3.7,

[4] Throughout this book the *underlying* linear system for a given nonlinear system is that obtained by deleting all nonlinear terms. Note that this system will be independent of the forcing amplitude as distinct from *linearized* systems which will only be defined with respect to a fixed operating level.

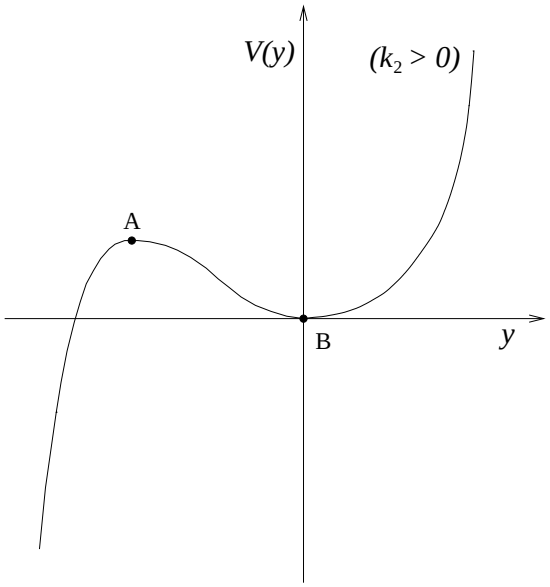


Figure 3.6. Potential energy of the quadratic oscillator with $k_2 > 0$.

arises. The system is again unstable for high enough excitation, with escape this time to ∞. However, in this case, $Y_0 > 0$; so $k_2 Y_0 < 0$ again.

This result indicates that the effective natural frequency for this system (given in (3.75)) always decreases with increasing excitation, i.e. the system is softening, independently of the sign of k_2. This is in contrast to the situation for cubic systems.

Although one cannot infer jumps from the FRF at this level of approximation, they are found to occur, always below the linear natural frequency as shown in figure 3.8 which is computed from a simulation—the numerical equivalent of a stepped-sine test. The equation of motion for the simulation was (3.71) with parameter values $m = 1$, $c = 20$, $k = 10^4$ and $k_2 = 10^7$.

Because of the unstable nature of the pure quadratic, 'second-order' behaviour is usually modelled with a term of the form $k_2 y|y|$. The FRF for a system with this nonlinearity is given by

$$\Lambda(\omega) = \frac{1}{k + \frac{8k_2 Y}{3\pi} - m\omega^2 + \mathrm{i}c\omega} \tag{3.78}$$

and the bifurcation analysis is similar to that in the cubic case, but a little more complicated as the equation for the response amplitude is a quartic,

$$X^2 = Y^2 \left[\left(k + \frac{8k_2 Y}{3\pi} - m\omega^2 \right)^2 + c^2\omega^2 \right]. \tag{3.79}$$

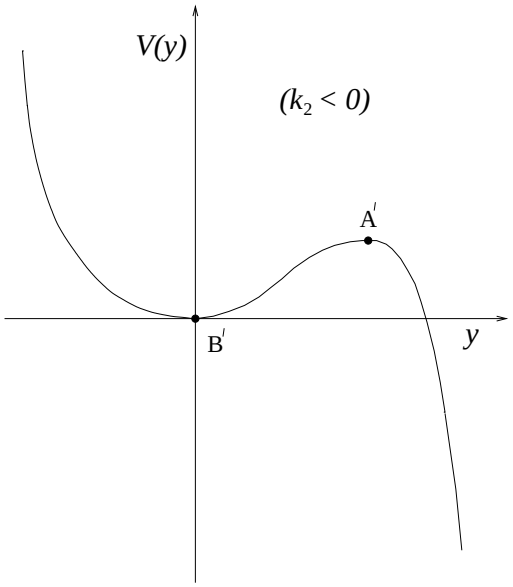


Figure 3.7. Potential energy of the quadratic oscillator with $k_2 < 0$.

3.7.2 Bilinear stiffness

Another system which is of physical interest is that with bilinear stiffness function of the form (figure 3.9)

$$f_{\mathrm{s}}(y) = \begin{cases} k, & \text{if } y < y_{\mathrm{c}} \\ k'y + (k - k')y_{\mathrm{c}}, & \text{if } y \geq y_{\mathrm{c}}. \end{cases} \tag{3.80}$$

Without loss of generality, one can specify that $y_{\mathrm{c}} > 0$. The equivalent stiffness is given by equation (3.43). There is a slight subtlety here, the integrand changes when the displacement $Y\sin(\omega t)$ exceeds y_{c}. This corresponds to a point in the cycle $\theta_{\mathrm{c}} = \omega t_{\mathrm{c}}$ where

$$\theta_{\mathrm{c}} = \sin^{-1}\left(\frac{y_{\mathrm{c}}}{Y}\right). \tag{3.81}$$

The integrand switches back when $\theta = \pi - \theta_{\mathrm{c}}$. A little thought shows that the equivalent stiffness must have the form

$$k_{\mathrm{eq}} = k + \frac{(k' - k)}{\pi}\int_{\theta_{\mathrm{c}}}^{\pi - \theta_{\mathrm{c}}} \mathrm{d}\theta\, \sin\theta \left(\sin\theta - \frac{y_{\mathrm{c}}}{Y}\right) \tag{3.82}$$

so, after a little algebra,

$$k_{\mathrm{eq}} = k + \frac{(k' - k)}{2\pi}\left(\pi - 2\theta_{\mathrm{c}} + \sin 2\theta_{\mathrm{c}} - \frac{4y_{\mathrm{c}}}{Y}\cos\theta_{\mathrm{c}}\right) \tag{3.83}$$

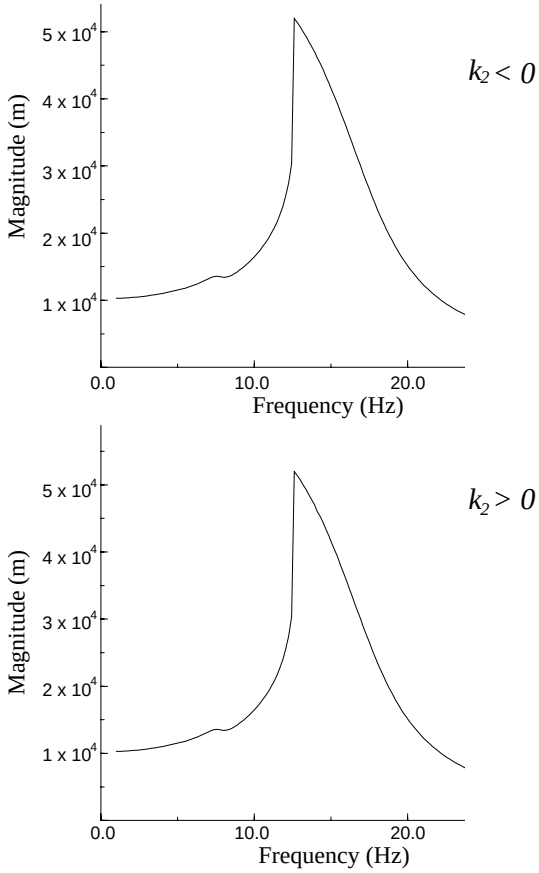


Figure 3.8. Response of the quadratic oscillator to a constant magnitude stepped-sine input.

or

$$k_{\mathrm{eq}} = k + \frac{(k'-k)}{2\pi}\left\{(\pi - 2\sin^{-1}\left(\frac{y_{\mathrm{c}}}{Y}\right) + \sin\left[2\sin^{-1}\left(\frac{y_{\mathrm{c}}}{Y}\right)\right] - \frac{4y_{\mathrm{c}}}{Y}\cos\left[\sin^{-1}\left(\frac{y_{\mathrm{c}}}{Y}\right)\right]\right\}. \tag{3.84}$$

As a check, substituting $k = k'$ or $Y = y_{\mathrm{c}}$ yields $k_{\mathrm{eq}} = k$ as necessary.

The FRF has the form

$$\Lambda(\omega) = \frac{1}{k + \frac{(k-k')}{2\pi}\left\{(\pi - 2\sin^{-1}\left(\frac{y_{\mathrm{c}}}{Y}\right) - \frac{2y_{\mathrm{c}}}{Y}\sqrt{Y^2 - y_{\mathrm{c}}^2}\right\} - m\omega^2 + \mathrm{i}c\omega}. \tag{3.85}$$

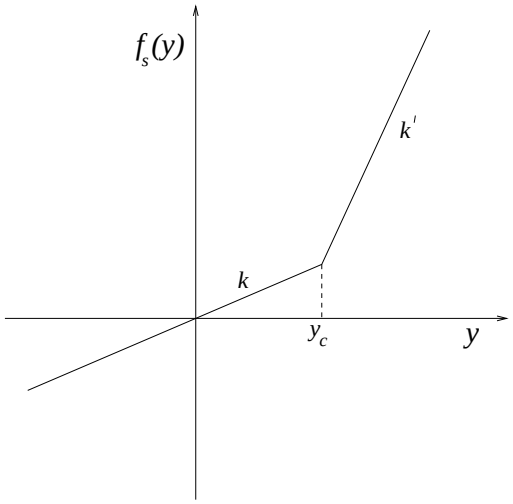


Figure 3.9. Bilinear stiffness characteristic with offset.

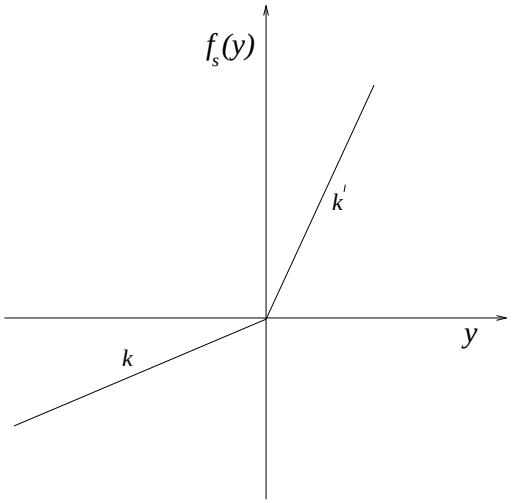


Figure 3.10. Bilinear stiffness characteristic without offset.

Now, let $y_c = 0$ (figure 3.10). The expression (3.84) collapses to

$$k_{\mathrm{eq}} = \tfrac{1}{2}(k + k') \tag{3.86}$$

which is simply the average stiffness. So the system has an effective natural frequency and FRF, independent of the size of Y and therefore, independent of X. The system is thus *homogeneous* as described in chapter 2. The homogeneity

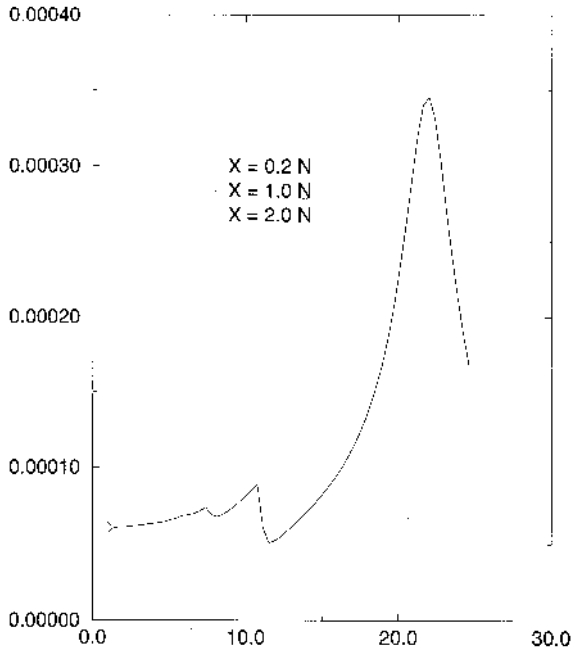


Figure 3.11. The stepped-sine FRF of a bilinear oscillator at different levels of the input force excitation showing independence of the output of the input, i.e. satisfying homogeneity.

test fails to detect that this system is nonlinear. That it *is* nonlinear is manifest; the Fourier expansion of $f_{\mathrm{s}}(y)$ (figure 3.10) contains *all* harmonics so the response of the system to a sinusoid will also contain all harmonics. The homogeneity of this system is a consequence of the fact that the stiffness function looks the same at all length scales. This analysis is only first order; however, figure 3.11 shows FRFs for different levels of excitation for the simulated system

$$\ddot{y} + 20\dot{y} + 10^4 y + 4 \times 10^4 y\Theta(y) = X\sin(30t). \tag{3.87}$$

The curves overlay and this demonstrates why homogeneity is a necessary but not sufficient condition for linearity.

3.8 Application of harmonic balance to an aircraft component ground vibration test

In the aircraft industry, one procedure for detecting nonlinearity during a ground vibration test is to monitor the resonant frequency of a given mode of vibration as the input force is increased. This is usually carried out using normal mode testing

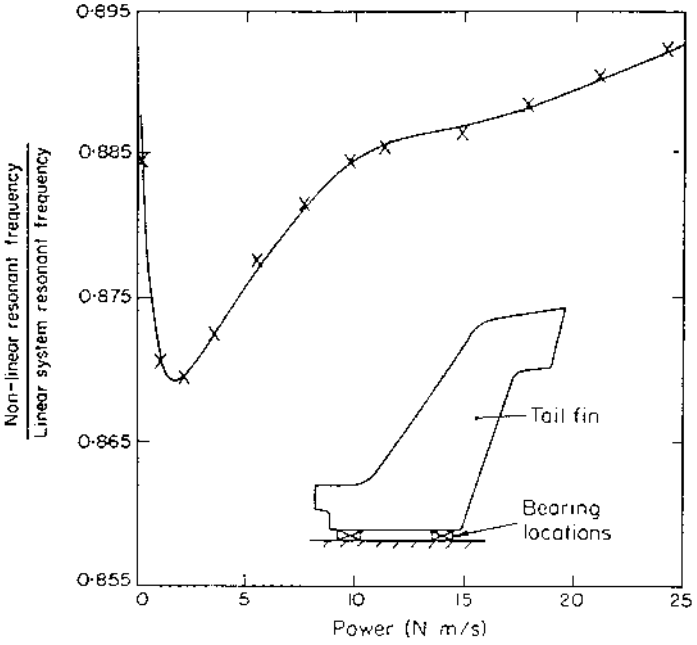


Figure 3.12. Experimental results from sine tests on an aircraft tail-fin showing the variation in resonant frequency of the first bending mode as a function of the increasing power input.

where force appropriation is used to calculate driving forces for multiple vibration exciters so that single modes of vibration are isolated. The response in a given mode then approximates to that from a single-degree-of-freedom (SDOF) system. By gradually increasing the input forces but maintaining the ratio of excitations at the various exciters, the same normal mode can be obtained and the corresponding natural frequency can be monitored. Note that in normal mode testing, the peak or resonant frequency coincides with the natural frequency, so the two terms can be used interchangeably.

If the system is linear, the normal mode natural frequency is invariant under changes in forcing level; any variations indicate the presence of nonlinearity. An example of the results from such a test is given in figure 3.12. This shows the variation in the first bending mode natural frequency for an aircraft tail-fin mounted on its bearing location pins as the input power is increased. The test shows nonlinearity. It was suspected that the nonlinearity was due to the bearing location pins being out of tolerance, this would result in a pre-loaded clearance nonlinearity at the bearing locations. The pre-load results from the self-weight of the fin loading the bearings and introduces an asymmetrical clearance. In order to test this hypothesis, a harmonic balance approach was adopted.

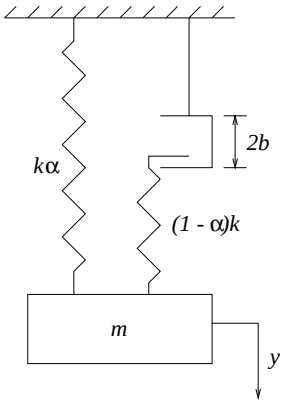


Figure 3.13. System with pre-loaded piecewise linear stiffness.

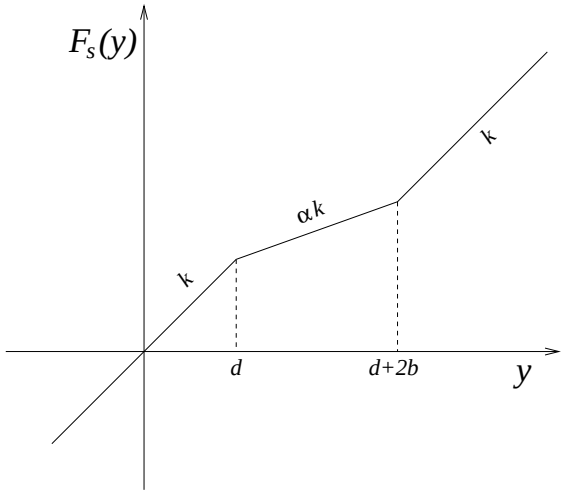


Figure 3.14. Pre-loaded piecewise linear stiffness curve.

Figure 3.13 shows the model used with stiffness curve as in figure 3.14. The equivalent stiffness is obtained from a harmonic balance calculation only a little more complicated than that for the bilinear stiffness already discussed,

$$k_{\rm eq} = k\left\{1 - \left(\frac{1-\alpha}{\pi}\right)\left[\sin^{-1}\left(\frac{2b+d}{Y}\right) - \sin^{-1}\left(\frac{d}{Y}\right) + \left(\frac{2b+d}{Y}\right)\left(1 - \left(\frac{2b+d}{Y}\right)^2\right)^{\frac{1}{2}} - \left(\frac{d}{Y}\right)\left(1 - \left(\frac{d}{Y}\right)^2\right)^{\frac{1}{2}}\right]\right\}. \tag{3.88}$$

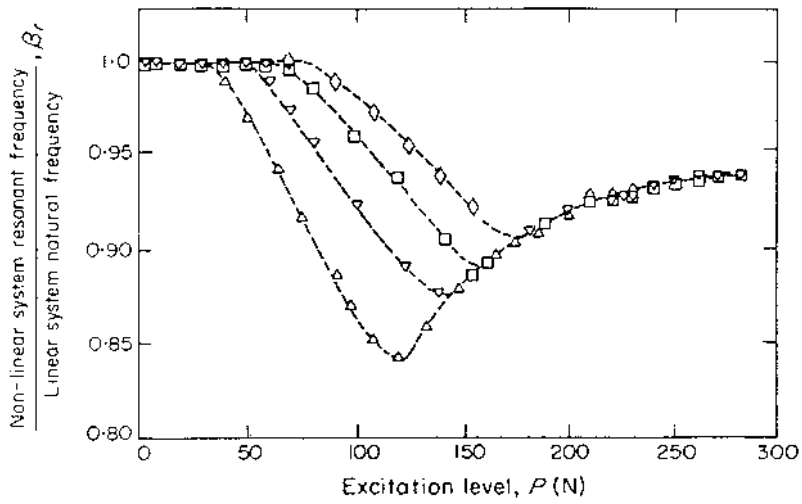


Figure 3.15. Variation in resonant frequency with excitation level for system with pre-loaded piecewise linear stiffness.

The FRF could have been obtained from (3.44); however, the main item of interest in this case was the variation in frequency with Y. Figure 3.12 actually shows the variation in β, the ratio of effective natural frequency to 'linear' natural frequency, i.e. the natural frequency at sufficiently low excitation that the clearance is not reached. The corresponding theoretical quantity is trivially obtained from (3.88) and is

$$\beta^2 = 1 - \left(\frac{1-\alpha}{\pi}\right)\left[\sin^{-1}\left(\frac{2b+d}{Y}\right) - \sin^{-1}\left(\frac{d}{Y}\right) + \left(\frac{2b+d}{Y}\right)\left(1-\left(\frac{2b+d}{Y}\right)^2\right)^{\frac{1}{2}} - \left(\frac{d}{Y}\right)\left(1-\left(\frac{d}{Y}\right)^2\right)^{\frac{1}{2}}\right]. \tag{3.89}$$

The form of the β–Y (actually β against power) curve is given in figure 3.15 for a number of d/b ratios. It admits a straightforward explanation in terms of the clearance parameters. As Y is increased from zero, at low values, the first break point at d is not reached and the system is linear with stiffness k. Over this range β is therefore unity. Once Y exceeds d a region of diminished stiffness αk is entered so β decreases with Y as more of the low stiffness region is covered. Once Y exceeds $d+2b$, the relative time in the stiffness k region begins to increase again and β increases correspondingly. β asymptotically reaches unity again as long as no other clearances are present. The clearance parameters can therefore be taken from the β–Y curve: $Y = d$ at the point when β first dips below unity,

and $Y = d + 2b$ at the minimum of the frequency ratio β[5].

This is a quite significant result, information is obtained from the FRF which yields physical parameters of the system which are otherwise difficult to estimate.

The characteristics of the β-power curves in figure 3.15 are very similar to the experimentally obtained curve of figure 3.12. In fact, the variation in β *was* due to a clearance in the bearing location pins and after adjustment the system behaved much more like the expected linear system.

This example shows how a simple analysis can be gainfully employed to investigate the behaviour of nonlinear systems.

3.9 Alternative FRF representations

In dynamic testing, it is very common to use different presentation formats for the FRF. Although the Bode plot (modulus and phase) is arguably the most common, the Nyquist plot or real and imaginary parts are often shown. For nonlinear systems, the different formats offer insights into different aspects of the nonlinear behaviour. For systems with nonlinear stiffness, the dominant effects are changes in the resonant frequencies and these are best observed in the Bode plot or real/imaginary plot. For systems with nonlinear damping, as shown later, the Argand diagram or Nyquist plot is often more informative.

3.9.1 Nyquist plot: linear system

For a linear system with viscous damping

$$\ddot{y} + 2\zeta\omega_n\dot{y} + \omega_n^2 y = \frac{x(t)}{m} \tag{3.90}$$

the Nyquist plot has different aspects, depending on whether the data are *receptance* (displacement), *mobility* (velocity) or *accelerance* (acceleration). In all cases, the plot approximates to a circle as shown in figure 3.16. The most interesting case is mobility, there the plot is a circle in the positive real half-plane, bisected by the real axis (figure 3.16(*b*)). The mobility FRF is given by

$$H_{\mathrm{M}}(\omega) = \frac{1}{m}\frac{\mathrm{i}\omega}{\omega_n^2 - \omega^2 + 2\mathrm{i}\zeta\omega_n\omega} \tag{3.91}$$

[5] In fact, the analysis of the situation is a little more subtle than this. In the first case, calculus shows that the minimum of the β–Y curve is actually at

$$Y = [(2b + d)^2 + d^2]^{\frac{1}{2}}.$$

In the second case, as the stiffness function is asymmetric it leads to a non-zero operating point for the motion $y_0 = S$, so the minimum will actually be at

$$Y = [(2b + d)^2 + d^2]^{\frac{1}{2}} + S.$$

Details of the necessary calculations can be found in [252].

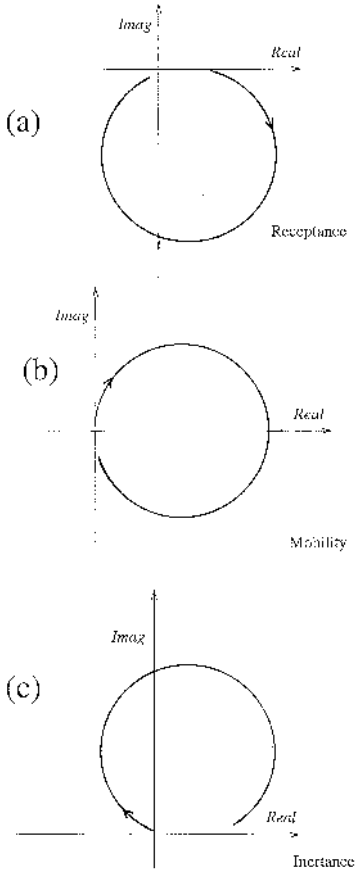


Figure 3.16. Nyquist plots for: (*a*) receptance; (*b*) mobility; (*c*) accelerance.

and it is a straightforward exercise to show that this curve in the Argand diagram is a circle, centre $(\frac{\omega_n}{4\zeta}, 0)$ and radius $\frac{\omega_n}{4\zeta}$.

For a system with hysteretic damping

$$\ddot{y} + \omega_n^2(1 + \mathrm{i}\eta)y = \frac{x(t)}{m}. \tag{3.92}$$

The Nyquist plots are also approximate to circles; however, it is the receptance FRF which is circular in this case, centred at $(0, -\frac{1}{2\eta})$ with radius $\frac{1}{2\eta}$. The receptance FRF is

$$H_{\mathrm{R}}(\omega) = \frac{1}{m}\frac{1}{\omega_n^2 - \omega^2 + \mathrm{i}\eta\omega_n^2}. \tag{3.93}$$

One approach to modal analysis, the *vector plot* method of Kennedy and Pancu [139] relies on fitting circular arcs from the resonant region of the Nyquist

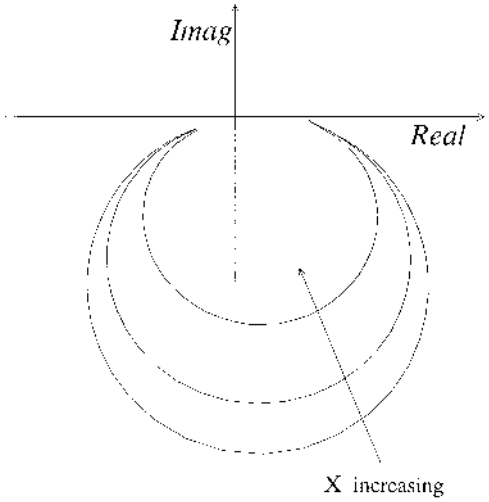


Figure 3.17. Nyquist plot distortion for a SDOF system with velocity-squared (quadratic) damping.

plot [212, 121]. Any deviations from circularity will introduce errors and this will occur for most nonlinear systems. However, if the deviations are characteristic of the type of nonlinearity, something at least is salvaged.

3.9.2 Nyquist plot: velocity-squared damping

Using a harmonic balance approach, the FRF for the system with quadratic damping (3.62) is given by (3.66). For mixed viscous–quadratic damping

$$f_{\mathrm{d}}(\dot{y}) = c\dot{y} + c_2\dot{y}|\dot{y}| \tag{3.94}$$

the FRF is

$$\Lambda(\omega) = \frac{1}{k - m\omega^2 + \mathrm{i}(c + \frac{8c_2\omega Y}{\pi})\omega}. \tag{3.95}$$

At low levels of excitation, the Nyquist (receptance) plot looks like the linear system. However, as the excitation level X, and hence the response amplitude Y, increases, characteristic distortions occur (figure 3.17); the FRF decreases in size and becomes elongated along the direction of the real axis.

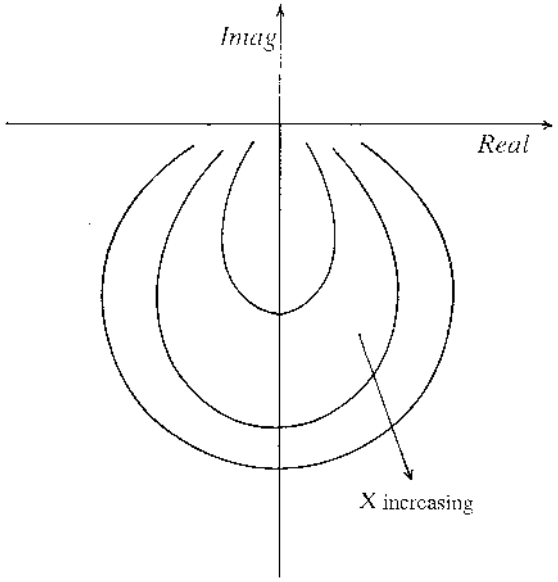


Figure 3.18. Nyquist plot distortion for a SDOF system with Coulomb friction.

3.9.3 Nyquist plot: Coulomb friction

In this case, the force–velocity relationship is

$$f_{\mathrm{d}}(\dot{y}) = c\dot{y} + c_{\mathrm{F}}\frac{\dot{y}}{|\dot{y}|} = c\dot{y} + c_{\mathrm{F}}\,\mathrm{sgn}(\dot{y}) \tag{3.96}$$

and the FRF is found to be

$$\Lambda(\omega) = \frac{1}{k - m\omega^2 + \mathrm{i}(c\omega + \frac{4c_{\mathrm{F}}}{\pi\omega Y})}. \tag{3.97}$$

The analysis in this case is supplemented by a condition

$$X > \frac{4c_{\mathrm{F}}}{\pi} \tag{3.98}$$

which is necessary to avoid stick-slip motion. Intermittent motion invalidates (3.98). Typical distortions of the receptance FRF as X, and hence, Y increases are given in figure 3.18. At low levels of excitation, the friction force is dominant and a Nyquist plot of reduced size is obtained, the curve is also elongated in the direction of the imaginary axis. As X increases, the friction force becomes relatively unimportant and the linear FRF is obtained in the limit.

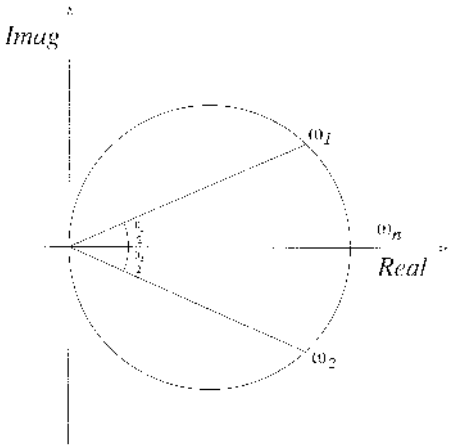


Figure 3.19. Reference points for circle fitting procedure: viscous damping.

3.9.4 Carpet plots

Suppose the Nyquist plot is used to estimate the damping in the system. Consider the geometry shown in figure 3.19 for the mobility FRF in the viscous damping case. Simple trigonometry yields

$$\tan\frac{\theta_1}{2} = \frac{\omega_n^2 - \omega_1^2}{2\zeta\omega_n\omega_1} \tag{3.99}$$

and

$$\tan\frac{\theta_2}{2} = \frac{\omega_2^2 - \omega_n^2}{2\zeta\omega_n\omega_2} \tag{3.100}$$

so

$$\zeta = \frac{\omega_2(\omega_n^2 - \omega_1^2) - \omega_1(\omega_n^2 - \omega_2^2)}{2\zeta\omega_1\omega_2\omega_n}\left(\frac{1}{\tan\frac{\theta_1}{2} + \tan\frac{\theta_2}{2}}\right) \tag{3.101}$$

and this estimate should be independent of the points chosen. If ζ is plotted over the (θ_1, θ_2) plane it should yield a flat constant plane. Any deviation from linearity produces a variation in the so-called *carpet plot* [87]. Figure 3.20 shows carpet plots for a number of common nonlinear systems. The method is very restricted in its usage, problems are: sensitivity to phase distortion and noise, lack of quantitative information about the nonlinearity, restriction to SDOF systems and the requirement of an *a priori* assumption of the damping model. On this last point, the plot can be defined for the hysteretic damping case by reference to the receptance FRF of figure 3.21, there

$$\tan\frac{\theta_1}{2} = \frac{\omega_1^2 - \omega_n^2}{\eta\omega_n^2} \tag{3.102}$$

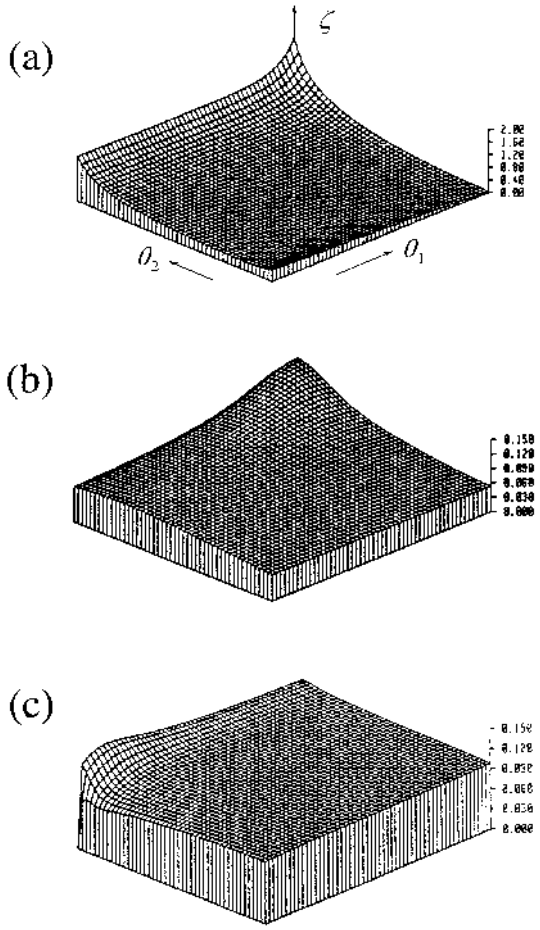


Figure 3.20. Carpet plots of SDOF nonlinear systems: (*a*) Coulomb friction; (*b*) quadratic damping; (*c*) hardening spring.

$$\tan\frac{\theta_2}{2} = \frac{\omega_n^2 - \omega_2^2}{\eta\omega_n^2} \tag{3.103}$$

and so

$$\eta = \frac{\omega_1^2 - \omega_2^2}{\omega_n^2}\left(\frac{1}{\tan\frac{\theta_1}{2} + \tan\frac{\theta_2}{2}}\right). \tag{3.104}$$

Note that this analysis only holds in the case of a constant magnitude harmonic excitation.

One comment applies to all the methods of this section: characteristic distortions are still produced by nonlinearities in multi-degree-of-freedom

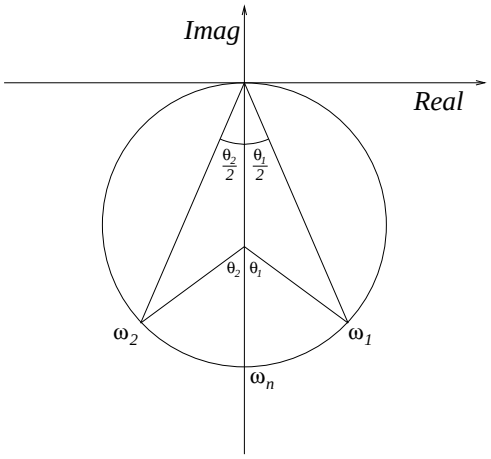


Figure 3.21. Reference points for circle fitting procedure: hysteretic damping.

(MDOF) systems. This analysis will still apply in some cases where the modal density is not high, i.e. the spacing between the modes is large.

3.10 Inverse FRFs

The philosophy of this approach is very simple. The *inverse* $\frac{1}{\Lambda(\omega)}$ of the SDOF system FRF[6] is much simpler to handle than the FRF itself, in the general case for mixed stiffness and damping nonlinearities:

$$I(\omega) = \frac{1}{\Lambda(\omega)} = k_{\mathrm{eq}}(\omega) - m\omega^2 + \mathrm{i}c_{\mathrm{eq}}(\omega). \tag{3.105}$$

In the linear case

$$\mathrm{Re}\, I(\omega) = k - m\omega^2 \tag{3.106}$$

and a plot of the real part against ω^2 yields a straight line with intercept k and gradient $-m$. The imaginary part

$$\mathrm{Im}\, I(\omega) = c\omega \tag{3.107}$$

is a line through the origin with gradient c. If the system is nonlinear, these plots will not be straight lines, but will contain distortions characteristic of the nonlinearity. It is usual to plot the IFRF (Inverse FRF) components with linear curve-fits superimposed to show more clearly the distortions. Figure 3.22 shows the IFRF for a linear system; the curves are manifestly linear. Figures 3.23 and 3.24 show the situation for stiffness nonlinearities—the distortions only occur in

[6] Note: *not* $\Lambda^{-1}(\omega)$.

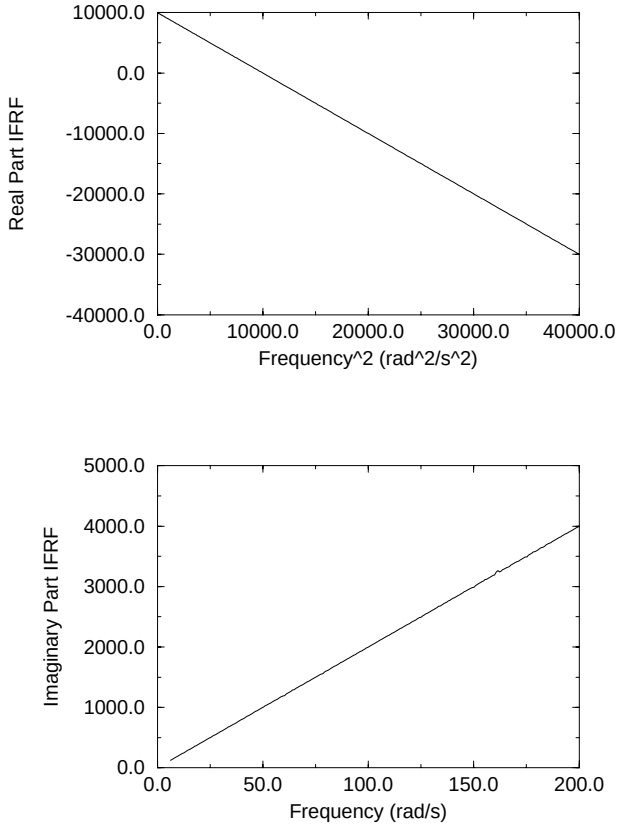


Figure 3.22. Inverse FRF (IFRF): SDOF linear system.

the real part. Conversely, for damping nonlinearities (figures 3.25), distortions only occur in the imaginary part. Mixed nonlinearities show the characteristics of both types.

Again, this analysis makes sense for MDOF systems as long as the modes are well spaced. On a practical note, measurement of the IFRFs is trivial. All that is required is to change over the input and output channels to a standard spectrum or FRF analyser so that the input enters channel A and the output, channel B.

3.11 MDOF systems

As discussed in chapter 1, the extension from SDOF to MDOF for linear systems is not trivial, but presents no real mathematical difficulties[7]. Linear MDOF

[7] Throughout this book, proportional damping is assumed so the problem of complex modes does not occur. In any case this appears to be a problem of interpretation rather than a difficulty with the mathematics.

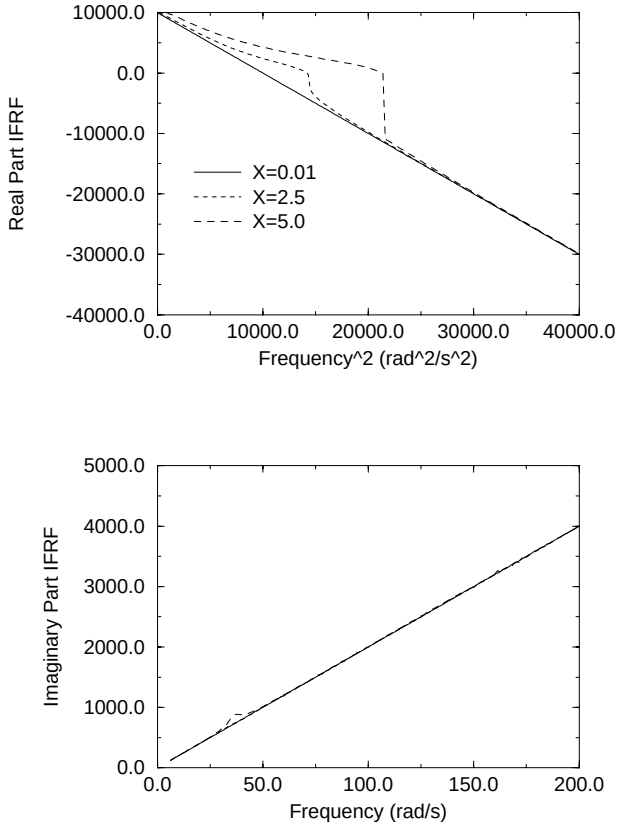


Figure 3.23. IFRF for SDOF hardening cubic system for a range of constant force sinusoidal excitation levels.

systems can be decomposed into a sequence of uncoupled SDOF systems by a linear transformation of coordinates to *modal* space. It is shown here that the situation for nonlinear systems is radically different; for generic systems, such uncoupling proves impossible.

However, first consider the 2DOF system shown in figure 3.26 and specified by the equations of motion

$$m\ddot{y}_1 + c\dot{y}_1 + 2ky_1 - ky_2 + k_3(y_1 - y_2)^3 = x_1(t) \tag{3.108}$$

$$m\ddot{y}_2 + c\dot{y}_2 + 2ky_2 - ky_1 + k_3(y_2 - y_1)^3 = x_2(t) \tag{3.109}$$

or, in matrix notation,

$$\begin{pmatrix} m & 0 \\ 0 & m \end{pmatrix} \begin{pmatrix} \ddot{y}_1 \\ \ddot{y}_2 \end{pmatrix} + \begin{pmatrix} c & 0 \\ 0 & c \end{pmatrix} \begin{pmatrix} \dot{y}_1 \\ \dot{y}_2 \end{pmatrix} + \begin{pmatrix} 2k & -k \\ -k & 2k \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

$$+\begin{pmatrix} k_3(y_1-y_2)^3 \\ -k_3(y_1-y_2)^3 \end{pmatrix} = \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix}. \tag{3.110}$$

The modal matrix for the underlying linear system is

$$[\psi] = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \tag{3.111}$$

corresponding to modal coordinates

$$u_1 = \frac{1}{\sqrt{2}}(y_1 + y_2) \tag{3.112}$$

$$u_2 = \frac{1}{\sqrt{2}}(y_1 - y_2). \tag{3.113}$$

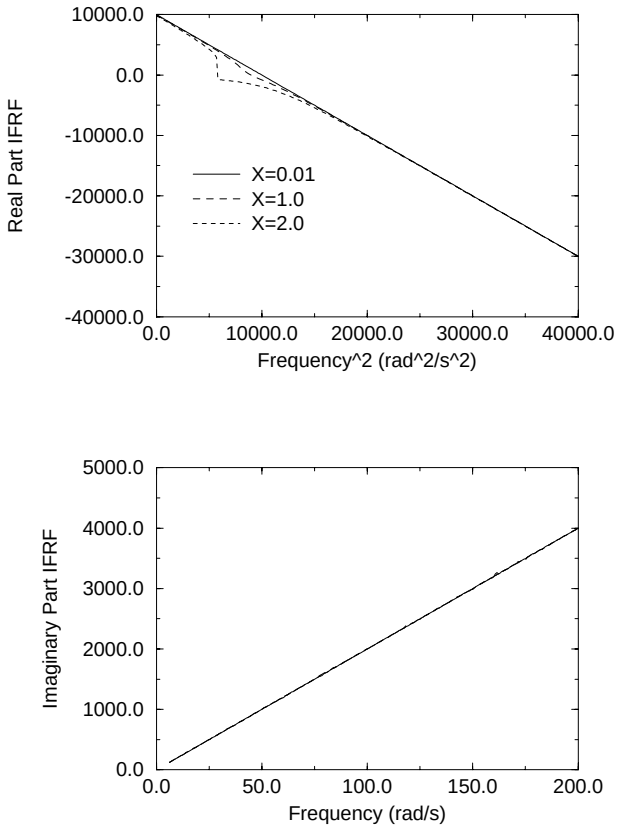


Figure 3.24. IFRF for SDOF softening cubic system for a range of constant force sinusoidal excitation levels.

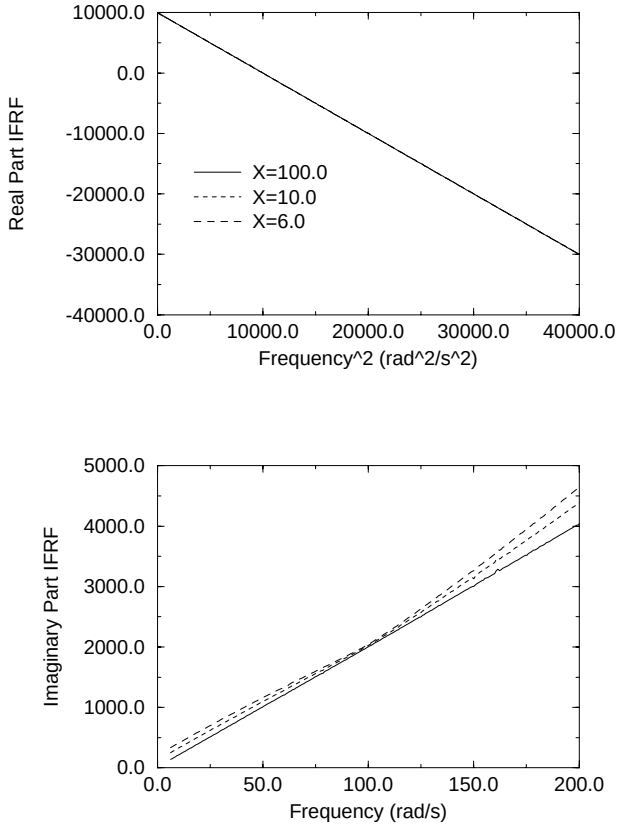


Figure 3.25. IFRF for SDOF Coulomb friction system for a range of constant force sinusoidal excitation levels.

Changing to these coordinates for the system (3.110) yields

$$m\ddot{u}_1 + c\dot{u}_1 + ku_1 = \frac{1}{\sqrt{2}}(x_1 + x_2) = p_1 \tag{3.114}$$

$$m\ddot{u}_2 + c\dot{u}_2 + 3ku_2 + \frac{1}{2}k_3u_2^3 = \frac{1}{\sqrt{2}}(x_1 - x_2) = p_2. \tag{3.115}$$

So the systems are decoupled, although one of them remains nonlinear. Assuming for the sake of simplicity that $x_1 = 0$, the FRF for the process $x_2 \longrightarrow u_1$ is simply the linear,

$$H_{x_2u_1}(\omega) = \frac{1}{\sqrt{2}}\left(\frac{1}{k - m\omega^2 + \mathrm{i}c\omega}\right) \tag{3.116}$$

and standard SDOF harmonic balance analysis suffices to extract the FRF for the

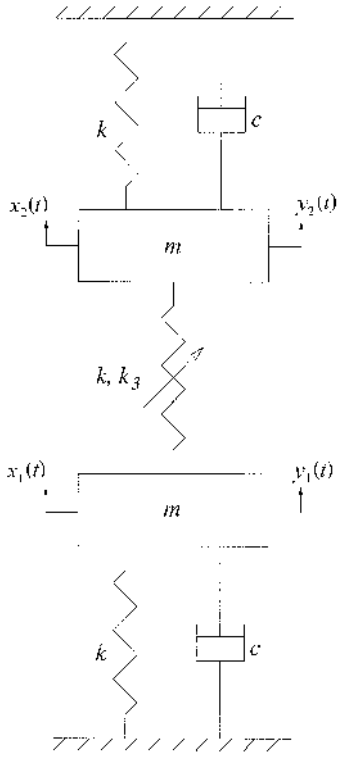


Figure 3.26. 2DOF symmetrical system with a nonlinear stiffness coupling the masses.

nonlinear process $x_2 \longrightarrow u_2$,

$$\Lambda_{x_2u_2}(\omega) = -\frac{1}{\sqrt{2}}\left(\frac{1}{3k + \frac{3}{8}k_3|U_2|^2 - m\omega^2 + \mathrm{i}c\omega}\right). \tag{3.117}$$

Dividing the inverse coordinate transformation,

$$Y_1(\omega) = \frac{1}{\sqrt{2}}(U_1(\omega) + U_2(\omega)) \tag{3.118}$$

in the frequency domain[8], by $X_2(\omega)$ yields

$$\Lambda_{21}(\omega) = \frac{1}{\sqrt{2}}(H_{x_2u_1}(\omega) + \Lambda_{x_2u_2}(\omega)) \tag{3.119}$$

[8] Here, Y_1, U_1 and U_2 are complex to encode the phases.

so that back in the physical coordinate system

$$\Lambda_{21}(\omega) = \frac{1}{2}\left(\frac{1}{k - m\omega^2 + \mathrm{i}c\omega}\right) - \frac{1}{2}\left(\frac{1}{3k + \frac{3}{8}k_3|U_2|^2 - m\omega^2 + \mathrm{i}c\omega}\right) \tag{3.120}$$

and, similarly,

$$\Lambda_{22}(\omega) = \frac{1}{2}\left(\frac{1}{k - m\omega^2 + \mathrm{i}c\omega}\right) + \frac{1}{2}\left(\frac{1}{3k + \frac{3}{8}k_3|U_2|^2 - m\omega^2 + \mathrm{i}c\omega}\right). \tag{3.121}$$

This shows that in the FRFs for the system (3.110), only the second mode is ever distorted as a result of the nonlinearity. Figure 3.27 shows the magnitudes of the FRFs in figures 1.16 and 1.18 for different levels of excitation (actually from numerical simulation). As in the SDOF case, the FRFs show discontinuities if the level of excitation exceeds a critical value.

The first natural frequency is

$$\omega_{n1} = \sqrt{\frac{k}{m}} \tag{3.122}$$

and is independent of the excitation. However, the second natural frequency,

$$\omega_{n2} = \sqrt{\frac{3k + \frac{3}{8}k_3U_2^2}{m}} \tag{3.123}$$

increases with increasing excitation if $k_3 > 0$ and decreases if $k_3 < 0$.

In this case, the decoupling of the system in modal coordinates manifests itself in physical space via the distortion of the second mode only, one can say that only the second mode is nonlinear. This situation is clearly very fragile; any changes in the system parameters will usually lead to distortion in both modes. Also, the position of the nonlinear spring is critical here. Physically, the first mode has the two masses moving in unison with identical amplitude. This means that the central nonlinear spring never extends and therefore has no effect. The central spring is the only component which can be nonlinear and still allow decoupling. Decoupling only occurs in systems which possess a high degree of symmetry. As another example, consider the linear 3DOF system which has equations of motion,

$$\begin{pmatrix} m & 0 & 0 \\ 0 & m & 0 \\ 0 & 0 & m \end{pmatrix}\begin{pmatrix} \ddot{y}_1 \\ \ddot{y}_2 \\ \ddot{y}_3 \end{pmatrix} + \begin{pmatrix} 2c & -c & 0 \\ -c & 2c & -c \\ 0 & -c & 2c \end{pmatrix}\begin{pmatrix} \dot{y}_1 \\ \dot{y}_2 \\ \dot{y}_3 \end{pmatrix} + \begin{pmatrix} 2k & -k & 0 \\ -k & 2k & -k \\ 0 & -k & 2k \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}. \tag{3.124}$$

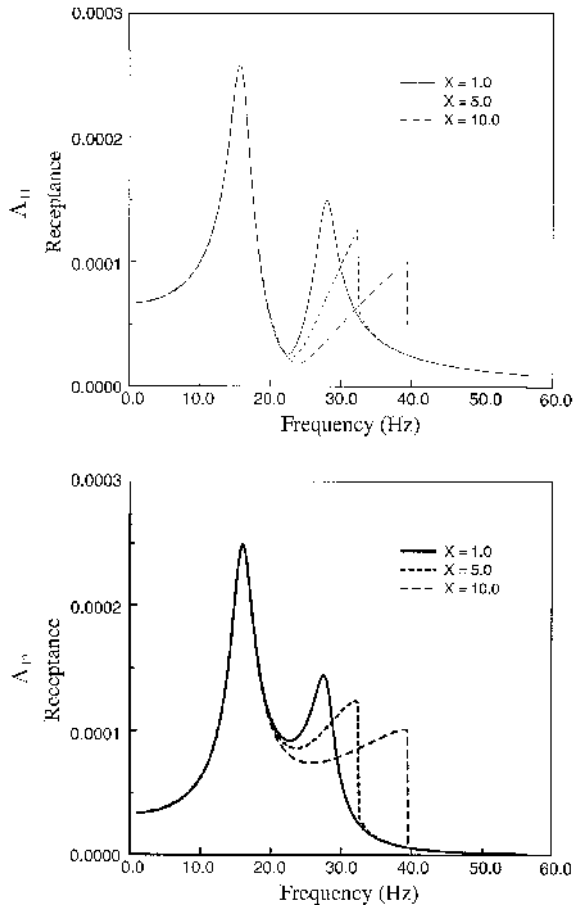


Figure 3.27. Stepped-sine FRFs Λ_{11} and Λ_{12} for 2DOF system with nonlinearity between masses.

In this system, one position for a nonlinearity which allows any decoupling is joining the centre mass to ground. This is because in the underlying linear system, the second mode has masses 1 and 3 moving in anti-phase while the centre mass remains stationary. As a result, the FRFs for this system would show the second mode remaining free of distortion as the excitation level was varied.

The equations for harmonic balance for the system in (3.124) would be complicated by the fact that modes 1 and 3 remain coupled even if the nonlinearity is at the symmetry point. This effect can be investigated in a simpler system; suppose the nonlinearity in figure 3.26 is moved to connect one of the masses, the

upper one say, to ground. The resulting equations of motion are

$$m\ddot{y}_1 + c\dot{y}_1 + 2ky_1 - ky_2 + k_3 y_1^3 = x_1(t) \tag{3.125}$$

$$m\ddot{y}_2 + c\dot{y}_2 + 2ky_2 - ky_1 = x_2(t). \tag{3.126}$$

The transformation to modal space is given by (3.112) and (3.113) as the new system has the same underlying linear system as (3.110). In modal space, the new system is

$$m\ddot{u}_1 + c\dot{u}_1 + ku_1 + \frac{k_3}{4}(u_1 - u_2)^3 = \frac{1}{\sqrt{2}}(x_1(t) + x_2(t)) = p_1(t) \tag{3.127}$$

$$m\ddot{u}_2 + c\dot{u}_2 + 3ku_2 + \frac{k_3}{4}(u_2 - u_1)^3 = \frac{1}{\sqrt{2}}(x_1(t) - x_2(t)) = p_2(t) \tag{3.128}$$

which is still coupled by the nonlinearity. Note that there is no *linear* transformation which completely uncouples the system as (3.111) is the unique (up to scale) transformation which uncouples the underlying linear part. Harmonic balance for this system now proceeds by substituting the excitations, $x_1(t) = X\sin(\omega t)$ and $x_2(t) = 0$ (for simplicity) and trial solutions $u_1(t) = U_1\sin(\omega t + \phi_1)$ and $u_2(t) = U_2\sin(\omega t + \phi_2)$ into equations (3.127) and (3.128). After a lengthy but straightforward calculation, the fundamental components of each equation can be extracted. This gives a system of equations

$$\begin{aligned}
&-m\omega^2 U_1\cos\phi_1 - c\omega U_1\sin\phi_1 + kU_1\cos\phi_1 \\
&\quad + \frac{3}{16}k_3\left\{U_1^3\cos\phi_1 + U_1^2U_2[2\cos\phi_1\cos(\phi_1-\phi_2) + \cos\phi_2]\right. \\
&\quad \left.\times\, U_1U_2^2[2\cos\phi_2\cos(\phi_1-\phi_2) + \cos\phi_1] + U_2^3\cos\phi_2\right\} = \frac{1}{\sqrt{2}}X
\end{aligned} \tag{3.129}$$

$$\begin{aligned}
&-m\omega^2 U_1\sin\phi_1 + c\omega U_1\cos\phi_1 + kU_1\sin\phi_1 \\
&\quad + \frac{3}{16}k_3\left\{U_1^3\sin\phi_1 + U_1^2U_2[2\sin\phi_1\cos(\phi_1-\phi_2) + \sin\phi_2]\right. \\
&\quad \left.\times\, U_1U_2^2[2\sin\phi_2\cos(\phi_1-\phi_2) + \sin\phi_1] + U_2^3\sin\phi_2\right\} = 0
\end{aligned} \tag{3.130}$$

$$\begin{aligned}
&-m\omega^2 U_2\cos\phi_2 - c\omega U_2\sin\phi_2 + kU_2\cos\phi_2 \\
&\quad - \frac{3}{16}k_3\left\{U_1^3\cos\phi_1 + U_1^2U_2[2\cos\phi_1\cos(\phi_1-\phi_2) + \cos\phi_2]\right. \\
&\quad \left.\times\, U_1U_2^2[2\cos\phi_2\cos(\phi_1-\phi_2) + \cos\phi_1] + U_2^3\cos\phi_2\right\} = 0
\end{aligned} \tag{3.131}$$

$$\begin{aligned}
&-m\omega^2 U_2\sin\phi_2 + c\omega U_2\cos\phi_2 + kU_2\sin\phi_2 \\
&\quad - \frac{3}{16}k_3\left\{U_1^3\sin\phi_1 + U_1^2U_2[2\sin\phi_1\cos(\phi_1-\phi_2) + \sin\phi_2]\right. \\
&\quad \left.\times\, U_1U_2^2[2\sin\phi_2\cos(\phi_1-\phi_2) + \sin\phi_1] + U_2^3\sin\phi_2\right\} = 0
\end{aligned} \tag{3.132}$$

which must be solved for U_1, U_2, ϕ_1 and ϕ_2 for each ω value required in the FRF. This set of equations is very complicated; to see if there is any advantage

in pursuing the modal approach, one should compare this with the situation if the system is studied in physical space. The relevant equations are (3.125) and (3.126). If the same excitation is used, but a trial solution of the form $y_1(t) = Y_1 \sin(\omega t + \theta_1)$, $y_2(t) = Y_2 \sin(\omega t + \theta_2)$ is adopted, a less lengthy calculation yields the system of equations

$$\begin{aligned}
&- m\omega^2 Y_1 \cos\theta_1 - c\omega Y_1 \sin\theta_1 + 2kY_1 \cos\theta_1 - kY_2 \cos\theta_2 \\
&\quad + \tfrac{3}{4}k_3 Y_1^3 \cos\theta_1 = X && (3.133)\\
&- m\omega^2 Y_1 \sin\theta_1 - c\omega Y_1 \cos\theta_1 + 2kY_1 \sin\theta_1 - kY_2 \sin\theta_2 \\
&\quad + \tfrac{3}{4}k_3 Y_1^3 \sin\theta_1 = X && (3.134)\\
&- m\omega^2 Y_2 \cos\theta_2 - c\omega Y_2 \sin\theta_2 + 2kY_2 \cos\theta_1 - kY_1 \cos\theta_1 = 0 && (3.135)\\
&- m\omega^2 Y_2 \sin\theta_2 - c\omega Y_2 \cos\theta_2 + 2kY_2 \sin\theta_2 - kY_1 \sin\theta_1 = 0 && (3.136)
\end{aligned}$$

which constitute a substantial simplification over the set (3.129)–(3.132) obtained in modal space. The moral of this story is that, for nonlinear systems, transformation to modal space is only justified if there is a simplification of the nonlinearity supplementing the simplification of the underlying linear system. If the transformation complicates the nonlinearity, one is better off in physical space.

Judging by previous analysis, there is a potential advantage in forsaking the symmetry of the trial solution above and shifting the time variable from t to $t - \theta_1/\omega$. So the excitation is now $x_1(t) = X \sin(\omega t - \theta_1)$ and the trial solution is $y_1(t) = Y_1 \sin(\omega t)$, $y_2(t) = Y_2 \sin(\omega t + \psi)$ where $\psi = \theta_2 - \theta_1$, the new set of equations is

$$\begin{aligned}
- m\omega^2 Y_1 + 2kY_1 - kY_2 \cos\psi + \tfrac{3}{4}k_3 Y_1^3 &= X \cos\theta_1 && (3.137)\\
c\omega Y_1 - kY_2 \sin\psi &= -X \sin\theta_1 && (3.138)\\
- m\omega^2 Y_2 \cos\psi - c\omega Y_2 \sin\psi + 2kY_2 \cos\psi - kY_1 &= 0 && (3.139)\\
- m\omega^2 Y_2 \sin\psi + c\omega Y_2 \cos\psi + 2kY_2 \sin\psi &= 0 && (3.140)
\end{aligned}$$

and if the trivial solution $Y_1 = Y_2 = 0$ is to be avoided, the last equation forces the condition

$$-m\omega^2 \sin\psi + c\omega \cos\psi + 2k \sin\psi = 0 \tag{3.141}$$

so

$$\psi = \tan^{-1}\left(\frac{c\omega}{2k - m\omega^2}\right) \tag{3.142}$$

and there are only three equations (3.137)–(3.139) to solve for the remaining three unknowns Y_1, Y_2 and θ_1. Equation (3.139) then furnishes a simple relationship between Y_1 and Y_2, i.e.

$$Y_2 = \left(\frac{k}{-m\omega^2 \cos\psi - c\omega \sin\psi + 2k \cos\psi}\right) Y_1 \tag{3.143}$$

and this can be used to 'simplify' (3.137) and (3.138). This yields

$$\left[\left(-m\omega^2+2k-\frac{k^2\cos\psi}{-m\omega^2\cos\psi-c\omega\sin\psi+2k\cos\psi}\right)+\frac{3}{4}k_3Y_1^2\right]Y_1 = X\cos\theta_2 \tag{3.144}$$

$$\left(c\omega-\frac{k^2\sin\psi}{-m\omega^2\cos\psi-c\omega\sin\psi+2k\cos\psi}\right)Y_1=-X\sin\theta_1. \tag{3.145}$$

Squaring and adding these last two equations gives

$$\left\{\left[\left(-m\omega^2+2k-\frac{k^2\cos\psi}{-m\omega^2\cos\psi-c\omega\sin\psi+2k\cos\psi}\right)+\frac{3}{4}k_3Y_1^2\right]^2 + \left(c\omega-\frac{k^2\sin\psi}{-m\omega^2\cos\psi-c\omega\sin\psi+2k\cos\psi}\right)^2\right\}Y_1^2=X^2 \tag{3.146}$$

and the problem has been reduced to a cubic in ${Y_1}^2$ in much the same way that the SDOF analysis collapsed in section 3.2. This can be solved quite simply analytically or in a computer algebra package. The same bifurcations can occur in (3.146) between the cases of one and three real roots, so jumps are observed in the FRF exactly as in the SDOF case. In principle, one could compute the discriminant of this cubic and therefore estimate the frequencies where the jumps occur. However, this would be a tedious exercise, and the calculation is not pursued here. Once Y_1 is known, θ_1 follows simply from the ratio of equations (3.144) and (3.145)

$$\tan\theta_1=\frac{\left(c\omega-\frac{k^2\sin\psi}{-m\omega^2\cos\psi-c\omega\sin\psi+2k\cos\psi}\right)}{\left[\left(-m\omega^2+2k-\frac{k^2\cos\psi}{-m\omega^2\cos\psi-c\omega\sin\psi+2k\cos\psi}\right)+\frac{3}{4}k_3Y_1^2\right]} \tag{3.147}$$

and the solution for Y_2 is known from (3.143).

Figure 3.28 shows the magnitude of the Λ_{11} FRF for this system, this has been obtained by the numerical equivalent of a stepped-sine test rather than using the expressions given here. Note that both modes show distortion as expected. Unlike the case of the centred nonlinearity, the expressions for Y_1 and Y_2 obtained here obscure the fact that both modes distort. This obscurity will be the general case in MDOF analysis.

Unfortunately, the 'exact' solution here arrived somewhat fortuitously. In general, harmonic balance analysis for nonlinear MDOF systems will yield systems of algebraic equations which are too complex for exact analysis. The method can still yield useful information via numerical or hybrid numerical-symbolic computing approaches.

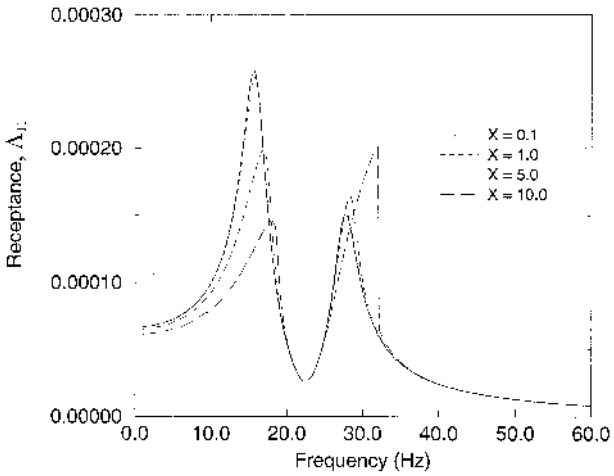


Figure 3.28. Stepped-sine FRF Λ_{11} for 2DOF system with nonlinearity connected to ground.

3.12 Decay envelopes

The FRF contains useful information about the behaviour of nonlinear systems under harmonic excitation. Stiffness nonlinearities produce characteristic changes in the resonant frequencies, damping nonlinearities typically produce distortions in the Nyquist plots. Under random excitation, the situation is somewhat different, the FRFs $\Lambda_{\mathrm{r}}(\omega)$ are considerably less distorted than their harmonic counterparts $\Lambda_{\mathrm{s}}(\omega)$ and usually prove less useful for the qualification of nonlinearity. This is discussed in some detail in chapter 8. The other member of the triumvirate of experimental excitations is impulse and the object of this section is to examine the utility of free decay data for the elucidation of system nonlinearity. This discussion sits aside from the rest of the chapter as it is not possible to define an FRF on the basis of decay data. However, in order to complete the discussion of different excitations, it is included here. It is shown in chapter 1 that the decay envelope for the linear system impulse response is a pure exponential whose characteristic time depends on the linear damping. For nonlinear systems, the envelope is modified according to the type of nonlinearity as shown here. In order to determine the envelopes a new technique is introduced.

3.12.1 The method of slowly varying amplitude and phase

This approach is particularly suited to the study of envelopes, as a motion of the form

$$y(t) = Y(t)\sin(\omega_n t + \phi(t)) \tag{3.148}$$

is assumed, where the envelope (amplitude) Y and phase ϕ vary with time, but slowly compared to the natural period of the system $\tau_n = \frac{2\pi}{\omega_n}$. Consider the system

$$\ddot{y} + f_{\mathrm{d}}(\dot{y}) + \omega_n^2 y = 0 \tag{3.149}$$

i.e. the free decay of a SDOF oscillator with nonlinear damping. (The extension to stiffness or mixed nonlinearities is straightforward.) A coordinate transformation $(y(t), \dot{y}(t)) \longrightarrow (Y(t), \phi(t))$ is defined using (3.148) supplemented by

$$\dot{y}(t) = Y(t)\omega_n \cos(\omega_n t + \phi(t)). \tag{3.150}$$

Now, this transformation is inconsistent as it stands. The required consistency condition is obtained by differentiating (3.148) with respect to t and equating to (3.150), the result is

$$\begin{aligned}\dot{Y}(t)\sin(\omega_n t + \phi(t)) + Y(t)\omega_n \cos(\omega_n t + \phi(t)) + Y(t)\dot{\phi}(t)\cos(\omega_n t + \phi(t)) \\ = Y(t)\omega_n \cos(\omega_n t + \phi(t))\end{aligned} \tag{3.151}$$

or

$$\dot{Y}(t)\sin(\omega_n t + \phi(t)) + Y(t)\dot{\phi}(t)\cos(\omega_n t + \phi(t)) = 0. \tag{3.152}$$

Once this equation is established, (3.150) can be differentiated to yield the acceleration

$$\begin{aligned}\ddot{y}(t) = \dot{Y}(t)\omega_n \cos(\omega_n t + \phi(t)) - Y(t)\omega_n^2 \sin(\omega_n t + \phi(t)) \\ - Y(t)\dot{\phi}(t)\omega_n \sin(\omega_n t + \phi(t)).\end{aligned} \tag{3.153}$$

Now, substituting (3.148), (3.150) and (3.153) into the equation of motion (3.149) yields

$$\begin{aligned}\dot{Y}(t)\omega_n \cos(\omega_n t + \phi(t)) - Y(y)\dot{\phi}(t)\omega_n \sin(\omega_n t + \phi(t)) \\ = -f_{\mathrm{d}}(\omega_n Y(t)\cos(\omega_n t + \phi(t)))\end{aligned} \tag{3.154}$$

and multiplying (3.152) by $\omega_n \sin(\omega_n t + \phi(t))$, (3.154) by $\cos(\omega_n t + \phi(t))$ and adding the results gives

$$\dot{Y}(t) = -\frac{1}{\omega_n} f_{\mathrm{d}}(\omega_n Y(t)\cos(\omega_n t + \phi(t)))\cos(\omega_n t + \phi(t)) \tag{3.155}$$

while multiplying (3.152) by $\omega_n \cos(\omega_n t + \phi(t))$, (3.154) by $\sin(\omega_n t + \phi(t))$ and differencing yields

$$\dot{\phi}(t) = -\frac{1}{\omega_n Y} f_{\mathrm{d}}(\omega_n Y(t)\cos(\omega_n t + \phi(t)))\sin(\omega_n t + \phi(t)). \tag{3.156}$$

These equations together are exactly equivalent to (3.149). Unfortunately, they are just as difficult to solve. However, if one makes use of the fact that $Y(t)$

and $\phi(t)$ are essentially constant over one period τ_n, the right-hand sides of the equations can be approximately replaced by an average over one cycle, so

$$\dot{Y}(t) = -\frac{1}{2\pi\omega_n}\int_0^{2\pi} \mathrm{d}\theta\, f_\mathrm{d}(\omega_n Y\cos(\theta+\phi))\cos(\theta+\phi) \tag{3.157}$$

$$\dot{\phi}(t) = -\frac{1}{2\pi\omega_n Y}\int_0^{2\pi} \mathrm{d}\theta\, f_\mathrm{d}(\omega_n Y\cos(\theta+\phi))\sin(\theta+\phi) \tag{3.158}$$

and it is understood that Y and ϕ are treated as constants when the integrals are evaluated. In order to see how these equations are used, two cases of interest will be examined.

3.12.2 Linear damping

In this case

$$f_\mathrm{d}(\dot{y}) = 2\zeta\omega_n\dot{y}. \tag{3.159}$$

Equation (3.157) gives

$$\dot{Y}(t) = -\frac{1}{2\pi\omega_n}\int_0^{2\pi} \mathrm{d}\theta\, 2\zeta\omega_n^2 Y\cos^2(\theta+\phi) \tag{3.160}$$

a simple integral, which yields

$$\dot{Y} = -\zeta\omega_n Y \tag{3.161}$$

so that

$$Y(t) = Y_0\mathrm{e}^{-\zeta\omega_n t}. \tag{3.162}$$

Equation (3.158) gives

$$\dot{\phi}(t) = -\frac{1}{2\pi\omega_n Y}\int_0^{2\pi} \mathrm{d}\theta\, \zeta\omega_n^2 Y\cos(\theta+\phi)\sin(\theta+\phi) = 0 \tag{3.163}$$

so

$$\phi(t) = \phi_0 \tag{3.164}$$

and the overall solution for the motion is

$$y(t) = Y_0\mathrm{e}^{-\zeta\omega_n t}\sin(\omega_n t+\phi_0) \tag{3.165}$$

which agrees with the exact solution for a linear system. The decay is exponential as required.

3.12.3 Coulomb friction

In this case

$$f_{\mathrm{d}}(\dot{y}) = c_{\mathrm{F}}\,\mathrm{sgn}(\dot{y}) \tag{3.166}$$

and equation (3.157) gives

$$\dot{Y}(t) = -\frac{1}{2\pi\omega_n}\int_0^{2\pi} \mathrm{d}\theta\, c_{\mathrm{F}}\,\mathrm{sgn}(\cos(\theta+\phi))\cos(\theta+\phi) \tag{3.167}$$

so (ignoring ϕ, as the integral is over a whole cycle)

$$\dot{Y}(t) = -\frac{c_{\mathrm{F}}}{2\pi\omega_n}\left\{2\int_0^{\frac{\pi}{2}} \mathrm{d}\theta\,\cos\theta - \int_{\frac{\pi}{2}}^{\frac{3\pi}{2}} \mathrm{d}\theta\,\cos\theta\right\} \tag{3.168}$$

and

$$\dot{Y} = -\frac{2c_{\mathrm{F}}}{\pi\omega_n} \tag{3.169}$$

which integrates trivially to give

$$Y(t) = Y_0 - \frac{2c_{\mathrm{F}}}{\pi\omega_n}t. \tag{3.170}$$

Equation (3.158) gives

$$\dot{\phi}(t) = -\frac{1}{2\pi\omega_n Y}\int_0^{2\pi} \mathrm{d}\theta\, c_{\mathrm{F}}\,\mathrm{sgn}(\cos\theta)\sin\theta = 0 \tag{3.171}$$

so the final solution has

$$\phi(t) = \phi_0. \tag{3.172}$$

Equation (3.170) shows that the expected form of the decay envelope for a Coulomb friction system is linear (figure 3.29). This is found to be the case by simulation or experiment.

It transpires that for SDOF systems at least, the form of the envelope suffices to fix the form of the nonlinear damping *and* stiffness functions. The relevant method of identification requires the use of the Hilbert transform, so the discussion is postponed until the next chapter.

3.13 Summary

Harmonic balance is a useful technique for deriving the describing functions or FRFs of nonlinear systems if the nonlinear differential equation of the system is known. The method of slowly varying amplitude and phase similarly suffices to estimate the decay envelopes. In fact, many techniques exist which agree with these methods to the first-order approximations presented in this chapter. Among them are: perturbation methods [197], multiple scales [196], Galerkin's method

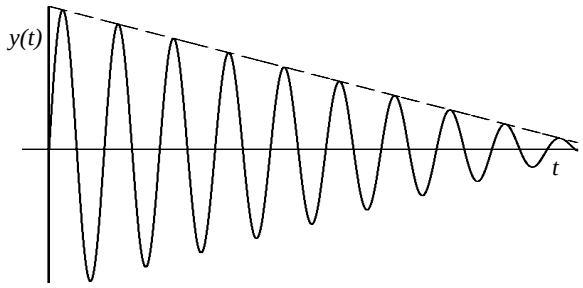


Figure 3.29. Envelope for SDOF Coulomb friction system.

[76] and normal forms [125]. Useful graphical techniques also exist like the method of isoclines or Liénard's method [196]. Other more convenient methods of calculating the strength of harmonics can be given, once the Volterra series is defined in chapter 8.

Chapter 4

The Hilbert transform—a practical approach

4.1 Introduction

The Hilbert Transform is a mathematical tool which allows one to investigate the causality, stability and linearity of passive systems. In this chapter its main application will be to the detection and identification of nonlinearity. The theory can be derived by two independent approaches: the first, which is the subject of this chapter, relies on the decomposition of a function into odd and even parts and the behaviour of this decomposition under Fourier transformation. The second method is more revealing but more complicated, relying as it does on complex analysis; discussion of this is postponed until the next chapter.

The Hilbert transform is an integral transform of the same family as the Fourier transform, the difference is in the kernel function. The complex exponential $\mathrm{e}^{\mathrm{i}\omega t}$ is replaced by the function $-1/\mathrm{i}\pi(\Omega-\omega)$, so if the Hilbert transform operator is denoted by $\mathcal{H}$, its action on functions [1] is given by[2]

$$\mathcal{H}\{G(\omega)\} = \tilde{G}(\omega) = -\frac{1}{\mathrm{i}\pi} PV \int_{-\infty}^{\infty} \mathrm{d}\Omega \, \frac{G(\Omega)}{\Omega-\omega} \tag{4.1}$$

where PV denotes the Cauchy principal value of the integral, and is needed as the integrand is singular, i.e. has a pole at $\omega = \Omega$. To maintain simplicity of notation, the PV will be omitted in the following discussions, as it will be clear from the integrands, which expressions need it. The tilde ˜ is used to denote the transformed function.

[1] In this chapter and the following the functions of interest will generally be denoted $g(t)$ and $G(\omega)$ to indicate that the objects are not necessarily from linear or nonlinear systems. Where it is important to make a distinction $h(t)$ and $H(\omega)$ will be used for linear systems and $\lambda(t)$ and $\Lambda(\omega)$ will be used for nonlinear.

[2] This differs from the original transform defined by Hilbert and used by mathematicians, by the introduction of a prefactor $-1/\mathrm{i} = \mathrm{i}$. It will become clear later why the additional constant is useful.

The Hilbert transform and Fourier transform also differ in their interpretation. The Fourier transform is considered to map functions of time to functions of frequency and *vice versa*. In contrast, the Hilbert transform is understood to map functions of time or frequency into the same *domain*, i.e.

$$\mathcal{H}\{G(\omega)\} = \tilde{G}(\omega) \tag{4.2}$$

$$\mathcal{H}\{g(t)\} = \tilde{g}(t). \tag{4.3}$$

The Hilbert transform has long been the subject of study by mathematicians, a nice pedagogical study can be found in [204]. In recent times it has been adopted as a useful tool in signal processing, communication theory and linear dynamic testing. A number of relevant references are [24, 43, 49, 65, 81, 89, 105, 116, 126, 130, 151, 210, 211, 247, 255]. The current chapter is intended as a survey of the Hilbert transform's recent use in the testing and identification of nonlinear structures.

4.2 Basis of the method

4.2.1 A relationship between real and imaginary parts of the FRF

The discussion begins with a function of time $g(t)$ which has the property that $g(t) = 0$ when $t < 0$. By a slight abuse of terminology, such functions will be referred to henceforth as *causal*.

Given *any* function $g(t)$, there is a decomposition

$$g(t) = g_{\mathrm{even}}(t) + g_{\mathrm{odd}}(t) = \tfrac{1}{2}(g(t) + g(-t)) + \tfrac{1}{2}(g(t) - g(-t)) \tag{4.4}$$

as depicted in figure 4.1. If, in addition, $g(t)$ is causal, it follows that

$$g_{\mathrm{even}}(t) = \begin{cases} g(|t|)/2, & t > 0 \\ g(|t|)/2, & t < 0 \end{cases} \tag{4.5}$$

and

$$g_{\mathrm{odd}}(t) = \begin{cases} g(|t|)/2, & t > 0 \\ -g(|t|)/2, & t < 0\,. \end{cases} \tag{4.6}$$

That this is only true for causal functions is shown by the simple counterexample in figure 4.2. It follows immediately from equations (4.5) and (4.6) that

$$g_{\mathrm{even}}(t) = g_{\mathrm{odd}}(t) \times \epsilon(t) \tag{4.7}$$

$$g_{\mathrm{odd}}(t) = g_{\mathrm{even}}(t) \times \epsilon(t) \tag{4.8}$$

where $\epsilon(t)$ is the *signum* function, defined by [3]

$$\epsilon(t) = \begin{cases} 1, & t > 0 \\ 0, & t = 0 \\ -1, & t < 0. \end{cases} \tag{4.9}$$

[3] The notation $\mathrm{sgn}(t)$ is often used.

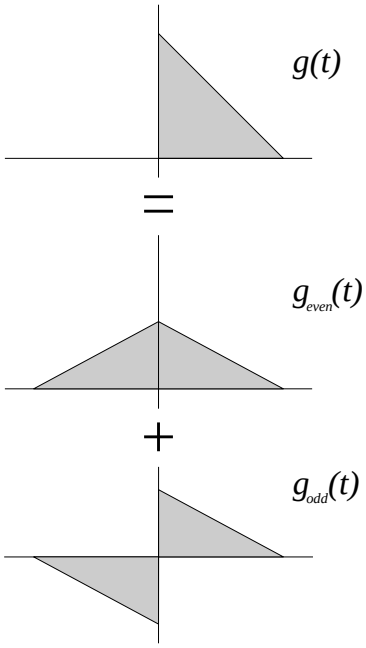


Figure 4.1. Decomposition of a causal function into odd and even parts.

Assuming that the Fourier transform of $g(t)$ is defined, it is straightforward to show that

$$\operatorname{Re} G(\omega) = \mathcal{F}\{g_{\mathrm{even}}(t)\} \tag{4.10}$$

and

$$\operatorname{Im} G(\omega) = \mathcal{F}\{g_{\mathrm{odd}}(t)\}. \tag{4.11}$$

Substituting equations (4.7) and (4.8) into this expression yields

$$\operatorname{Re} G(\omega) = \mathcal{F}\{g_{\mathrm{odd}}(t) \times \epsilon(t)\} \tag{4.12}$$

$$\operatorname{Im} G(\omega) = \mathcal{F}\{g_{\mathrm{even}}(t) \times \epsilon(t)\}. \tag{4.13}$$

Now, noting that multiplication of functions in the time domain corresponds to convolution in the frequency domain, and that $\mathcal{F}\{\epsilon(t)\} = -\mathrm{i}/\pi\omega$ (see appendix D), equations (4.12) and (4.13) become

$$\operatorname{Re} G(\omega) = \mathrm{i} \operatorname{Im} G(\omega) * -\frac{\mathrm{i}}{\pi\omega} \tag{4.14}$$

$$\operatorname{Im} G(\omega) = \operatorname{Re} G(\omega) * -\frac{\mathrm{i}}{\pi\omega}. \tag{4.15}$$

Using the standard definition of convolution,

$$X(\omega) * Y(\omega) = \int_{-\infty}^{\infty} \mathrm{d}\Omega\, X(\Omega) Y(\omega - \Omega). \tag{4.16}$$

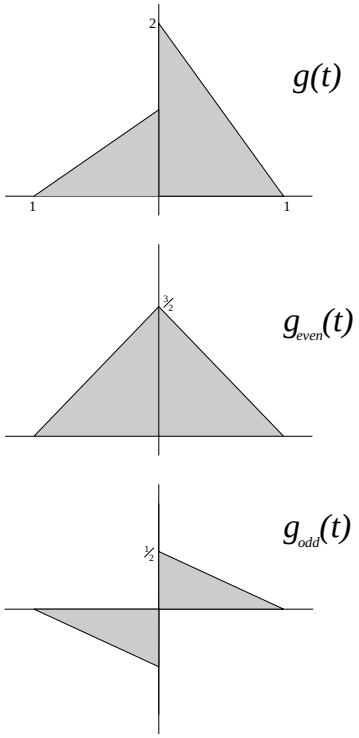


Figure 4.2. Counterexample decomposition for a non-causal function.

Equations (4.14) and (4.15) can be brought into the final forms

$$\operatorname{Re} G(\omega) = -\frac{1}{\pi}\int_{-\infty}^{\infty} \mathrm{d}\Omega\, \frac{\operatorname{Im} G(\Omega)}{\Omega - \omega} \tag{4.17}$$

$$\operatorname{Im} G(\omega) = +\frac{1}{\pi}\int_{-\infty}^{\infty} \mathrm{d}\Omega\, \frac{\operatorname{Re} G(\Omega)}{\Omega - \omega}. \tag{4.18}$$

It follows from these expressions that the real and imaginary parts of a function $G(\omega)$, the Fourier transform of a causal function $g(t)$, are not independent. Given one quantity, the other is uniquely specified. (Recall that these integrals are *principal value* integrals.) Equations (4.17) and (4.18) can be combined into a single complex expression by forming $G(\omega) = \operatorname{Re} G(\omega) + \mathrm{i} \operatorname{Im} G(\omega)$, the result is

$$G(\omega) = -\frac{1}{\mathrm{i}\pi}\int_{-\infty}^{\infty} \mathrm{d}\Omega\, \frac{G(\Omega)}{\Omega - \omega}. \tag{4.19}$$

Now, applying the definition of the Hilbert transform in equation (4.1) yields

$$G(\omega) = \tilde{G}(\omega) = \mathcal{H}\{G(\omega)\}. \tag{4.20}$$

So $G(\omega)$, the Fourier transform of a causal $g(t)$, is invariant under the Hilbert transform and $\mathrm{Re}\,G(\omega)$ and $\mathrm{Im}\,G(\omega)$ are said to form a *Hilbert transform pair*. Now, recall from chapter 1 that the impulse response function $h(t)$ of a linear system is causal, this implies that the Fourier transform of $h(t)$—the FRF $H(\omega)$—is invariant under Hilbert transformation. It is this property which will be exploited in later sections in order to detect nonlinearity as FRFs from nonlinear systems are not guaranteed to have this property.

Further simplifications to these formulae follow from a consideration of the parity (odd or even) of the functions $\mathrm{Re}\,G(\omega)$ and $\mathrm{Im}\,G(\omega)$. In fact, $\mathrm{Re}\,G(\omega)$ is even

$$\mathrm{Re}\,G(-\omega) = \int_{-\infty}^{\infty} \mathrm{d}t\, g(t)\cos(-\omega t) = \int_{-\infty}^{\infty} \mathrm{d}t\, g(t)\cos(\omega t) = \mathrm{Re}\,G(\omega t) \tag{4.21}$$

and $\mathrm{Im}\,G(\omega)$ is odd or *conjugate-even*

$$\begin{aligned} \mathrm{Im}\,G(-\omega) &= \int_{-\infty}^{\infty} \mathrm{d}t\, g(t)\sin(-\omega t) = -\int_{-\infty}^{\infty} \mathrm{d}t\, g(t)\sin(\omega t) \\ &= -\mathrm{Im}\,G(\omega) = \mathrm{Im}\,\overline{G(\omega)} \end{aligned} \tag{4.22}$$

where the overline denotes complex conjugation.

Using the parity of $\mathrm{Im}\,G(\omega)$, equation (4.17) can be rewritten:

$$\begin{aligned} \mathrm{Re}\,G(\omega) &= -\frac{1}{\pi}\int_{-\infty}^{\infty} \mathrm{d}\Omega\, \frac{\mathrm{Im}\,G(\Omega)}{\Omega-\omega} \\ &= -\frac{1}{\pi}\left[\int_{-\infty}^{0} \mathrm{d}\Omega\, \frac{\mathrm{Im}\,G(\Omega)}{\Omega-\omega} + \int_{0}^{\infty} \mathrm{d}\Omega\, \frac{\mathrm{Im}\,G(\Omega)}{\Omega-\omega}\right] \\ &= -\frac{1}{\pi}\left[-\int_{0}^{-\infty} \mathrm{d}\Omega\, \frac{\mathrm{Im}\,G(\Omega)}{\Omega-\omega} + \int_{0}^{\infty} \mathrm{d}\Omega\, \frac{\mathrm{Im}\,G(\Omega)}{\Omega-\omega}\right] \\ &= -\frac{1}{\pi}\left[\int_{0}^{\infty} \mathrm{d}\Omega\, \frac{\mathrm{Im}\,G(-\Omega)}{-\Omega-\omega} + \int_{0}^{\infty} \mathrm{d}\Omega\, \frac{\mathrm{Im}\,G(\Omega)}{\Omega-\omega}\right] \\ &= -\frac{1}{\pi}\left[\int_{0}^{\infty} \mathrm{d}\Omega\, \frac{\mathrm{Im}\,G(\Omega)}{\Omega+\omega} + \int_{0}^{\infty} \mathrm{d}\Omega\, \frac{\mathrm{Im}\,G(\Omega)}{\Omega-\omega}\right] \\ &= -\frac{2}{\pi}\int_{0}^{\infty} \mathrm{d}\Omega\, \frac{\mathrm{Im}\,G(\Omega)\Omega}{\Omega^2-\omega^2} \end{aligned} \tag{4.23}$$

and similarly

$$\mathrm{Im}\,G(\omega) = \frac{2\omega}{\pi}\int_{0}^{\infty} \mathrm{d}\Omega\, \frac{\mathrm{Re}\,G(\Omega)}{\Omega^2-\omega^2}. \tag{4.24}$$

These equations are often referred to as the Kramers–Kronig relations [154]. The advantage of these forms over (4.17) and (4.18) is simply that the range of integration is halved and one of the infinite limits is removed.

4.2.2 A relationship between modulus and phase

Suppose $G(\omega)$, the Fourier transform of causal $g(t)$, is expressed in terms of gain and phase:

$$G(\omega) = |G(\omega)|\mathrm{e}^{\mathrm{i}\phi(\omega)} \tag{4.25}$$

where

$$|G(\omega)| = \sqrt{(\mathrm{Re}\,G(\omega))^2 + (\mathrm{Im}\,G(\omega))^2} \tag{4.26}$$

and

$$\phi(\omega) = \tan^{-1}\left(\frac{\mathrm{Im}\,G(\omega)}{\mathrm{Re}\,G(\omega)}\right). \tag{4.27}$$

Taking the natural logarithm of (4.25) yields[4]

$$\log G(\omega) = \log|G(\omega)| + \mathrm{i}\phi(\omega). \tag{4.28}$$

Unfortunately, $\log|G(\omega)|$ and $\phi(\omega)$, as they stand, do not form a Hilbert transform pair. However, it can be shown that the function $(\log G(\omega) - \log G(0))/\omega$ is invariant under the transform and so the functions $(\log|G(\omega)| - \log|G(0)|)/\omega$ and $(\phi(\omega) - \phi(0))/\omega$ do form such a pair. If in addition, the minimum phase condition, $\phi(0) = 0$, is assumed, the Hilbert transform relations can be written:

$$\log|G(\omega)| - \log|G(0)| = -\frac{2\omega^2}{\pi}\int_0^\infty \mathrm{d}\Omega\,\frac{\phi(\Omega)}{\Omega(\Omega^2 - \omega^2)} \tag{4.29}$$

$$\phi(\omega) = \frac{2\omega}{\pi}\int_0^\infty \mathrm{d}\Omega\,\frac{\log|G(\omega)| - \log|G(0)|}{\Omega^2 - \omega^2}. \tag{4.30}$$

The effort involved in deriving these equations rigorously is not justified as they shall play no further part in the development; they are included mainly for completeness. They are of some interest as they allow the derivation of FRF phase from FRF modulus information, which is available if one has some means of obtaining auto-power spectra as

$$|H(\omega)| = \sqrt{\frac{S_{yy}(\omega)}{S_{xx}(\omega)}}. \tag{4.31}$$

4.3 Computation

Before proceeding to applications of the Hilbert transform, some discussion of how to compute the transform is needed. Analytical methods are not generally applicable; nonlinear systems will provide the focus of the following discussion and closed forms for the FRFs of nonlinear systems are not usually available. Approximate FRFs, e.g. from harmonic balance (see chapter 3), lead to integrals

[4] Assuming the principal sheet for the log function.

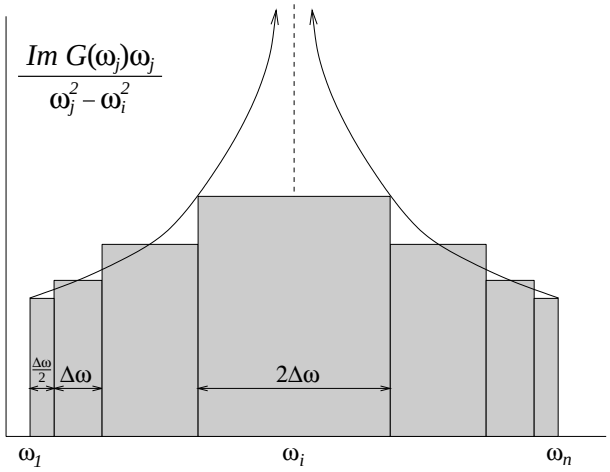


Figure 4.3. Integration mesh for direct Hilbert transform evaluation.

(4.1) which cannot be evaluated in closed form. It is therefore assumed that a vector of sampled FRF values $G(\omega_i); i = 1, \ldots, N$, will constitute the available data, and numerical methods will be applied. For simplicity, equal spacing $\Delta\omega$, of the data will be assumed. A number of methods for computing the transform are discussed in this section.

4.3.1 The direct method

This, the most direct approach, seeks to estimate the frequency-domain integrals (4.17) and (4.18). In practice, the Kramers–Kronig relations (4.23) and (4.24) are used as the range of integration is simplified. Converting these expressions to discrete sums yields

$$\mathrm{Re}\,\tilde{G}(\omega_i) = -\frac{2}{\pi}\sum_{j=1}^{N}\frac{\mathrm{Im}\,G(\omega_j)\omega_j}{\omega_j^2-\omega_i^2}\Delta\omega \tag{4.32}$$

$$\mathrm{Im}\,\tilde{G}(\omega_i) = -\frac{2\omega_i}{\pi}\sum_{j=1}^{N}\frac{\mathrm{Re}\,G(\omega_j)}{\omega_j^2-\omega_i^2}\Delta\omega \tag{4.33}$$

and some means of avoiding the singularity at $\omega_i = \omega_j$ is needed. This approximation is the well-known *rectangle rule*. It can be lifted in accuracy to the *trapezium rule* with very little effort. The rectangular sub-areas should be summed as in figure 4.3 with half-width rectangles at the ends of the range. The singularity is avoided by taking a double-width step. The effect of the latter strategy can be ignored if $\Delta\omega$ is appropriately small.

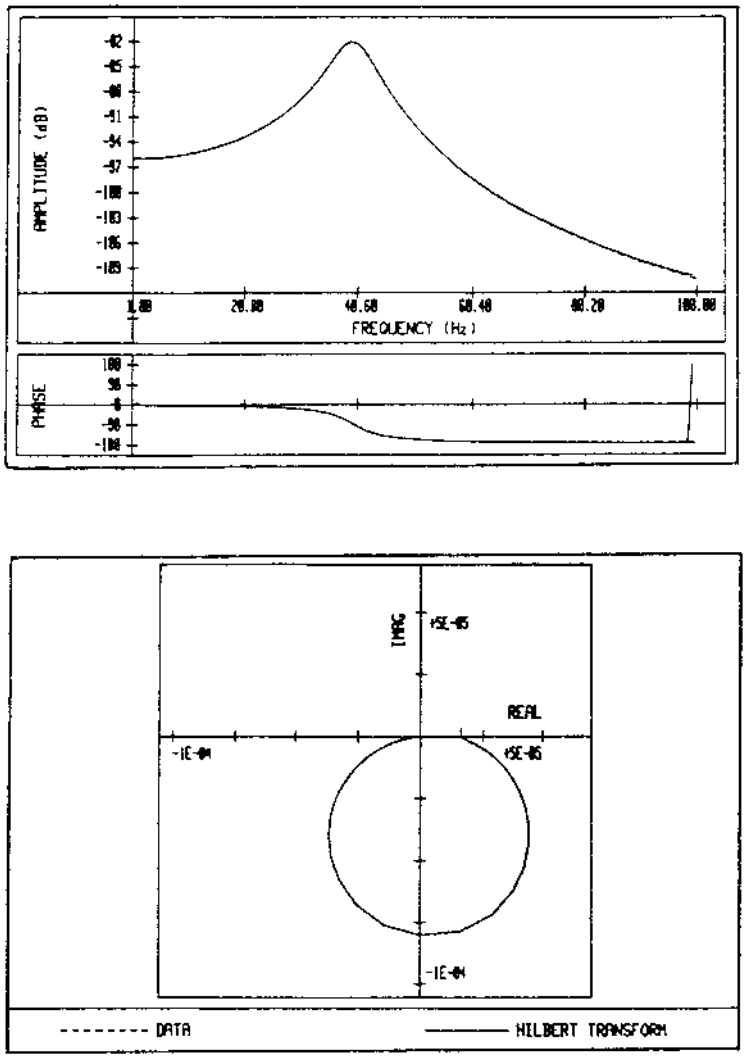


Figure 4.4. Hilbert transform of a simulated SDOF linear system showing perfect overlay.

Figure 4.4 shows a linear system FRF with the Hilbert transform superimposed. Almost perfect overlay is obtained. However, there is an important assumption implicit in this calculation, i.e. that $\omega_1 = 0$ and that ω_N can be substituted for the infinite upper limit of the integral with impunity. If the integrals from 0 to ω_1 or ω_N to infinity in (4.23) and (4.24) are non-zero, the estimated Hilbert transform is subject to *truncation errors*. Figure 4.5 shows the effect of truncation on the Hilbert transform of a zoomed linear system FRF.

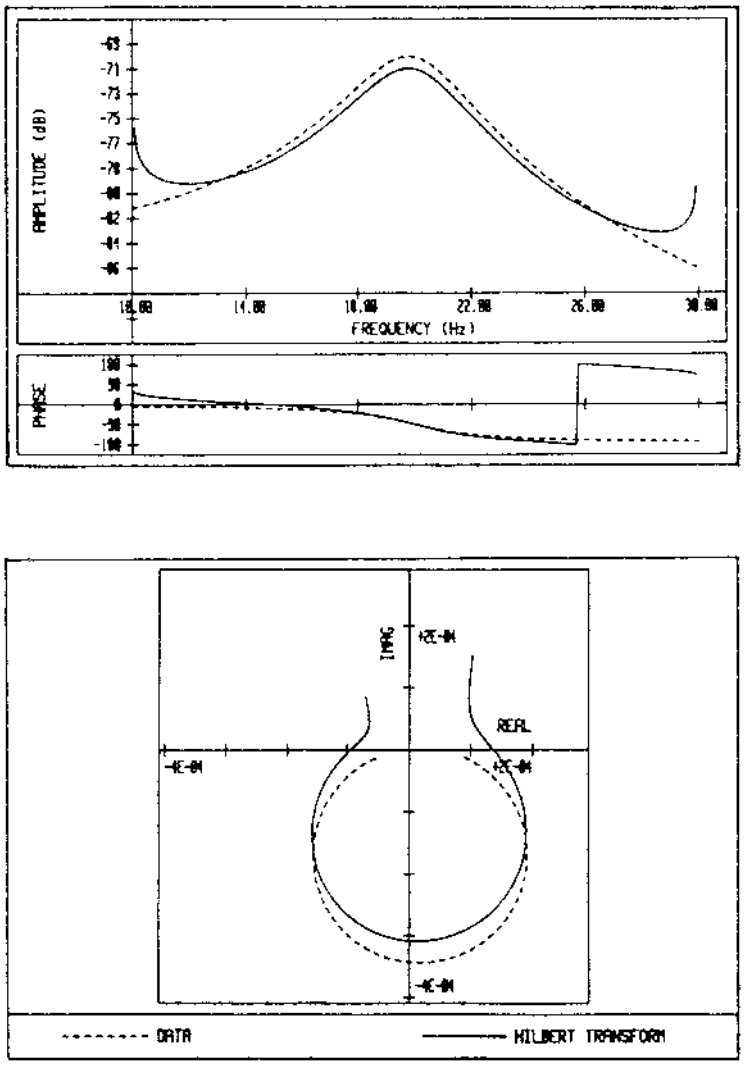


Figure 4.5. Hilbert transform of a simulated SDOF linear system showing truncation problems.

4.3.2 Correction methods for truncated data

There are essentially five methods of correcting Hilbert transforms for truncation errors, they will now be described in order of complexity.

4.3.2.1 Conversion to receptance

This correction is only applicable to data with $\omega_1 = 0$, commonly referred to as *baseband* data. The principle is very simple; as the high-frequency decay of receptance FRF data is faster ($\mathrm{O}(\omega^{-2})$) than mobility or accelerance data ($\mathrm{O}(\omega^{-1})$ and O(1) respectively), the high-frequency truncation error for the latter forms of the FRF is reduced by initially converting them to receptance, carrying out the Hilbert transform, and converting them back. The relations between the forms are

$$H_{\mathrm{I}}(\omega) = \mathrm{i}\omega H_{\mathrm{M}}(\omega) = -\omega^2 H_{\mathrm{R}}(\omega). \tag{4.34}$$

4.3.2.2 The Fei correction term

This approach was developed by Fei [91] for baseband data and is based on the asymptotic behaviour of the FRFs of linear systems. The form of the correction term is entirely dependent on the FRF type; receptance, mobility or accelerance. As each of the correction terms is similar in principle, only the term for mobility will be described.

The general form of the mobility function for a linear system with proportional damping is

$$H_{\mathrm{M}}(\omega) = \sum_{k=1}^{N} \frac{\mathrm{i}\omega A_k}{\omega_k^2 - \omega^2 + \mathrm{i}2\zeta_k\omega_k\omega} \tag{4.35}$$

where A_k is the complex modal amplitude of the kth mode; ω_k is the undamped natural frequency of the kth mode and ζ_k is its viscous damping ratio. By assuming that the damping is small and that the truncation frequency, $\omega_{\max}$, is much higher than the natural frequency of the highest mode, equation (4.35) can be reduced to (for $\omega > \omega_{\max}$)

$$H_{\mathrm{M}}(\omega) = -\mathrm{i}\sum_{k=1}^{N} \frac{A_k}{\omega} \tag{4.36}$$

which is an approximation to the 'out-of-band' FRF. This term is purely imaginary and thus provides a correction for the real part of the Hilbert transform via equation (4.32). No correction term is applied to the imaginary part as the error is assumed to be small under the specified conditions.

The actual correction is the integral in equation (4.1) over the interval $(\omega_{\max}, \infty)$. Hence the correction term, denoted $C^{\mathrm{R}}(\omega)$, for the real part of the Hilbert transform is

$$C^{\mathrm{R}}(\omega) = -\frac{2}{\pi}\int_{w_{\max}}^{\infty} \mathrm{d}\Omega\, \frac{\Omega\,\mathrm{Im}(G(\Omega))}{\Omega^2 - \omega^2} = -\frac{2}{\omega}\sum_{k=1}^{N} A_k \int_{\omega_{\max}}^{\infty} \frac{\mathrm{d}\Omega}{\Omega^2 - \omega^2} \tag{4.37}$$

which, after a little algebra [91], leads to

$$C^{\mathrm{R}}(\omega) = \frac{-\omega_{\max}\,\mathrm{Im}(G(\omega_{\max}))}{\pi\omega}\ln\left(\frac{\omega_{\max}+\omega}{\omega_{\max}-\omega}\right). \tag{4.38}$$

4.3.2.3 The Haoui correction term

The second correction term, which again, caters specifically for baseband data, is based on a different approach. The term was developed by Haoui [130], and unlike the Fei correction has a simple expression independent of the type of FRF data used. The correction for the real part of the Hilbert transform is

$$C^{\mathrm{R}}(\omega) = -\frac{2}{\pi}\int_{w_{\max}}^{\infty}\mathrm{d}\Omega\,\frac{\Omega\,\mathrm{Im}(G(\Omega))}{\Omega^2-\omega^2}. \tag{4.39}$$

The analysis proceeds by assuming a Taylor expansion for $G(\omega)$ about $\omega_{\max}$ and expanding the term $(1-\omega^2/\Omega^2)^{-1}$ using the binomial theorem. If it is assumed that $\omega_{\max}$ is not close to a resonance so that the slope $\mathrm{d}G(\omega)/\mathrm{d}\omega$ (and higher derivatives) can be neglected, a straightforward calculation yields

$$C^{\mathrm{R}}(\omega) = C^{\mathrm{R}}(0) - \frac{\mathrm{Im}(G(\omega_{\max}))}{\pi}\left[\frac{\omega^2}{\omega_{\max}^2}+\frac{\omega^4}{2\omega_{\max}^4}+\cdots\right] \tag{4.40}$$

where $C^{\mathrm{R}}(0)$ is estimated from

$$C^{\mathrm{R}}(0) = \mathrm{Re}(G(0)) - \left[-\frac{2}{\pi}\int_{0+\epsilon}^{w_{\max}}\mathrm{d}\Omega\,\frac{\mathrm{Im}(G(\Omega))}{\Omega}\right]. \tag{4.41}$$

Using the same approach, the correction term for the imaginary part, denoted by $C^{\mathrm{I}}(\omega)$, can be obtained:

$$C^{\mathrm{I}}(\omega_{\max}) = \frac{2}{\pi}\,\mathrm{Re}(G(\omega_{\max}))\left[\frac{\omega}{\omega_{\max}}+\frac{\omega^3}{3\omega_{\max}^3}+\frac{\omega^5}{5\omega_{\max}^5}+\cdots\right]. \tag{4.42}$$

4.3.2.4 The Simon correction method

This method of correction was proposed by Simon [229]; it allows for truncation at a low frequency, $\omega_{\min}$ and a high frequency $\omega_{\max}$. It is therefore suitable for use with zoomed data. This facility makes the method the most versatile so far. As before, it is based on the behaviour of the linear FRF, say equation (4.35) for mobility data. Splitting the Hilbert transform over three frequency ranges: $(0,\omega_{\min})$, $(\omega_{\min},\omega_{\max})$ and $(\omega_{\max},\infty)$, the truncation errors on the real part of the Hilbert transform, $B^{\mathrm{R}}(\omega)$ at low frequency and the now familiar $C^{\mathrm{R}}(\omega)$ at high frequency, can be written as

$$B^{\mathrm{R}}(\omega) = -\frac{2}{\pi}\int_{0}^{\omega_{\min}}\mathrm{d}\Omega\,\frac{\Omega\,\mathrm{Im}(G(\omega))}{\Omega^2-\omega^2} \tag{4.43}$$

and

$$C^{\mathrm{R}}(\omega) = -\frac{2}{\pi}\int_{\omega_{\max}}^{\infty} \mathrm{d}\Omega\, \frac{\Omega\,\mathrm{Im}(G(\omega))}{\Omega^2 - \omega^2}. \tag{4.44}$$

If the damping can be assumed to be small, then rewriting equations (4.40) and (4.44) using the mobility form (4.35) yields

$$B^{\mathrm{R}}(\omega) = \frac{2}{\pi}\int_{0}^{\omega_{\min}} \mathrm{d}\Omega \sum_{k=1}^{N} \frac{\Omega^2 A_k}{(\Omega^2 - \omega_k^2)(\Omega^2 - \omega^2)} \tag{4.45}$$

and

$$C^{\mathrm{R}}(\omega) = \frac{2}{\pi}\int_{\omega_{\max}}^{\infty} \mathrm{d}\Omega \sum_{k=1}^{N} \frac{\Omega^2 A_k}{(\Omega^2 - \omega_k^2)(\Omega^2 - \omega^2)}. \tag{4.46}$$

Evaluating these integrals gives

$$\begin{aligned} B^{\mathrm{R}}(\omega) + C^{\mathrm{R}}(\omega) = & -\sum_{k=1}^{N} \frac{A_k}{\pi(\omega^2 - \omega_k^2)} \left\{ \omega_k \ln\left(\frac{(\omega_{\max} + \omega_k)(\omega_k - \omega_{\min})}{(\omega_{\max} - \omega_k)(\omega_k + \omega_{\min})} \right) \right. \\ & \left. + \omega \ln\left(\frac{(\omega + \omega_{\min})(\omega_{\max} - \omega)}{(\omega - \omega_{\min})(\omega_{\max} + \omega)} \right) \right\}. \end{aligned} \tag{4.47}$$

The values of the modal parameters A_k and ω_k are obtained from an initial modal analysis.

4.3.2.5 The Ahmed correction term

This is the most complex correction term theoretically, but also the most versatile. It is applicable to zoomed data and, like the Simon correction term, assumes that the FRF takes the linear form away from resonance. The form of the correction depends on the FRF type; to illustrate the theory the mobility form (4.35) will be assumed. The form (4.35) gives real and imaginary parts:

$$\mathrm{Re}\, H_{\mathrm{M}}(\omega) = \sum_{k=1}^{N} \frac{2A_k \zeta_k \omega_k \omega^2}{(\omega_k^2 - \omega^2)^2 + 4\zeta_k^2 \omega_k^2 \omega^2} \tag{4.48}$$

$$\mathrm{Im}\, H_{\mathrm{M}}(\omega) = \sum_{k=1}^{N} \frac{A_k \omega (\omega_k^2 - \omega^2)}{(\omega_k^2 - \omega^2)^2 + 4\zeta_k^2 \omega_k^2 \omega^2}. \tag{4.49}$$

So, assuming that the damping can be neglected away from resonant regions,

$$\mathrm{Re}\, H_{\mathrm{M}}(\omega) = \sum_{k=1}^{N} \frac{2A_k \zeta_k \omega_k \omega^2}{(\omega_k^2 - \omega^2)^2} \tag{4.50}$$

$$\mathrm{Im}\, H_{\mathrm{M}}(\omega) = \sum_{k=1}^{N} \frac{A_k \omega}{(\omega_k^2 - \omega^2)^2}. \tag{4.51}$$

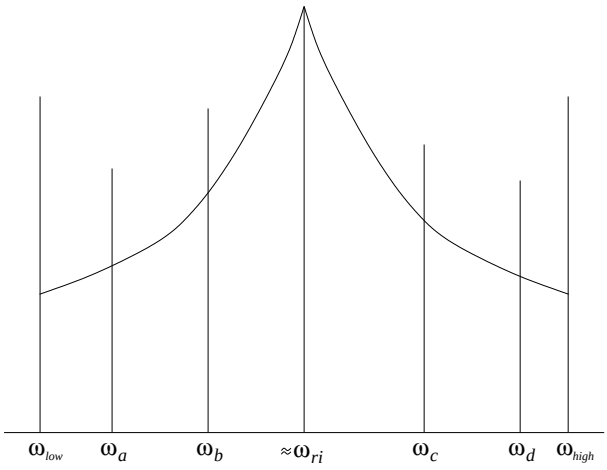


Figure 4.6. Frequency grid for the Ahmed correction term.

Suppose mode i is the lowest mode in the measured region with resonant frequency ω_{ri} and therefore has the greatest effect on the low-frequency truncation error $B^{\mathrm{R}}(\omega)$, the relevant part of $\operatorname{Im} H_m$ can be decomposed:

$$\operatorname{Im} H_{\mathrm{M}}^{m}(\omega) = \sum_{k=1}^{i-1} \frac{A_k \omega}{(\omega_k^2 - \omega^2)^2} + \frac{A_i \omega}{(\omega_i^2 - \omega^2)^2} \tag{4.52}$$

where the superscript m indicates that this is the *mass asymptote* of the FRF. In the lower part of the frequency range ω/ω_k is small and the first term can be expanded:

$$\sum_{k=1}^{i-1} \frac{A_k \omega}{(\omega_k^2 - \omega^2)^2} = \sum_{k=1}^{i-1} A_k \frac{\omega}{\omega_k} \left[1 + \left(\frac{\omega}{\omega_k} \right)^2 + \cdots \right] = \mathrm{O}\left(\left(\frac{\omega}{\omega_k} \right) \right) \tag{4.53}$$

and neglected, so

$$\operatorname{Im} H_{\mathrm{M}}^{m}(\omega) \approx \frac{A_i \omega}{(\omega_i^2 - \omega^2)^2}. \tag{4.54}$$

Now, Ahmed estimates the unknown coefficient A_i by curve-fitting function (4.54) to the data in the range ω_{a} to ω_{b} where $\omega_{\mathrm{a}} > \omega_{\min}$ and $\omega_{\mathrm{b}} < \omega_{\mathrm{high}}$ (figure 4.6). (An appropriate least-squares algorithm can be found in [7].) The low-frequency correction to the Hilbert transform is then found by substituting (4.51) into the appropriate Kramers–Kronig relation, so

$$B^{\mathrm{R}}(\omega) = \frac{2}{\pi} \int_0^{\omega_{\min}} \mathrm{d}\Omega \, \frac{\Omega^2 A_i}{(\Omega^2 - \omega_i^2)(\Omega^2 - \omega^2)} \tag{4.55}$$

and this can be evaluated using partial fractions

$$B^{\mathrm{R}}(\omega) = \frac{A_i}{\pi(\omega_i^2 - \omega^2)} \left\{ \omega_i \ln \left(\frac{\omega_i - \omega_{\min}}{\omega_i + \omega_{\min}} \right) + w \ln \left(\frac{\omega + \omega_{\min}}{\omega - \omega_{\min}} \right) \right\}. \quad (4.56)$$

The high-frequency correction term depends on the *stiffness asymptote* of the FRF,

$$\operatorname{Im} H_{\mathrm{M}}^{\mathrm{s}}(\omega) = \sum_{k=j+1}^{N} \frac{A_k \omega}{(\omega_k^2 - \omega^2)^2} + \frac{A_j \omega}{(\omega_j^2 - \omega^2)^2} \quad (4.57)$$

where mode j is the highest mode in the measured region which is assumed to contribute most to the high-frequency truncation error $C^{\mathrm{R}}(\omega)$. In the higher part of the frequency range ω_k/ω is small and the first term can now be expanded:

$$\sum_{k=j+1}^{N} \frac{A_k \omega}{(\omega_k^2 - \omega^2)^2} = \sum_{k=j+1}^{N} \frac{A_k}{\omega_k} \frac{\omega_k}{\omega} \left[1 + \left(\frac{\omega_k}{\omega} \right)^2 + \cdots \right] = \mathrm{O}\left(\left(\frac{\omega_k}{\omega} \right) \right) \quad (4.58)$$

and neglected, so

$$\operatorname{Im} H_{\mathrm{M}}^{\mathrm{s}}(\omega) \approx \frac{A_j \omega}{(\omega_j^2 - \omega^2)^2} \quad (4.59)$$

and a_j is estimated by fitting the function (4.59) to the data over the range ω_{c} to ω_{d} (figure 4.6). The high-frequency correction term is obtained by substituting (4.59) into the Kramers–Kronig relation:

$$C^{\mathrm{R}}(\omega) = \frac{2}{\pi} \int_{\omega_{\min}}^{\infty} \mathrm{d}\Omega \, \frac{\Omega^2 A_j}{(\Omega^2 - \omega_j^2)(\Omega^2 - \omega^2)} \quad (4.60)$$

and this integral can also be evaluated by partial fractions:

$$C^{\mathrm{R}}(\omega) = -\frac{A_j}{\pi(\omega_j^2 - \omega^2)} \left\{ \omega_j \ln \left(\frac{\omega_{\max} - \omega_j}{\omega_{\max} + \omega_j} \right) + w \ln \left(\frac{\omega_{\max} + \omega}{\omega_{\max} - \omega} \right) \right\}. \quad (4.61)$$

Note that in this particular case, Ahmed's correction term is simply a reduced form of the Simon correction term (4.47). This is not the case for the correction to the imaginary part. This depends on the asymptotic behaviour of the real part of $H_{\mathrm{M}}(\omega)$ (4.50). The mass asymptote for the real part takes the form

$$\operatorname{Re} H_{\mathrm{M}}^{m}(\omega) = \sum_{k=1}^{i-1} \frac{2A_k \zeta_k \omega_k \omega^2}{(\omega_k^2 - \omega^2)^2} + \frac{2A_i \zeta_i \omega_i \omega^2}{(\omega_i^2 - \omega^2)^2}. \quad (4.62)$$

As before, the sum term can be neglected where ω/ω_k is small, so

$$\operatorname{Re} H_{\mathrm{M}}^{m}(\omega) \approx \frac{2A_i \zeta_i \omega_i \omega^2}{(\omega_i^2 - \omega^2)^2} = \frac{a_i \omega^2}{(\omega_i^2 - \omega^2)^2} \quad (4.63)$$

and the a_i coefficient is estimated as before by curve-fitting.

The correction term for the imaginary part of the Hilbert transform is, therefore,

$$B^{\mathrm{I}}(\omega) = -\frac{2}{\pi}\int_0^{\omega_{\min}} \mathrm{d}\Omega\, \frac{\Omega^3 a_i}{(\Omega^2-\omega_i^2)^2(\Omega^2-\omega^2)}. \tag{4.64}$$

Evaluation of this expression is a little more involved, but leads to

$$B^{\mathrm{I}}(\omega) = \frac{2a_1\omega}{\pi}\left\{\gamma_{i1}(\omega)\ln\left(\frac{\omega+\omega_{\min}}{\omega-\omega_{\min}}\right) + \gamma_{i2}(\omega)\ln\left(\frac{\omega_i+\omega_{\min}}{\omega_i-\omega_{\min}}\right) + \gamma_{i3}(\omega)\left(\frac{2\omega_{\min}}{\omega_i^2-\omega_{\min}^2}\right)\right\} \tag{4.65}$$

where

$$\gamma_{i1}(\omega) = \frac{\omega}{2(\omega_i^4-\omega^4)}, \quad \gamma_{i2}(\omega) = -\frac{1}{4\omega_i(\omega_i^2-\omega^2)}, \quad \gamma_{i3}(\omega) = -\frac{1}{4(\omega_i^2-\omega^2)}. \tag{4.66}$$

Finally, to evaluate the high-frequency correction to the imaginary part of the Hilbert transform, the stiffness asymptote of the real part is needed. The starting point is

$$\mathrm{Re}\, H_{\mathrm{M}}^{\mathrm{s}}(\omega) = \sum_{k=j+1}^{N} \frac{\alpha_k\omega^2}{(\omega_k^2-\omega^2)^2} + \frac{\alpha_j\omega^2}{(\omega_j^2-\omega^2)^2} \tag{4.67}$$

where $\alpha_k = 2A_k\zeta_k\omega_k$. Expanding the first term yields

$$\sum_{k=j+1}^{N}\alpha_k\left\{1+\left(\frac{\omega_k}{\omega}\right)^2+\cdots\right\} \approx \sum_{k=j+1}^{N}\alpha_k \tag{4.68}$$

as ω_k/ω is considered to be small. The final form for the asymptote is

$$\mathrm{Re}\, H_{\mathrm{M}}^{\mathrm{s}}(\omega) \approx \frac{b_1+b_2\omega^2+b_3\omega^4}{(\omega_j^2-\omega^2)^2} \tag{4.69}$$

where the coefficients

$$b_1 = \omega_j^4\sum_{k=j+1}^{N}\alpha_k, \quad b_2 = \alpha_j - 2\omega_j^2\sum_{k=j+1}^{N}\alpha_k, \quad b_3 = \sum_{k=j+1}^{N}\alpha_k \tag{4.70}$$

are once again obtained by curve-fitting.

The high-frequency correction is obtained by substituting (4.69) into the Kramers–Kronig integral. The calculation is a little involved and yields

$$C^{\mathrm{I}}(\omega) = \frac{2\omega}{\pi}\left\{\gamma_{j1}(\omega)\ln\left(\frac{\omega_{\max}-\omega}{\omega_{\max}+\omega}\right) + \gamma_{j2}(\omega)\ln\left(\frac{\omega_{\max}-\omega_j}{\omega_{\max}-\omega_j}\right) + \gamma_{j3}(\omega)\left(\frac{2\omega_{\max}}{\omega_{\max}^2-\omega_j^2}\right)\right\} \tag{4.71}$$

where

$$\gamma_{j1}(\omega) = \frac{b_3\omega(2\omega_j^2 + \omega^2) + b_2\omega - b_1\omega_j^2}{2(\omega_j^2 - \omega^2)},$$
$$\gamma_{j2}(\omega) = \frac{b_1\omega - b_2 + 3b_3\omega_j^2}{4(\omega_j^2 - \omega^2)}, \tag{4.72}$$
$$\gamma_{j3}(\omega) = -\frac{a_1\omega + b_2 + b_3\omega_j^2}{4(\omega_j^2 + \omega^2)}.$$

Note that these results only apply to mobility FRFs, substantially different correction terms are needed for the other FRF forms. However, they are derived by the same procedure as the one described here.

Although the Ahmed correction procedure is rather more complex than the others, it produces excellent results. Figure 4.7 shows the Hilbert transform in figure 4.5 recomputed using the Ahmed correction terms; an almost perfect overlay is obtained.

4.3.2.6 Summary

None of the correction methods can claim to be faultless; truncation near to a resonance will always give poor results. Considerable care is needed to obtain satisfactory results. The conversion to receptance, Fei and Haoui techniques are only suitable for use with baseband data and the Simon and Ahmed corrections require *a priori* curve-fitting. The next sections and the next chapter outline approaches to the Hilbert transform which do not require correction terms and in some cases overcome the problems.

Note also that the accelerance FRF tends to a constant non-zero value as $\omega \longrightarrow \infty$. As a consequence the Hilbert transform will always suffer from truncation problems, no matter how high $\omega_{\max}$ is taken. The discussion of this problem requires complex analysis and is postponed until the next chapter.

4.3.3 Fourier method 1

This method relies on the fact that the Hilbert transform is actually a convolution of functions and can therefore be factored into Fourier operations. Consider the basic Hilbert transform,

$$\mathcal{H}\{G(\omega)\} = \tilde{G}(\omega) = -\frac{1}{\mathrm{i}\pi}\int_{-\infty}^{\infty} \mathrm{d}\Omega\, \frac{G(\Omega)}{\Omega - \omega}. \tag{4.73}$$

Recalling the definition of the convolution product $*$,

$$f_1(t) * f_2(t) = \int_{-\infty}^{\infty} \mathrm{d}\tau\, f_1(\tau) f_2(t - \tau) \tag{4.74}$$

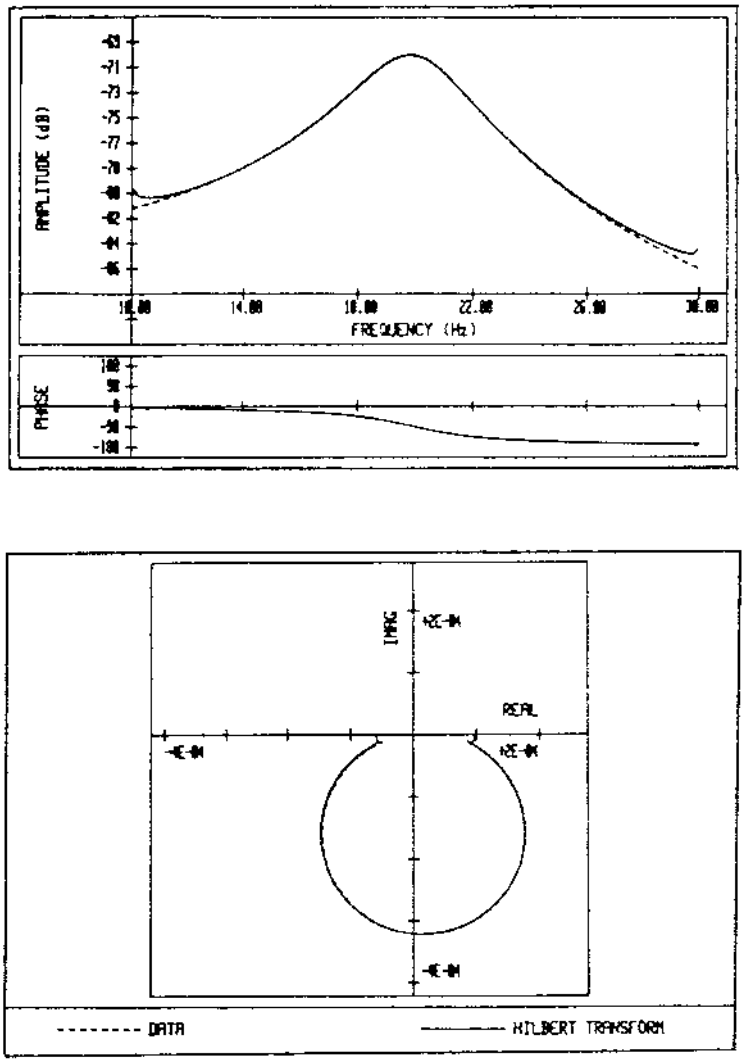


Figure 4.7. Hilbert transform with Ahmed's correction of zoomed linear data.

it is clear that

$$\tilde{G}(\omega) = G(\omega) * -\frac{\mathrm{i}}{\pi\omega}. \tag{4.75}$$

Now, a basic theorem of Fourier transforms states that

$$\mathcal{F}\{f_1(t)f_2(t)\} = \mathcal{F}\{f_1(t)\} * \mathcal{F}\{f_2(t)\}. \tag{4.76}$$

It therefore follows from (4.75) that

$$\mathcal{F}^{-1}\{\tilde{G}(\omega)\} = \mathcal{F}^{-1}\{G(\omega)\}\mathcal{F}^{-1}\left\{-\frac{\mathrm{i}}{\pi\omega}\right\} = g(t)\epsilon(t) \tag{4.77}$$

where $\epsilon(t)$ is the signum function defined in (4.9). ($\mathcal{F}\{\epsilon(t)\} = -\mathrm{i}/(\pi\omega)$ is proved in appendix D.)

It immediately follows from (4.77) that

$$\tilde{G}(\omega) = \mathcal{F}\circ \in \circ\mathcal{F}^{-1}\{G(\omega)\} \tag{4.78}$$

where the operator $\in$ represents multiplication by $\epsilon(t)$, i.e. $\in\{g(t)\} = g(t)\epsilon(t)$ and composition is denoted by $\circ$, i.e. $(f_1 \circ f_2)(t) = f_1(f_2(t))$. In terms of operators,

$$\mathcal{H} = \mathcal{F}\circ \in \circ\mathcal{F}^{-1} \tag{4.79}$$

and the Hilbert transform can therefore be implemented in terms of the Fourier transform by the three-step procedure:

(1) Take the inverse Fourier transform of $G(\omega)$. This yields the time domain $g(t)$.
(2) Multiply $g(t)$ by the signum function $\epsilon(t)$.
(3) Take the Fourier transform of the product $g(t)\epsilon(t)$. This yields the required Hilbert transform $\tilde{G}(\omega)$.

In practice these operations will be carried out on sampled data, so the discrete Fourier transform (DFT) or fast Fourier transform will be used. In the latter case, the number of points should usually be 2^N for some N.

The advantage of this method over the direct method described in the previous section is its speed (if the FFT is used). A comparison was made in [170]. (The calculations were made on a computer which was extremely slow by present standards. As a consequence, only ratios of the times have any meaning.)

Number of points	256	512
Direct method	6.0 min	24.1 min
Fourier method 1	1.0 min	2.0 min

The disadvantages of the method arise from the corrections needed. Both result from the use of the FFT, an operation based on a finite data set.

The first problem arises because the FFT forces periodicity onto the data outside the measured range, so the function $\epsilon(t)$ which should look like figure 4.8(*a*), is represented by the square-wave function $\mathrm{sq}(t)$ of figure 4.8(*b*). This means that the function $G(\omega)$ is effectively convolved with the function $-\mathrm{i}\cot(\pi\omega) = \mathcal{F}\{\mathrm{sq}(t)\}$ instead of the desired $-\mathrm{i}/(\pi\omega)$. (See [260] for the appropriate theory.) The effective convolving functions is shown in figure 4.9(*b*).

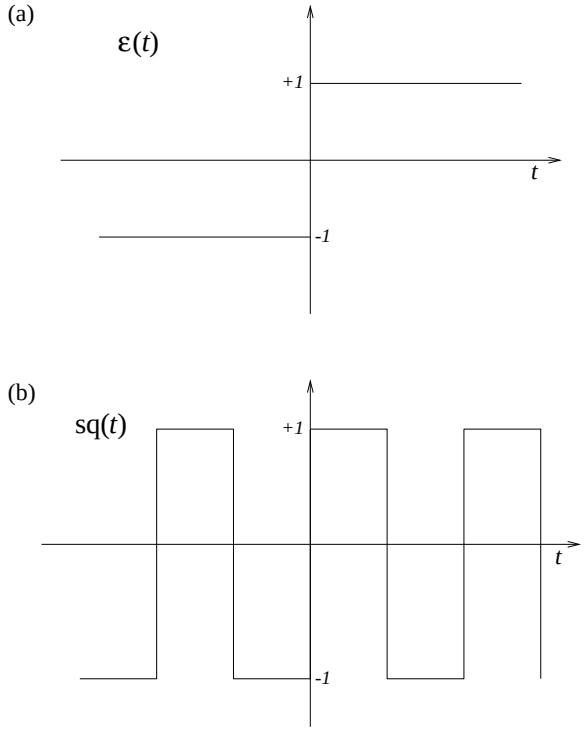


Figure 4.8. Effect of the discrete Fourier transform on the signum function.

As $\omega \longrightarrow 0$, $-\mathrm{i}\cot(\pi\omega) \longrightarrow -\mathrm{i}/(\pi\omega)$, so for low frequencies or high sampling rates, the error in the convolution is small. If these conditions are not met, a correction should be made. The solution is simply to compute the *discrete inverse DFT* of the function $-\mathrm{i}/(\pi\omega)$ and multiply by that in the time-domain in place of $\epsilon(t)$. The problem is that $-\mathrm{i}/(\pi\omega)$ is singular at $\omega = 0$. A naive approach to the problem is to zero the singular point and take the discrete form [5] of $-\mathrm{i}/(\pi\omega)$:

$$U_k = \begin{cases} 0, & k = 1 \\ -\dfrac{\mathrm{i}}{\pi(k-1)}, & 2 \le k \le \dfrac{N}{2} \\ \dfrac{\mathrm{i}}{\pi(N+1-k)}, & \dfrac{N}{2}+1 \le k \le N. \end{cases} \tag{4.80}$$

The corresponding time function, often called a *Hilbert window*, is shown in figure 4.10 (only points $t > 0$ are shown). It is clear that this is a poor representation of $\epsilon(t)$. The low-frequency component of the signal between

[5] There are numerous ways of coding the data for an FFT, expression (4.80) follows the conventions of [209].

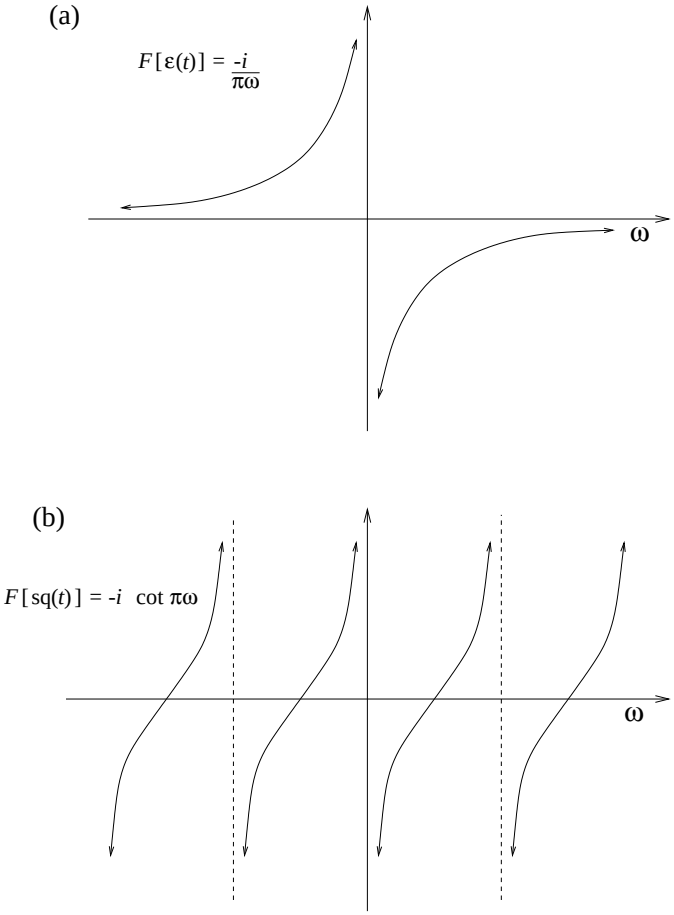


Figure 4.9. Desired Hilbert window and periodic form from the discrete FFT.

$-\Delta\omega/2$ and $\Delta\omega/2$ has been discarded. This can be alleviated by transferring energy to the neighbouring lines and adopting the definition

$$U_k = \begin{cases} 0, & k = 1 \\ -\dfrac{3\mathrm{i}}{2\pi}, & k = 2 \\ -\dfrac{\mathrm{i}}{\pi(k-1)}, & 3 \le k \le \dfrac{N}{2} \\ \dfrac{\mathrm{i}}{\pi(N+1-k)}, & \dfrac{N}{2}+1 \le k \le N-1 \\ \dfrac{3\mathrm{i}}{2\pi}, & k = N. \end{cases} \tag{4.81}$$

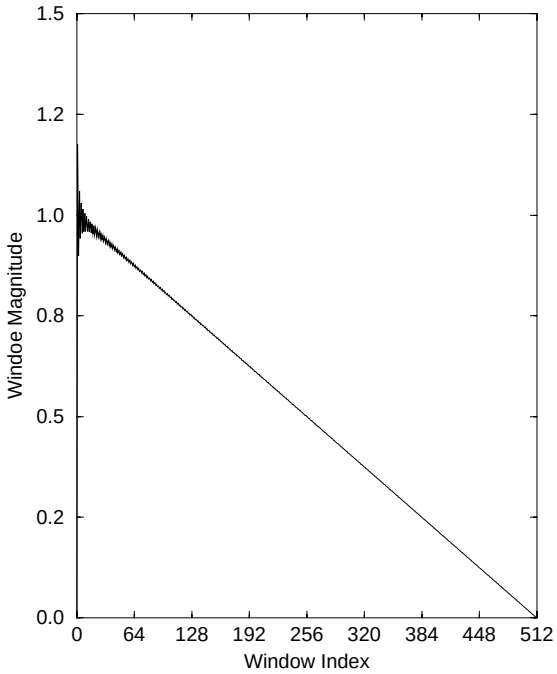


Figure 4.10. Naive discrete Hilbert window.

The Hilbert window corresponding to this definition is shown in figure 4.11. There is a noticeable improvement.

The next problem is of *circular convolution*. The ideal convolution is shown in figure 4.12. The actual convolution implemented using the FFT is depicted in figure 4.13. The error occurs because the function $G(\omega)$ should vanish in region B but does not because of the imposed periodicity. The solution is straightforward. The sampled function $G(\omega)$, defined at N points, is extended to a $2N$-point function by translating region B by N points and padding by zeros. The corresponding Hilbert window is computed from the $2N$-point discretization of $-1/(\pi\omega)$. The resulting calculation is illustrated in figure 4.14.

Finally, the problem of truncation should be raised. The Fourier method can only be used with baseband data. In practice, $G(\omega)$ will only be available for positive ω, the negative frequency part needed for the inverse Fourier transform is obtained by using the known symmetry properties of FRFs which follow from the reality of the impulse response. Namely, $\mathrm{Re}\, G(-\omega) = \mathrm{Re}\, G(\omega)$ and $\mathrm{Im}\, G(-\omega) = -\mathrm{Im}\, G(\omega)$. If one naively completes the FRF of zoomed data by these reflections, the result is as shown in figure 4.15(*b*), instead of the desired figure 4.15(*a*). This leads to errors in the convolution. One way of overcoming this problem is to pad the FRF with zeros from $\omega = 0$ to $\omega = \omega_{\min}$. This is

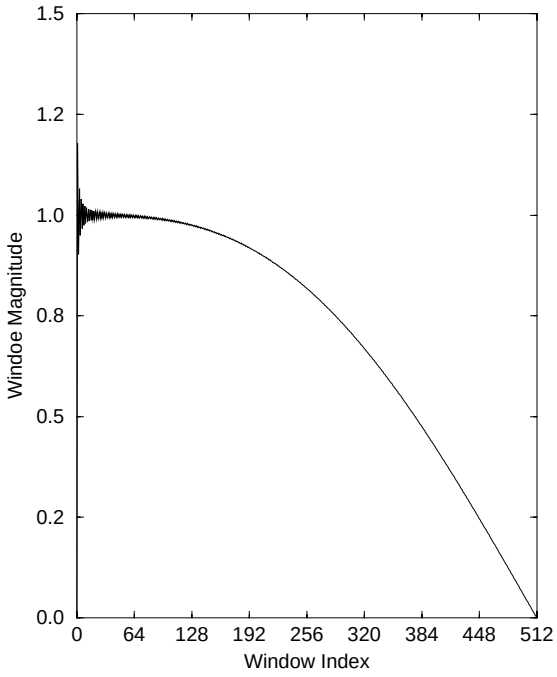


Figure 4.11. Corrected discrete Hilbert window.

inefficient if the zoom range is small or at high frequency and will clearly lead to errors if low-frequency modes have been discarded.

Of the correction methods described in section 4.4.2, the only one applicable is conversion to receptance and this should be stressed. This is only effective for correcting the high-frequency error. However, as previously discussed, the data should always be baseband in any case.

In summary then, the modified Fourier method 1 proceeds as follows.

(1) Convert the measured $\frac{1}{2}N$-point positive-frequency FRF $G(\omega)$ to an N-point positive-frequency FRF by translation, reflection and padding.
(2) Complete the FRF by generating the negative-frequency component. The real part is reflected about $\omega = 0$, the imaginary part is reflected with a sign inversion. The result is a $2N$-point function.
(3) Take the inverse Fourier transform of the discretized $-\mathrm{i}/(\pi\omega)$ on $2N$ points. This yields the Hilbert window $\mathrm{hi}(t)$.
(4) Take the inverse Fourier transform of the $2N$-point FRF. This yields the impulse response $g(t)$.
(5) Form the product $g(t)\,\mathrm{hi}(t)$.
(6) Take the Fourier transform of the product. This yields the desired Hilbert transform $\tilde{G}(\omega)$.

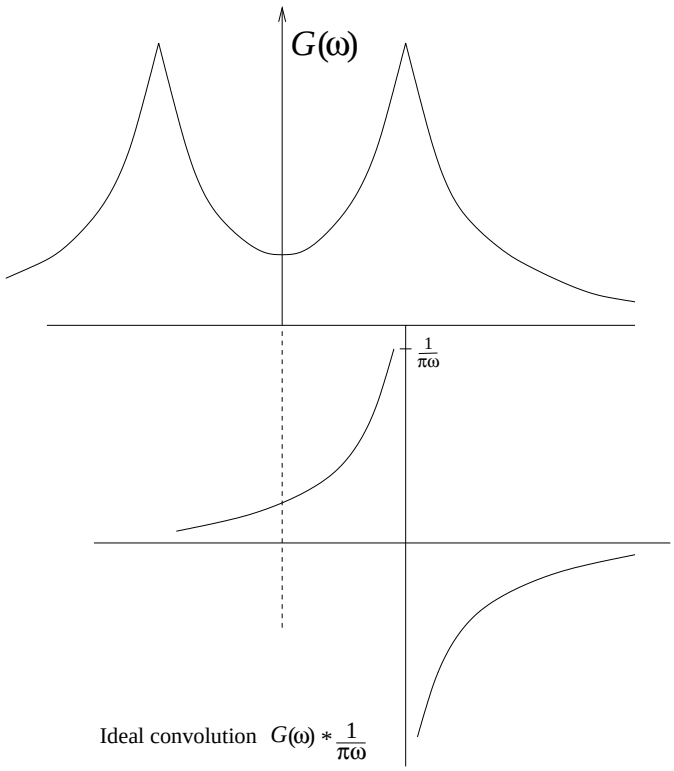


Figure 4.12. Ideal convolution for the Hilbert transform.

4.3.4 Fourier method 2

Fourier method 1 was discussed as it was the first Hilbert transform method to exploit Fourier transformation. However, it is rather complicated to implement and the method discussed in this section is to be preferred in practice.

The implementation of this method is very similar to Fourier method 1; however, the theoretical basis is rather different. This method is based on the properties of *analytic*[6] signals and is attributed to Bendat [24]. Given a time

[6] This terminology is a little unfortunate, as the word analytic will have two different meanings in this book. The first meaning is given by equation (4.82). The second meaning relates to the pole-zero structure of complex functions—a function is analytic in a given region of the complex plane if it has no poles in that region. (Alternatively, the function has a convergent Taylor series.) The appropriate meaning will always be clear from the context.

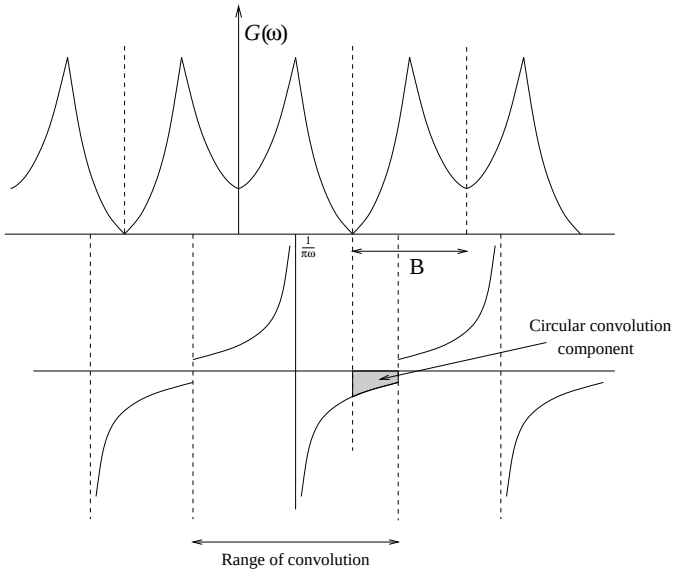


Figure 4.13. The problem of circular convolution.

signal $g(t)$, the corresponding analytic signal, $a(t)$, is given by[7]

$$a(t) = g(t) - \tilde{g}(t) = g(t) - \mathcal{H}\{g(t)\}. \tag{4.82}$$

Taking the Fourier transform of this equation yields

$$A(\omega) = G(\omega) - \mathcal{F} \circ \mathcal{H}\{g(t)\} = G(\omega) - \mathcal{F} \circ \mathcal{H} \circ \mathcal{F}^{-1}\{G(\omega)\}. \tag{4.83}$$

Now, recall that the Hilbert transform factors into Fourier operations. The decomposition depends on whether the operator acts on time- or frequency-domain functions. The appropriate factorization in the frequency domain is given by (4.79). Essentially the same derivation applies in the time domain and the result is

$$\mathcal{H} = -\mathcal{F}^{-1} \circ \in \circ \mathcal{F}. \tag{4.84}$$

[7] This definition differs from convention

$$a(t) = g(t) + \mathrm{i}\tilde{g}(t).$$

The reason is that the conventional definition of the Hilbert transform of a *time* signal omits the imaginary i, and reverses the sign to give a true convolution, i.e.

$$\mathcal{H}\{g(t)\} = \tilde{g}(t) = \frac{1}{\pi}\int_{-\infty}^{\infty} \mathrm{d}\tau \, \frac{g(\tau)}{t-\tau}.$$

Modifying the definition of the analytic signal avoids the unpleasant need to have different Hilbert transforms for different signal domains.

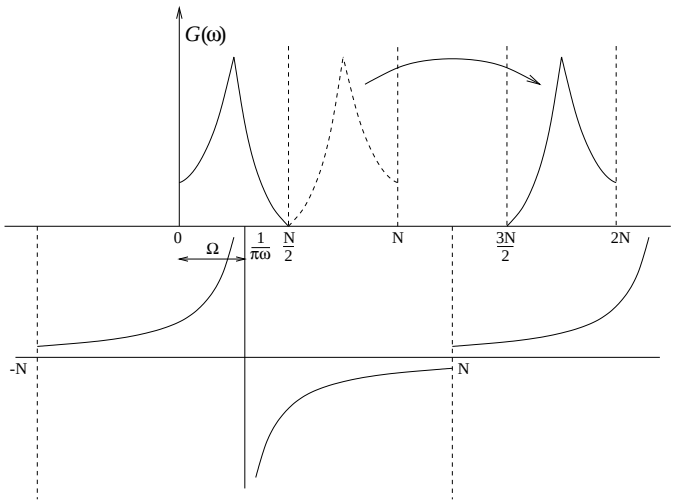


Figure 4.14. Solution to the circular convolution problem using translation and zero-padding.

Substituting this expression into (4.83) yields

$$A(\omega) = G(\omega) + \in\{G(\omega)\} = G(\omega)[1 + \epsilon(\omega)] \tag{4.85}$$

so

$$A(\omega) = \begin{cases} 2G(\omega), & \omega > 0 \\ G(\omega), & \omega = 0 \\ 0, & \omega < 0 \end{cases} \tag{4.86}$$

thus, *the spectrum of an analytic signal depends only on the spectrum of the real part*. This fact is the basis of the method.

Any function of frequency has a trivial decomposition

$$G(\omega) = \mathrm{Re}\, G(\omega) + \mathrm{i}\, \mathrm{Im}\, G(\omega). \tag{4.87}$$

However, if $G(\omega)$ has a causal inverse Fourier transform, $\mathrm{i}\, \mathrm{Im}\, G(\omega) = \mathcal{H}\{\mathrm{Re}\, G(\omega)\}$ by (4.17). Therefore

$$G(\omega) = \mathrm{Re}\, G(\omega) + \mathcal{H}\{\mathrm{Re}\, G(\omega)\} \tag{4.88}$$

so $G(\omega)$ is analytic, *provided* that ω is considered to be a time-like variable. If the Fourier transform, (*not* the inverse transform) is applied

$$\mathcal{G}(\tau) = \mathcal{F}_\tau\{G(\omega)\} = \int_{-\infty}^{\infty} \mathrm{d}\omega\, \mathrm{e}^{-\mathrm{i}\tau\omega} G(\omega) \tag{4.89}$$

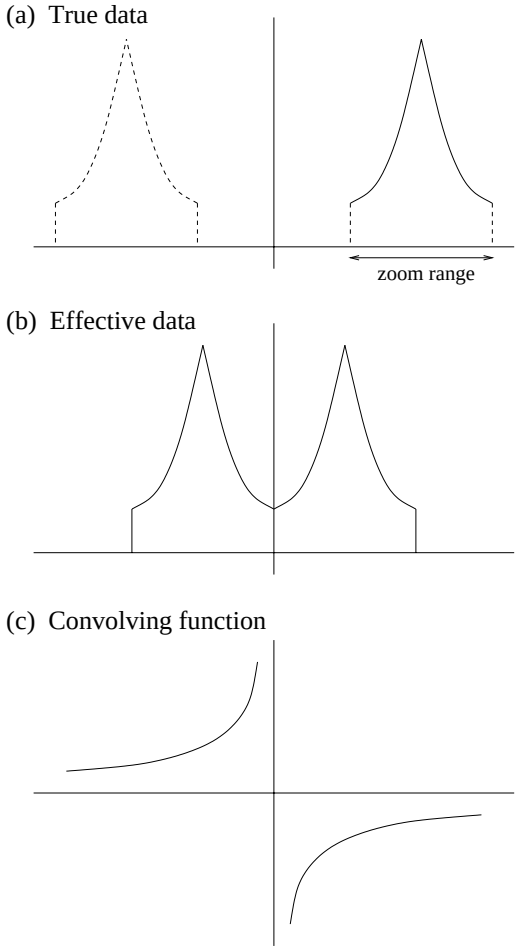


Figure 4.15. Convolution problem for zoomed data.

the result is

$$\mathcal{G}(\tau) = \begin{cases} 2GR(\tau), & \tau > 0 \\ GR(\tau), & \tau = 0 \\ 0, & \tau < 0 \end{cases} \tag{4.90}$$

where

$$GR(\tau) = \mathcal{F}_\tau\{\mathrm{Re}\, G(\omega)\} \tag{4.91}$$

so *the Fourier transform of the FRF is completely specified by the Fourier*

transform of the real part[8]. This fact provides a means of computing the FRF imaginary part from the real part. In principle, three steps are required:

(1) Take the Fourier transform $\mathcal{F}_\tau$ of the FRF real part $\mathrm{Re}\,G(\omega)$, i.e. $GR(\tau)$.
(2) Form the transform $\mathcal{G}(\tau)$ using (4.89).
(3) Take the inverse Fourier transform $\mathcal{F}_\tau^{-1}$ of $\mathcal{G}(\tau)$. The result is $G(\omega)$, i.e. the desired Hilbert transform, $\tilde{\mathrm{Re}}\,G(\omega)$, has been obtained as the imaginary part.

A trivial modification—exchange $\mathrm{Im}\,G$ and $\mathrm{Re}\,G$—in this argument leads to the means of computing $\tilde{\mathrm{Im}}\,G(\omega)$.

One advantage of the method is its speed, the timings are essentially those of Fourier method 1. Also, because the FFT is applied to a spectrum, which has already been obtained by FFT and is periodic, there are no leakage effects. The method is subject to the same truncation problems that afflict all the methods and the only applicable correction is conversion to receptance. The implementation of the method is now illustrated by a case study [142].

4.3.5 Case study of the application of Fourier method 2

The structure used to obtain the experimental data was a composite $\frac{1}{23}$ scale aircraft wing used for wind tunnel tests. The wing was secured at its root to a rigid support, effectively producing a cantilever boundary condition. Excitation of the wing was via an electrodynamic exciter attached to the wing via a push rod (stinger) and a force transducer. The excitation was a band-limited random signal in the range 0–512 Hz. The response of the wing was measured using lightweight accelerometers. (Note that random excitation is not optimal for nonlinear structures—this will be discussed later. This study is intended to show how the Hilbert transform is computed, and one can only validate the method on a linear structure.)

Figure 4.16 shows the accelerance FRF measured by the experiment. At least seven modes are visible. For information, the resonance at 76 Hz was identified as first wing bending, that at 215 Hz was identified as first wing torsion.

[8] Note that

$$\mathrm{Re}\,G(\omega) = \int_{-\infty}^{\infty} \mathrm{d}t\, \cos(\omega t) g(t) = \int_{-\infty}^{\infty} \mathrm{d}t\, \mathrm{e}^{-\mathrm{i}\omega t} g_{\mathrm{even}}(t)$$

so,

$$\begin{aligned} GR(\tau) &= \int_{-\infty}^{\infty} \mathrm{d}\omega\, \mathrm{e}^{-\mathrm{i}\tau\omega} \int_{-\infty}^{\infty} \mathrm{d}t\, \mathrm{e}^{-\mathrm{i}\omega t} g_{\mathrm{even}}(t) \\ &= \int_{-\infty}^{\infty} \mathrm{d}t\, g_{\mathrm{even}}(t) \int_{-\infty}^{\infty} \mathrm{d}\omega\, \mathrm{e}^{-\mathrm{i}\omega(\tau+t)} = 2\pi \int_{-\infty}^{\infty} \mathrm{d}t\, g_{\mathrm{even}}(t)\delta(\tau+t) \\ &= 2\pi g_{\mathrm{even}}(-\tau). \end{aligned}$$

So GR is essentially the even component of the original time signal. This fact does not help with the development of the algorithm. However, it does show that the terminology 'pseudo spectrum' for GR, which is sometimes used, is probably inappropriate.

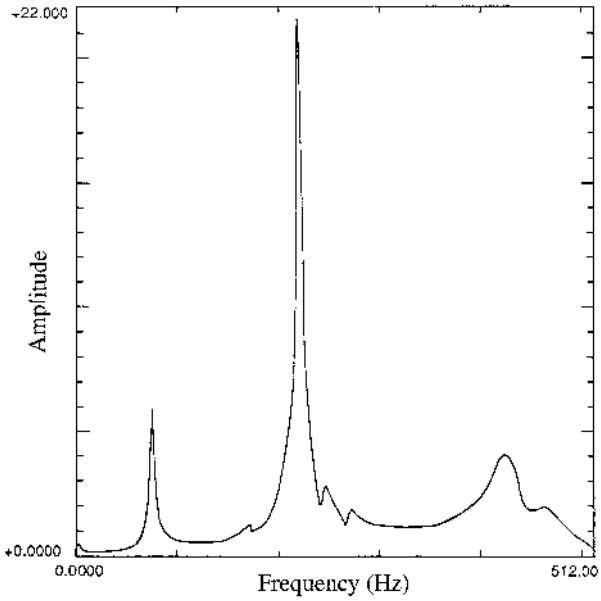


Figure 4.16. A typical experimental cross-accelerance FRF measured from a scaled wing model.

The first step in the procedure is to correct for truncation; the FRF is converted to receptance by dividing by $-\omega^2$ (avoiding the division at $\omega = 0$). The result is shown in figure 4.17. To further reduce truncation errors, the FRF was extended to $2N$ points by padding with zeroes (figure 4.18).

The next stage was the completion of the FRF, i.e. the conversion to a double-sided form. The negative frequency parts were obtained by assuming even symmetry for the real part and odd symmetry for the imaginary part. The double-sided signals are given in figure 4.19.

The function $GR(\tau)$ was formed by Fourier transforming the real part (figure 4.20(*a*)). This was converted to $\mathcal{G}(\tau)$ by zeroing the negative-τ component and doubling the positive-τ part. The $\tau = 0$ line was left untouched. Taking the inverse Fourier transform then gave $\widetilde{\mathrm{Re}\, G}$ as the imaginary part.

The function $GI(\tau)$ was formed by Fourier transforming the imaginary part of the FRF (figure 4.20(*b*)). This was also converted to the full $\mathcal{G}(\tau)$ as before. Taking the inverse FFT gave $\widetilde{\mathrm{Im}\, G}$ as the real part.

Both the real and imaginary parts of the Hilbert transform have now been obtained. The next stage was simply to convert back to the accelerance form. In order to evaluate the results, the Hilbert transform is shown overlaid on the original FRF in figure 4.21, the two curves should match. Both the Bode magnitude and Nyquist plots are given. The somewhat poor quality of the Nyquist

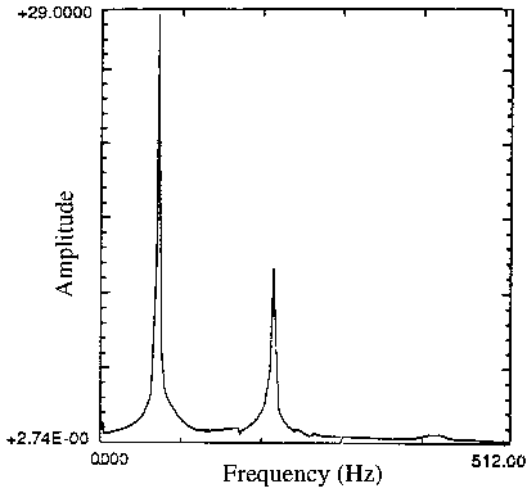


Figure 4.17. Receptance FRF converted from accelerance FRF in figure 4.16.

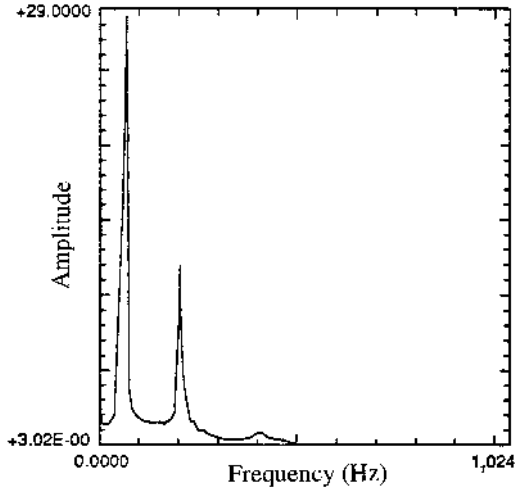


Figure 4.18. Receptance FRF padded with zeroes to $2f_{\max}$.

comparison is due to the limited frequency resolution.

The method clearly produces an excellent Hilbert transform and indicates, for the excitation used, that the system is nominally linear.

Having established methods of computing the transform, it is now finally time to show how the method allows the detection and identification of nonlinearity.

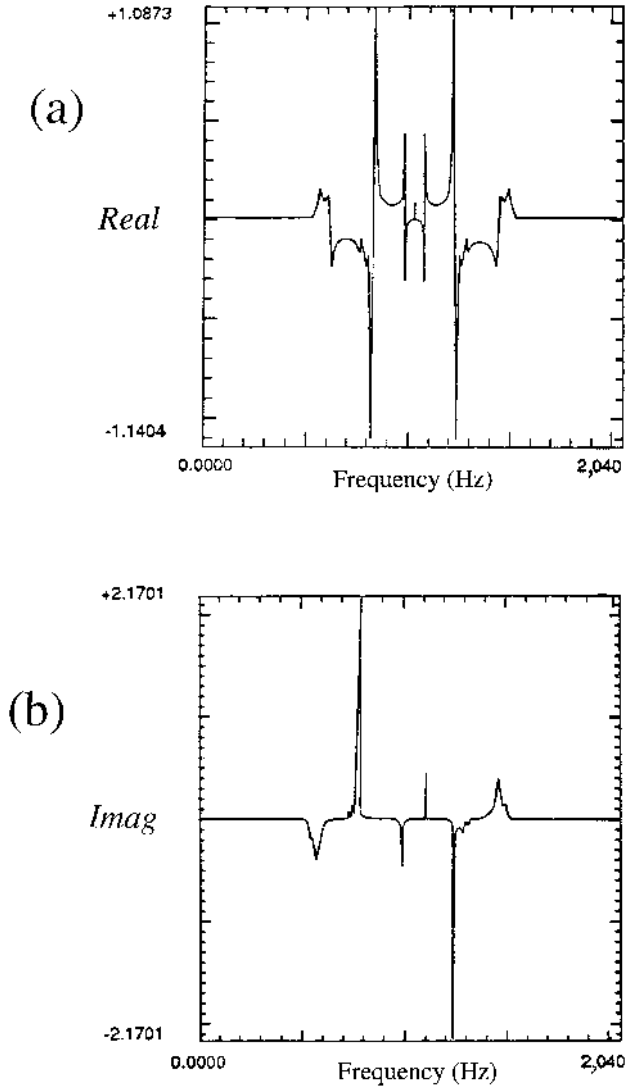


Figure 4.19. (*a*) Double-sided (even function) real part of the FRF of figure 4.18. (*b*) Double-sided (odd function) imaginary part of the FRF of figure 4.18.

4.4 Detection of nonlinearity

The basis of the Hilbert transform as a nonlinearity detection method is equation (4.20) which asserts that the Hilbert transform acts as the identity on

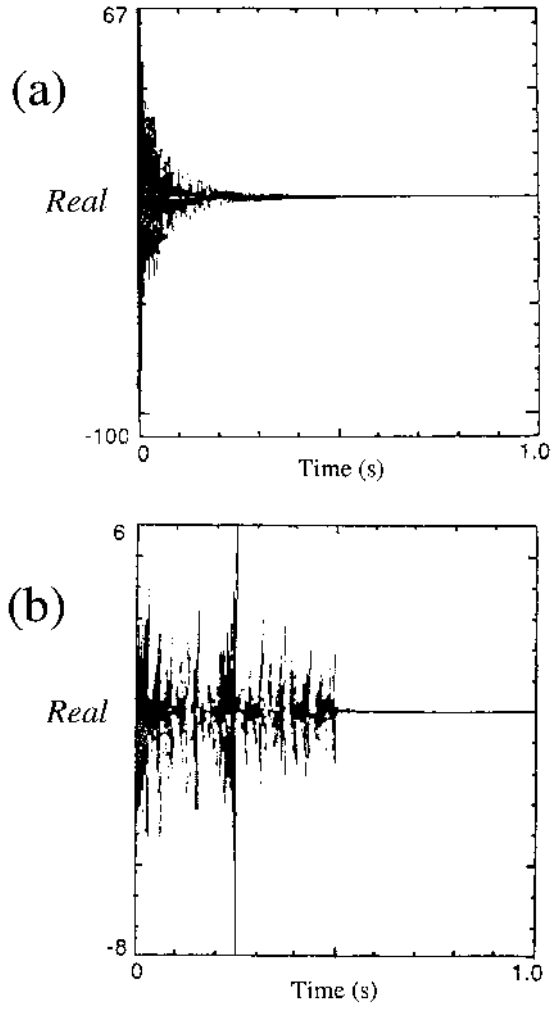


Figure 4.20. (*a*) Pseudo-spectrum from the Fourier transform of the curve in figure 4.19(*a*). (*b*) Pseudo-spectrum from the Fourier transform of the curve in figure 4.19(*b*).

functions $G(\omega)$ which have causal inverse Fourier transforms, i.e.

$$G(\omega) = \mathcal{H}\{G(\omega)\} \Leftrightarrow \mathcal{F}^{-1}\{G(\omega)\} = g(t) = 0, \quad \forall t < 0. \tag{4.92}$$

The inverse Fourier transform of a linear system FRF $H(\omega)$, is the system impulse response $h(t)$ which is always zero for negative times by the principle of causality (see chapter 1). This means that the FRF $H(\omega)$ is invariant under the Hilbert transform. There is no compelling reason why this condition should hold

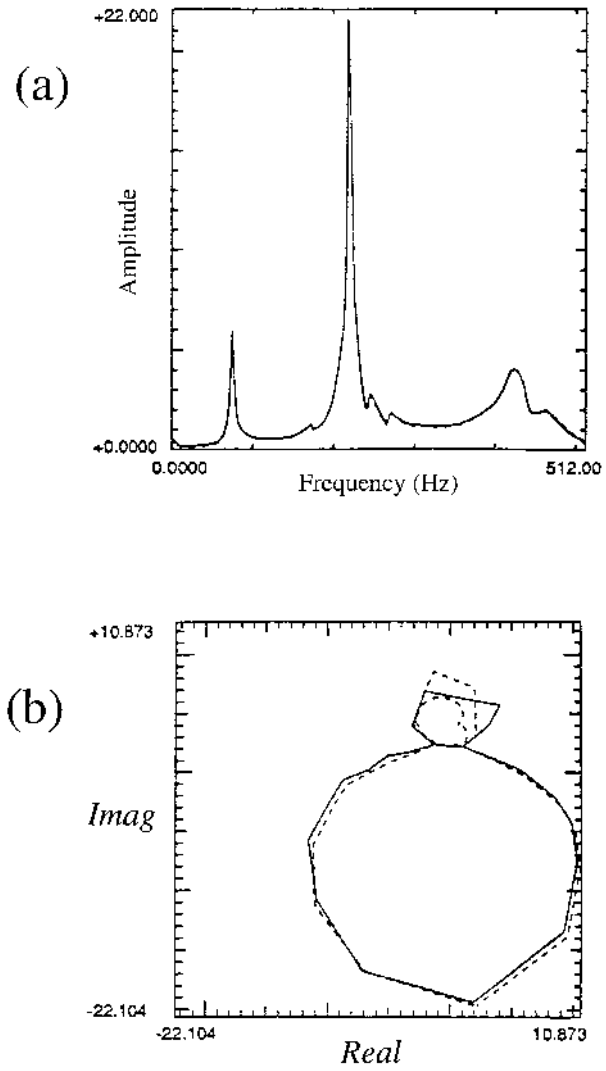


Figure 4.21. Overlay of the experimental (——) and Hilbert transformed (– – –) data in (*a*) Bode plot, (*b*) Nyquist plot.

for the FRF of a nonlinear system.

Consider the FRF of a generic nonlinear system $G(\omega)$. It is impossible to show that $\mathcal{F}^{-1}\{G(\omega)\} = g(t)$ will

(1) be real and
(2) be causal.

In practice reality is imposed because the one-sided FRF is often converted to a double-sided FRF by imposing evenness and oddness conditions on the real and imaginary parts respectively. This forces a real $g(t)$. This, in turn, means that the usual consequence of nonlinearity is non-causality of the 'impulse response' function, i.e. the inverse Fourier transform of the FRF. This *does not* mean that the system is non-causal in the physical sense; cause must always precede effect. It simply means that the inverse Fourier transform of a nonlinear system FRF must not be interpreted as an impulse response. The specification and calculation of nonlinear system impulse responses is more complicated and will be discussed in a later chapter. The fact that $g(t) \neq 0$ for all negative t is often referred to as *artificial non-causality*.

As a result the Hilbert transform will not act as the identity on $G(\omega)$: $G(\omega) \neq \mathcal{H}\{G(\omega)\}$. It is possible to see this directly using the factorization (4.79) of $\mathcal{H}$,

$$\mathcal{H}\{G(\omega)\} = \mathcal{F}\{\epsilon(t)g(t)\}. \tag{4.93}$$

If $g(t)$ is causal, $\epsilon(t)g(t) = g(t)$ and $\mathcal{H}$ is the identity. If not $\epsilon(t)g(t) \neq g(t)$ and $\mathcal{H} \neq \mathrm{Id}$. The argument is summarized diagrammatically in figure 4.22.

The question arises: If $\mathcal{H}$ is not the identity, what is its effect on nonlinear system FRFs? Consider the hardening Duffing oscillator,

$$m\ddot{y} + c\dot{y} + ky + k_3y^3 = x(t), \quad k_3 > 0. \tag{4.94}$$

Suppose an FRF is obtained from this system with $x(t)$ a low-amplitude signal (the appropriate form for $x(t)$, i.e. whether stepped-sine or random etc. is discussed later.) At low levels of excitation, the linear term dominates and the FRF is essentially that of the underlying linear system. In that case, the Hilbert transform will overlay the original FRF. If the level of excitation is increased, the Hilbert transform will start to depart from the original FRF; however because the operator $\mathcal{H}$ is continuous, the main features of the FRF—resonances etc—are retained but in a distorted form. Figure 4.23 shows the FRF of a Duffing oscillator and the corresponding Hilbert transform, the level of excitation is set so that the Hilbert transform is just showing mild distortion.

A number of points are worth noting about figure 4.23. First, it is sometimes helpful to display the FRF and transform in different formats as each conveys different information: the Bode plot and Nyquist plot are given here. The figure also shows that the Hilbert transform is a sensitive indicator of nonlinearity. The FRF shows no discernible differences from the linear form, so using FRF distortion as a diagnostic fails in this case. The Hilbert transform, however, clearly shows the effect of the nonlinearity, particularly in the Nyquist plot. Finally, experience shows that the form of the distortion is actually characteristic of the type of nonlinearity, so the Hilbert transform can help in identifying the system. In the case of the hardening cubic stiffness, the following observations apply. In the Bode plot the peak of the Hilbert transform curve appears at a higher frequency than in the FRF. The peak magnitude of the Hilbert transform is higher.

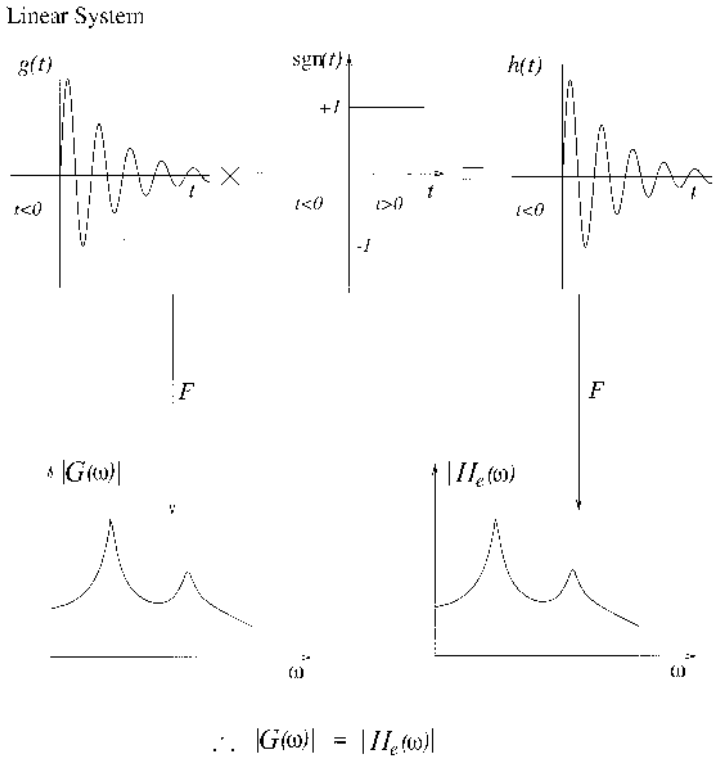


Figure 4.22. Demonstration of artificial non-causality for a nonlinear system.

In the Nyquist plot, the characteristic circle is rotated clockwise and elongated into a more elliptical form. Figure 4.24 shows the FRF and transform in a more extreme case where the FRF actually shows a jump bifurcation. The rotation and elongation of the Nyquist plot are much more pronounced.

The characteristic distortions for a number of common nonlinearities are summarized next (in all cases the FRFs are obtained using sine excitation).

4.4.1 Hardening cubic stiffness

The equation of motion of the typical SDOF system is given in (4.94). The FRF and Hilbert transform in the two main formats are given in figure 4.23. The FRF is given by the dashed line and the transform by the solid line.

In the Bode plot the peak of the Hilbert transform curve appears at a higher frequency than in the FRF. The peak magnitude of the Hilbert transform is higher. In the Nyquist plot, the characteristic circle is rotated clockwise and elongated into a more elliptical form.

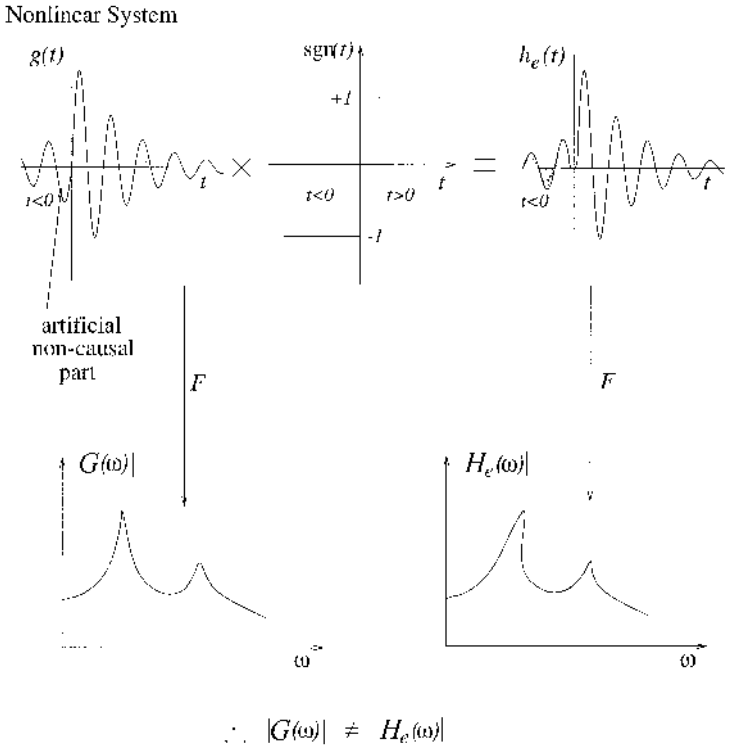


Figure 4.22. (Continued)

4.4.2 Softening cubic stiffness

The equation of motion is

$$m\ddot{y} + c\dot{y} + ky + k_3 y^3 = x(t), \quad k_3 < 0. \tag{4.95}$$

The FRF and Hilbert transform are given in figure 4.25. In the Bode plot the peak of the Hilbert transform curve appears at a lower frequency than in the FRF. The peak magnitude of the Hilbert transform is higher. In the Nyquist plot, the characteristic circle is rotated anti-clockwise and elongated into a more elliptical form.

4.4.3 Quadratic damping

The equation of motion is

$$m\ddot{y} + c\dot{y} + c_2\dot{y}|\dot{y}| + ky + = x(t), \quad c_2 > 0. \tag{4.96}$$

The FRF and Hilbert transform are given in figure 4.26. In the Bode plot the peak of the Hilbert transform curve stays at the same frequency as in the FRF, but

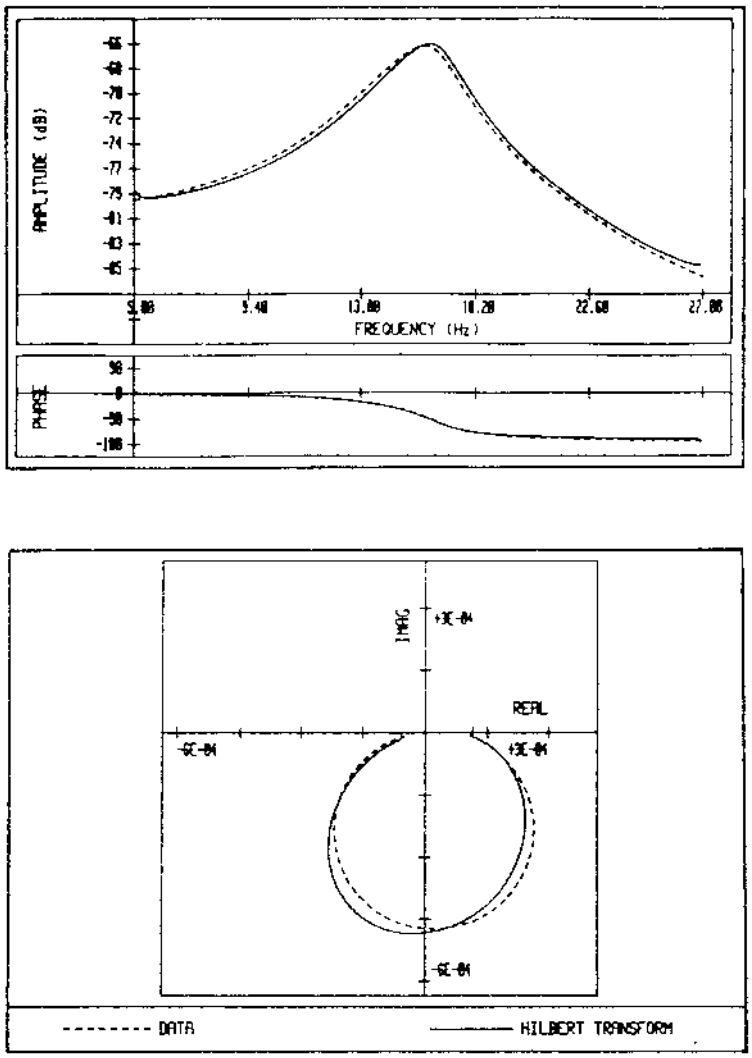


Figure 4.23. Hilbert transform of a hardening cubic spring FRF at a low sine excitation level.

increases in magnitude. In the Nyquist plot, the characteristic circle is elongated into an ellipse along the imaginary axis.

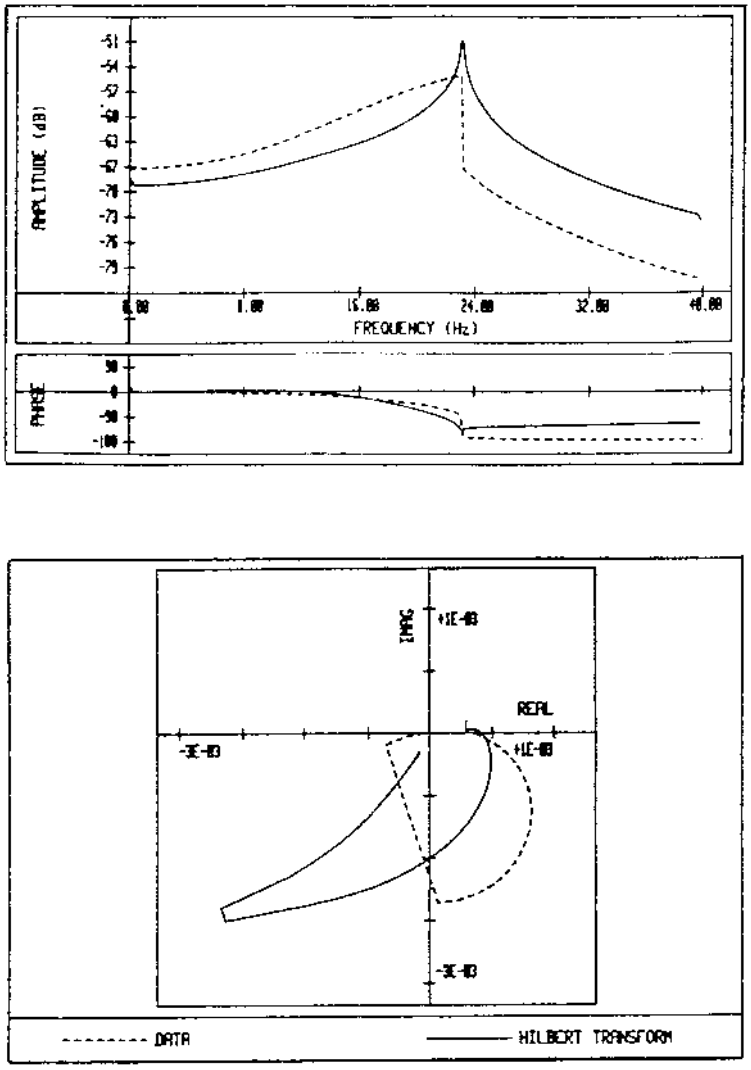


Figure 4.24. Hilbert transform of a hardening cubic spring FRF at a high sine excitation level.

4.4.4 Coulomb friction

The equation of motion is

$$m\ddot{y} + c\dot{y} + c_{\mathrm{F}}\dot{y}|\dot{y}| + ky + = x(t), \quad c_{\mathrm{F}} > 0. \tag{4.97}$$

The FRF and Hilbert transform are given in figure 4.27. In the Bode plot

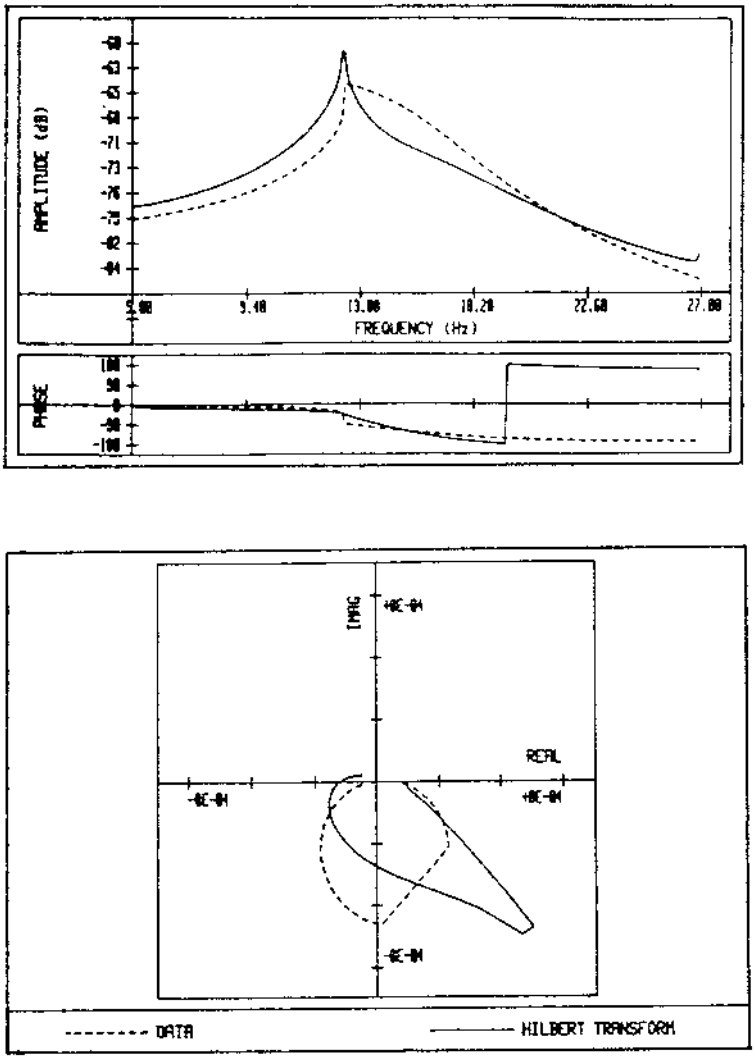


Figure 4.25. Hilbert transform of a softening cubic spring FRF at a high sine excitation level.

the peak of the Hilbert transform curve stays at the same frequency as in the FRF, but decreases in magnitude. In the Nyquist plot, the characteristic circle is compressed into an ellipse along the imaginary axis.

Note that in the case of Coulomb friction, the nonlinearity is only visible if the level of excitation is *low*. Figure 4.28 shows the FRF and transform at a high level of excitation where the system is essentially linear.

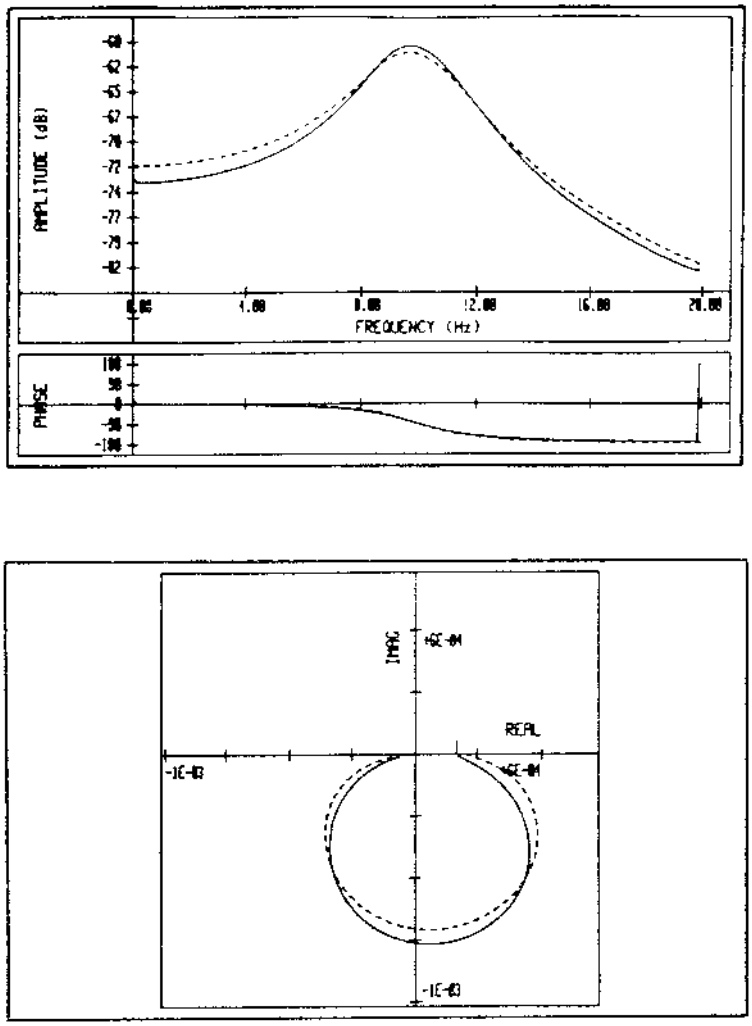


Figure 4.26. Hilbert transform of a velocity-squared damping FRF.

4.5 Choice of excitation

As discussed in the first two chapters, there are essentially four types of excitation which can be used to produce a FRF: impulse, stepped-sine, chirp and random. Figure 2.17 shows the resulting FRFs. The question arises as to which of the FRFs generates the inverse Fourier transform with the most marked non-causality; this will be the optimal excitation for use with the Hilbert transform.

Roughly speaking, the FRFs with the most marked distortion will transform

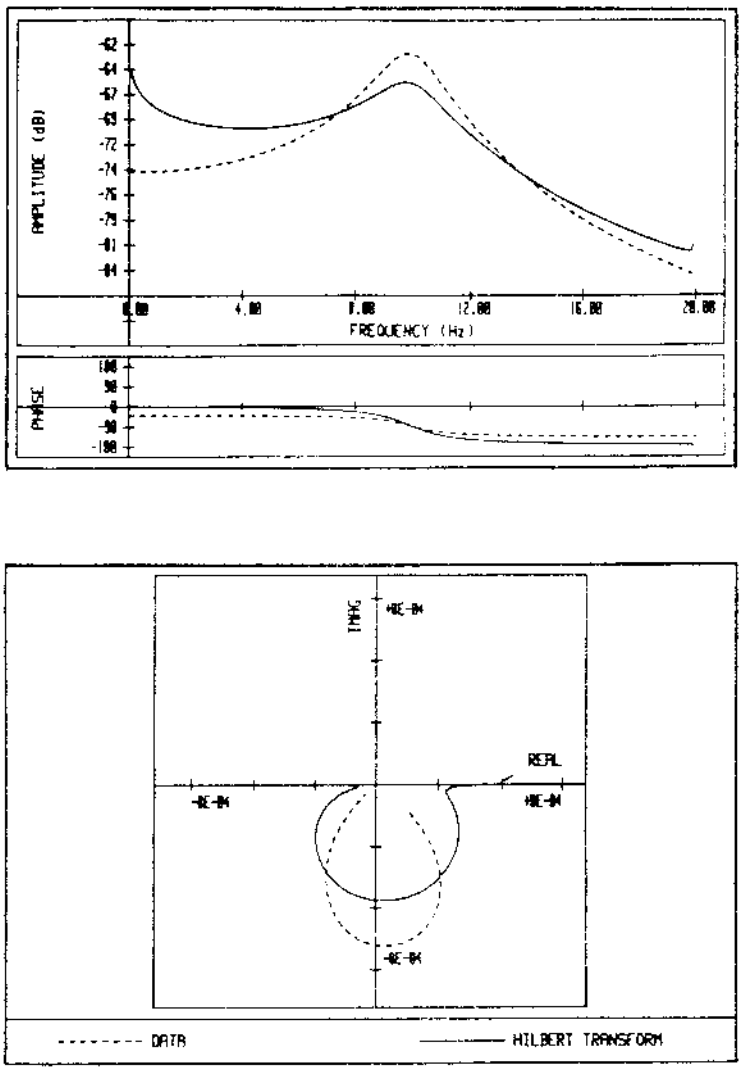


Figure 4.27. Hilbert transform of a Coulomb friction system FRF at a low sine excitation level.

to the most non-causal time functions. Recalling the discussion of chapter 2, the most distorted FRFs are obtained from stepped-sine excitation and, in fact, it will be proved later that such FRFs for nonlinear systems will generically show Hilbert transform distortions. (The proof requires the use of the Volterra series and is therefore postponed until chapter 8 where the appropriate theory is introduced.)

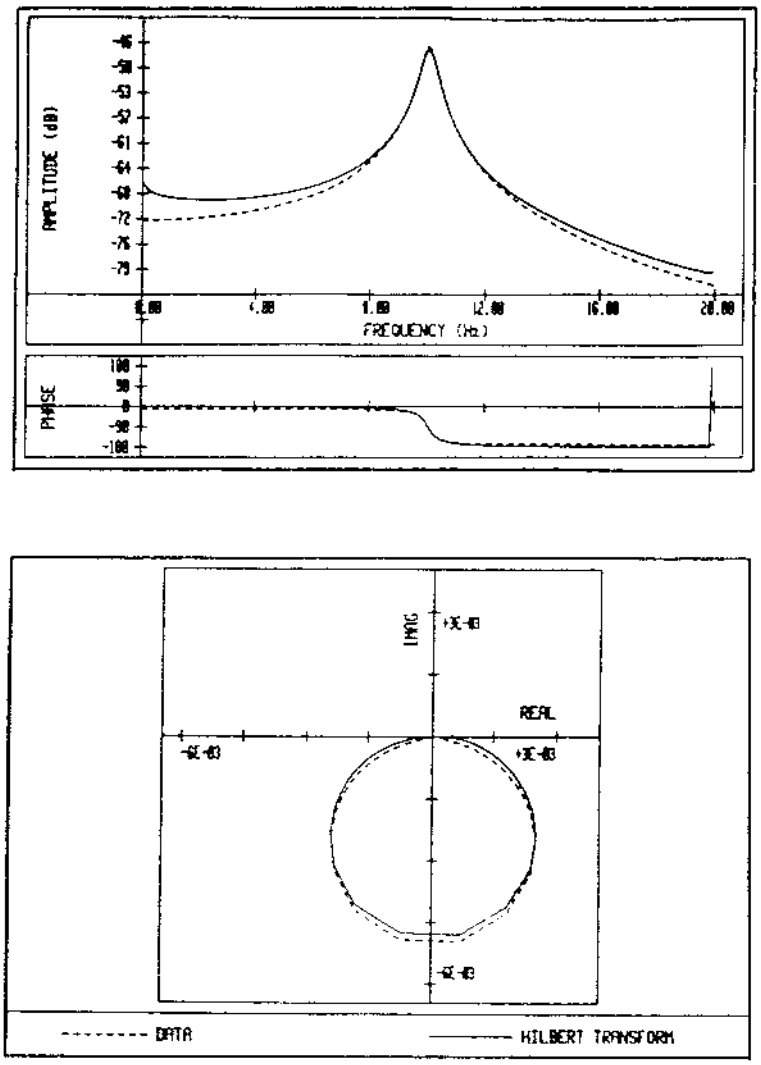


Figure 4.28. Hilbert transform of a Coulomb friction system FRF at a high sine excitation level.

This form of excitation is therefore recommended. The main disadvantage is its time-consuming nature.

At the other end of the spectrum is random excitation. As discussed in chapter 2, random excitation has the effect of producing a FRF which appears to be linearized about the operating level. For example, as the level of excitation is increased for a hardening cubic system, the resonant frequency increases,

but the characteristic linear Lorentzian shape appears to be retained. In fact, Volterra series techniques (chapter 8) provide a compelling argument that random excitation FRFs *do* change their form for nonlinear systems, *but* they still do not show Hilbert transform distortions. Random excitation should not, therefore, be used if the Hilbert transform is to be used as a diagnostic for detecting nonlinearity.

The impulse and chirp excitations are intermediate between these two extremes. They can be used if the test conditions dictate accordingly. Both methods have the advantage of giving broadband coverage at reasonable speed.

4.6 Indicator functions

The Hilbert transform operations described earlier give a diagnosis of nonlinearity with a little qualitative information available to those with appropriate experience. There has in the past been some effort at making the method quantitative. The FREEVIB approach discussed later actually provides an estimate of the stiffness or damping functions under certain conditions. There are also a number of less ambitious attempts which are usually based on computing some statistic or *indicator function* which sheds light on the type or extent of nonlinearity. Some of the more easily computable or interpretable are discussed in the following.

4.6.1 NPR: non-causal power ratio

This statistic was introduced in [141]. It does not make direct use of the Hilbert transform, but it is appropriate to discuss it here as it exploits the artificial non-causality of nonlinear system 'impulse responses'. The method relies on the decomposition

$$g(t) = \mathcal{F}^{-1}\{G(\omega)\} = g_{\mathrm{n}}(t) + g_{\mathrm{c}}(t) \tag{4.98}$$

where $g_{\mathrm{c}}(t)$ is the *causal part* defined by

$$g_{\mathrm{c}}(t) = \begin{cases} g(t), & t \geq 0 \\ 0, & t < 0 \end{cases} \tag{4.99}$$

and $g_{\mathrm{n}}(t)$ is the *non-causal part*

$$g_{\mathrm{n}}(t) = \begin{cases} 0, & t \geq 0 \\ g(t), & t < 0. \end{cases} \tag{4.100}$$

The non-causal power ratio (NPR) is then defined as the ratio of non-causal power P_{n} to the total system power P as encoded in the FRF

$$\mathrm{NPR} = \frac{P_{\mathrm{n}}}{P} = \frac{\int_{-\infty}^{0} \mathrm{d}t\, |g_{\mathrm{n}}(t)|^2}{\int_{-\infty}^{\infty} \mathrm{d}t\, |g(t)|^2}. \tag{4.101}$$

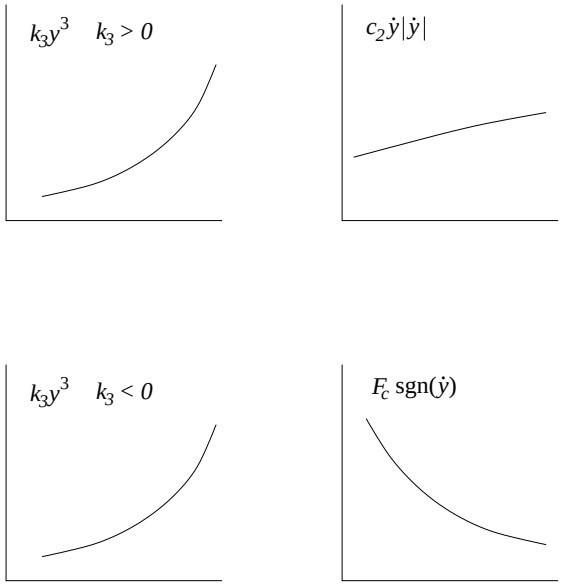


Figure 4.29. Non-causal power ratio plots for various SDOF nonlinear systems.

By Parseval's theorem, this also has a representation as

$$\mathrm{NPR} = \frac{P_{\mathrm{n}}}{P} = \frac{\int_{-\infty}^{0} \mathrm{d}t\, |g_{\mathrm{n}}(t)|^2}{\frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{d}\omega\, |G(\omega)|^2}. \tag{4.102}$$

This index is readily computed using an inverse FFT.

The NPR is, of course, a function of excitation amplitude (the form of the excitation being dictated by the considerations of the previous section). Kim and Park [141] compute this function for a number of common nonlinearities: hardening and softening cubic springs and quadratic and Coulomb damping. It is argued that the functions are characteristic of the nonlinearity as shown in figure 4.29, the cubic nonlinearities show NPRs which increase quickly with amplitude as expected. The NPR for quadratic damping shows a much more gentle increase, and the Coulomb friction function decreases with amplitude—again in agreement with intuition. The function certainly gives an indication of nonlinearity, but claims that it can suggest the type are probably rather optimistic.

The method is not restricted to SDOF systems. A case study is presented in [141] and it is suggested that computing the NPRs for all elements of the FRF matrix can yield information about the probable location of the nonlinearity.

4.6.2 Corehence

This measure of nonlinearity, based on the Hilbert transform, was introduced in [213] as an adjunct to the coherence function described in chapter 2. The basis of the theory is the *operator of linearity* P, defined by[9]

$$\tilde{G}(\omega) = P(\omega)G(\omega). \tag{4.103}$$

The operator is the identity $P(\omega) = 1\ \forall\omega$ if the system is linear (i.e. $G(\omega)$ has a causal inverse Fourier transform). Deviations of P from unity indicate nonlinearity. Note that P is a function of the level of excitation. As in the case of the coherence γ^2 (chapter 2), it is useful to have a normalized form for the operator, this is termed the *corehence* and denoted by λ^2. The definition is[10]

$$\lambda(\omega)^2 = \frac{|E\{\tilde{G}(\omega)G(\omega)^*\}|^2}{E\{|\tilde{G}(\omega)|^2\}E\{|G(\omega)|^2\}}. \tag{4.104}$$

There appears to be one major advantage of corehence over coherence. Given a coherence which departs from unity, it is impossible to determine whether the departure is the result of nonlinearity or measurement noise. It is claimed in [213] that this is not the case for corehence, it only responds to nonlinearity. It is also stated that a coherence of unity does not imply that the system is nonlinear. However, a rather unlikely type of nonlinearity is needed to create this condition. It is suggested that the corehence is more sensitive than the coherence.

4.6.3 Spectral moments

Consider a generic time signal $x(t)$; this has a representation

$$x(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{d}\omega\, \mathrm{e}^{\mathrm{i}\omega t} X(\omega) \tag{4.105}$$

where $X(\omega)$ is the spectrum. It follows that, if $x(t)$ is n-times differentiable,

$$\frac{\mathrm{d}^n x}{\mathrm{d}t^n} = \frac{\mathrm{i}^n}{2\pi}\int_{-\infty}^{\infty} \mathrm{d}\omega\, \omega^n \mathrm{e}^{\mathrm{i}\omega t} X(\omega) \tag{4.106}$$

[9] There are actually a number of P operators, each associated with a different FRF estimator, i.e. H_1, H_2 etc. The results in the text are for the estimator $H_1(\omega) = S_{yx}(\omega)/S_{xx}(\omega)$.

[10] The actual definition in [213] is

$$\lambda(\omega)^2 = \frac{|\tilde{G}(\omega)G(\omega)^*|^2}{|\tilde{G}(\omega)|^2\,|G(\omega)|^2}.$$

However, the expectation operators are implied; if the $G(\omega)$ and$\tilde{G}(\omega)$ are themselves expectations, expression (4.104) collapses to unity. There, is therefore, an implicit assumption that the form of excitation must be random as it is in the case of the coherence. Now, it is stated above that the Hilbert transform of an FRF obtained from random excitation does not show distortions. This does not affect the utility of the corehence as that statement only applies to the expectation of the FRF, i.e. the FRF after averaging. Because $E\{\tilde{G}G^*\} \neq E\{\tilde{G}\}E\{G\}$, the corehence departs from unity for nonlinear systems.

so

$$\left.\frac{\mathrm{d}^n x}{\mathrm{d}t^n}\right|_{t=0} = \frac{\mathrm{i}^n}{2\pi}\int_{-\infty}^{\infty} \mathrm{d}\omega\, \omega^n X(\omega) = \frac{\mathrm{i}^n}{2\pi} M^{(n)} \tag{4.107}$$

where $M^{(n)}$ denotes the nth moment integral of $X(\omega)$ or the nth spectral moment—$\int_{-\infty}^{\infty} \mathrm{d}\omega\, \omega^n X(\omega)$. Now it follows from the Taylor's series

$$x(t) = \sum_{n=1}^{\infty} \frac{1}{n!} \left.\frac{\mathrm{d}^n x}{\mathrm{d}t^n}\right|_{t=0} t^n = \frac{1}{2\pi} \sum_{n=1}^{\infty} M^{(n)} \frac{(\mathrm{i}t)^n}{n!} \tag{4.108}$$

that the function $x(t)$ is specified completely by the set of spectral moments. As a result, $X(\omega)$ is also specified by this set of numbers. The moments offer a means of characterizing the shape of the FRF or the corresponding Hilbert transform in terms of a small set of parameters. Consider the analogy with statistical theory: there, the mean and standard deviation (first- and second-order moments) of a probability distribution establish the gross features of the curve. The third- and fourth-order moments describe more subtle features—the skewness and the 'peakiness' (kurtosis). The latter features are considered to be measures of the distortion from the ideal Gaussian form. The zeroth moment is also informative; this is the energy or area under the curve.

Assuming that the moments are estimated for a single resonance between $\omega_{\min}$ and $\omega_{\max}$, the spectral moments of an FRF $G(\omega)$ are

$$M_G^{(n)} = \int_{\omega_{\min}}^{\omega_{\max}} \mathrm{d}\omega\, \omega^n G(\omega). \tag{4.109}$$

Note that they are complex, and in general depend on the limits; for consistency, the half-power points are usually taken. The moments are approximated in practice by

$$M_G^{(n)} \approx \sum_{k=\omega_{\min}}^{\omega_{\max}} \omega_k^n G(\omega_k) \Delta\omega \tag{4.110}$$

where $\Delta\omega$ is the spectral line spacing.

So-called Hilbert transform describers—HTDs—are then computed from

$$\mathrm{HTD}^{(n)} = 100 \frac{M_{\tilde{G}}^{(n)} - M_G^{(n)}}{M_G^{(n)}} \tag{4.111}$$

and these are simply the percentage differences between the Hilbert transform moments and the original FRF moments.

In practice, only the lowest-order moments have been investigated; in the terminology of [145], they are

$$\begin{aligned} \text{real energy ratio (RER)} &= \mathrm{Re}\,\mathrm{HTD}^{(0)} \\ \text{imaginary energy ratio (IER)} &= \mathrm{Im}\,\mathrm{HTD}^{(0)} \\ \text{real frequency ratio (RFR)} &= \mathrm{Re}\,\mathrm{HTD}^{(1)}. \end{aligned}$$

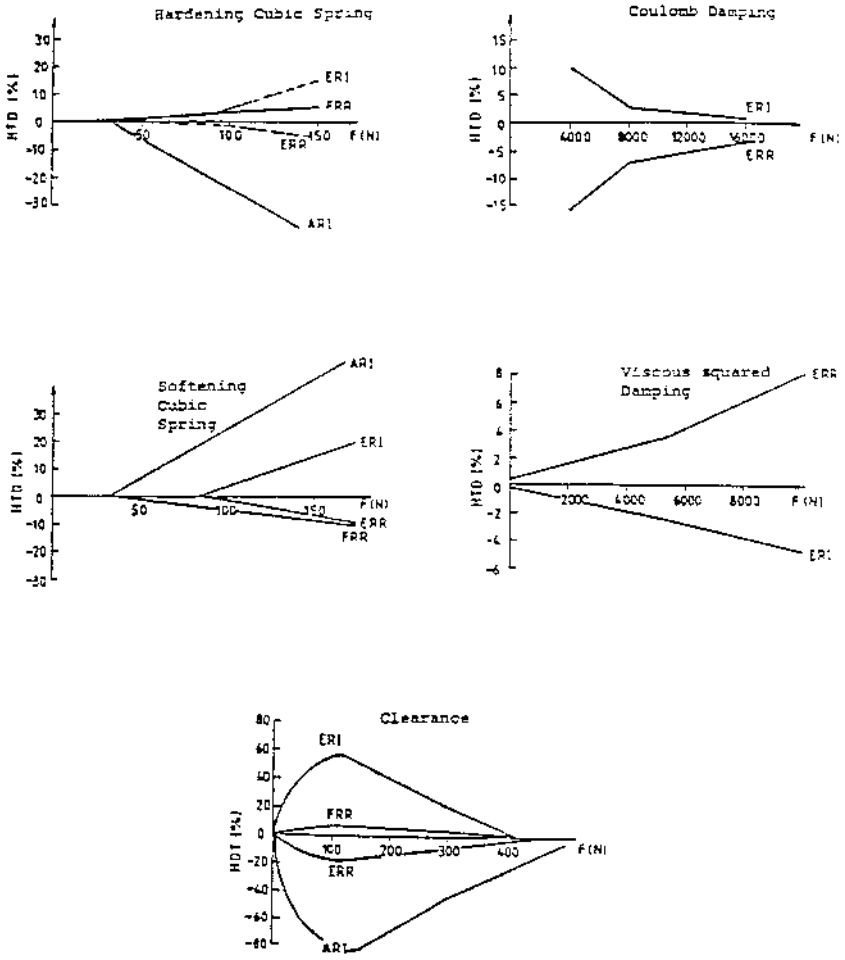


Figure 4.30. The variation in Hilbert transform describers (HTDs) for various SDOF nonlinear systems.

They are supplemented by

$$\text{imaginary amplitude ratio (IAR)} = \text{Im}\, 100 \frac{N_{\bar{G}} - N_G}{N_G}$$

where

$$N_G = \int_{\omega_{\min}}^{\omega_{\max}} \mathrm{d}\omega \, G(\omega)^2 \tag{4.112}$$

(which is essentially the centroid of the FRF about the ω-axis).

Figure 4.30 shows the plots of the HTD statistics as a function of applied force for several common nonlinearities. The parameters appear to separate stiffness and damping nonlinearities very effectively. Stiffness nonlinearity is identified from the changes in the RFR and IAR, while damping nonlinearity is indicated by changes in the energy statistics without change in the other describers. Note that the describers tend to zero at low forcing for the polynomial nonlinearities as expected; $\tilde{G} \longrightarrow G$ in this region. For the discontinuous nonlinearities, clearance and friction, the describers tend to zero at high forcing as the behaviour near the discontinuities becomes less significant. The describers therefore indicate the level of forcing at which the FRF of the underlying linear system can be extracted.

4.7 Measurement of apparent damping

It is well known that the accurate estimation of damping for lightly damped and/or nonlinear structures presents a difficult problem. In the first case, traditional methods of curve-fitting to FRFs break down due to low resolution of the peaks. In the second case, the damping ratio $c/2\sqrt{km}$ is not constant, whether the nonlinearity is in stiffness or damping (as a result, the term *apparent damping ratio* is used). However, it transpires that there is an effective procedure based on the Hilbert transform [245], which has actually been implemented on several commercial FRF analyers. The application to light damping is discussed in [4, 5]. Investigations of nonlinear systems are presented in [187, 188].

The basis of the method is the analytic signal. Consider the function $\mathrm{e}^{(-\lambda+\mathrm{i}\mu)t}$ with $\lambda > 0$. It is shown in appendix C that there are relations between the real and imaginary parts:

$$\mathcal{H}\{\mathrm{e}^{-\lambda t}\sin(\mu t)\} = \mathrm{i}\mathrm{e}^{-\lambda t}\cos(\mu t) \tag{4.113}$$

and

$$\mathcal{H}\{\mathrm{e}^{-\lambda t}\cos(\mu t)\} = -\mathrm{i}\mathrm{e}^{-\lambda t}\sin(\mu t) \tag{4.114}$$

provided λ is small. These relations therefore apply to the impulse response of a linear system provided the damping ratio ζ is small (overall constant factors have no effect):

$$h(t) = \frac{1}{m\omega_{\mathrm{d}}}\mathrm{e}^{-\zeta\omega_n t}\sin(\omega_{\mathrm{d}} t), \quad t > 0 \tag{4.115}$$

which can be interpreted as the real part of an analytic signal,

$$\begin{aligned} a_h(t) = h(t) - \tilde{h}(t) &= \frac{1}{m\omega_{\mathrm{d}}}\mathrm{e}^{-\zeta\omega_n t}\sin(\omega_{\mathrm{d}} t) - \mathrm{i}\frac{1}{m\omega_{\mathrm{d}}}\mathrm{e}^{-\zeta\omega_n t}\cos(\omega_{\mathrm{d}} t) \\ &= \frac{-\mathrm{i}}{m\omega_{\mathrm{d}}}\mathrm{e}^{(-\zeta\omega_n+\mathrm{i}\omega_{\mathrm{d}})t}. \end{aligned} \tag{4.116}$$

Now, the magnitude of this analytic signal is given by

$$|a_h(t)| = \sqrt{h^2 - \tilde{h}^2} = \frac{1}{m\omega_{\mathrm{d}}} \mathrm{e}^{-\zeta\omega_n t} \tag{4.117}$$

and this is revealed as the *envelope* of the impulse response (see section 3.12)[11]. Taking the natural logarithm of this expression yields

$$\log|a_h(t)| = -\zeta\omega_n t - \log(m\omega_{\mathrm{d}}) \tag{4.118}$$

and this provides a new *time-domain* algorithm for estimating the damping of a system, given the linear system FRF $H(\omega)$:

(1) Take the inverse Fourier transform of $H(\omega)$ to get the impulse response $h(t)$.
(2) Take the Hilbert transform of $h(t)$ and form the *analytic impulse response* $a_h(t)$ as in (4.116).
(3) Plot the log magnitude of $a_h(t)$ against time; the gradient (extracted by a linear regression) is $\sigma = -\zeta\omega_n$.
(4) If ω_{d} is measured, $\omega_n = \sqrt{\zeta^2\omega_n^2 + \omega_{\mathrm{d}}^2} = \sqrt{\sigma^2 + \omega_{\mathrm{d}}^2}$ and $\zeta = \sigma/\omega_n$.

There are no real subtleties involved in applying the method to a nonlinear system. The only critical factor is choice of excitation. It can be shown that random excitation properly represents the apparent damping (in the sense that the FRF S_{yx}/S_{xx} correctly represents the amount of power dissipated), this is the appropriate excitation. Note that curve-fitting to the FRF would also characterize the damping; this method is of interest because it extends to light damping, is more insensitive to noise and also because it makes neat use of the Hilbert transform.

To illustrate the procedure, random excitation FRFs were obtained for the Duffing oscillator system

$$\ddot{y} + 5\dot{y} + 10^4 y + 10^9 y^3 = x(t) \tag{4.119}$$

at low and high levels of excitation. Figure 4.31 shows the corresponding log envelopes. Extremely clear results are obtained in both cases. In contrast, the corresponding FRFs with curve-fits are shown in figure 4.32. The high excitation FRF is significantly noisier.

[11] Note that using the conventional definition of analytic signal and Hilbert transform given in footnote 4.6, equation (4.116) is modified to

$$a_h(t) = h(t) + \mathrm{i}\tilde{h}(t) = \frac{1}{m\omega_{\mathrm{d}}} \mathrm{e}^{-\zeta\omega_n t} \sin(\omega_{\mathrm{d}} t) - \mathrm{i}\frac{1}{m\omega_{\mathrm{d}}} \mathrm{e}^{-\zeta\omega_n t} \cos(\omega_{\mathrm{d}} t) = \frac{-\mathrm{i}}{m\omega_{\mathrm{d}}} \mathrm{e}^{(-\zeta\omega_n + \mathrm{i}\omega_{\mathrm{d}})t}$$

and equation (4.117) becomes

$$|a_h(t)| = \sqrt{h^2 + \tilde{h}^2} = \frac{1}{m\omega_{\mathrm{d}}} \mathrm{e}^{-\zeta\omega_n t}$$

and the argument then proceeds unchanged.

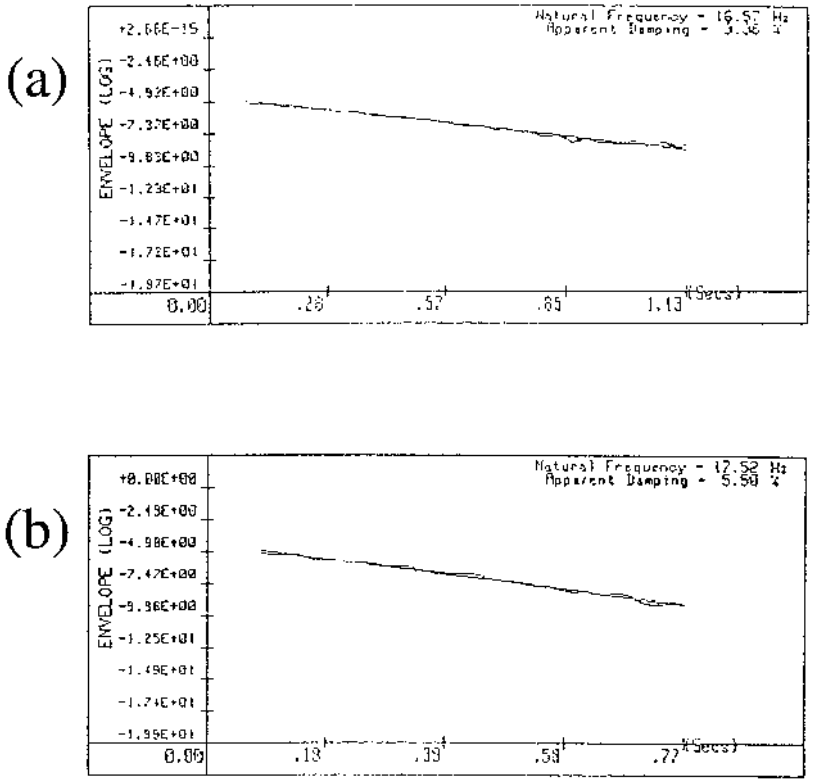


Figure 4.31. Impulse response and envelope function for a nonlinear system under random excitation: (*a*) low level; (*b*) high level.

An experimental example for an impacting cantilever beam (figure 4.33) also shows the utility of the method. Figure 4.34 shows the FRF, impulse response and log envelope for the low excitation case where the system does not impact. Figure 4.35 shows the corresponding plots for the high-excitation contacting case—note that the FRF is considerably noisier. If the initial, linear, portions of the log envelope curves are used for regression, the resulting natural frequencies and damping ratios are given in figure 4.36.

The apparent variation in damping ratio ζ is due to the fact that the definition $\zeta = c/\sqrt{km}$ depends on the nonlinear stiffness. The corresponding value of c should be constant (by linearization arguments presented in chapter 2).

4.8 Identification of nonlinear systems

The method described in this section is the result of a programme of research by Feldman [92, 93, 94]. It provides a means of obtaining the stiffness and damping

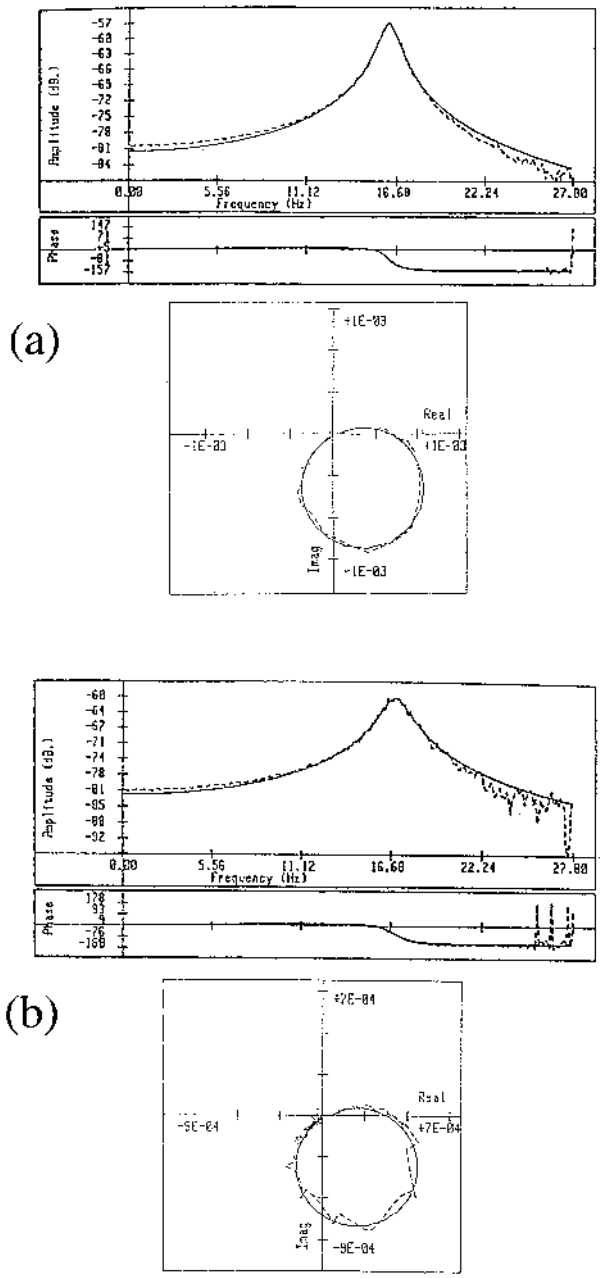


Figure 4.32. Result of curve-fitting FRFs for data in figure 4.31.

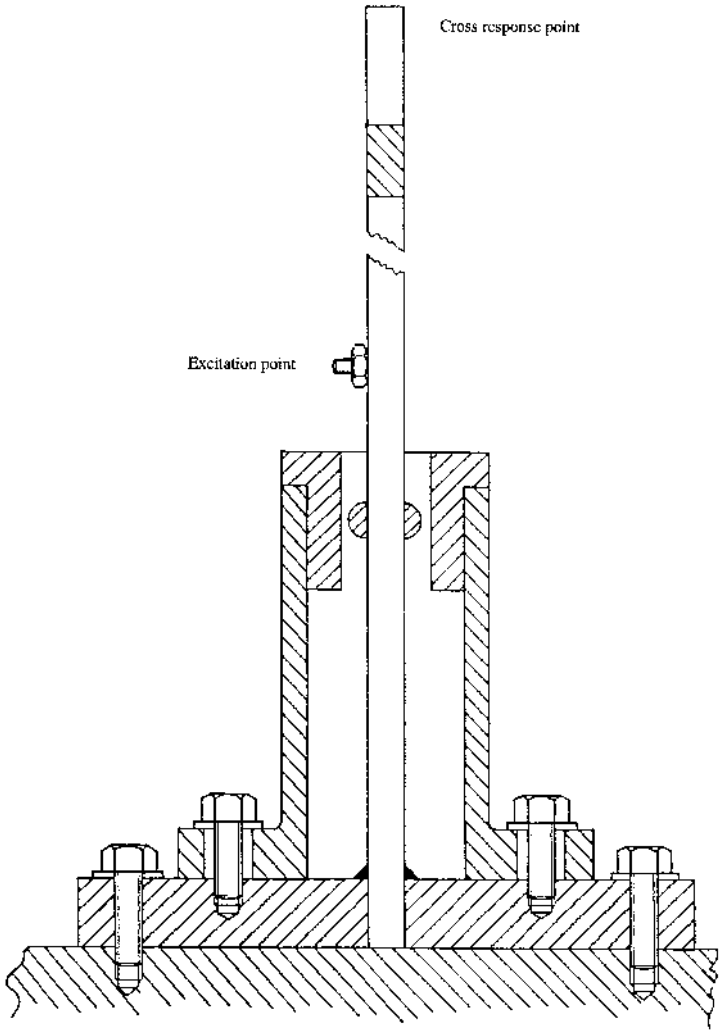


Figure 4.33. Nonlinear (impacting) cantilever beam test rig.

characteristics of SDOF systems. There are essentially two approaches, one based on free vibration *FREEVIB* and one on forced vibration *FORCEVIB*. They will be discussed separately. Note that Feldman uses the traditional definition of the analytic signal and time-domain Hilbert transform throughout his analysis.

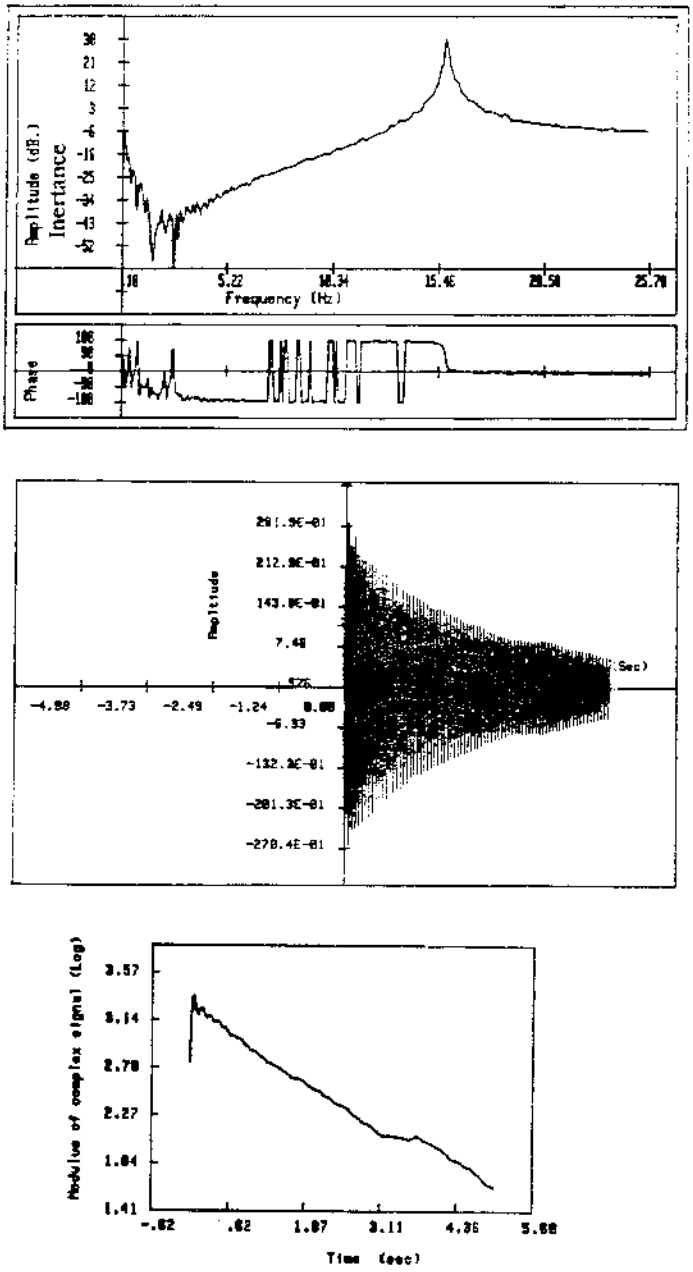


Figure 4.34. Data from the nonlinear beam in non-impacting condition: (*a*) measured FRF; (*b*) calculated impulse response; (*c*) calculated envelope.

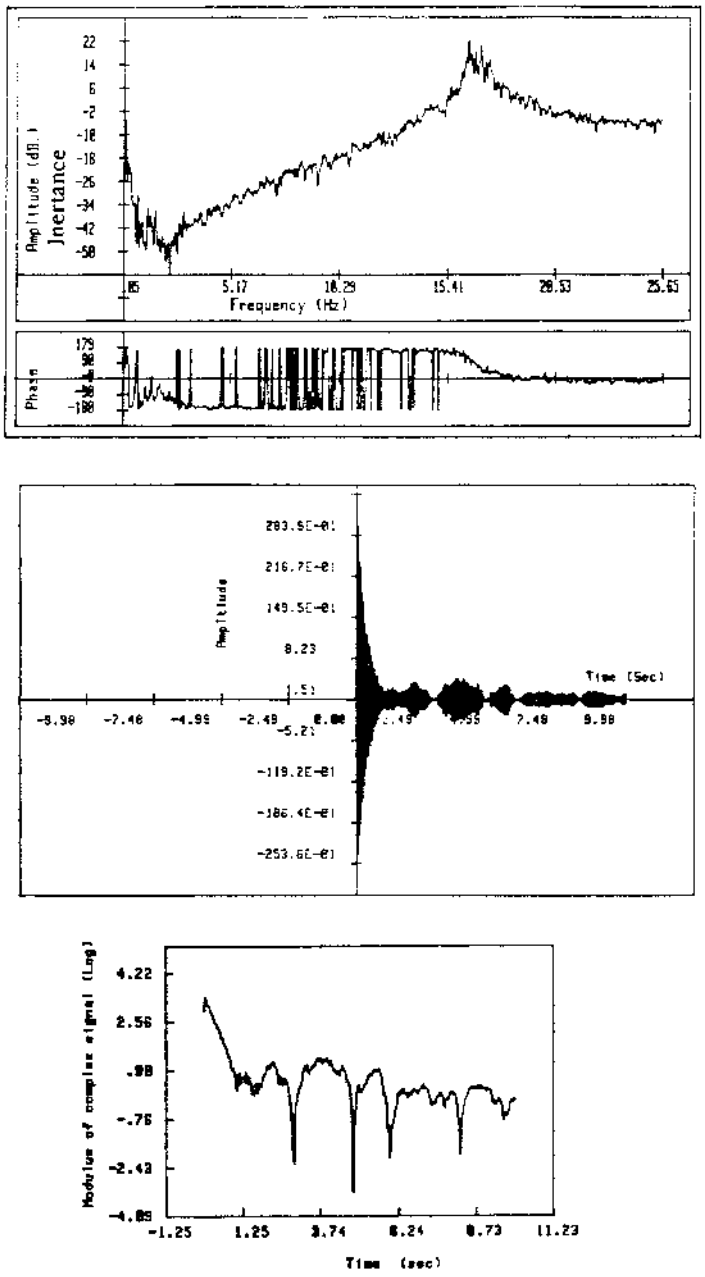


Figure 4.35. Data from the nonlinear beam impacting condition: (*a*) measured FRF; (*b*) calculated impulse response; (*c*) calculated envelope.

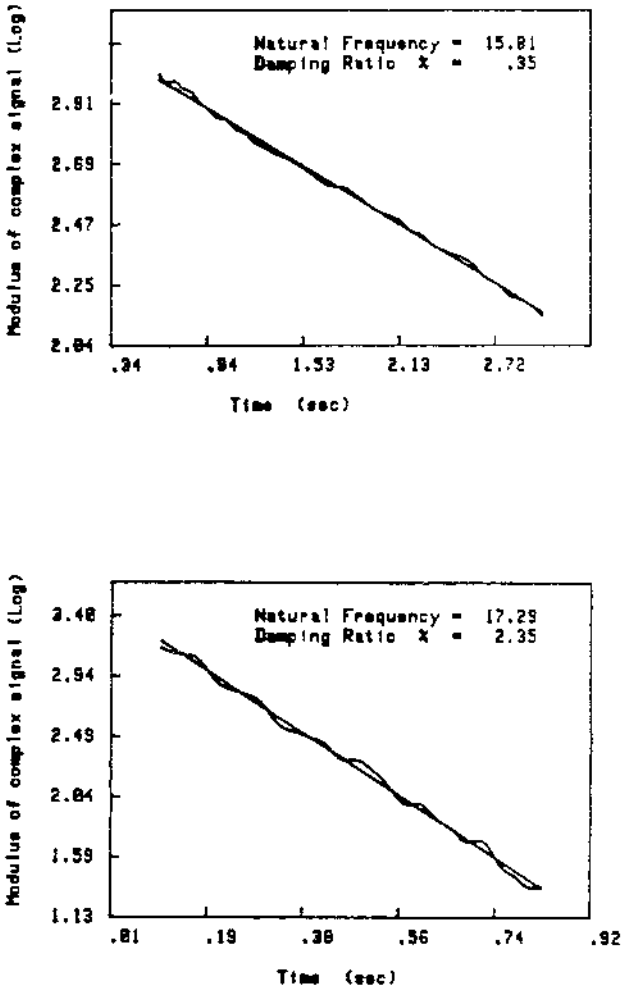


Figure 4.36. Results of the estimated natural frequency and apparent damping ratio for the impacting cantilever: (*a*) linear regime; (*b*) nonlinear regime.

4.8.1 FREEVIB

Consider a SDOF nonlinear system under free vibration:

$$\ddot{y} + h(\dot{y})\dot{y} + \omega_0^2(y)y = 0. \tag{4.120}$$

The object of the exercise of identification is to use measured data, say $y(t)$, and deduce the forms of the nonlinear damping function $h(\dot{y})$ and nonlinear stiffness $k(y) = \omega_0^2(y)$.

The method is based on the analytic signal defined in (4.82)

$$Y(t) = y(t) - \tilde{y}(t) \tag{4.121}$$

and uses the magnitude and phase representation

$$Y(t) = A(t)\mathrm{e}^{\mathrm{i}\psi(t)} \tag{4.122}$$

where $A(t)$ is the instantaneous magnitude or envelope, and $\psi(t)$ is the instantaneous phase. Both are real functions so

$$y(t) = A(t)\cos(\psi(t)), \quad \tilde{y} = -\mathrm{i}A(t)\sin(\psi(t)) \tag{4.123}$$

and

$$A(t) = \sqrt{y(t)^2 - \tilde{y}(t)^2} \tag{4.124}$$

$$\psi(t) = \tan^{-1}\left(\frac{\tilde{y}(t)}{\mathrm{i}y(t)}\right). \tag{4.125}$$

So both envelope and phase are available as functions of time if $y(t)$ is known and $\tilde{y}(t)$ can be computed. The derivatives can also be computed, either directly or using the relations

$$\dot{A}(t) = \frac{y(t)\dot{y}(t) - \tilde{y}(t)\dot{\tilde{y}}(t)}{\sqrt{y(t)^2 - \tilde{y}(t)^2}} = A(t)\,\mathrm{Re}\left[\frac{\dot{Y}(t)}{Y(t)}\right] \tag{4.126}$$

$$\omega(t) = \dot{\psi}(t) = \frac{\mathrm{i}(y(t)\dot{\tilde{y}}(t) - \dot{y}(t)\tilde{y}(t))}{y(t)^2 - \tilde{y}(t)^2} = \mathrm{Im}\left[\frac{\dot{Y}(t)}{Y(t)}\right] \tag{4.127}$$

where $\omega(t)$ is the instantaneous frequency, again a real signal. The last two equations can be used to generate the first two derivatives of the analytic signal

$$\dot{Y}(t) = Y(t)\left[\frac{\dot{A}(t)}{A(t)} + \mathrm{i}\omega(t)\right] \tag{4.128}$$

$$\ddot{Y}(t) = Y(t)\left[\frac{\ddot{A}(t)}{A(t)} - \omega(t)^2 + 2\mathrm{i}\frac{\dot{A}(t)\omega(t)}{A(t)} + \mathrm{i}\dot{\omega}(t)\right]. \tag{4.129}$$

Now, consider the equation of motion (4.120), with $h(\dot{y}(t)) = h(t)$ and $\omega_0^2(y(t)) = \omega_0^2(t)$ considered purely as functions of time (there is a slight abuse of notation here). Because the functions h and ω_0^2 will generally be low-order polynomials of the envelope A, they will have a lowpass characteristic. If the resonant frequency of the system is high, $y(t)$ will, roughly speaking, have a highpass characteristic. This means that h and y can be considered as non-overlapping signals (see appendix C) as can ω_0^2 and y. If the Hilbert transform

is taken of (4.120), it will pass through the functions h and ω_0^2. Further, the transform commutes with differentiation (appendix C again), so

$$\ddot{\tilde{y}} + h(t)\dot{\tilde{y}} + \omega_0^2(t)\tilde{y} = 0. \tag{4.130}$$

Adding (4.120) and (4.130) yields a differential equation for the analytic signal Y, i.e.

$$\ddot{Y} + h(t)\dot{Y} + \omega_0^2(t)Y = 0 \tag{4.131}$$

or, the quasi-linear form,

$$\ddot{Y} + h(A)\dot{Y} + \omega_0^2(A)Y = 0. \tag{4.132}$$

Now, the derivatives $\ddot{Y}$ and $\dot{Y}$ are known functions of A and ω by (4.128) and (4.129). Substitution yields

$$Y\left[\frac{\ddot{A}}{A} - \omega^2 + \omega_0^2 + h\frac{\dot{A}}{A} + \mathrm{i}\left(2\omega\frac{\dot{A}}{A} + \dot{\omega} + h\omega\right)\right] = 0. \tag{4.133}$$

Separating out the real and imaginary parts gives

$$h(t) = -2\frac{\dot{A}}{A} - \frac{\dot{\omega}}{\omega} \tag{4.134}$$

$$\omega_0^2(t) = \omega^2 - \frac{\ddot{A}}{A} - h\frac{\dot{A}}{A} \tag{4.135}$$

or

$$\omega_0^2(t) = \omega^2 - \frac{\ddot{A}}{A} + 2\frac{\dot{A}^2}{A^2} + \frac{\dot{A}\dot{\omega}}{A\omega} \tag{4.136}$$

and these are the basic equations of the theory.

On to practical matters. Suppose the free vibration is induced by an impulse, the subsequent response of the system will take the form of a decay. $y(t)$ can be measured and $\tilde{y}$ can then be computed[12]. This means that $A(t)$ and $\omega(t)$ are available by using (4.124) and (4.125) and numerically differentiating $\psi(t)$.

Now, consider how the damping function is obtained. $h(t)$ is known from (4.134). As $A(t)$ is monotonically decreasing (energy is being dissipated), the inverse function $t(A)$ is single-valued and can be obtained from the graph of $A(t)$ against time (figure 4.37). The value of $h(A)$ is simply the value of $h(t)$ at $t(A)$ (figure 4.38). Similarly, the stiffness function is obtained via the sequence $A \longrightarrow t(A) \longrightarrow \omega_0^2(t(A)) = \omega_0^2(A)$. The inverse of the latter mapping $A(\omega)$ is sometimes referred to as the *backbone* curve of the system. (For fairly simple systems like the Duffing oscillator, the backbone curves can be calculated [41].)

[12] As in the frequency-domain case, there are a number of methods of computing $\tilde{y}$, the decomposition $\mathcal{H} = \mathcal{F}^{-1}\circ \in \circ\mathcal{F}$ provides one.

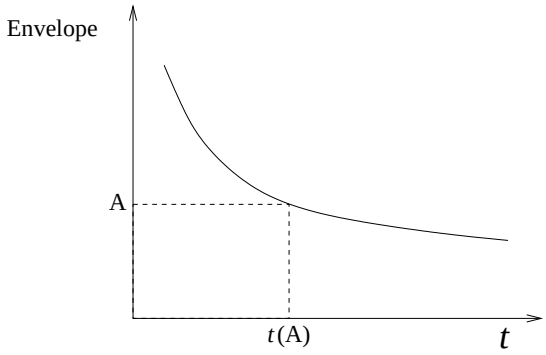


Figure 4.37. Envelope used in Feldman's method.

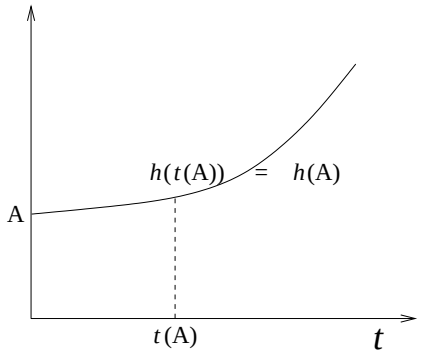


Figure 4.38. Damping curve for Feldman's method.

Once $h(A)$ and $\omega_0^2(A)$ are known, the damper and spring characteristics $f_{\mathrm{d}}(A)$ and $f_{\mathrm{s}}(A)$ can be obtained trivially

$$f_{\mathrm{d}}(A) = \omega(A)Ah(A) \tag{4.137}$$

$$f_{\mathrm{s}}(A) = A\omega_0^2(A). \tag{4.138}$$

Note that as there are no assumptions on the forms of f_{d} and f_{s}, the method is truly non-parametric. However, once the graphs $A \longrightarrow f_{\mathrm{d}}$ etc have been obtained, linear least-squares methods (as described in chapter 6) suffice to estimate parameters.

The method can be readily illustrated using data from numerical simulation[13]. The first system is a Duffing oscillator with equation of motion

$$\ddot{y} + 10\dot{y} + 10^4 y + 5 \times 10^4 y^3 = 0 \tag{4.139}$$

[13] The results for figures 4.39–4.41 were obtained by Dr Michael Feldman—the authors are very grateful for permission to use them.

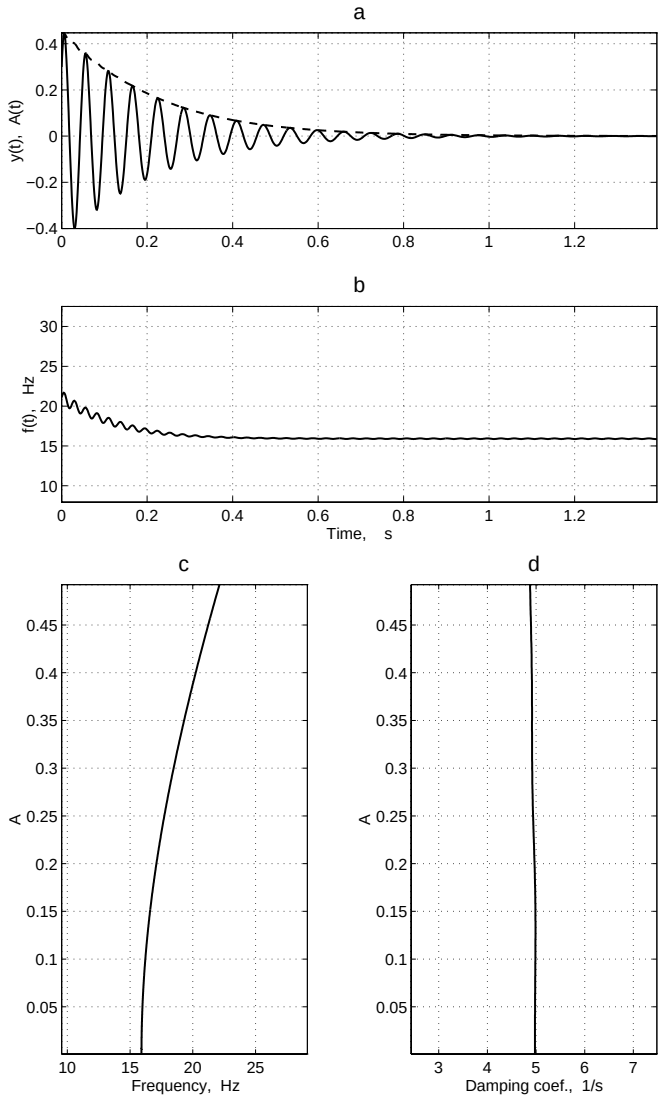


Figure 4.39. Identification of cubic stiffness system: (*a*) impulse response; (*b*) envelope; (*c*) backbone curve; (*d*) damping curve; (*e*) stiffness characteristic; (*f*) damping characteristic.

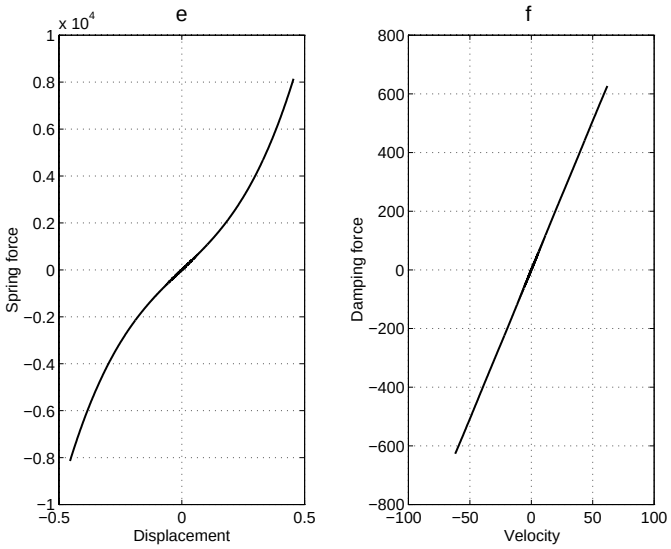


Figure 4.39. (Continued)

and initial condition $\dot{y}(0) = 200$. Figure 4.39(*a*) shows the decaying displacement and the envelope computed via equation (4.124). Figure 4.39(*b*) shows the corresponding instantaneous frequency obtained from (4.127). The backbone and damping curve are given in figures 4.39(*c*) and (*d*) respectively. As expected for a stiffening system, the natural frequency increases with the amplitude of excitation. Apart from a high-frequency modulation, the damping curve shows constant behaviour. Using equations (4.138) and (4.139), the stiffness and damping curves can be obtained and these are shown in figures 4.39(*e*) and (*f*).

The second example shows the utility of the method for non-parametric system identification. The system has a stiffness deadband, the equation of motion is

$$\ddot{y} + 5\dot{y} + f_{\mathrm{s}}(y) = 0 \tag{4.140}$$

where

$$f_{\mathrm{s}}(y) = \begin{cases} 10^4(y - 0.1), & y > 0.1 \\ 0, & |y| < 0.1 \\ 10^4(y + 0.1), & y < -0.1 \end{cases} \tag{4.141}$$

and the motion began with $\dot{y}(0) = 200$ once more. The sequence of figures 4.40(*a*)–(*f*) show the results of the analysis. The backbone curve (figure 4.40(*c*)) shows the expected result that the natural frequency is only sensitive to the nonlinearity for low levels of excitation. The stiffness curve (figure 4.40(*e*)) shows the size of the deadband quite clearly. (This is useful information, if the clearance is specified, the parameter estimation problem becomes linear and

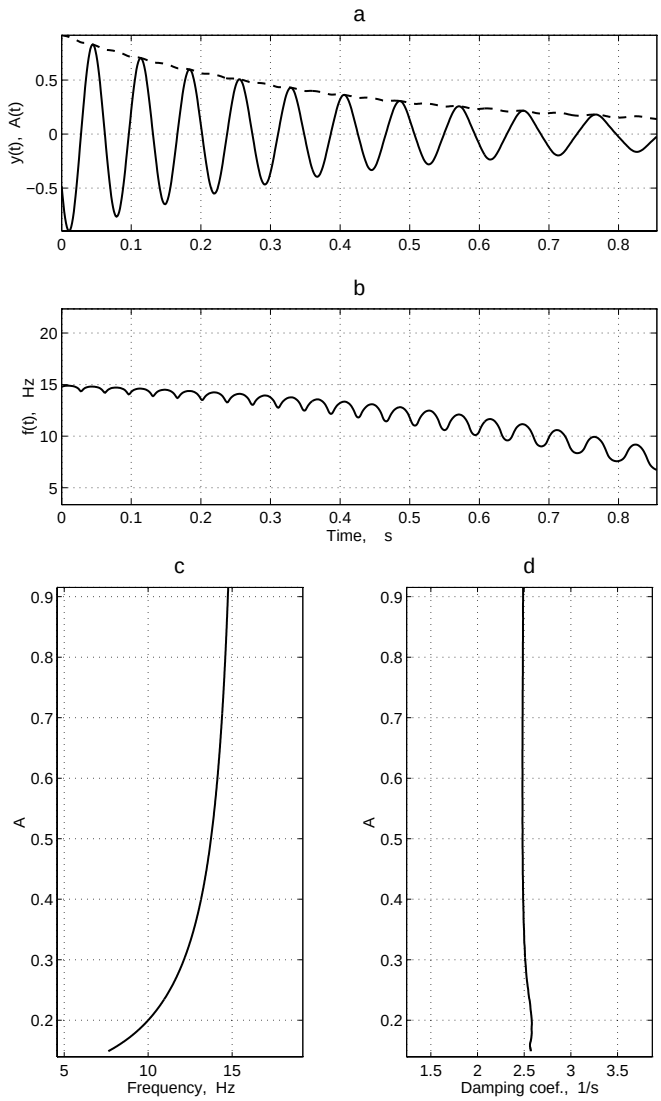


Figure 4.40. Identification of backlash system: (*a*) impulse response; (*b*) envelope; (*c*) backbone curve; (*d*) damping curve; (*e*) stiffness characteristic, (*f*) damping characteristic.

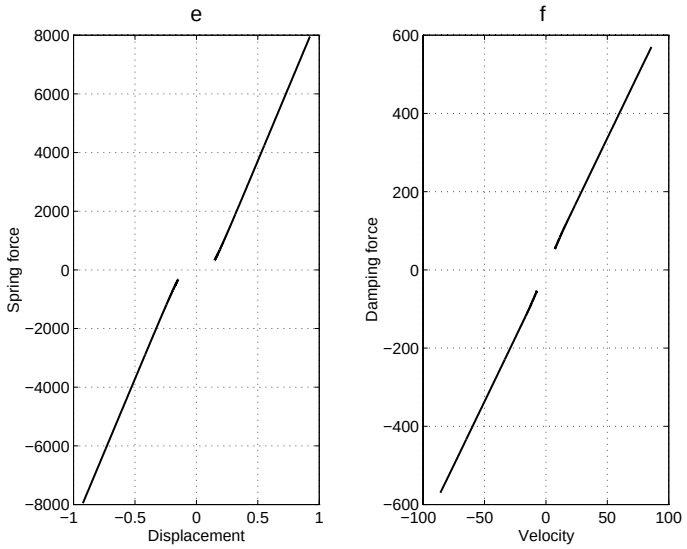


Figure 4.40. (Continued)

simple methods suffice to estimate the stiffness function.) Note that because the initial displacement did not decay away completely, there are gaps in the stiffness and damping functions at low amplitude.

The final example shows a damping nonlinearity. The system has equation of motion

$$\ddot{y} + 300\,\mathrm{sgn}(\dot{y}) + 10^4 y = 0 \tag{4.142}$$

so Coulomb friction is present. The decay began with the same initial conditions as before and the resulting anlysis is shown in figures 4.41(*a*)–(*f*). Note the characteristic linear decay envelope for this type of nonlinear system as shown in figure 4.41(*a*). In this case, the backbone (figure 4.41(*c*)) shows no variation of natural frequency with amplitude as expected. The coefficient of friction can be read directly from the damping function (figure 4.41(*f*)).

Further examples of nonlinear systems can be found in [93, 95]. A practical application to a nonlinear ocean mooring system is discussed in [120].

All of these examples have viscous damping models. It is a simple matter to modify the theory for structural (hysteretic) damping, the equation of motion for the analytic signal becomes

$$\ddot{Y} + \omega_0^2(A)\left[1 + \frac{\mathrm{i}}{\pi}\delta(A)\right]Y = 0 \tag{4.143}$$

where $\delta(A)$ is the loss factor or logarithmic decrement. The basic equations are

$$\omega_0^2(t) = \omega^2 - \frac{\ddot{A}}{A} \tag{4.144}$$

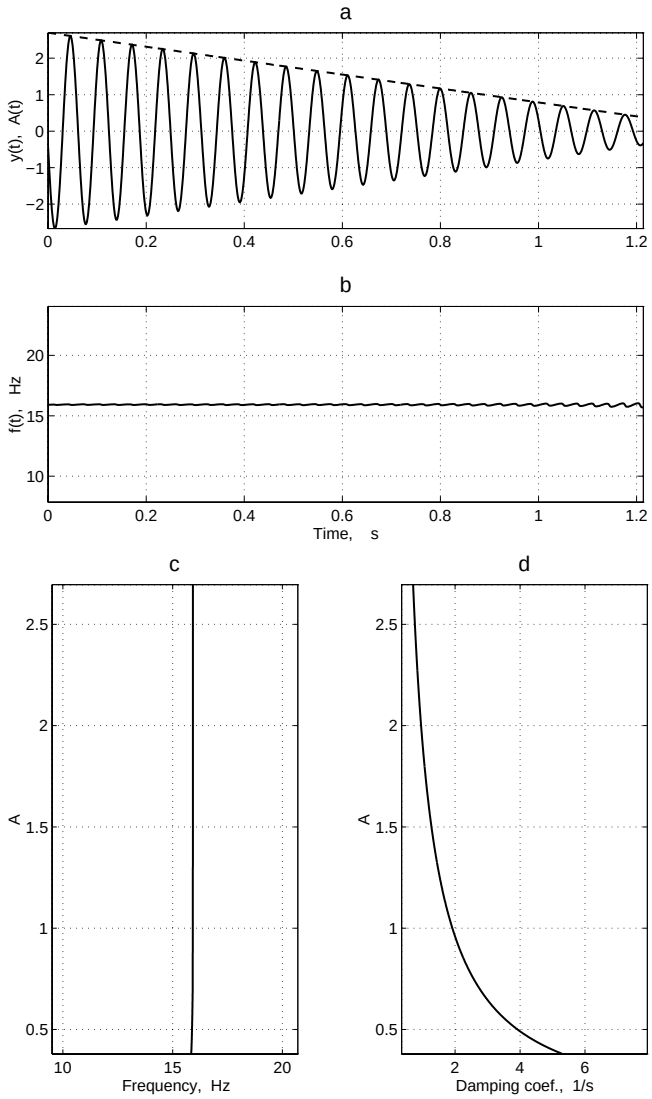


Figure 4.41. Identification of Coulomb friction system: (*a*) impulse response; (*b*) envelope; (*c*) backbone curve; (*d*) damping curve; (*e*) stiffness characteristic; (*f*) damping characteristic.

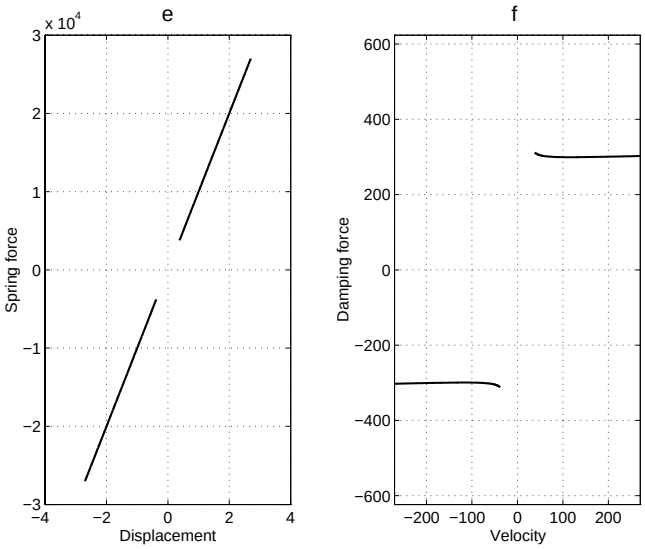


Figure 4.41. (Continued)

and

$$\delta(t) = -\frac{2\pi\dot{A}\omega}{A\omega_0^2} - \frac{\pi\dot{\omega}}{\omega_0^2}. \tag{4.145}$$

The method described here is only truly suitable for monocomponent signals, i.e. those with a single dominant frequency. The extension to two-component signals is discussed in [96].

4.8.2 FORCEVIB

The analysis for the forced vibration case is very similar to FREEVIB; the presence of the excitation complicates matters very little. Under all the same assumptions as before, the quasi-linear equation of motion for the analytic signal can be obtained:

$$\ddot{Y} + h(A)\dot{Y} + \omega_0^2(A)Y = \frac{X}{m}. \tag{4.146}$$

Carrying out the same procedures as before which lead to equations (4.134) and (4.135) yields

$$h(t) = \frac{\beta(t)}{\omega m} - 2\frac{\dot{A}}{A} - \frac{\dot{\omega}}{\omega} \tag{4.147}$$

and

$$\omega_0^2(t) = \omega^2 + \frac{\alpha(t)}{m} - \frac{\beta(t)\dot{A}}{A\omega m} - \frac{\ddot{A}}{A} + 2\frac{\dot{A}^2}{A^2} + \frac{\dot{A}\dot{\omega}}{A\omega} \tag{4.148}$$

where $\alpha(t)$ and $\beta(t)$ are, respectively, the real and imaginary parts of the input/output ratio X/Y, i.e.

$$\frac{X(t)}{Y(t)} = \alpha(t) + \mathrm{i}\beta(t) = \frac{x(t)y(t) + \tilde{x}(t)\tilde{y}(t)}{y^2(t) + \tilde{y}^2(t)} + \mathrm{i}\frac{\tilde{x}(t)y(t) - x(t)\tilde{y}(t)}{y^2(t) + \tilde{y}^2(t)} \quad (4.149)$$

where $x(t)$ is the real part of $X(t)$, i.e. the original physical excitation.

Implementation of this method is complicated by the fact that an estimate of the mass m is needed. This problem is discussed in detail in [94].

4.9 Principal component analysis (PCA)

This is a classical method of multivariate statistics and its theory and use are documented in any textbook from that field (e.g. [224]). Only the briefest description will be given here. Given a set of p-dimensional vectors $\{x\} = (x_1, \ldots, x_p)$, the principal components algorithm seeks to project, by a linear transformation, the data into a new p-dimensional set of Cartesian coordinates $(z_1, z_2, \ldots, z_p)$. The new coordinates have the following property: z_1 is the linear combination of the original x_i with *maximal* variance, z_2 is the linear combination which explains most of the remaining variance and so on. It should be clear that, if the p-coordinates are actually a linear combination of $q < p$ variables, the first q principal components will completely characterize the data and the remaining $p - q$ will be zero. In practice, due to measurement uncertainty, the principal components will all be non-zero and the user should select the number of *significant* components for retention.

Calculation is as follows: given data $\{x\}_i = (x_{1i}, x_{2i}, \ldots, x_{ip}), i = 1, \ldots, N$, form the covariance matrix $[\Sigma]$ (see appendix A—here the factor $1/(N-1)$ is irrelevant)

$$[\Sigma] = \sum_{i=1}^{N} (\{x\}_i - \{\overline{x}\})(\{x\}_i - \{\overline{x}\})^{\mathrm{T}} \quad (4.150)$$

(where $\{\overline{x}\}$ is the vector of means of the x data) and decompose so

$$[C] = [A][\Lambda][A]^{\mathrm{T}} \quad (4.151)$$

where $[\Lambda]$ is diagonal. (Singular value decomposition can be used for this step [209].) The transformation to principal components is then

$$\{z\}_i = [A]^{\mathrm{T}}(\{x\}_i - \{\overline{x}\}). \quad (4.152)$$

Considered as a means of dimension reduction then, PCA works by discarding those linear combinations of the data which contribute least to the overall variance or range of the data set. Another way of looking at the transformation is to consider it as a means of identifying correlations or

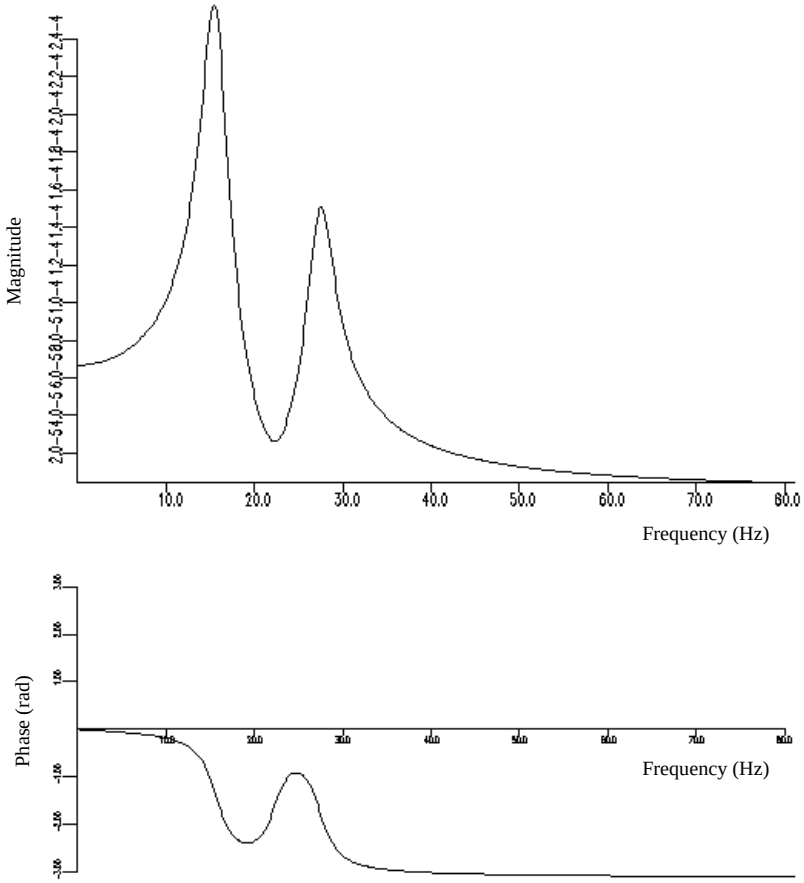


Figure 4.42. FRF H_1 for symmetric 2DOF linear system.

redundancy in data. The transformation to principal components results in uncorrelated vectors and thus eliminates the redundancy.

The first applications of the method in dynamics date back to the early 1980s. One of the first references is by Moore [191]. The first applications in modal testing or structural dynamics are due to Leuridan [163, 164]. In both cases, the object of the exercise was model reduction.

Consider a structure instrumented with p sensors, say measuring displacement. At each time instant t, the instrumentation returns a vector of measurements $\{y(t)\} = (y(t)_1, \ldots, y(t)_p)$. Because of the dynamical interactions between the coordinates there will be some correlation and hence redundancy; using PCA this redundancy can potentially be eliminated leaving a lower dimensional vector of 'pseudo-sensor' measurements which are linear

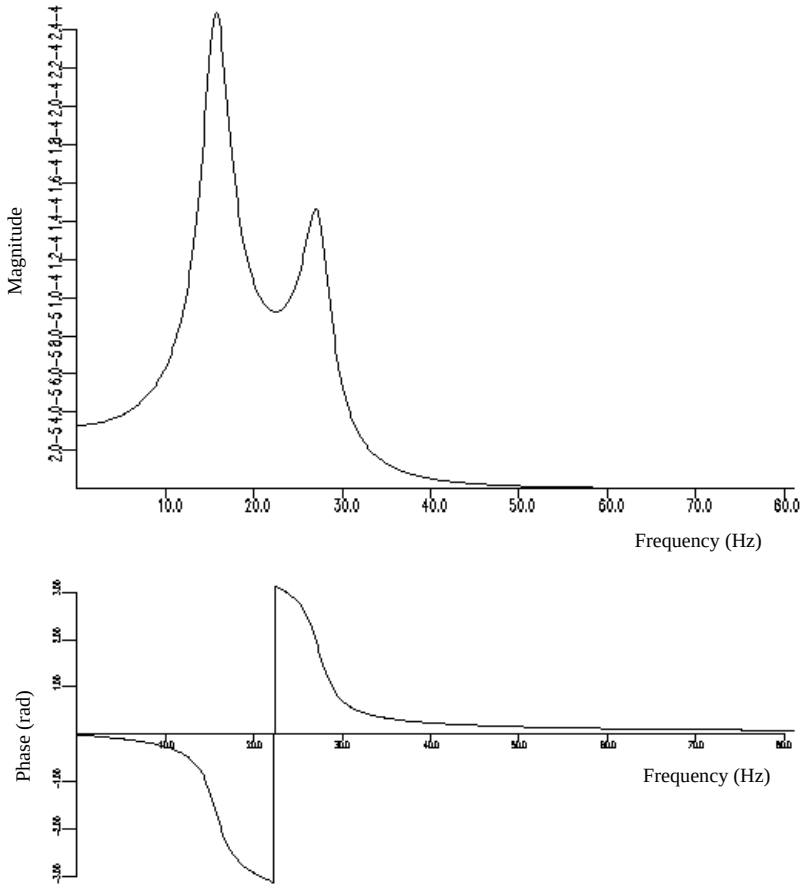


Figure 4.43. FRF H_2 for symmetric 2DOF linear system.

combinations of the original, yet still encode all the dynamics. This was the idea of Leuridan.

In terms of sampled data, there would be N samples of $\{y(t)\}$ taken at regular intervals Δt. These will be denoted $\{y(t_i)\}, i = 1, \ldots, N$. The signals observed from structures are usually zero-mean, so the covariance matrix for the system is

$$[\Sigma] = \sum_{i=1}^{N} \{y(t_i)\}\{y(t_i)\}^{\mathrm{T}}. \tag{4.153}$$

It is not particularly illuminating to look at the principal time signals. Visualization is much simpler in the frequency domain. The passage from time to frequency is accomplished using the multi-dimensional version of Parseval's

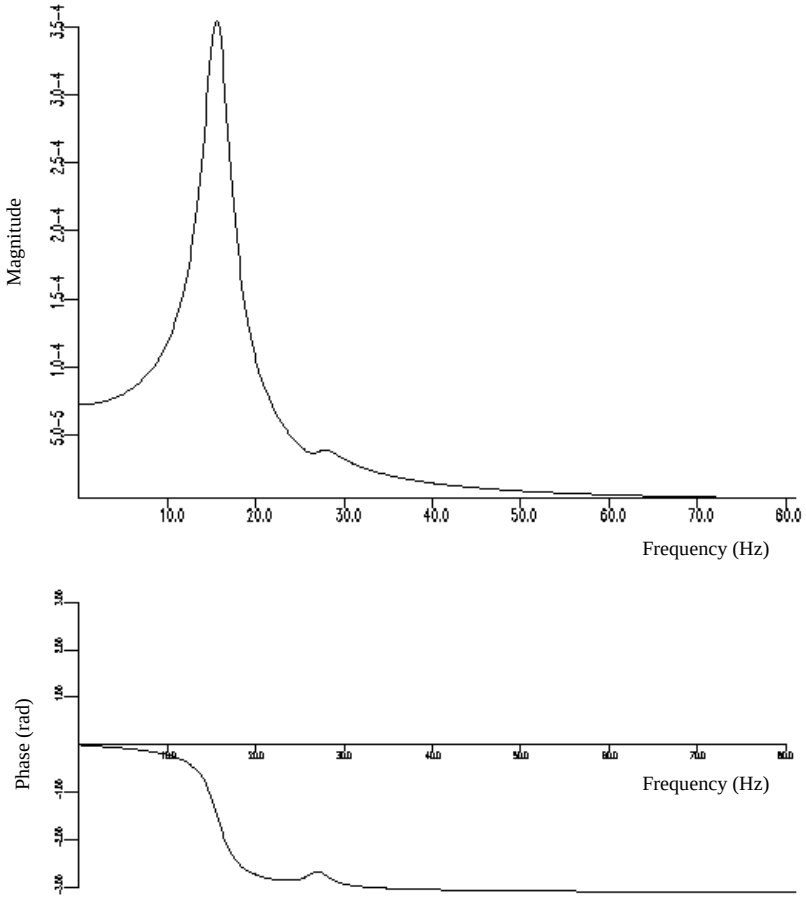


Figure 4.44. Principal FRF PH_1 for symmetric 2DOF linear system.

Theorem. For simplicity consider the continuous-time analogue of (4.153)

$$[\Sigma] = \int_{-\infty}^{\infty} \mathrm{d}t \, \{y(t)\}\{y(t)\}^{\mathrm{T}}. \tag{4.154}$$

Taking Fourier transforms gives

$$[\Sigma] = \int_{-\infty}^{\infty} \mathrm{d}t \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{d}\omega_1 \, \mathrm{e}^{\mathrm{i}\omega_1 t} \{Y(\omega_1)\} \right) \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{d}\omega_2 \, \mathrm{e}^{-\mathrm{i}\omega_2 t} \{\overline{Y}(\omega_2)\}^{\mathrm{T}} \right) \tag{4.155}$$

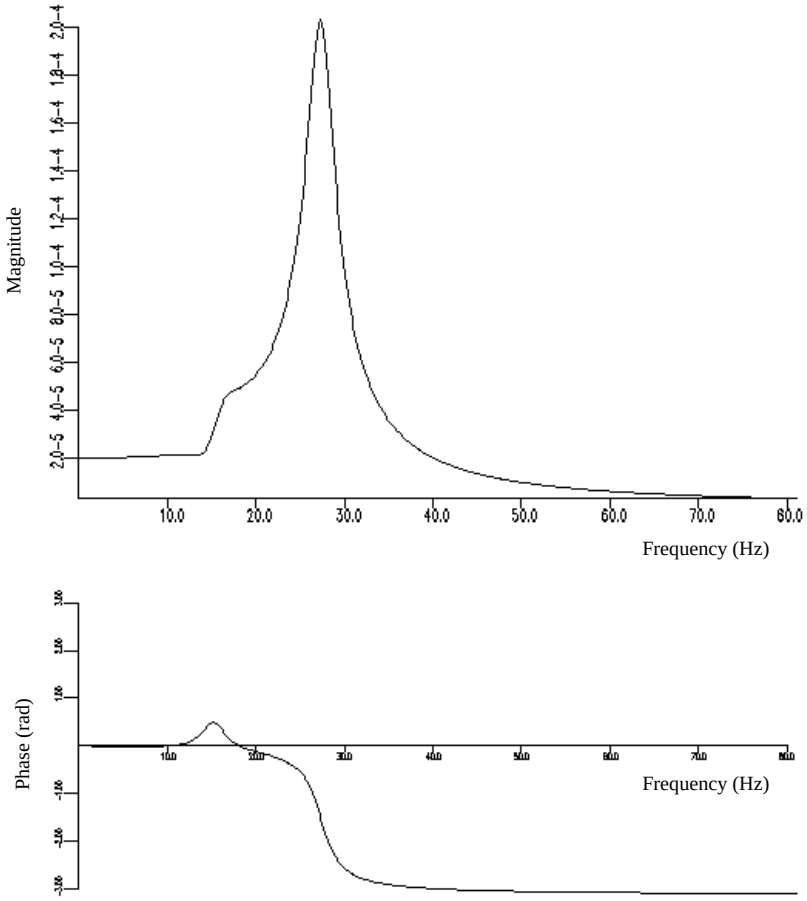


Figure 4.45. Principal FRF PH_2 for symmetric 2DOF linear system.

where the reality of the time signals has been used. Rearranging yields

$$[\Sigma] = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{d}\omega_1\,\mathrm{d}\omega_2\,\{Y(\omega_1)\}\{\overline{Y}(\omega_2)\}^{\mathrm{T}}\left(\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{d}t\mathrm{e}^{\mathrm{i}(\omega_1-\omega_2)t}\right). \tag{4.156}$$

Now, using the integral representation of the δ-function from appendix D, one finds

$$[\Sigma] = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{d}\omega_1\,\mathrm{d}\omega_2\,\{Y(\omega_1)\}\{\overline{Y}(\omega_2)\}^{\mathrm{T}}\delta(\omega_1-\omega_2) \tag{4.157}$$

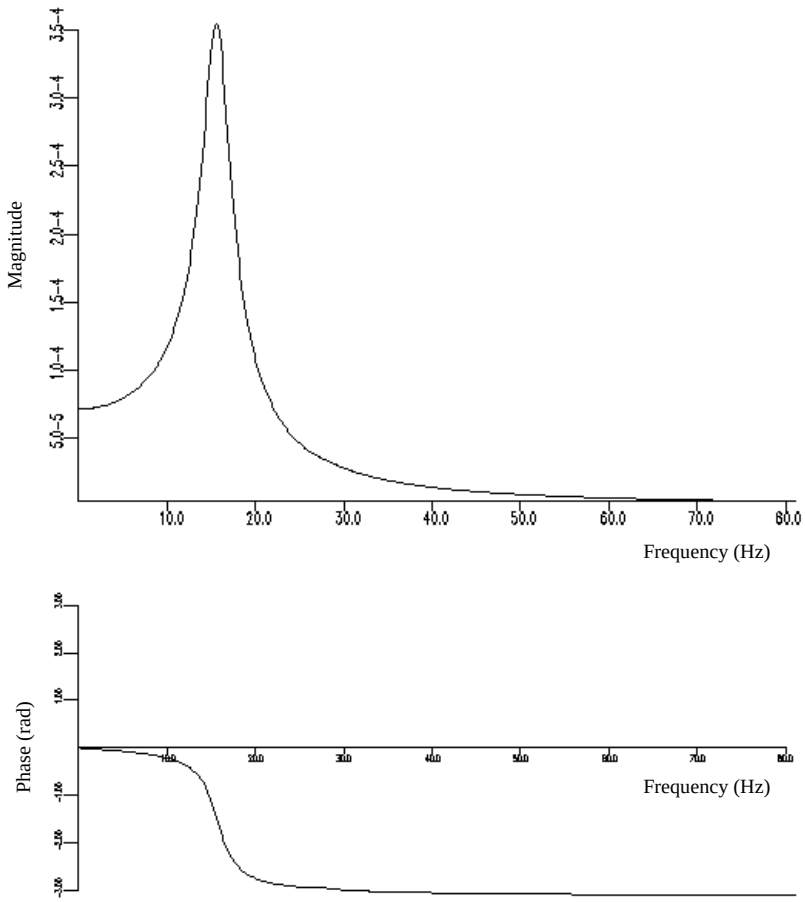


Figure 4.46. Corrected principal FRF PH_1 for symmetric 2DOF linear system.

and the projection property of $\delta(\omega)$ (again—appendix D) gives the final result

$$[\Sigma] = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{d}\omega_1 \, \{Y(\omega_1)\}\{\overline{Y}(\omega_1)\}^{\mathrm{T}} \tag{4.158}$$

and the transformation which decorrelates the time signals also decorrelates the spectra. (In (4.158) the overline refers to the complex conjugate and not the mean. In order to avoid confusion with complex quantities, the mean will be expressed in the rest of this section using the expectation operator, i.e. $\overline{x} = E[x]$.)

Now suppose the system is excited at a single point with a white excitation so

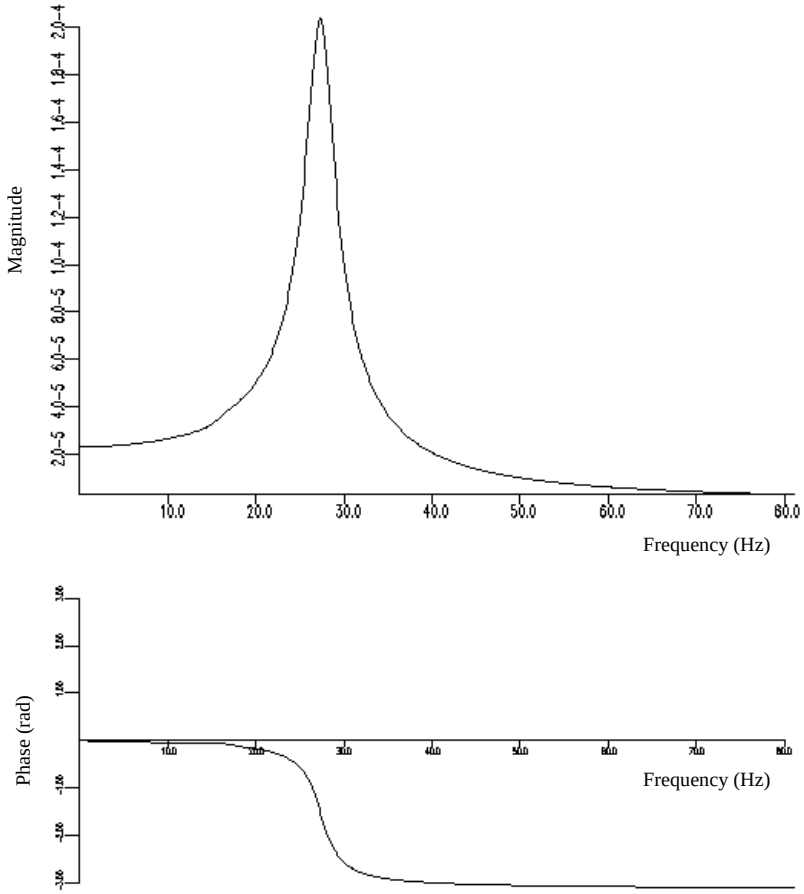


Figure 4.47. Corrected principal FRF PH_2 for symmetric 2DOF linear system.

that $X(\omega) = P$. This defines a vector of FRFs $\{H(\omega)\} = \{Y(\omega)\}/P$. Because

$$[\Sigma] = P^2 \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{d}\omega \, \{H(\omega)\}\{\overline{H}(\omega)\}^{\mathrm{T}} \tag{4.159}$$

the same principal component transformation as before also decorrelates the FRFs. (A similar result occurs for systems excited by sinusoidal excitation.) This offers the possibility of defining *principal FRFs*.

At this point it is useful to look at a concrete example. Consider the 2DOF linear system,

$$m\ddot{y}_1 + c\dot{y}_1 + 2ky_1 - ky_2 = X\sin(\omega t) \tag{4.160}$$

$$m\ddot{y}_2 + c\dot{y}_2 + 2ky_2 - ky_1 = 0. \tag{4.161}$$

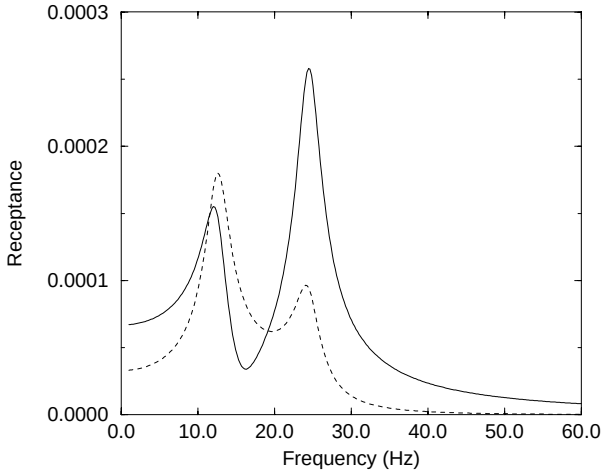


Figure 4.48. Principal FRFs for asymmetric 2DOF linear system.

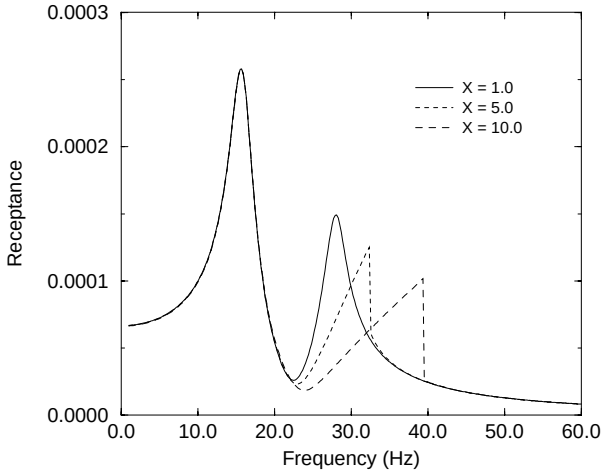


Figure 4.49. FRF Λ_1 for symmetric 2DOF nonlinear system at low medium and high excitation.

This defines a vector of FRFs $(H_1(\omega), H_2(\omega)) = (Y_1(\omega)/X, Y_2(\omega)/X)$. The FRFs H_1 and H_2 are shown in figures 4.42 and 4.43.

If the principal FRFs $PH_1(\omega)$ and $PH_2(\omega)$ are computed by the PCA procedure of (4.150)–(4.152) using the discrete version of (4.159)

$$[\Sigma] = \sum_{i=1}^{N/2} \{H(\omega_i)\}\{\overline{H}(\omega_i)\}^{\mathrm{T}} \tag{4.162}$$

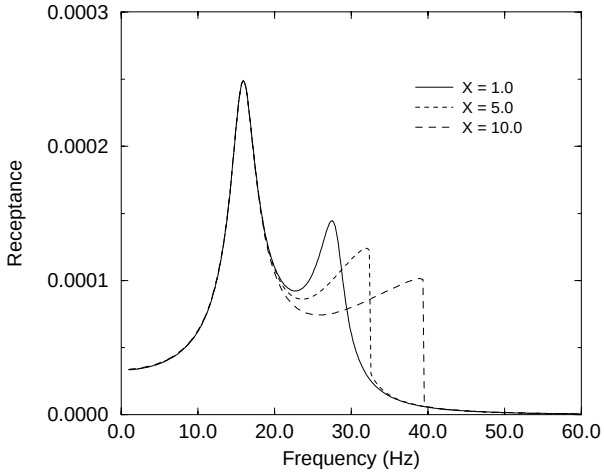


Figure 4.50. FRF Λ_2 for symmetric 2DOF nonlinear system at low medium and high excitation.

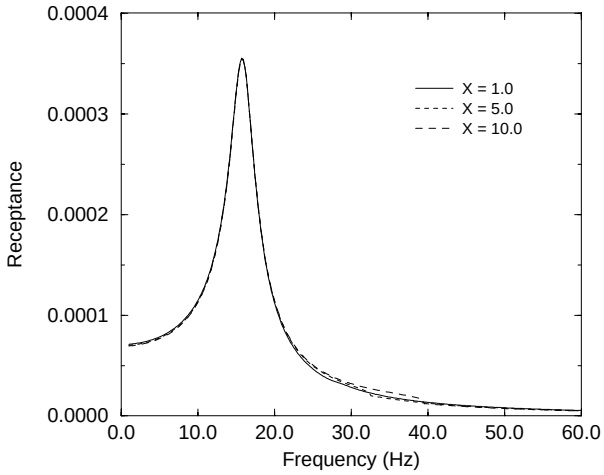


Figure 4.51. Principal FRF $P\Lambda_1$ for symmetric 2DOF nonlinear system at low medium and high excitation.

the results are as shown in figures 4.44 and 4.45. The decomposition appears to have almost produced a transformation to modal coordinates, both FRFs are only mildly distorted versions of SDOF FRFs. In fact in this case, the distortions are simple to explain.

The previous argument showed that the principal component transformation for time data also decorrelated the FRF vector. However, this proof used integrals

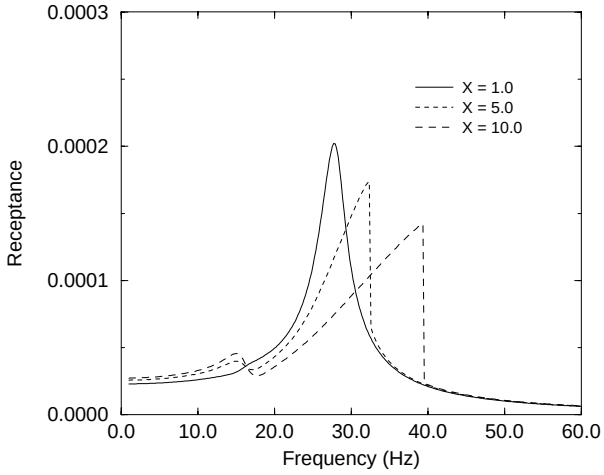


Figure 4.52. Principal FRF $P\Lambda_2$ for symmetric 2DOF nonlinear system at low medium and high excitation.

with infinite ranges. In practice, the covariance matrices are computed using finite summations. In the time-domain case, this presents no serious problems in applying (4.153) as long as the records are long enough that the means of the signals approximate to zero. However, in the frequency domain, the FRFs are *not* zero-mean due to the finite frequency range. This means that the covariance matrix in (4.162) is inappropriate to decorrelate the FRF vector. The remedy is simply to return to equation (4.150) and use the covariance matrix

$$[\Sigma] = \sum_{i=1}^{N/2} (\{H(\omega_i)\} - E[\{H(\omega_i)\}])(\{H(\omega_i)\} - E[\{H(\omega_i)\}])^{\mathrm{T}}. \qquad (4.163)$$

Using this prescription gives the principal FRFs shown in figures 4.46 and 4.47. This time the principal component transformation has produced modal FRFs. Unfortunately, this situation is not generic. It is the result here of considering a system with a high degree of symmetry; also the mass matrix is unity and this appears to be critical. Figure 4.48 shows the principal FRFs for a system identical to (4.160) and (4.161) except that the two equations have different mass values—the decoupling property has been lost even though the modal transformation can still achieve this. However, throughout the development of the PCA method it was hoped that the principal FRFs would generally exhibit *some* simplification.

In terms of nonlinear systems, the aim of PCA (or as it is sometimes called—the Karhunen–Loeve expansion [257]) is to hopefully localize the nonlinearity in a subset of the responses. By way of illustration consider the system in (4.160) and (4.161) supplemented by a cubic stiffness nonlinearity connecting the two

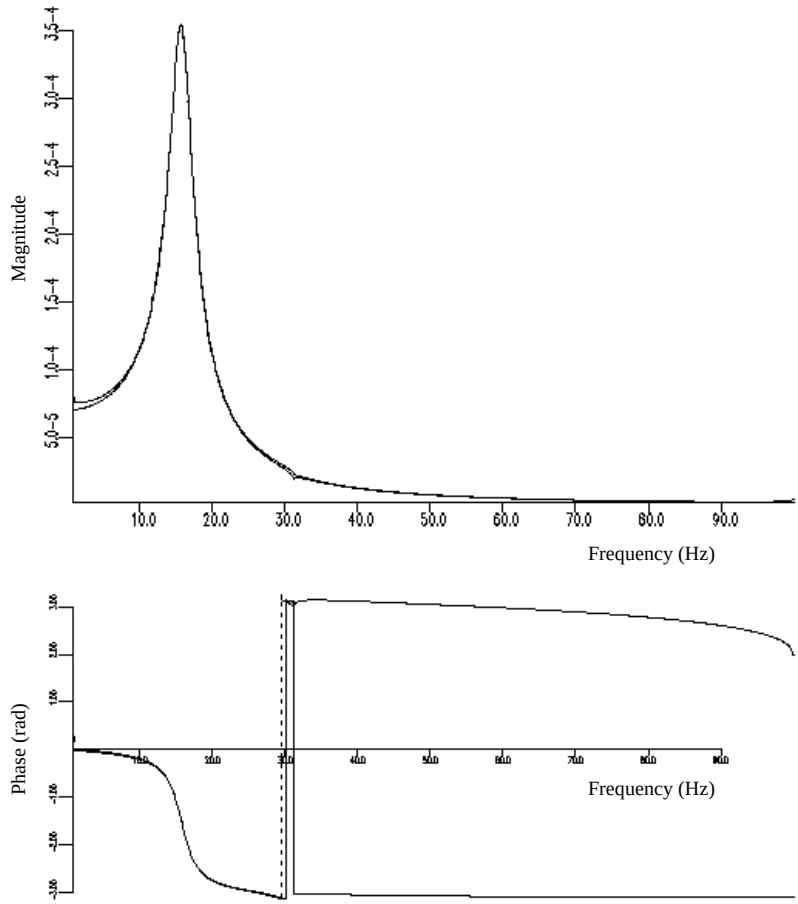


Figure 4.53. Principal FRF $P\Lambda_1$ for symmetric 2DOF nonlinear system with Hilbert transform.

masses

$$m\ddot{y}_1 + c\dot{y}_1 + 2ky_1 - ky_2 + k_3(y_1 - y_2)^3 = X\sin(\omega t) \tag{4.164}$$

$$m\ddot{y}_2 + c\dot{y}_2 + 2ky_2 - ky_1 + k_3(y_2 - y_1)^3 = 0. \tag{4.165}$$

The FRFs for the system at a number of different levels of excitation are given in figures 4.49 and 4.50. The distortion is only shown on the second mode as this is the only nonlinear mode (as discussed in section 3.1). When the principal FRFs are computed (figures 4.51 and 4.52), only the second principal FRF shows the distortion characteristic of nonlinearity. Again one should not overemphasize these results due to the high symmetry of the system.

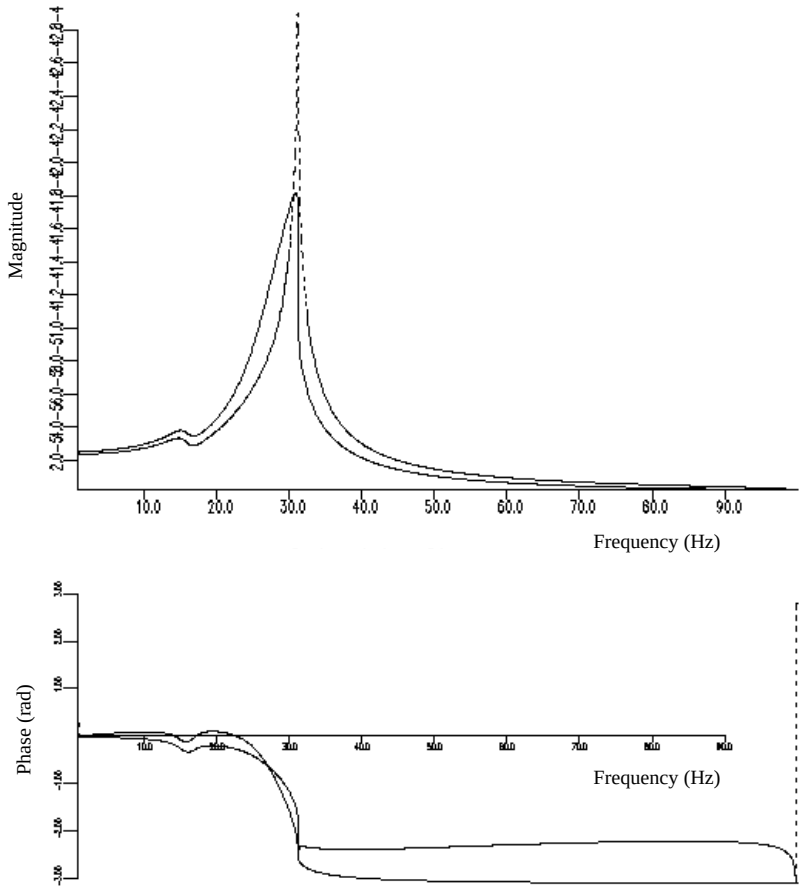


Figure 4.54. Principal FRF $P\Lambda_2$ for symmetric 2DOF nonlinear system with Hilbert transform.

The reason for the presence of this section in this chapter is that *any* test for nonlinearity can be applied to the principal FRFs including of course the Hilbert transform. This has been studied in the past by Ahmed [7] amongst others. Figures 4.53 and 4.54 show the result of applying the Hilbert transform to the principal FRFs for the system discussed earlier. As one might expect, the nonlinearity is only flagged for the second mode.

With that brief return to the Hilbert transform the chapter is concluded. The Hilbert transform has been seen to be a robust and sensitive indicator of nonlinearity. It is a little surprising that it has not yet been adopted by suppliers of commercial FRF analysers. The next chapter continues the Hilbert transform theme by considering an approach to the analysis which uses complex function theory.

Chapter 5

The Hilbert transform—a complex analytical approach

5.1 Introduction

The previous chapter derived the Hilbert transform and showed how it could be used in a number of problems in engineering dynamics and in particular how it could be used to detect and identify nonlinearity. It was clear from the analysis that there is a relationship between causality of the impulse response function and the occurrence of Hilbert transform pairs in the FRF. In fact, this relationship is quite deep and can only be fully explored using the theory of complex functions. Because of this, the mathematical background needed for this chapter is more extensive than for any other in the book with the exception of chapter 8. However, the effort is worthwhile as many useful new results become available. There are many textbooks on complex analysis which provide the prerequisites for this chapter: [6] is a classic text which provides a rigorous approach, while [234] provides a more relaxed introduction. Many texts on engineering mathematics cover the relevant material; [153] is a standard.

5.2 Hilbert transforms from complex analysis

The starting point for this approach is Cauchy's theorem [234], which states:

> given a function $G : \mathbb{C} \to \mathbb{C}$ (where $\mathbb{C}$ denotes the complex plane) and a simple closed contour C such that G is analytic[1] on and inside C, then
>
> $$\frac{1}{2\pi \mathrm{i}} \int_C \mathrm{d}\Omega \, \frac{G(\Omega)}{\Omega - \omega} = 0 \tag{5.1}$$
>
> if and only if ω lies outside C.

[1] Not analytic in the *signal* sense, meaning that the function G has no poles, i.e. singularities.

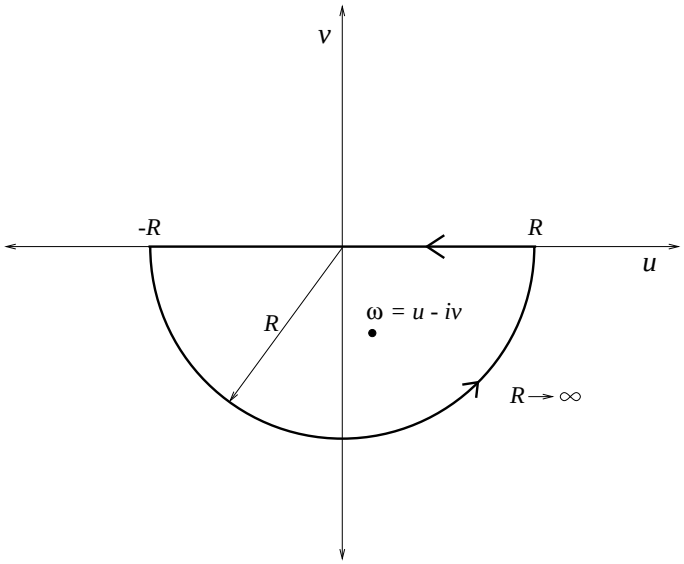


Figure 5.1. Main contour for deriving the Hilbert transform relation.

The derivation requires that the value of the integral be established (1) when ω is inside C and (2) when ω is on C:

(1) ω *inside* C. In this case one can use Cauchy's calculus of residues [234] to find the value of the integral, i.e.

$$\frac{1}{2\pi \mathrm{i}} \int_C \mathrm{d}\Omega \, \frac{G(\Omega)}{\Omega - \omega} = \sum_{\text{Poles}} \text{Res} \left[\frac{G(\Omega)}{\Omega - \omega} \right] \tag{5.2}$$

and, in this case, there is a single simple pole at $\Omega = \omega$, so the residue is given by

$$\lim_{\Omega \to \omega} (\Omega - \omega) \times \frac{G(\Omega)}{\Omega - \omega}. \tag{5.3}$$

So

$$\frac{1}{2\pi \mathrm{i}} \int_C \mathrm{d}\Omega \, \frac{G(\Omega)}{\Omega - \omega} = G(\omega). \tag{5.4}$$

(2) ω *on* C. In all the cases of interest for studying the Hilbert transform, only one type of contour is needed; so, for the sake of simplicity, the results that follow are established using that contour. The argument follows closely that of [193]. Consider the contour in figure 5.1. Initially $\omega = u - \mathrm{i}v$ is below the real axis and the residue theorem gives

$$G(\omega) = G(u - \mathrm{i}v) = \frac{1}{2\pi \mathrm{i}} \int_{-R}^{\mathrm{R}} \mathrm{d}\Omega \, \frac{G(\Omega)}{\Omega - u + \mathrm{i}v} + I_C \tag{5.5}$$

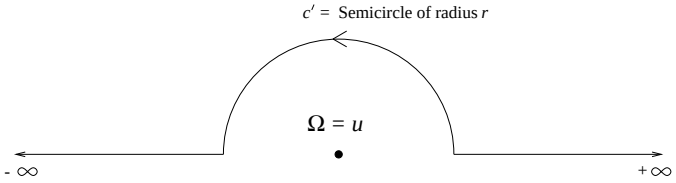


Figure 5.2. Contour deformation used to avoid the pole on the real axis.

where I_C is the semi-circular part of the contour. If now $R \to \infty$ under the additional assumption that $G(\Omega)/(\Omega - \omega)$ tends to zero as $\Omega \to \infty$ fast enough to make I_C vanish[2], the result is

$$G(\omega) = G(u - \mathrm{i}v) = \frac{1}{2\pi\mathrm{i}} \int_{-\infty}^{\infty} \mathrm{d}\Omega \, \frac{G(\Omega)}{\Omega - u + \mathrm{i}v}. \tag{5.6}$$

In order to restrict the integrand in (5.5) to real values, one must have $v \to 0$ or $\omega \to u$. However, in order to use the results previously established, ω should lie off the contour—in this case the real axis. The solution to this problem is to deform the contour by adding the section C' as shown in figure 5.2. C' is essentially removed by allowing its radius r to tend to zero after ω has moved onto the real axis. Equation (5.5) becomes (on taking the integration anticlockwise around the contour)

$$2\pi\mathrm{i}G(\omega) = 2\pi\mathrm{i} \lim_{v \to 0} G(u - \mathrm{i}v) \tag{5.7}$$

$$= \lim_{\mathrm{r} \to 0} \left\{ \lim_{v \to 0} \left\{ \int_{\infty}^{\omega + r} \mathrm{d}\Omega \frac{G(\Omega)}{\Omega - u + \mathrm{i}v} + \int_{\omega - r}^{-\infty} \mathrm{d}\Omega \frac{G(\Omega)}{\Omega - u + \mathrm{i}v} + \int_{C'} \mathrm{d}\Omega \frac{G(\Omega)}{\Omega - u + \mathrm{i}v} \right\} \right\}. \tag{5.8}$$

Taking the first limit and changing to polar coordinates on the small semi-circle yields

$$2\pi\mathrm{i}G(\omega) = \lim_{\mathrm{r} \to 0} \left\{ \int_{\infty}^{\omega + r} \mathrm{d}\Omega \frac{G(\Omega)}{\Omega - \omega} + \int_{\omega - r}^{-\infty} \mathrm{d}\Omega \frac{G(\Omega)}{\Omega - \omega} + \int_{\theta = 0}^{\theta = \pi} \frac{\mathrm{d}(\omega + r\mathrm{e}^{\mathrm{i}\theta})}{r\mathrm{e}^{\mathrm{i}\theta}} G(\omega + r\mathrm{e}^{\mathrm{i}\theta}) \right\} \tag{5.9}$$

$$= -PV \int_{-\infty}^{\infty} \mathrm{d}\Omega \, \frac{G(\Omega)}{\Omega - \omega} + \mathrm{i}\pi G(\omega) \tag{5.10}$$

[2] For example, suppose that $G(\Omega)$ is $\mathrm{O}(R^{-1})$ as $R \to \infty$, then the integrand is $\mathrm{O}(R^{-2})$ and the integral I_C is $\pi R \times \mathrm{O}(R^{-2}) = \mathrm{O}(R^{-1})$ and therefore tends to zero as $R \to \infty$. This is by no means a rigorous argument, consult [234] or any introductory book on complex analysis.

where PV denotes the *Cauchy principal value* defined by

$$PV\int_{-\infty}^{\infty} \mathrm{d}\Omega\, G(\Omega) = \lim_{\mathrm{r}\to 0}\left\{\int_{-\infty}^{\omega-r} \mathrm{d}\Omega\, G(\Omega) + \int_{\omega+r}^{\infty} \mathrm{d}\Omega\, G(\Omega)\right\} \tag{5.11}$$

in the case that $G(\Omega)$ has a pole at $\Omega = \omega$.
The final result of this analysis is

$$\pi i G(\omega) = -PV\int_{-\infty}^{\infty} \mathrm{d}\Omega\, \frac{G(\Omega)}{\Omega - \omega}, \quad \omega, \Omega \in \mathbb{R}. \tag{5.12}$$

In pure mathematics, as discussed in the previous chapter, the Hilbert Transform $\mathcal{H}[F]$ of a function G is defined by

$$\mathcal{H}[F](\omega) = PV\frac{1}{\pi}\int_{-\infty}^{\infty} \mathrm{d}\Omega\, \frac{G(\Omega)}{\Omega - \omega} \tag{5.13}$$

so equation (5.12) can be written in the more compact form

$$G(\omega) = -\mathrm{i}\mathcal{H}\{G(\omega)\}. \tag{5.14}$$

Equation (5.14) is the desired result. It is important to bear in mind the assumptions made in its derivation, namely

(1) G is analytic in the area bounded by the contour C. In the limit above as $R \to \infty$, this is the lower complex half-plane.
(2) $G(\omega)$ tends to zero fast enough as $R \to \infty$ for the integral I_C to vanish.

It is convenient (and also follows the conventions introduced somewhat arbitrarily in the last chapter) to absorb the factor $-\mathrm{i}$ into the definition of the Hilbert transform. In which case equation (5.14) becomes

$$G(\omega) = \mathcal{H}\{G(\omega)\} \tag{5.15}$$

as in equation (4.20). This is a fascinating result—the same condition is obtained on the class of functions analytic in the lower half-plane as was derived for transfer functions whose impulse responses are causal. This is not a coincidence; the reasons for this correspondence will be given in the next section.

5.3 Titchmarsh's theorem

The arguments of the previous section are expressed rigorously by Titchmarsh's theorem which is stated here in its most abstract form (taken from [118]).

Theorem. *If $G(\omega)$ is the Fourier transform of a function which vanishes for $t < 0$ and*

$$\int_{-\infty}^{\infty} |G(\omega)|^2 < \infty \tag{5.16}$$

then $G(\omega)$ is the boundary value of a function $G(\omega - \mathrm{i}\gamma)$, $\gamma > 0$, which is analytic in the lower half-plane. Further

$$\int_{-\infty}^{\infty} |G(\omega - \mathrm{i}\gamma)|^2 < \infty. \tag{5.17}$$

The previous section showed that conditions—(i) analycity in the lower half-plane and (ii) fast fall-off of $G(\omega)$—are necessary for the Hilbert transform relations to hold. Titchmarsh's theorem states that they are sufficient and that $G(\omega)$ need only tend to zero as $\omega \to \infty$ fast enough to ensure the existence of $\int \mathrm{d}\omega\, |G(\omega)|^2$.

The conditions on the integrals simply ensure that the functions considered are Lesbesgue square-integrable. Square-integrability is, in any case, a necessary condition for the existence of Fourier transforms. If it is assumed that all relevant transforms and inverses exist, then the theorem can be extended and stated in a simpler, more informative form:

Theorem. *If one of (i), (ii) or (iii) is true, then so are the other two.*

(i) $G(\omega)$ *satisfies the Hilbert transform relation (5.15).*
(ii) $G(\omega)$ *has a causal inverse Fourier transform, i.e. if* $t < 0$, $g(t) = \mathcal{F}^{-1}\{G(\omega)\} = 0$.
(iii) $G(\omega)$ *is analytic in the lower half-plane.*

The simple arguments of the previous section showed that (i) $\Longleftrightarrow$ (iii). A fairly simple demonstration that (i) $\Longleftrightarrow$ (ii) follows, and this establishes the theorem.

(i) $\Longrightarrow$ *(ii).* Assume that[3]

$$G(\omega) = -\frac{1}{\pi \mathrm{i}} \int_{-\infty}^{\infty} \mathrm{d}\Omega\, \frac{G(\Omega)}{\Omega - \omega}. \tag{5.18}$$

Then as

$$g(t) = \mathcal{F}^{-1}\{G(\omega)\} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{d}\omega\, \mathrm{e}^{\mathrm{i}\omega t} G(\omega) \tag{5.19}$$

it follows that

$$g(t) = -\frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{d}\omega\, \mathrm{e}^{\mathrm{i}\omega t} \frac{1}{\pi \mathrm{i}} \int_{-\infty}^{\infty} \mathrm{d}\Omega \frac{G(\Omega)}{\Omega - \omega}. \tag{5.20}$$

Assuming that it is valid to interchange the order of integration, this becomes

$$g(t) = +\frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{d}\Omega\, G(\Omega)\, \frac{1}{\pi \mathrm{i}} \int_{-\infty}^{\infty} \mathrm{d}\omega\, \frac{\mathrm{e}^{\mathrm{i}\omega t}}{\omega - \Omega}. \tag{5.21}$$

[3] In most cases, the principal value restriction can be understood from the context, in which case the letters PV will be omitted

It is shown in appendix D that

$$\frac{1}{\pi \mathrm{i}} \int_{-\infty}^{\infty} \mathrm{d}\omega \, \frac{\mathrm{e}^{\mathrm{i}\omega t}}{\omega - \Omega} = \mathrm{e}^{\mathrm{i}\Omega t} \epsilon(t) \tag{5.22}$$

where $\epsilon(t)$ is the sign function, $\epsilon(t) = 1$ if $t > 0$, $\epsilon(t) = -1$ if $t < 0$. This implies that

$$g(t) = +\frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{d}\Omega \, G(\Omega) \mathrm{e}^{\mathrm{i}\Omega t} = g(t), \quad \text{if } t > 0 \tag{5.23}$$

and

$$g(t) = -\frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{d}\Omega \, G(\Omega) \mathrm{e}^{\mathrm{i}\Omega t} = -g(t), \quad \text{if } t < 0. \tag{5.24}$$

The first of these equations says nothing; however, the second can only be true if $g(t) = 0$ for all $t < 0$, and this is the desired result.

(i) $\Longrightarrow$ *(ii).* Suppose that $g(t) = \mathcal{F}^{-1}\{G(\omega)\} = 0$ if $t < 0$. It follows trivially that

$$g(t) = g(t)\epsilon(t). \tag{5.25}$$

Fourier transforming this expression gives the convolution

$$G(\omega) = -\frac{1}{\pi \mathrm{i}} \int_{-\infty}^{\infty} \mathrm{d}\Omega \, \frac{G(\Omega)}{\Omega - \omega} \tag{5.26}$$

which is the desired result.

This discussion establishes the connection between causality and the Hilbert transform relation (5.15). It is important to point out that the theorems hold *only* if the technicalities of Titchmarsh's theorem are satisfied. The next section shows how the Hilbert transform relations are applied to functions which do not satisfy the necessary conditions.

5.4 Correcting for bad asymptotic behaviour

The crucial point in Titchmarsh's theorem is that $G(\omega)$ should be square-integrable, i.e. $\int \mathrm{d}\omega \, |G(\omega)|^2 < \infty$. It happens that in some cases of interest this condition is not satisfied; however, there is a way of circumnavigating this problem.

Arguably the least troublesome function which is not square-integrable is one which tends to a constant value at infinity, i.e. $G(\omega) \longrightarrow G_\infty$ as $\omega \longrightarrow \infty$. A sufficiently general function for the purposes of this discussion is a rational function

$$G(\omega) = \frac{A(\omega)}{B(\omega)} = \frac{a_0 + a_1\omega + \cdots + a_n\omega^n}{b_0 + b_1\omega + \cdots + b_n\omega^n} \tag{5.27}$$

where $A(\omega)$ and $B(\omega)$ are polynomials of the same order n and all the zeroes of $B(\omega)$ are in the upper half-plane. Clearly

$$\lim_{\omega \longrightarrow \infty} G(\omega) = G_\infty = \frac{a_n}{b_n}. \tag{5.28}$$

Carrying out a long division on (5.27) yields

$$G(\omega) = \frac{a_n}{b_n} + \frac{A'(\omega)}{B(\omega)} \tag{5.29}$$

where A' is a polynomial of order $n-1$. In other words,

$$G(\omega) - G_\infty = G(\omega) - \frac{a_n}{b_n} = \frac{A'(\omega)}{B(\omega)} \tag{5.30}$$

and $A'(\omega)/B(\omega)$ is $\mathrm{O}(\omega^{-1})$ as $\omega \longrightarrow \infty$. This means that $A'(\omega)/B(\omega)$ is square-integrable and therefore satisfies the conditions of Titchmarsh's theorem. Hence,

$$\frac{A'(\omega)}{B(\omega)} = -\frac{1}{\pi \mathrm{i}} \int_{-\infty}^{\infty} \mathrm{d}\Omega \, \frac{A'(\Omega)}{B(\Omega)} \frac{1}{\Omega - \omega} \tag{5.31}$$

or

$$G(\omega) - G_\infty = -\frac{1}{\pi \mathrm{i}} \int_{-\infty}^{\infty} \mathrm{d}\Omega \, \frac{G(\Omega) - G_\infty}{\Omega - \omega}. \tag{5.32}$$

So if a function fails to satisfy the conditions required by Titchmarsh's theorem because of asymptotically constant behaviour, subtracting the limiting value produces a valid function. The relations between real and imaginary parts (4.17) and (4.18) are modified as follows:

$$\mathrm{Re}\, G(\omega) - \mathrm{Re}\, G_\infty = -\frac{1}{\pi} \int_{-\infty}^{\infty} \mathrm{d}\Omega \, \frac{\mathrm{Im}\, G(\Omega) - \mathrm{Im}\, G_\infty}{\Omega - \omega} \tag{5.33}$$

$$\mathrm{Im}\, G(\omega) - \mathrm{Im}\, G_\infty = +\frac{1}{\pi} \int_{-\infty}^{\infty} \mathrm{d}\Omega \, \frac{\mathrm{Re}\, G(\Omega) - \mathrm{Re}\, G_\infty}{\Omega - \omega}. \tag{5.34}$$

These equations are well known in physical optics and elementary particle physics. The first of the pair produces the Kramers–Kronig dispersion relation if $G(\omega)$ is taken as $n(\omega)$—the complex refractive index of a material. The term 'dispersion' refers to the variation of the said refractive index with frequency of incident radiation [77].

One possible obstruction to the direct application of equations (5.32)–(5.34) is that $G(\omega)$ is usually an experimentally measured quantity. It is clear that G_∞ will not usually be available. However, this problem can be solved by using a *subtraction* scheme as follows. Suppose for the sake of simplicity that the limiting value of $G(\omega)$ as $\omega \longrightarrow \infty$ is real and that a measurement of G is available at $\omega = \alpha < \infty$. Equation (5.33) yields

$$\mathrm{Re}\, G(\omega) - \mathrm{Re}\, G_\infty = -\frac{1}{\pi} \int_{-\infty}^{\infty} \mathrm{d}\Omega \, \frac{\mathrm{Im}\, G(\Omega)}{\Omega - \omega} \tag{5.35}$$

and at $\omega = \alpha$ this becomes

$$\mathrm{Re}\, G(\alpha) - \mathrm{Re}\, G_\infty = -\frac{1}{\pi} \int_{-\infty}^{\infty} \mathrm{d}\Omega \, \frac{\mathrm{Im}\, G(\Omega)}{\Omega - \alpha} \tag{5.36}$$

and subtracting (5.36) from (5.35) yields

$$\mathrm{Re}\,G(\omega) - \mathrm{Re}\,G(\alpha) = -\frac{1}{\pi}\int_{-\infty}^{\infty} \mathrm{d}\Omega \left\{ \frac{1}{\Omega - \omega} - \frac{1}{\Omega - \alpha} \right\} \mathrm{Im}\,G(\Omega) \qquad (5.37)$$

or

$$\mathrm{Re}\,G(\omega) - \mathrm{Re}\,G(\alpha) = -\frac{(\omega - \alpha)}{\pi}\int_{-\infty}^{\infty} \mathrm{d}\Omega \, \frac{\mathrm{Im}\,G(\Omega)}{(\Omega - \omega)(\Omega - \alpha)}. \qquad (5.38)$$

Note that in compensating for lack of knowledge of G_∞, the analysis has produced a more complicated integral. In general if $G(\omega)$ behaves as some polynomial as $\omega \longrightarrow \infty$, a subtraction strategy will correct for the bad asymptotic behaviour in much the same way as before. Unfortunately, each subtraction complicates the integral further.

The application of these formulae will now be demonstrated in a number of case studies.

5.4.1 Simple examples

The first pair of case studies allow all the relevant calculations to be carried out by hand.

The first example calculation comes from [215]. The object of the paper was to demonstrate a nonlinear system which was nonetheless causal and therefore satisfied Hilbert transform relations. The system under study was a simple squaring device[4], i.e. $y(t) = x(t)^2$. The excitation was designed to give no response at negative times, i.e.

$$x(t) = \begin{cases} A\mathrm{e}^{-at}, & t > 0, a > 0 \\ 0, & t < 0. \end{cases} \qquad (5.39)$$

A type of FRF was defined by dividing the spectrum of the output by the spectrum of the input:

$$\Lambda(\omega) = \frac{Y(\omega)}{X(\omega)} = \frac{\mathcal{F}\{A^2\mathrm{e}^{-2at}\}}{\mathcal{F}\{A\mathrm{e}^{-at}\}} = \frac{A(\omega - \mathrm{i}a)}{(\omega - 2\mathrm{i}a)} \qquad (5.40)$$

so

$$\mathrm{Re}\,\Lambda(\omega) = \frac{A(\omega^2 + 2a^2)}{\omega^2 + 4a^2}, \quad \mathrm{Im}\,\Lambda(\omega) = \frac{Aa\omega}{\omega^2 + 4a^2}. \qquad (5.41)$$

[4] As a remark for the sophisticate or person who has read later chapters first, it does not really make sense to consider this system for this purpose as it does not possess a linear FRF. If the system is excited with a pure harmonic $\mathrm{e}^{\mathrm{i}\omega t}$ the response consists of a purely second order part $\mathrm{e}^{2\mathrm{i}\omega t}$; thus $H_2(\omega_1, \omega_2) = 1$ and $H_n = 0\ \forall n \neq 2$. As the system has no H_1, it has no impulse response h_1 and therefore discussions of causality do not apply.

Now, despite the fact that Λ is manifestly analytic in the lower half-plane, $\mathrm{Re}\,\Lambda(\omega)$ and $\mathrm{Im}\,\Lambda(\omega)$ do not form a Hilbert transform pair, i.e. they are not related by the equations (4.17) and (4.18). In fact, directly evaluating the integrals gives

$$\frac{1}{\pi}\int_{-\infty}^{\infty} \mathrm{d}\Omega\, \frac{\mathrm{Re}\,\Lambda(\Omega)}{\Omega-\omega} = \frac{Aa\omega}{\omega^2+4a^2} = \mathrm{Im}\,\Lambda(\omega) \tag{5.42}$$

as required, while

$$-\frac{1}{\pi}\int_{-\infty}^{\infty} \mathrm{d}\Omega\, \frac{\mathrm{Im}\,\Lambda(\Omega)}{\Omega-\omega} = \frac{-2Aa^2}{\omega^2+4a^2} \neq \mathrm{Re}\,\Lambda(\omega). \tag{5.43}$$

The reason for the breakdown is that

$$\lim_{\omega\longrightarrow\infty} \Lambda(\omega) = A \neq 0 \tag{5.44}$$

so $\Lambda(\omega)$ is not square-integrable and Titchmarsh's theorem does not hold. However, the modified dispersion relations (5.33) and (5.34) can be used with $\mathrm{Re}\,\Lambda(\infty) = A$ and $\mathrm{Im}\,\Lambda(\infty) = 0$. The appropriate relation is

$$\mathrm{Re}\,\Lambda(\omega) - A = -\frac{1}{\pi}\int_{-\infty}^{\infty} \mathrm{d}\Omega\, \frac{\mathrm{Im}\,\Lambda(\Omega)}{\Omega-\omega} \tag{5.45}$$

i.e.

$$\mathrm{Re}\,\Lambda(\omega) = A - \frac{2Aa^2}{\omega^2+4a^2} = \frac{A(\omega^2+2a^2)}{\omega^2+4a^2} \tag{5.46}$$

as required[5].

The problem also shows up in the time domain, taking the inverse Fourier transform of Λ

$$\mathcal{F}^{-1}\{\Lambda(\omega)\} = \lambda(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{d}\omega\, \mathrm{e}^{\mathrm{i}\omega t}\left\{\frac{A(\omega-\mathrm{i}a)}{(\omega-2\mathrm{i}a)}\right\} \tag{5.47}$$

yields

$$\begin{aligned}\lambda(t) &= \frac{A}{2\pi}\int_{-\infty}^{\infty} \mathrm{d}\omega\, \mathrm{e}^{\mathrm{i}\omega t}\left\{1+\frac{\mathrm{i}a}{(\omega-2\mathrm{i}a)}\right\} \\ &= \frac{A}{2\pi}\int_{-\infty}^{\infty} \mathrm{d}\omega\, \mathrm{e}^{\mathrm{i}\omega t} + \frac{\mathrm{i}aA}{2\pi}\int_{-\infty}^{\infty} \mathrm{d}\omega\, \frac{\mathrm{e}^{\mathrm{i}\omega t}}{(\omega-2\mathrm{i}a)}.\end{aligned} \tag{5.48}$$

Using the results of appendix D, the first integral gives a δ-function; the second integral is easily evaluated by contour integration. Finally,

$$\lambda(t) = A\delta(t) + aA\mathrm{e}^{-2at}\Theta(t) \tag{5.49}$$

where $\Theta(t)$ is the Heaviside function. This shows that the 'impulse response' λ contains a δ-function in addition to the expected causal part. Removal of

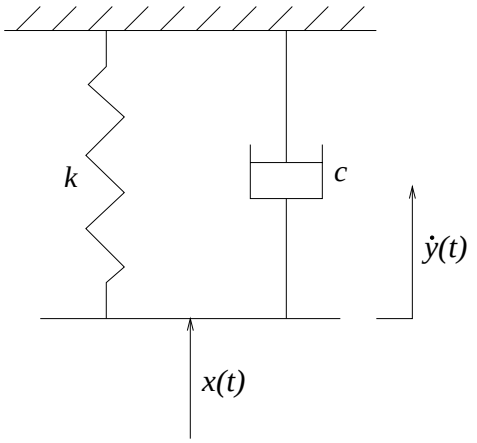


Figure 5.3. A first-order dynamical system.

the δ-function is the time-domain analogue of correcting for the bad asymptotic behaviour in the frequency domain.

Another example of this type of calculation occurs in [122]. The first-order linear system depicted in figure 5.3 is used to illustrate the theory. The system has the FRF

$$H(\omega) = \frac{\mathrm{i}\omega}{\mathrm{i}c\omega + k} = \frac{-c\omega^2}{c^2\omega^2 + k^2} - \mathrm{i}\frac{k\omega}{c^2\omega^2 + k^2}. \tag{5.50}$$

It is correctly stated that

$$\operatorname{Im} H(\omega) = \frac{1}{\pi}\int_{-\infty}^{\infty} \mathrm{d}\Omega\, \frac{\operatorname{Re} H(\omega)}{\Omega - \omega} = \frac{-k\omega}{c^2\omega^2 + k^2} \tag{5.51}$$

i.e. the relation in (4.18) applies. However, because

$$\lim_{\omega \longrightarrow \infty} H(\omega) = -\frac{1}{c} \neq 0 \tag{5.52}$$

the appropriate formula for calculating $\operatorname{Re} H(\omega)$ from $\operatorname{Im} H(\omega)$ is (5.33), i.e.

$$\operatorname{Re} H(\omega) + \frac{1}{c} = -\frac{1}{\pi}\int_{-\infty}^{\infty} \mathrm{d}\Omega\, \frac{\operatorname{Im} H(\omega)}{\Omega - \omega}. \tag{5.53}$$

5.4.2 An example of engineering interest

Consider the linear system

$$m\ddot{y} + c\dot{y} + ky = x(t). \tag{5.54}$$

[5] The integrals involve terms of the form $\int \mathrm{d}\Omega/(\Omega - \omega)$ which are proportional to $\log(-1)$. If the principal sheet of the $\log$ function is specified, these terms can be disregarded.

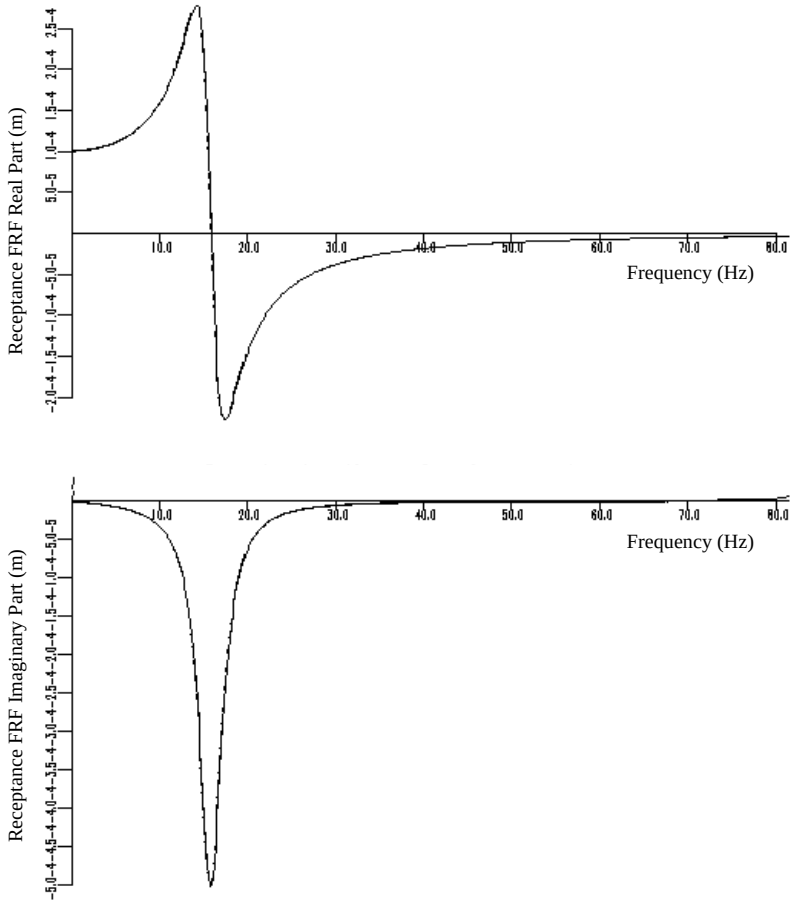


Figure 5.4. Real and imaginary parts of the receptance FRF and the corresponding Hilbert transform.

Depending on which sort of output data is measured, the system FRF can take essentially three forms. If force and displacement are measured, the receptance form is obtained as discussed in chapter 1:

$$H_{\mathrm{R}}(\omega) = \frac{\mathcal{F}\{y(t)\}}{\mathcal{F}\{x(t)\}} = \frac{1}{-m\omega^2 + \mathrm{i}c\omega + k} \tag{5.55}$$

and

$$\lim_{\omega \longrightarrow \infty} H_{\mathrm{R}}(\omega) = 0. \tag{5.56}$$

Measuring the output velocity yields the mobility form

$$H_{\mathrm{M}}(\omega) = \frac{\mathcal{F}\{\dot{y}(t)\}}{\mathcal{F}\{x(t)\}} = \frac{\mathrm{i}\omega}{-m\omega^2 + \mathrm{i}c\omega + k} \tag{5.57}$$

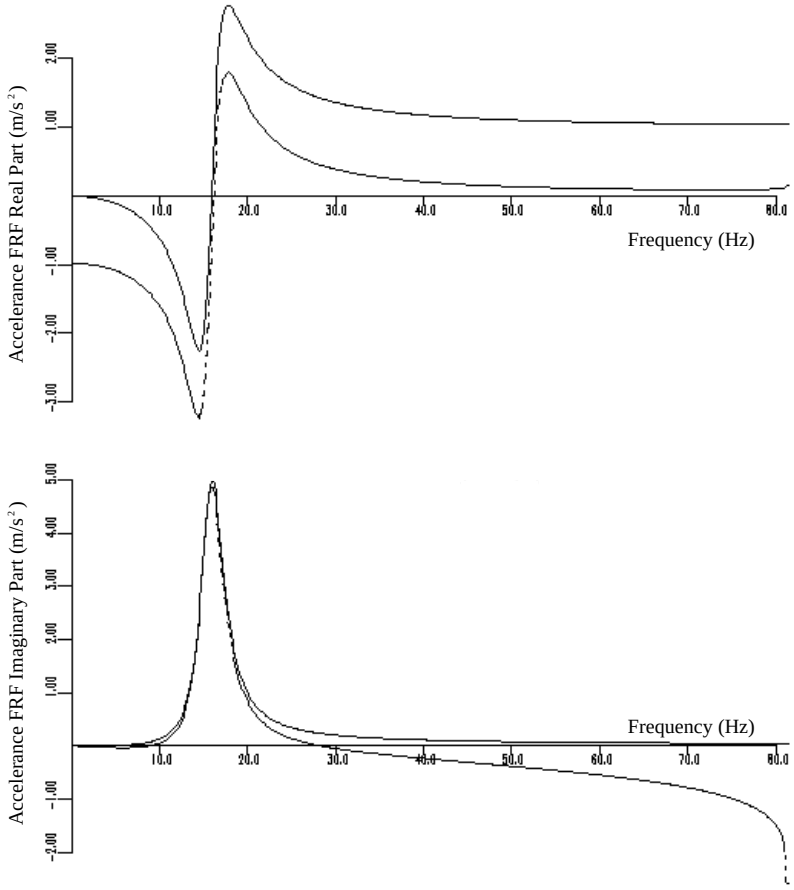


Figure 5.5. Real and imaginary parts of the accelerance FRF and the corresponding Hilbert transform.

and

$$\lim_{\omega \longrightarrow \infty} H_{\mathrm{M}}(\omega) = 0. \tag{5.58}$$

Finally, measuring the output acceleration gives the accelerance form

$$H_{\mathrm{A}}(\omega) = \frac{\mathcal{F}\{\ddot{y}(t)\}}{\mathcal{F}\{x(t)\}} = \frac{-\omega^2}{-m\omega^2 + \mathrm{i}c\omega + k} \tag{5.59}$$

and, in this case,

$$\lim_{\omega \longrightarrow \infty} H_{\mathrm{A}}(\omega) = \frac{1}{m} \neq 0. \tag{5.60}$$

This means that if the Hilbert transform is used to test for nonlinearity, the appropriate Hilbert transform pair is $(\mathrm{Re}\, H(\omega), \mathrm{Im}\, H(\omega))$ if the FRF is

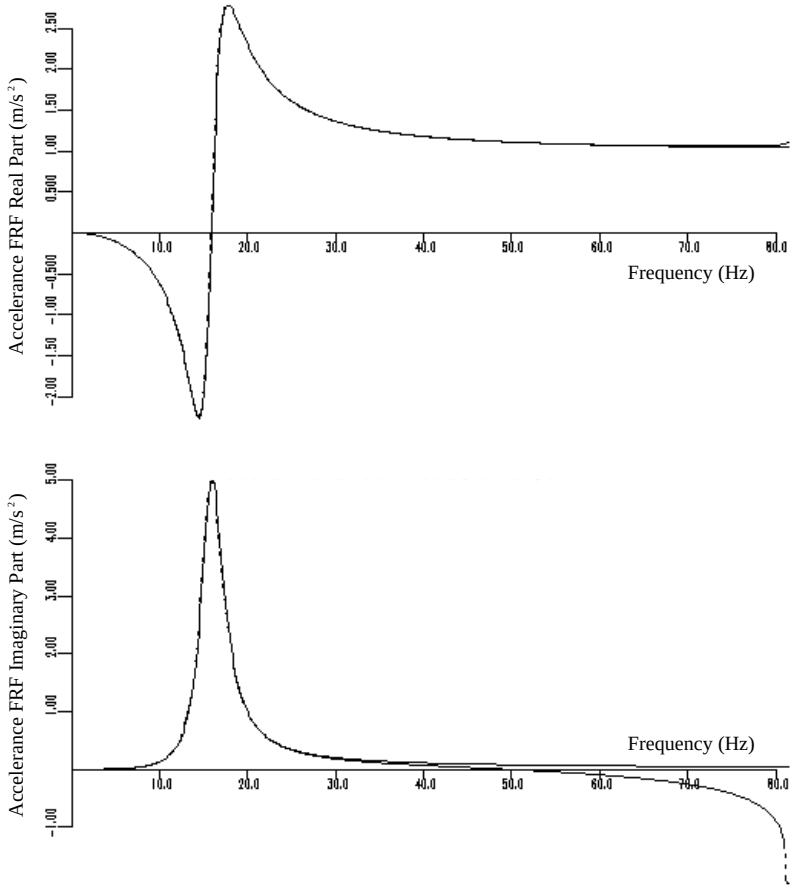


Figure 5.6. Real and imaginary parts of the accelerance FRF and the Hilbert transform. The transform was carried out by converting the FRF to receptance and then converting back to accelerance after the transform.

receptance or mobility but $(\mathrm{Re}\, H(\omega) - 1/m, \mathrm{Im}\, H(\omega))$ if it is accelerance. Figure 5.4 shows the receptance FRF and the corresponding Hilbert transform for the linear system described by the equation

$$\ddot{y} + 20\dot{y} + 10^4 y = x(t). \tag{5.61}$$

As expected, the two curves overlay perfectly. Figure 5.5 shows the corresponding accelerance FRF and the uncorrected Hilbert transform as obtained from equations (4.17) and (4.18). Overlay could be obtained (apart from errors due to the restriction of the integral to a finite frequency range) by using a subtraction as in equation (5.37); a much simpler method is to convert the FRF to receptance

form using (section 4.3)

$$H_{\mathrm{R}}(\omega) = \frac{H_{\mathrm{A}}(\omega)}{-\omega^2} \tag{5.62}$$

carry out the Hilbert transform and convert back to receptance. Figure 5.6 shows the result of this procedure.

In the case of a MDOF system (with proportional damping)

$$H_{\mathrm{A}}(\omega) = \sum_{i=1}^{N} \frac{A_i \omega^2}{\omega_i^2 - \omega^2 + \mathrm{i}\zeta_i \omega_i \omega} \tag{5.63}$$

the appropriate Hilbert transform pair is

$$\left(\mathrm{Re}\, H_{\mathrm{A}}(\omega) + \sum_{i=1}^{N} A_i, \mathrm{Im}\, H_{\mathrm{A}}(\omega) \right). \tag{5.64}$$

5.5 Fourier transform conventions

Throughout this book, the following conventions are used for the Fourier transform:

$$G(\omega) = \mathcal{F}\{g(t)\} = \int_{-\infty}^{\infty} \mathrm{d}t\, \mathrm{e}^{-\mathrm{i}\omega t} g(t) \tag{5.65}$$

$$g(t) = \mathcal{F}^{-1}\{G(\omega)\} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{d}\omega\, \mathrm{e}^{\mathrm{i}\omega t} G(\omega). \tag{5.66}$$

It is equally valid to choose

$$G(\omega) = \int_{-\infty}^{\infty} \mathrm{d}t\, \mathrm{e}^{\mathrm{i}\omega t} g(t) \tag{5.67}$$

$$g(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{d}\omega\, \mathrm{e}^{-\mathrm{i}\omega t} G(\omega). \tag{5.68}$$

These conventions shall be labelled $\mathcal{F}_-$ and $\mathcal{F}_+$ respectively. As would be expected, the Hilbert transform formulae depend critically on the conventions used. The results for $\mathcal{F}_-$ have already been established. The formulae for $\mathcal{F}_+$ can be derived as follows.

In the proof that (i) $\Longleftrightarrow$ (ii) in section 5.2, the result

$$\frac{1}{\pi \mathrm{i}} \int_{-\infty}^{\infty} \mathrm{d}\omega\, \frac{\mathrm{e}^{\mathrm{i}\omega t}}{\omega - \Omega} = \mathrm{e}^{\mathrm{i}\Omega t} \epsilon(t) \tag{5.69}$$

was used from appendix D. If $\mathcal{F}_+$ conventions had been adopted, the result would have been

$$\frac{1}{\pi \mathrm{i}} \int_{-\infty}^{\infty} \mathrm{d}\omega\, \frac{\mathrm{e}^{\mathrm{i}\omega t}}{\omega - \Omega} = -\mathrm{e}^{\mathrm{i}\Omega t} \epsilon(t). \tag{5.70}$$

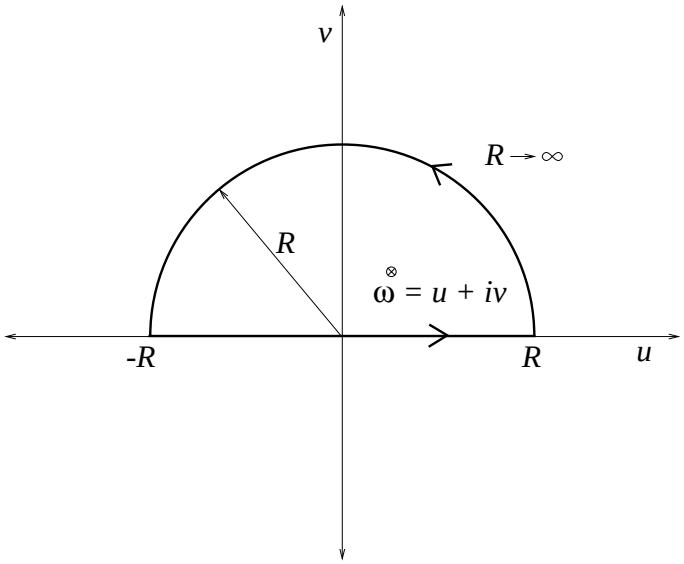


Figure 5.7. Contour for deriving the $\mathcal{F}_+$ Hilbert transform.

In order to cancel the negative sign, a different definition is needed for the Hilbert transform

$$\mathcal{H}\{G(\omega)\} = +\frac{1}{\pi \mathrm{i}}\int_{-\infty}^{\infty} \mathrm{d}\Omega\, \frac{G(\Omega)}{\Omega - \omega} \tag{5.71}$$

or the dispersion relations

$$\mathrm{Re}\, G(\omega) = +\frac{1}{\pi}\int_{-\infty}^{\infty} \mathrm{d}\Omega\, \frac{\mathrm{Im}\, G(\Omega)}{\Omega - \omega} \tag{5.72}$$

$$\mathrm{Im}\, G(\omega) = -\frac{1}{\pi}\int_{-\infty}^{\infty} d\Omega\, \frac{\mathrm{Re}\, G(\Omega)}{\Omega - \omega}. \tag{5.73}$$

To obtain these expressions from the contour integral of section 5.1, it is necessary for the section of contour on the real line to go from $-\infty$ to $+\infty$. As the contour must be followed anticlockwise, it should be completed in the upper half-plane as shown in figure 5.7. As a consequence of choosing this contour, analycity in the *upper* half-plane is needed. The result of these modifications is the $\mathcal{F}_+$ version of the second theorem of section 5.2, i.e. if one of (i)′, (ii)′ or (iii)′ is true, then so are the other two.

(i)′ $G(\omega)$ satisfies the Hilbert transform relations (5.71).
(ii)′ $G(\omega)$ has a causal inverse Fourier transform.
(iii)′ $G(\omega)$ is analytic in the upper half-plane.

The statements about testing FRFs for linearity made in the last chapter apply equally well to both $\mathcal{F}_-$ and $\mathcal{F}_+$. Suppose that an FRF has poles only in the upper

half-plane and therefore satisfies the conditions of Titchmarsh's theorem in $\mathcal{F}_-$. This means that the zeros of the denominator (assume a SDOF system)

$$d_-(\omega) = -m\omega^2 + \mathrm{i}c\omega + k \tag{5.74}$$

are in the upper half-plane. If the conventions are changed to $\mathcal{F}_+$, the denominator changes to

$$d_+(\omega) = -m\omega^2 - \mathrm{i}c\omega + k \tag{5.75}$$

i.e. the product of the roots remains the same while their sum changes sign. Clearly the roots of $d_+(\omega)$ are in the lower half-plane as required by the $\mathcal{F}_+$ Titchmarsh theorem.

5.6 Hysteretic damping models

Having established the connection between causality and FRF pole structure, now is a convenient time to make some observations about the different damping models used with FRFs. The two main models in use are the *viscous* damping model as discussed in chapter 1, where the receptance FRF takes the form

$$H(\omega) = \frac{1}{m(\omega_n^2 - \omega^2 + 2i\zeta\omega_n\omega)} \tag{5.76}$$

and the *hysteretic* or *structural* damping model whose FRF has the form [87]

$$H(\omega) = \frac{1}{m(\omega_n^2(1 + \mathrm{i}\eta) - \omega^2)} \tag{5.77}$$

where η is the hysteretic or structural damping *loss factor*. (The discussion can be restricted to SDOF systems without losing generality.)

It is shown in chapter 1 that the viscous damping model results in a causal impulse response and therefore constitutes a physically plausible approximation to reality. The corresponding calculations for the hysteretic damping model follow.

Before explicitly calculating the impulse response, the question of its causality can be settled by considering the pole structure of (5.77). The poles are at $\omega = \pm\lambda$, where

$$\lambda = \omega_n(1 + \mathrm{i}\eta)^{\frac{1}{2}}. \tag{5.78}$$

A short calculation shows that

$$\operatorname{Im}\lambda = \omega_n(-\tfrac{1}{2} + \tfrac{1}{2}(1 + \eta^2)^{\frac{1}{2}})^{\frac{1}{2}} \tag{5.79}$$

so if $\eta > 0$, it follows that

$$\tfrac{1}{2}(1 + \eta^2)^{\frac{1}{2}} - \tfrac{1}{2} > 0 \tag{5.80}$$

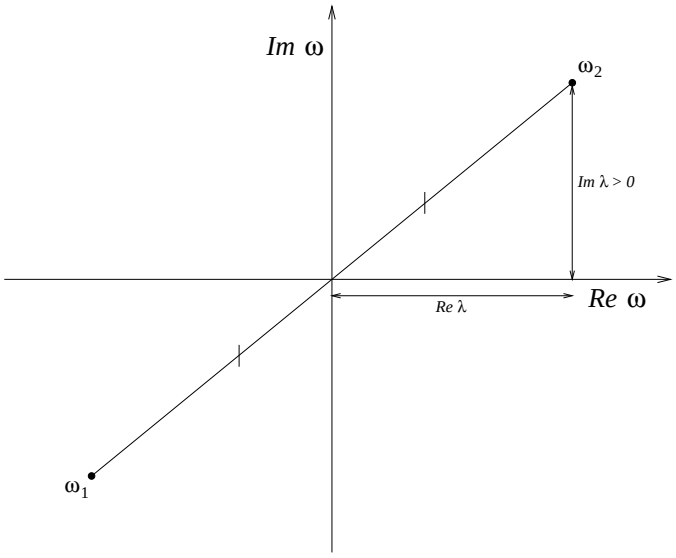


Figure 5.8. Poles of an FRF with hysteretic damping.

and λ has a non-zero imaginary part. This gives the pole structure shown in figure 5.8. $H(\omega)$ in equation (5.77) therefore fails to be analytic in either half-plane. It can be concluded that the impulse response corresponding to this $H(\omega)$ is non-causal. The next question concerns the extent to which causality is violated; if the impulse response is small for all $t < 0$, the hysteretic damping model may still prove useful.

The next derivation follows that of [7], which in turn follows [185]. Recall that the impulse response is defined by

$$h(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{d}\omega\, \mathrm{e}^{\mathrm{i}\omega t} H(\omega). \tag{5.81}$$

It is a simple matter to show that reality of $h(t)$ implies a conjugate symmetry constraint on the FRF

$$H(-\omega) = H^*(\omega). \tag{5.82}$$

On making use of this symmetry, it is possible to cast the impulse response expression in a slightly different form

$$h(t) = \frac{1}{\pi}\,\mathrm{Re}\int_{0}^{\infty} \mathrm{d}\omega\, \mathrm{e}^{\mathrm{i}\omega t} H(\omega) \tag{5.83}$$

which will prove useful. Note that the expression (5.77) *does not* satisfy the conjugate symmetry constraint. To obtain an expression valid on the interval

$\infty > \omega > -\infty$, a small change is made; (5.77) becomes

$$H(\omega) = \frac{1}{m(\omega_n^2(1 + \mathrm{i}\epsilon(\omega)\eta) - \omega^2)} \tag{5.84}$$

where ϵ is the signum function [69].

The impulse response corresponding to the FRF in (5.77) is, from (5.83)

$$h(t) = \frac{1}{\pi m} \operatorname{Re} \int_0^\infty \mathrm{d}\omega \, \mathrm{e}^{\mathrm{i}\omega t} \frac{1}{\omega_n^2(1 + \mathrm{i}\eta) - \omega^2} \tag{5.85}$$

or

$$h(t) = \frac{1}{\pi m} \operatorname{Re} \int_0^\infty \mathrm{d}\omega \, \mathrm{e}^{\mathrm{i}\omega t} \frac{1}{\lambda^2 - \omega^2} \tag{5.86}$$

where λ is as defined before. The partial fraction decompostion of this expression is

$$h(t) = \frac{1}{2\pi\lambda m} \operatorname{Re} \left\{ \int_0^\infty \mathrm{d}\omega \, \frac{\mathrm{e}^{\mathrm{i}\omega t}}{\omega + \lambda} + \int_0^\infty \mathrm{d}\omega \, \frac{\mathrm{e}^{\mathrm{i}\omega t}}{\omega - \lambda} \right\} \tag{5.87}$$

and the integrals can now be expressed in terms of the exponential integral $\mathrm{Ei}(x)$ where [209]

$$\mathrm{Ei}(x) = \int_{-\infty}^x \mathrm{d}t \, \frac{\mathrm{e}^t}{t} = -\int_{-x}^\infty \mathrm{d}t \, \frac{\mathrm{e}^{-t}}{t}, \quad x > 0. \tag{5.88}$$

In fact, a slightly more general form is needed [123]:

$$\int_{-\infty}^x \mathrm{d}t \, \frac{\mathrm{e}^{at}}{t} = \int_{-x}^\infty \mathrm{d}t \, \frac{\mathrm{e}^{-at}}{t} = \mathrm{Ei}(ax). \tag{5.89}$$

The first integral in (5.87) is now straightforward:

$$\int_0^\infty \mathrm{d}\omega \, \frac{\mathrm{e}^{\mathrm{i}\omega t}}{\omega + \lambda} = \int_\lambda^\infty \mathrm{d}\omega \, \frac{\mathrm{e}^{\mathrm{i}(\omega - \lambda)t}}{\omega} = \mathrm{e}^{-\mathrm{i}\lambda t} \int_\lambda^\infty \mathrm{d}\omega \, \frac{\mathrm{e}^{\mathrm{i}\omega t}}{\omega} = -\mathrm{e}^{-\mathrm{i}\lambda t} \, \mathrm{Ei}(i\lambda t) \tag{5.90}$$

and this is valid for all t.

The second integral is a little more complicated. For negative time t,

$$\int_0^\infty \mathrm{d}\omega \, \frac{\mathrm{e}^{\mathrm{i}\omega t}}{\omega - \lambda} = \int_{-\lambda}^\infty \mathrm{d}\omega \, \frac{\mathrm{e}^{\mathrm{i}(\omega + \lambda)t}}{\omega} = \mathrm{e}^{\mathrm{i}\lambda t} \int_{-\lambda}^\infty \mathrm{d}\omega \, \frac{\mathrm{e}^{\mathrm{i}\omega t}}{\omega} = -\mathrm{e}^{\mathrm{i}\lambda t} \, \mathrm{Ei}(-\mathrm{i}\lambda t) \tag{5.91}$$

and for positive time t,

$$\int_0^\infty \mathrm{d}\omega \, \frac{\mathrm{e}^{\mathrm{i}\omega t}}{\omega - \lambda} = \int_{-\infty}^\infty \mathrm{d}\omega \, \frac{\mathrm{e}^{\mathrm{i}\omega t}}{\omega - \lambda} - \int_{-\infty}^0 \mathrm{d}\omega \, \frac{\mathrm{e}^{\mathrm{i}\omega t}}{\omega - \lambda} \tag{5.92}$$

$$= 2\pi\mathrm{i}\mathrm{e}^{\mathrm{i}\lambda t} - \mathrm{e}^{\mathrm{i}\lambda t} \int_{-\infty}^{-\lambda} \mathrm{d}\omega \, \frac{\mathrm{e}^{\mathrm{i}\omega t}}{\omega} = \mathrm{e}^{\mathrm{i}\lambda t}[2\pi\mathrm{i} - \mathrm{Ei}(-\mathrm{i}\lambda t)]. \tag{5.93}$$

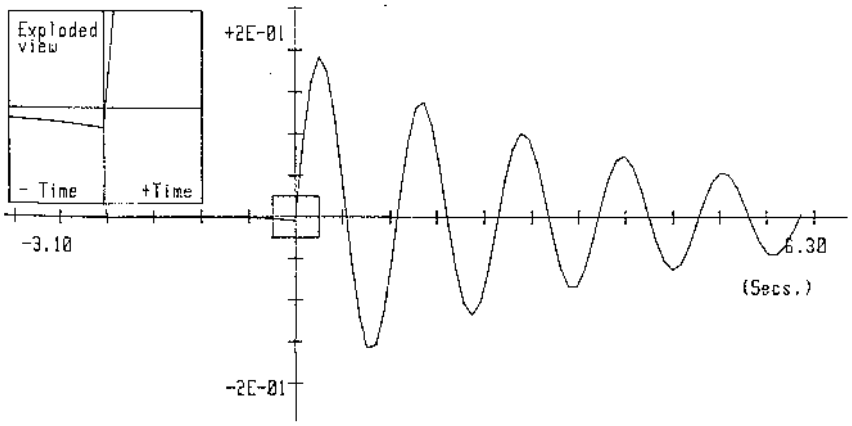


Figure 5.9. Impulse response of a SDOF system with 10% hysteretic damping showing a non-causal response.

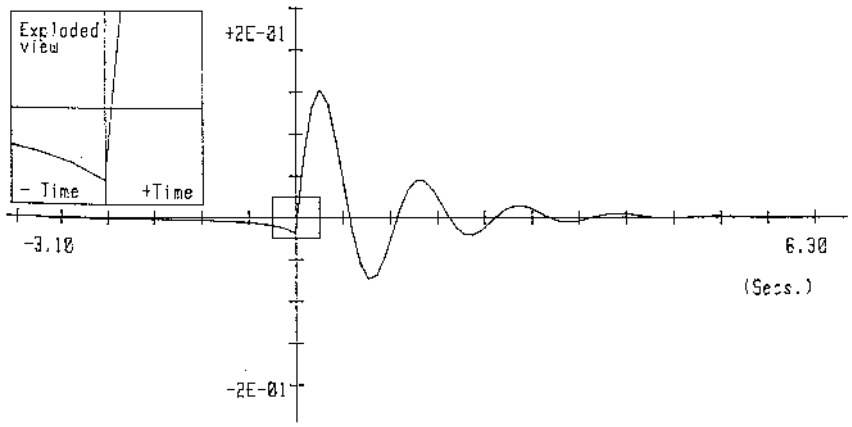
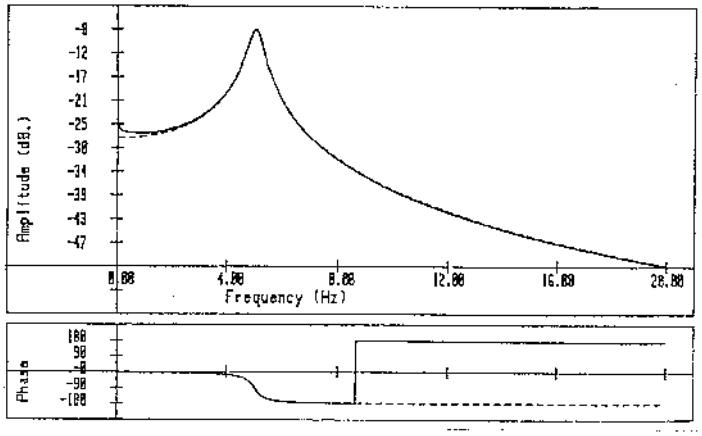


Figure 5.10. Impulse response of a SDOF system with 40% hysteretic damping showing an increased non-causal response.

So the overall expression for the impulse response is

$$h(t) = \frac{1}{2\pi\lambda m}\,\mathrm{Re}[-\mathrm{e}^{-\mathrm{i}\lambda t}\,\mathrm{Ei}(\mathrm{i}\lambda t) + \mathrm{e}^{\mathrm{i}\lambda t}\,\mathrm{Ei}(-\mathrm{i}\lambda t) - 2\pi\mathrm{i}\mathrm{e}^{\mathrm{i}\lambda t}\Theta(t)]. \qquad (5.94)$$

In order to display this expression, it is necessary to evaluate the exponential integrals. For small t, the most efficient means is to use the rapidly convergent



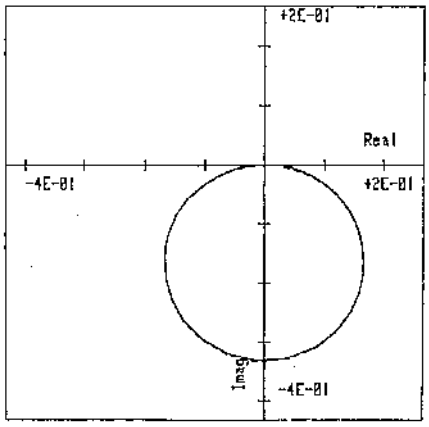


Figure 5.11. The FRF and Hilbert transform for a SDOF system with 10% hysteretic damping showing deviations at low frequency.

power series [209]

$$\mathrm{Ei}(x) = \gamma + \log x + \sum_{i=1}^{\infty} \frac{x^i}{i \cdot i!} \tag{5.95}$$

where γ is Euler's constant $0.5772157\ldots$. For large t, the asymptotic expansion [209]

$$\mathrm{Ei}(x) \sim \frac{\mathrm{e}^x}{x}\left(1 + \frac{1!}{x} + \frac{2!}{x^2} + \cdots\right) \tag{5.96}$$

can be used. Alternatively for large t, there is a rapidly convergent representation

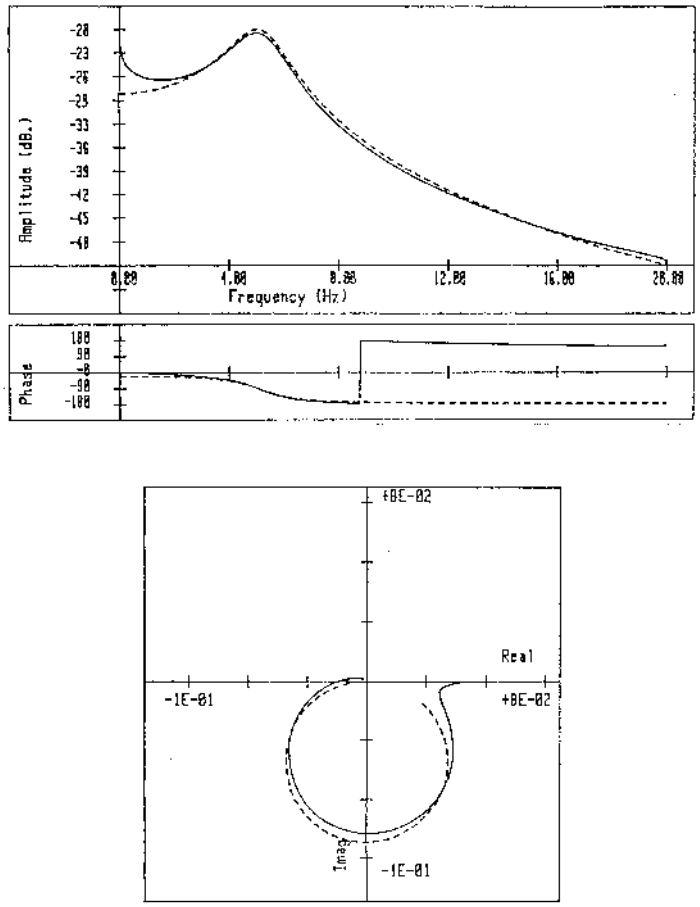


Figure 5.12. The FRF and Hilbert transform for a SDOF system with 40% hysteretic damping showing deviations even at resonance.

of the related function $\mathrm{E}_1(x) = -\,\mathrm{Ei}(-x)$, in terms of continued fractions[6], i.e.

$$\mathrm{E}_1(x) = \mathrm{e}^{-x}\left(\frac{1|}{|x+1} - \frac{1|}{|x+3} - \frac{4|}{|z+5} - \frac{9|}{|z+7} - \frac{16|}{|z+9}\cdots\right). \qquad (5.97)$$

Press *et al* [209] provide FORTRAN routines for all these purposes.

Figures 5.9 and 5.10 show the impulse response for 10% and 40% hysteretic damping (i.e. $\eta = 0.1$ and $\eta = 0.4$ respectively). The non-causal nature of these functions is evident, particularly for the highly damped system. Figures 5.11 and 5.12 show the extent to which the Hilbert transforms are affected, there

[6] The authors would like to thank Dr H Milne for pointing this out.

is noticeable distortion at low frequencies, and around resonance for higher damping. It can be concluded that hysteretic damping should only be used with caution in simulations where the object is to investigate Hilbert transform distortion as a result of nonlinearity.

5.7 The Hilbert transform of a simple pole

It has been previously observed that a generic linear dynamical system will have a rational function FRF. In fact, according to standard approximation theorems, any function can be approximated arbitrarily closely by a rational function of some order. It is therefore instructive to consider such functions in some detail. Assume a rational form for the FRF $G(\omega)$:

$$G(\omega) = \frac{A(\omega)}{B(\omega)} \tag{5.98}$$

with A and B polynomials in ω. It will be assumed throughout that the order of B is greater than the order of A. This can always be factorized to give a pole–zero decompositon:

$$G(\omega) = \alpha \frac{\prod_{i=1}^{N_z}(\omega - z_i)}{\prod_{i=1}^{N_p}(\omega - p_i)} \tag{5.99}$$

where α is a constant, N_z is the number of zeros z_i and N_p is the number of poles p_i. As $N_p > N_z$, the FRF has a partial fraction decomposition

$$G(\omega) = \sum_{i=1}^{N_p} \frac{C_i}{\omega - p_i} \tag{5.100}$$

(assuming for the moment that there are no repeated poles). Because the Hilbert transform is a linear operation, the problem of transforming G has been reduced to the much simpler problem of transforming a simple pole. Now, if the pole is in the upper-half plane, the results of the previous sections suffice to show that (assuming $\mathcal{F}_-$ conventions)

$$\mathcal{H}\left\{\frac{1}{\omega - p_i}\right\} = \frac{1}{\omega - p_i}. \tag{5.101}$$

A straightforward modification of the analysis leads to the result

$$\mathcal{H}\left\{\frac{1}{\omega - p_i}\right\} = -\frac{1}{\omega - p_i} \tag{5.102}$$

if p_i is in the lower half-plane. In fact, the results are the same for repeated poles $1/(\omega - p_i)^n$. Now, equation (5.100) provides a decomposition

$$G(\omega) = G^+(\omega) + G^-(\omega) \tag{5.103}$$

where $G^+(\omega)$ is analytic in the lower half-plane and $G^-(\omega)$ is analytic in the upper half-plane. It follows from these equations that

$$\mathcal{H}\{G(\omega)\} = G^+(\omega) - G^-(\omega). \tag{5.104}$$

This equation is fundamental to the discussion of the following section and will be exploited in other parts of this book.

Consider the effect of applying the Hilbert transform twice. This operation is made trivial by using the Fourier decompositions of the Hilbert operator, i.e.

$$\mathcal{H}^2 = (\mathcal{F} \circ \in \circ \mathcal{F}^{-1})^2 = \mathcal{F} \circ \in \circ \mathcal{F}^{-1} \circ \mathcal{F} \circ \in \circ \mathcal{F}^{-1} = \mathcal{F} \circ \in^2 \circ \mathcal{F}^{-1}. \tag{5.105}$$

Now, recall from chapter 4 that $\in \{g(t)\} = \epsilon(t)g(t)$, ($\epsilon(t)$ being the signum function) so $\in^2 \{g(t)\} = \epsilon(t)^2 g(t) = g(t)$, and $\in^2$ is the identity, and expression (5.105) collapses to

$$\mathcal{H}^2 = \text{Identity} \tag{5.106}$$

or, acting on a function $G(\omega)$

$$\mathcal{H}^2\{G(\omega)\} = G(\omega) \tag{5.107}$$

which shows that *any* function which is twice-transformable is an eigenvector or eigenfunction of the operator $\mathcal{H}^2$ with eigenvalue unity. It is a standard result of linear functional analysis that the eigenvalues of $\mathcal{H}$ must therefore be ± 1. This discussion therefore shows that the simple poles are eigenfunctions of the Hilbert transform with eigenvalue $+1$ if the pole is in the upper half-plane and -1 if the pole is in the lower half-plane.

5.8 Hilbert transforms without truncation errors

As discussed in the previous chapter, there are serious problems associated with computation of the Hilbert transform if the FRF data are truncated. The analysis of the previous section allows an alternative method to those discussed in chapter 4. More detailed discussions of the 'new' method can be found in [142] or [144].

The basis of the approach is to establish the position of the FRF poles in the complex plane and thus form the decomposition (5.103). This is achieved by formulating a Rational Polynomial (RP) model of the FRF of the form (5.98) over the chosen frequency range and then converting this into the required form via a pole–zero decomposition.

Once the RP model G_{RP} is established, it can be converted into a pole-zero form (5.99). The next stage is a long division and partial-fraction analysis in order to produce the decomposition (5.103). If p_i^+ are the poles in the upper half-plane and p_i^- are the poles in the lower half-plane, then

$$G^+_{\mathrm{RP}}(\omega) = \sum_{i=1}^{N_+} \frac{C_i^+}{\omega - p_i^+}, \quad G^-_{\mathrm{RP}}(\omega) = \sum_{i=1}^{N_-} \frac{C_i^-}{\omega - p_i^-} \tag{5.108}$$

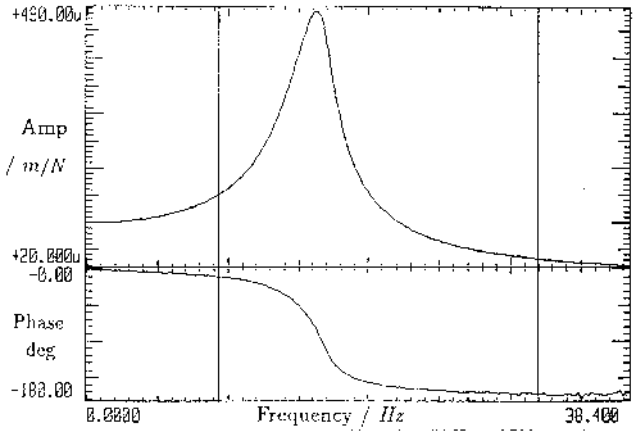


Figure 5.13. Bode plot of Duffing oscillator FRF with a low excitation level.

where C_i^+ and C_i^- are coefficients fixed by the partial fraction analysis. N_+ is the number of poles in the upper half-plane and N_- is the number of poles in the upper lower half-plane. Once this decomposition is established, the Hilbert transform follows from (5.104). (Assuming again that the RP model has more poles than zeros. If this is not the case, the decomposition (5.103) is supplemented by a term $G^0(\omega)$ which is analytic. This has no effect on the analysis.)

This procedure can be demonstrated using data from numerical simulation. The system chosen is a Duffing oscillator with equation of motion

$$\ddot{y} + 20\dot{y} + 10\,000y + 5 \times 10^9 y^3 = X \sin(\omega t). \tag{5.109}$$

Data were generated over 256 spectral lines from 0–38.4 Hz in a simulated stepped-sine test based on a standard fourth-order Runge–Kutta scheme [209]. The data were truncated by removing data above and below the resonance leaving 151 spectral lines in the range 9.25–32.95 Hz.

Two simulations were carried out. In the first, the Duffing oscillator was excited with $X = 1.0$ N giving a change in the resonant frequency from the linear conditions of 15.9 to 16.35 Hz and in amplitude from 503.24×10^{-6} m N^{-1} to 483.0×10^{-6} m N^{-1}. The FRF Bode plot is shown in figure 5.13, the cursor lines indicate the range of the FRF which was used. The second simulation took $X = 2.5$ N which was high enough to produce a jump bifurcation in the FRF. In this case the maximum amplitude of 401.26×10^{-6} m N^{-1} occurred at a frequency of 19.75 Hz. Note that in the case of this nonlinear system the term 'resonance' is being used to indicate the position of maximum gain in the FRF.

The first stage in the calculation process is to establish the RP model of the FRF data. On the first data set with $X = 1$, in order to obtain an accurate model of the FRF, 24 denominator terms and 25 numerator terms were used. The number

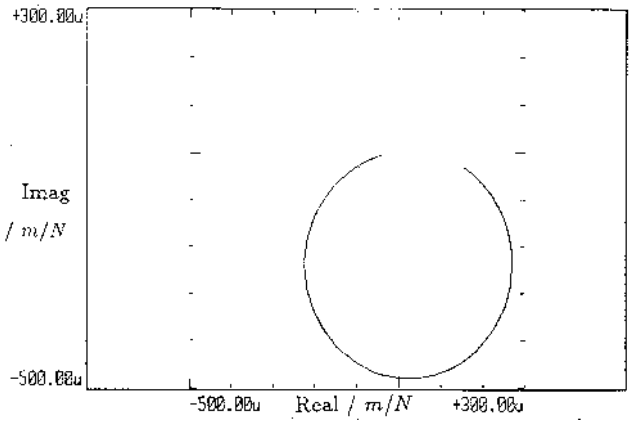


Figure 5.14. Overlay of RP model FRF $G_{\mathrm{RP}}(\omega)$ and original FRF $G(\omega)$ for the Duffing oscillator at a low excitation level. (The curves overlay with no distinction.)

of terms in the polynomial required to provide an accurate model of the FRF will depend on several factors including the number of modes in the frequency range, the level of distortion in the data and the amount of noise present. The accuracy of the RP model is evident from figure 5.14 which shows a Nyquist plot of the original FRF, $G(\omega)$ with the model $G_{\mathrm{RP}}(\omega)$ overlaid on the frequency range 10–30 Hz[7].

The next stage in the calculation is to obtain the pole–zero decomposition (5.99). This is accomplished by solving the numerator and denominator polynomials using a computer algebra package.

The penultimate stage of the procedure is to establish the decomposition (5.103). Given the pole-zero form of the model, the individual pole contributions are obtained by carrying out a partial fraction decomposition, because of the complexity of the model, a computer algebra package was used again.

Finally, the Hilbert transform is obtained by flipping the sign of $G^{-}(\omega)$, the sum of the pole terms in the lower half-plane. The result of this calculation for the low excitation data is shown in figure 5.15 in a Bode amplitude format. The overlay of the original FRF data and the Hilbert transform calculated by the RP method are given; the frequency range has been limited to 10–30 Hz.

A simple test of the accuracy of the RP Hilbert transform was carried out. A Hilbert transform of the low excitation data was calculated using the fast FFT-based technique (section 4.4.4) on an FRF using a range of 0–50 Hz in order to minimize truncation errors in the calculation. Figure 5.16 shows an overlay of the RP Hilbert transform (from the truncated data) with that calculated from the FFT

[7] The authors would like to thank Dr Peter Goulding of the University of Manchester for carrying out the curve-fit. The method was based on an instrumental variables approach and details can be found in [86].

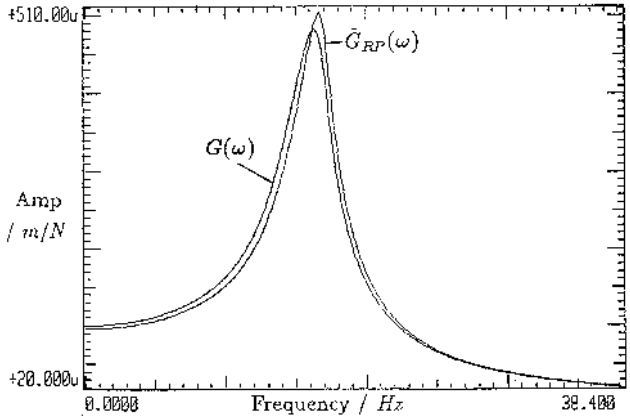


Figure 5.15. Original FRF $G(\omega)$ and RP Hilbert transform $\tilde{G}_{\mathrm{RP}}(\omega)$ for the Duffing oscillator at a low excitation level.

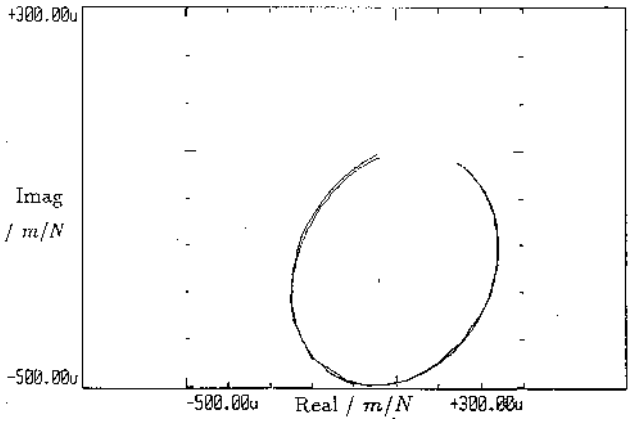


Figure 5.16. Nyquist plot comparison of RP and FFT Hilbert transform for the Duffing oscillator at a low excitation level.

technique. The Nyquist format is used.

The second, high-excitation, FRF used to illustrate the approach contained a bifurcation or 'jump' and thus offered a more stringent test of the RP curve-fitter. A greater number of terms in the RP model were required to match the FRF. Figure 5.17 shows the overlay achieved using 32 terms in the denominator and 33 terms in the numerator. There is no discernible difference. Following the same calculation process as above leads to the Hilbert transform shown in figure 5.18, shown with the FRF.

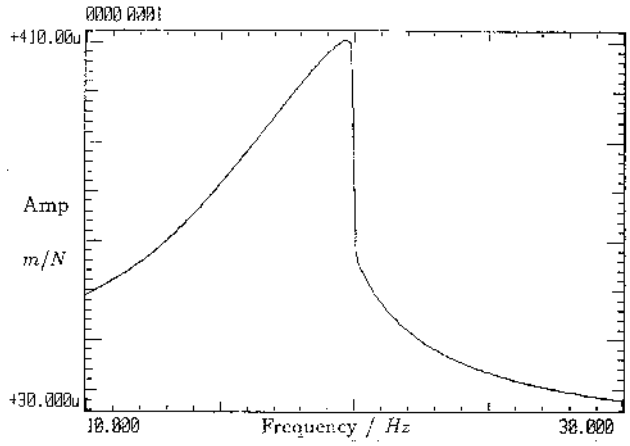


Figure 5.17. Overlay of RP model FRF $G_{\mathrm{RP}}(\omega)$ and original FRF $G(\omega)$ for the Duffing oscillator at a high excitation level.

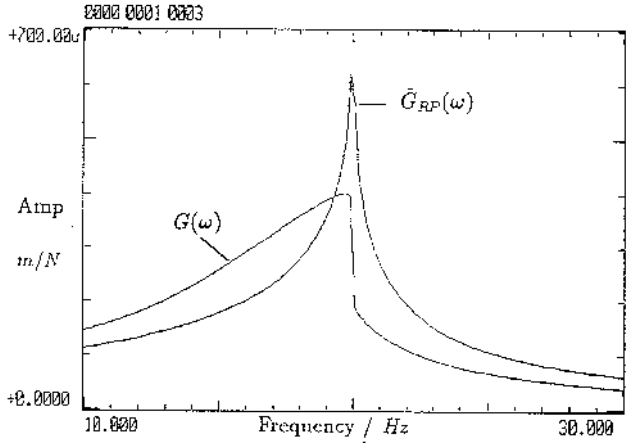


Figure 5.18. Original FRF $G(\omega)$ and RP Hilbert transform $\tilde{G}_{\mathrm{RP}}(\omega)$ for the Duffing oscillator at high excitation.

5.9 Summary

The end of this chapter not only concludes the discussion of the Hilbert transform, but suspends the main theme of the book thus far. With the exception of Feldman's method (section 4.8), the emphasis has been firmly on the problem of *detecting* nonlinearity. The next two chapters are more ambitious; methods of *system identification* are discussed which can potentially provide estimates of an unknown nonlinear system's equations of motion given measured data. Another

important difference is that the next two chapters concentrate almost exclusively on the time domain in contrast to the frequency-domain emphasis thus far. The reason is fairly simple: in order to identify the true nonlinear structure of the system, there must be no loss of information through linearization. Unfortunately, all the frequency-domain objects discussed so far correspond to linearizations of the system. This does not mean that the frequency domain has no place in detailed system identification; in chapter 8, an exact frequency-domain representation for nonlinear systems will be considered.

Chapter 7

System identification—continuous time

7.1 Introduction

The last chapter discussed a number of approaches to system identification based on discrete-time models. Once the structure of the model was fixed, the system identification (ID) problem was reduced to parameter estimation as only the coefficients of the model terms remained unspecified. For obvious reasons, such identification schemes are often referred to as *parametric*. The object of this chapter is to describe approaches to system ID based on the assumption of a continuous-time model. Such schemes can be either *parametric* or *non-parametric*. Unfortunately, there appears to be confusion in the literature as to what these terms mean. The following definitions are adopted here:

Parametric identification. This term shall be reserved for methods where a model structure is specified and the coefficients of the terms are obtained by some estimation procedure. Whether the parameters are physical (i.e. m, c and k for a SDOF continuous-time system) or unphysical (i.e. the coefficients of a discrete-time model) shall be considered irrelevant, the distinguishing feature of such approaches is that equations of motion are obtained.

Non-parametric identification. This term shall be reserved for methods of identification, where the primary quantities obtained do not directly specify equations of motion. One such approach, the restoring-force surface method discussed in this chapter, results in a visual representation of the internal forces in the system. The Volterra series of the following chapter is another such approach.

In many cases, this division is otiose. It will soon become evident that the restoring force surfaces are readily converted from non-parametric to parametric models. In some respects the division of models into *physical* and *non-physical* is more meaningful. The reader should, however, be aware of the terminology to be found in the literature.

The current chapter is not intended to be a comprehensive review of continuous-time approaches to system ID. Rather, the evolution of a particular class of models is described. The curious reader can refer to [152] and [287] for references to more general literature. The thread followed in this chapter begins with the identification procedure of Masri and Caughey.

7.2 The Masri–Caughey method for SDOF systems

7.2.1 Basic theory

The simple procedure described in this section allows a direct non-parametric identification for SDOF nonlinear systems. The only *a priori* information required is an estimate of the system mass. The basic procedures described in this section were introduced by Masri and Caughey [174]; developments discussed later arise from a parallel approach proposed independently by Crawley and Aubert [70, 71]; the latter method was referred to by them as 'force-state mapping'.

The starting point is the equation of motion as specified by Newton's second law

$$m\ddot{y} + f(y, \dot{y}) = x(t) \tag{7.1}$$

where m is the mass (or an effective mass) of the system and $f(y, \dot{y})$ is the internal restoring force which acts to return the absorber to equilibrium when disturbed. The function f can be a quite general function of position $y(t)$ and velocity $\dot{y}(t)$. In the special case when the system is linear

$$f(y, \dot{y}) = c\dot{y} + ky \tag{7.2}$$

where c and k are the damping constant and stiffness respectively. Because f is assumed to be dependent only on y and $\dot{y}$ it can be represented by a surface over the phase plane, i.e. the $(y, \dot{y})$-plane. A trivial re-arrangement of equation (7.1) gives

$$f(y(t), \dot{y}(t)) = x(t) - m\ddot{y}(t). \tag{7.3}$$

If the mass m is known and the excitation $x(t)$ and acceleration $\ddot{y}(t)$ are measured, all the quantities on the right-hand side of this equation are known and hence so is f. As usual, measurement of a time signal entails sampling it at regularly spaced intervals Δt. (In fact, such is the generality of the method that regular sampling is not essential; however, if any preprocessing is required for the measured data, regular sampling is usually required.) If $t_i = (i-1)\Delta t$ denotes the ith sampling instant, then at t_i, equation (7.3) gives

$$f_i = f(y_i, \dot{y}_i) = x_i - m\ddot{y}_i \tag{7.4}$$

where $x_i = x(t_i)$ and $\ddot{y}_i = \ddot{y}(t_i)$ and hence f_i are known at each sampling instant. If the velocities $\dot{y}_i$ and displacements y_i are also known (i.e. from direct

measurement or from numerical integration of the sampled acceleration data), at each instant $i = 1, \ldots, N$ a triplet $(y_i, \dot{y}_i, f_i)$ is specified. The first two values indicate a point in the phase plane, the third gives the height of the restoring force surface above that point. Given this scattering of force values above the phase plane there are a number of methods of interpolating a continuous surface on a regular grid; the procedures used here are discussed a little later.

Once the surface is obtained, Masri and Caughey [174] construct a parametric model of the restoring force in the form of a double Chebyshev series; formally

$$f(y, \dot{y}) = \sum_{i=0}^{m} \sum_{j=0}^{n} C_{ij} T_i(y) T_j(\dot{y}) \tag{7.5}$$

where $T_i(y)$ is the Chebyshev polynomial of order i. The use of these polynomials was motivated by a number of factors:

- They are orthogonal polynomials. This means that one can estimate coefficients for a double summation or series of order (m, n) and the truncation of the sum to order (i, j), where $i < m$ and $j < n$ is the best approximation of order (i, j). This means that one need not re-estimate coefficients if a lower-order model is acceptable. This is not the case for simple polynomial models. Similarly, if the model needs to be extended, the coefficients for the lower-order model will still stand.
- The estimation method for the coefficients used by Masri and Caughey required the evaluation of a number of integrals. In the case of the Chebyshev expansion, a change of variables exists which makes the numerical integrals fairly straightforward. This is shown later.
- In the family of polynomial approximations to a given function over a given interval, there will be one which has the smallest maximum deviation from that function over the interval. This approximating polynomial—the *minimax* polynomial has so far eluded discovery. However, one of the nice properties of the Chebyshev expansion is that it is very closely related to the required minimax expansion. The reason for this is that the error in the Chebyshev expansion on a given interval oscillates between almost equal upper and lower bounds. This property is sometimes referred to as the *equal-ripple* property.

Although more convenient approaches are now available which make use of ordinary polynomial expansions, the Masri–Caughey technique is still sometimes used for MDOF systems, so the estimation procedure for the Chebyshev series will be given. The various properties of Chebyshev polynomials used in this study are collected together in appendix H. A comprehensive reference can be found in [103]. A number of useful numerical routines relating to Chebyshev approximation can be found in [209].

The first problem encountered in fitting a model of the form (7.5) relates to the overall scale of the data y and $\dot{y}$. In order to obtain the coefficients C_{ij},

the orthogonality properties of the polynomials are needed (see appendix H). The $T_n(y)$ are orthogonal on the interval $[-1, 1]$, i.e.

$$\int_{-1}^{+1} \mathrm{d}y\, w(y) T_i(y) T_j(y) = \pi \delta_{ij} - \frac{\pi}{2} \delta_{0i} \delta_{0j} \tag{7.6}$$

where δ_{ij} is the Kronecker delta. The weighting factor $w(y)$ is

$$w(y) = (1 - y^2)^{\frac{1}{2}}. \tag{7.7}$$

It is a straightforward matter to show that the coefficients of the model (7.5) are given by

$$C_{ij} = X_i X_j \int_{-1}^{+1} \int_{-1}^{+1} \mathrm{d}y\, \mathrm{d}\dot{y}\, w(y) w(\dot{y}) T_i(y) T_j(\dot{y}) f(y, \dot{y}) \tag{7.8}$$

where

$$X_i = \frac{1}{\pi}(1 + \delta_{0i}) \tag{7.9}$$

as shown in appendix H. The scale or normalization problem arises from the fact that the measured data will not be confined to the region $[-1, 1] \times [-1, 1]$ in the phase plane, but will occupy part of the region $[y_{\min}, y_{\max}] \times [\dot{y}_{\min}, \dot{y}_{\max}]$, where $y_{\min}$ etc. specify the bounds of the data. Clearly if $y_{\max} \ll 1$, the data will not span the appropriate interval for orthogonality, and if $y_{\max} \gg 1$, very little of the data would be usable. Fortunately, the solution is very straightforward; the data is mapped onto the appropriate region $[-1, 1] \times [-1, 1]$ by the linear transformations

$$\zeta(y) = \overline{y} = \frac{y - \frac{1}{2}(y_{\max} + y_{\min})}{\frac{1}{2}(y_{\max} - y_{\min})} \tag{7.10}$$

$$\dot{\zeta}(\dot{y}) = \overline{\dot{y}} = \frac{\dot{y} - \frac{1}{2}(\dot{y}_{\max} + \dot{y}_{\min})}{\frac{1}{2}(\dot{y}_{\max} - \dot{y}_{\min})} \tag{7.11}$$

and in this case $\dot{\zeta}$ *does not* mean $\mathrm{d}\zeta/\mathrm{d}t$. This means that the model actually estimated is

$$\begin{aligned} f(y, \dot{y}) &= \overline{f}(\overline{y}, \overline{\dot{y}}) \\ &= \sum_{i=0}^{m} \sum_{j=0}^{n} C_{ij}^{\zeta} T_i(\overline{y}) T_j(\overline{\dot{y}}) = \sum_{i=0}^{m} \sum_{j=0}^{n} C_{ij}^{\zeta} T_i(\zeta(y)) T_j(\dot{\zeta}(\dot{y})) \end{aligned} \tag{7.12}$$

where the first of the three equations is simply the transformation law for a scalar function under a change of coordinates. It is clear from this expression that the model coefficients will be sample-dependent. The coefficients follow from a modified form of (7.8):

$$C_{ij}^{\zeta} = X_i X_j \int_{-1}^{+1} \int_{-1}^{+1} \mathrm{d}u\, \mathrm{d}v\, w(u) w(v) T_i(u) T_j(v) \overline{f}(u, v) \tag{7.13}$$

and

$$\overline{f}(u,v) = f(\zeta^{-1}(u), \dot{\zeta}^{-1}(v)) \tag{7.14}$$

Following a change of coordinates

$$\begin{aligned} \theta &= \cos^{-1}(u) \\ \psi &= \cos^{-1}(v) \end{aligned} \tag{7.15}$$

the integral (7.13) becomes

$$C_{ij}^{\zeta} = X_i X_j \int_0^{\pi} \int_0^{\pi} \mathrm{d}\theta \, \mathrm{d}\psi \, \cos(i\theta) \cos(j\psi) \overline{f}(\cos(\theta), \cos(\psi)) \tag{7.16}$$

and the troublesome singular functions $w(u)$ and $w(v)$ have been removed. The simplest approach to evaluating this integral is to use a rectangle rule. The θ-range $(0, \pi)$ is divided into n_θ intervals of length $\Delta\theta = \pi/n_\theta$ and the ψ-range into n_ψ intervals of length $\Delta\psi = \pi/n_\psi$ and the integral is approximated by the summation

$$C_{ij}^{\zeta} = X_i X_j \sum_{k=1}^{n_\theta} \sum_{l=1}^{n_\psi} \Delta\theta \Delta\psi \, \cos(i\theta_k) \cos(j\psi_l) \overline{f}(\cos(\theta_k), \cos(\psi_l)) \tag{7.17}$$

where $\theta_k = (k-1)\Delta\theta$ and $\psi_l = (l-1)\Delta\psi$.

At this point, the question of interpolation is raised again. The values of the force function f on a regular grid in the $(y, \dot{y})$-plane must be transformed into values of the function $\overline{f}$ on a regular grid in the (θ, ψ). This matter will be discussed shortly.

Once the coefficients C_{ij}^{ζ} have been obtained, the model for the restoring force is established. To recap

$$f(y, \dot{y}) = \sum_{i=0}^{m} \sum_{j=0}^{n} C_{ij}^{\zeta} T_i(\zeta(y)) T_j(\dot{\zeta}(\dot{y})) \tag{7.18}$$

and this is valid on the rectangle $[y_{\min}, y_{\max}] \times [\dot{y}_{\min}, \dot{y}_{\max}]$. As long as the true form of the restoring force $f(y, \dot{y})$ *is* multinomial and the force $x(t)$ driving the system excites the highest-order terms in f, the approximation will be valid throughout the phase plane. If either of these conditions do not hold, the model will only be valid on the rectangle containing the sample data. If the force $x(t)$ has not excited the system adequately, the model is input-dependent and may well lose its predictive power if radically different inputs are used to excite the system.

There is a class of systems for which the restoring force method cannot be used in the simple form described here, i.e. systems with memory or *hysteretic* systems. In this case, the internal force does not depend entirely on the instantaneous position of the system in the phase plane. As an illustration, consider the Bouc–Wen model [263]

$$m\ddot{y} + f(y, \dot{y}) + z = x(t) \tag{7.19}$$

$$\dot{z} = -\alpha|\dot{y}|z \cdot |z|^{n-1} - \beta\dot{y}|z^n| + A\dot{y} \tag{7.20}$$

which can represent a broad range of hysteresis characteristics. The restoring force surface would fail here because the internal force is a function of y, $\dot{y}$ *and* z; this means that the force surface over $(y, \dot{y})$ would appear to be multi-valued. A smooth surface can be obtained by exciting the system at a single frequency over a range of amplitudes; however, the surfaces would be different for each frequency. Extensions of the method to cover hysteretic systems have been devised [27, 169]; models of the type

$$\dot{f} = g(f, \dot{y}) \tag{7.21}$$

are obtained which also admit a representation as a surface over the $(f, \dot{y})$ plane. A parametric approach to modelling hysteretic systems was pursued in [285] where a Bouc–Wen model (7.20) was fitted to measured data; this approach is complicated by the fact that the model (7.20) is nonlinear in the parameters and a discussion is postponed until section 7.6 of this chapter.

7.2.2 Interpolation procedures

The problem of interpolating a continuous surface from values specified on a regular grid is well-known and documented [209]. In this case it is a straightforward matter to obtain an interpolated value or *interpolant* which is many times differentiable. The restoring force data are required on a regular grid in order to facilitate plotting of the surface. Unfortunately, the data used to construct a restoring force surface will generally be randomly or irregularly placed in the phase plane and this makes the interpolation problem considerably more difficult. A number of approaches are discussed in [182] and [160]. One method in particular, the *natural neighbour method* of Sibson [225], is attractive as it can produce a continuous and differentiable interpolant. The workings of the methods are rather complicated and involve the construction of a triangulation of the phase plane, the reader is referred to [225] for details. The software TILE4 [226] was used throughout this study in order to construct the Masri–Caughey restoring force surfaces.

The advantage of having a higher-order differentiable surface is as follows. The continuous or C^0 interpolant essentially assumes linear variations in the interpolated function between the data points, i.e. the interpolant is exact only for a linear restoring force surface:

$$f(y, \dot{y}) = \alpha + \beta y + \gamma \dot{y}. \tag{7.22}$$

As a consequence, it can only *grow* linearly in regions where there are very little data. As the functions of interest here are nonlinear, this is a disadvantage. The undesirable effects of this will be shown by example later.

The surfaces produced by natural neighbour interpolation, can be continuous or differentiable (designated C^1). Such functions are generally specified by

quadratic functions[1]

$$f(y,\dot{y}) = \alpha + \beta y + \gamma\dot{y} + \epsilon y^2 + \eta y\dot{y} + \sigma\dot{y}^2. \tag{7.23}$$

The natural neighbour method is used to solve the first interpolation problem in the Masri–Caughey approach. The second interpolation is concerned with going from a regular grid in the phase plane to a regular grid in the (θ,ψ)-plane. The natural neighbour method could be used again, but it is rather computationally expensive and as long as a reasonably fine mesh is used, simpler methods suffice. Probably the simplest is the C^0 *bilinear* interpolation [209].

If arrays of values y_i, $i = 1,\ldots,N$ and $\dot{y}_j$, $j = 1,\ldots,M$ specify the locations of the grid points and an array f_{ij} holds the corresponding values of the force function, the bilinear interpolant at a general point $(y,\dot{y})$, is obtained as follows.

(1) Identify the grid-square containing the point $(y,\dot{y})$, i.e. find (m,n) such that

$$\begin{aligned} y_m \le y \le y_{m+1} \\ \dot{y}_n \le \dot{y} \le \dot{y}_{n+1}. \end{aligned} \tag{7.24}$$

(2) Define

$$\begin{aligned} f_1 &= f_{mn} \\ f_2 &= f_{m+1\ n} \\ f_3 &= f_{m+1\ n+1} \\ f_4 &= f_{m\ n+1} \end{aligned} \tag{7.25}$$

and

$$\begin{aligned} t &= (y - y_m)/(y_{m+1} - y_m) \\ u &= (\dot{y} - \dot{y}_n)/(\dot{y}_{n+1} - \dot{y}_n). \end{aligned} \tag{7.26}$$

(3) Evaluate the interpolant:

$$f(y,\dot{y}) = (1-t)(1-u)f_1 + t(1-u)f_2 + tuf_3 + (1-t)uf_4. \tag{7.27}$$

All the machinery required for the basic Masri–Caughey procedure is now in place and the method can be illustrated on a number of simple systems.

[1] In fact, the natural neighbour method is exact for a slightly more restricted class of functions, namely the *spherical quadratics*:

$$f(y,\dot{y}) = \alpha + \beta y + \gamma\dot{y} + y^2 + \dot{y}^2.$$

7.2.3 Some examples

The Masri–Caughey procedure is demonstrated in this section on a number of computer-simulated SDOF systems. In each case, a fourth-order Runge–Kutta scheme [209], is used to integrate the equations of motion. Where the excitation is random, it is generated by filtering a Gaussian white-noise sequence onto the range 0–200 Hz. The sampling frequency is 1000 Hz (except for the Van der Pol oscillator). The simulations provide a useful medium for discussing problems with the procedure and how they can be overcome.

7.2.3.1 A linear system

The first illustration concerns a linear system with equation of motion:

$$\ddot{y} + 40\dot{y} + 10^4 y = x(t). \tag{7.28}$$

The system was excited with a random excitation with rms 1.0 and 10 000 points of data were collected. The distribution of the points in the phase plane is shown in figure 7.1. This figure shows the first problem associated with the method. Not only are the points randomly distributed as discussed earlier, they

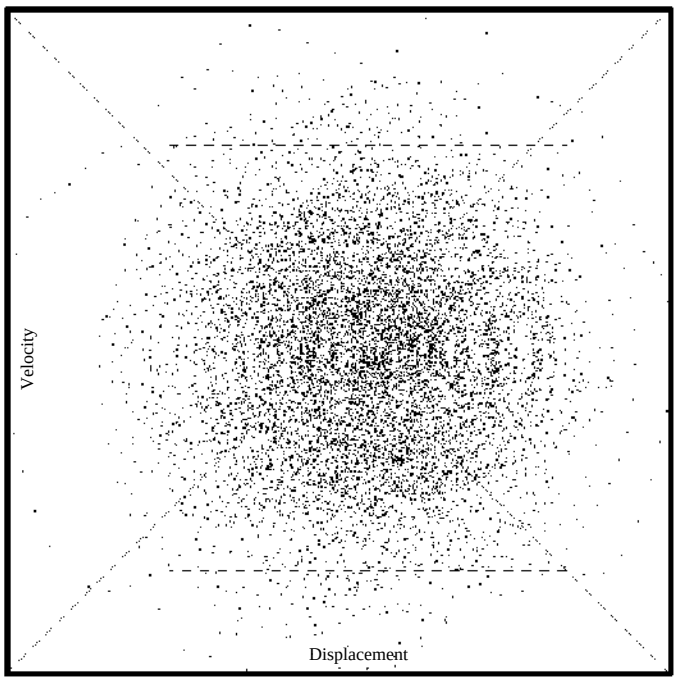


Figure 7.1. Distribution of points in the phase plane for a randomly excited linear system.

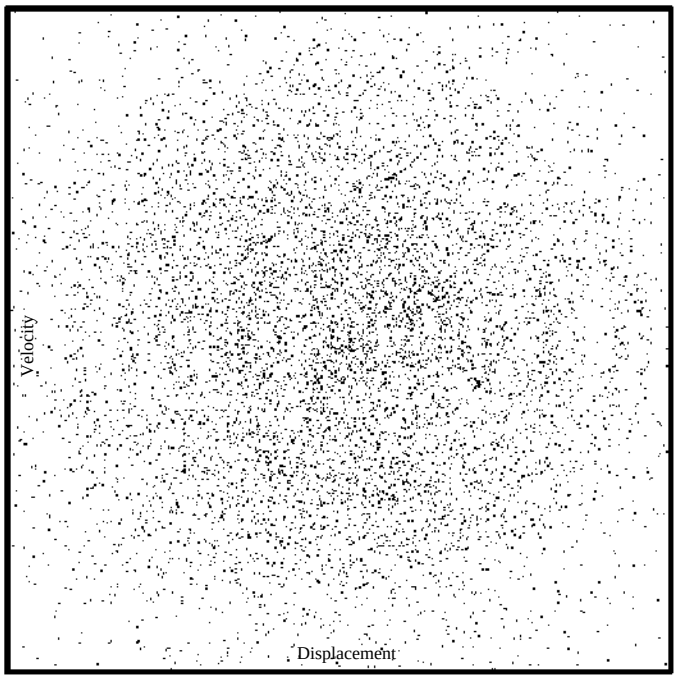


Figure 7.2. Zoomed region of figure 7.1.

have an irregular coverage or density. The data are mainly concentrated in an elliptical region (this appears circular as a result of the normalization imposed by plotting on a square) centred on the equilibrium. There are no data in the corners of the rectangle $[y_{\min}, y_{\max}] \times [\dot{y}_{\min}, \dot{y}_{\max}]$. The problem there is that the interpolation procedure can only estimate a value at a point surrounded by data, it cannot *extrapolate*. This is not particularly serious for the linear system data under investigation, as the interpolation procedure reproduces a linear or quadratic rate of growth away from the data. However, it will prove a serious problem with nonlinear data governed by functions of higher order than quadratic.

The solution to the problem adopted here is very straightforward, although it does involve a little wastage. As shown in figure 7.1, one can choose a rectangular sub-region of the phase plane which is more uniformly covered by data and carry out the analysis on this subset. (There is, of course, a subsequent renormalization of the data, which changes the ζ and $\dot{\zeta}$ transformations; however, the necessary algebra is straightforward.) The main caveat concerns the fact that the data lost correspond to the highest observed displacements and velocities. The experimenter must take care that the system is adequately excited even on the sub-region used for identification, otherwise there is a danger of concentrating on data which is nominally linear. The reduced data set in the case of the linear system is

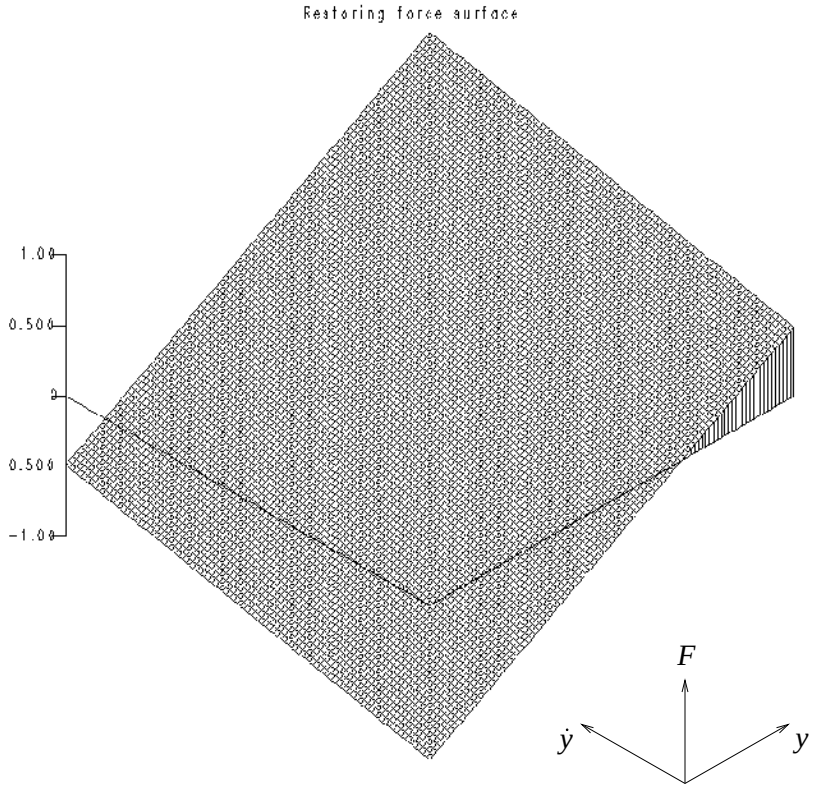


Figure 7.3. Identified restoring force surface for the linear system.

shown in figure 7.2, the coverage of the rectangle is more uniform.

Figure 7.3 shows the restoring force surface over the reduced region of the phase space as produced using C^1 natural neighbour interpolation. A perfect planar surface is obtained as required. The smoothness is due to the fact that the data are noise-free. Some of the consequences of measurement noise will be discussed later (in appendix I). Note that the data used here, i.e. displacement, velocity acceleration and force were all available from the simulation. Even if the acceleration and force could be obtained without error, the other data would usually be obtained by numerical integration and this process is approximate. Again, the consequences of this fact are investigated later. Using the data from the interpolation grid, the Chebyshev model coefficients are obtained with ease using (7.17). The results are given in table 7.1 together with the expected results obtained using theory given in appendix H.

The estimated coefficients show good agreement with the exact results. The

Table 7.1. Chebyshev coefficients for model of linear system.

Coefficient	Exact	Estimated	% Error
C^{ζ}_{00}	−0.0050	0.0103	−1840.9
C^{ζ}_{01}	0.3007	0.3004	−0.10
C^{ζ}_{10}	0.7899	0.7895	−0.06
C^{ζ}_{11}	0.0000	0.0218	—

Table 7.2. Model errors for various Chebyshev models of the linear system.

	0	1	2	3
0	100.05	87.38	87.43	87.43
1	12.71	0.07	0.11	0.12
2	12.90	0.29	0.32	0.33
3	12.90	0.28	0.32	0.33

only apparent exception is C^{ζ}_{00}. In fact a significance analysis would show that the coefficient can, in fact, be neglected. This will become apparent when the model predictions are shown a little later. This analysis assumes that the correct polynomial orders for the expansion are known. As this may not be the case, it is an advantage of the Chebyshev expansion that the initial model may be deliberately overfitted. The errors for the submodels can be evaluated and the optimum model can be selected. The coefficients of the optimal sub-model need not be re-evaluated because of the orthogonality discussed earlier. To illustrate this, a $(3,3)$ Chebyshev model was estimated and the MSE for the force surface was computed in each case (recall the definition of MSE from (6.108)). The results are given in table 7.2.

As expected the minimum error is for the $(1,1)$ model. Note that the addition of further terms is not guaranteed to lower the error. This is because, although the Chebyshev approximation is a least-squares procedure (as shown in appendix H), it is not implemented here as such. The model errors for overfitted models will generally fluctuate within some small interval above the minimum. Figure 7.4 shows a comparison between the force surface from the interpolation and that regenerated from the $(1,1)$ Chebyshev model. The difference is negligible. Although this comparison gives a good indication of the model, the final arbiter should be the error in reproducing the time data. In order to find this, the original Runge–Kutta simulation was repeated with the restoring force from the Chebyshev model. The results of comparing the displacement signal obtained

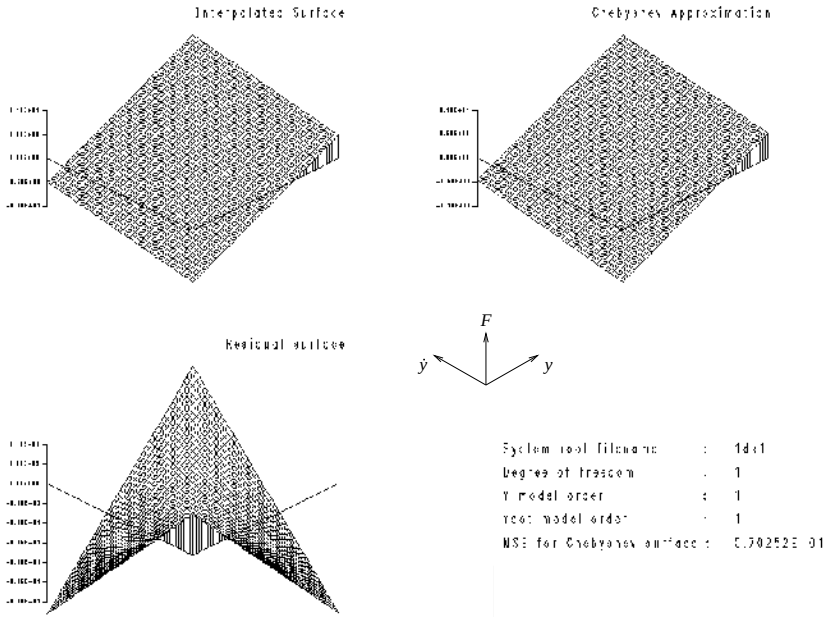


Figure 7.4. Comparison of the linear system Chebyshev model with the restoring force surface from interpolation.

with the exact signal is shown in figure 7.5. The MSE is 0.339 indicating excellent agreement.

One disadvantage of the method is that the model is unphysical, the coefficients obtained for the expansion do not directly yield information about the damping and stiffness of the structure. However, in the case of simple expansions (see appendix H), it is possible to reconstruct the ordinary polynomial coefficients. In the case of the linear system model, the results are

$$\hat{f}(y,\dot{y}) = 39.96\dot{y} + 9994.5y \tag{7.29}$$

which shows excellent agreement with the exact values in (7.28). Note that the conversion-back-to-a-physical-model generates constant and $y\dot{y}$ terms also which should not occur. These have been neglected here because of their low significance as witnessed by the model error. Note that there is a systematic means for estimating the significance of terms described in the last chapter. The significance factor would be particularly effective in the Chebyshev basis because the polynomials are orthogonal and therefore uncorrelated.

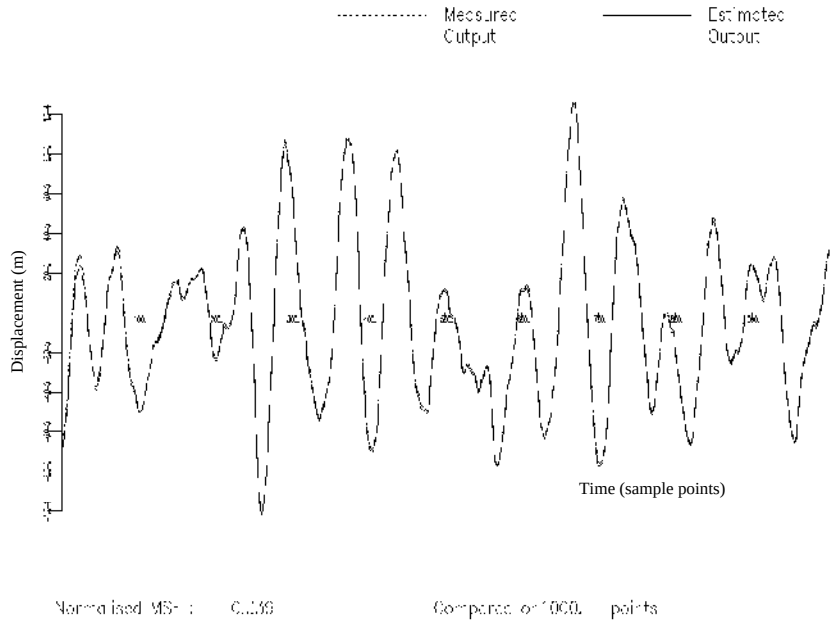


Figure 7.5. Comparison of measured response with that predicted by the linear Chebyshev model for the linear system.

7.2.3.2 A Van der Pol oscillator

This example is the first nonlinear system, a Van der Pol oscillator (vdpo) with the equation of motion,

$$\ddot{y} + 0.2(y^2 - 1)\dot{y} + y = 10\sin\left(\frac{t^2}{200}\right). \tag{7.30}$$

10 000 points were simulated with a sampling frequency of 10 Hz. The chirp excitation ranges from 0–10 rad s^{-1} over the period of simulation. The phase trajectory is shown in figure 7.6. In the early stages, the behaviour is very regular. However, as the trajectory spirals inward, it eventually reaches the region $y^2 < 1$, where the effective linear damping is negative. At this point, there is a transition to a very irregular motion. This behaviour will become important later when comparisons are made between the model and the true displacements. The distribution of points in the phase plane is shown in figure 7.7. Because of the particular excitation used, coverage of the plane is restricted to be within an envelope specified by a low-frequency periodic orbit (or *limit cycle*). There are no data whatsoever in the corners of the sampling rectangle. This is very serious in this case, because the force surface grows like y^3 on the diagonals $y = \pm\dot{y}$

If the natural neighbour method is used on the full data set, the force surface

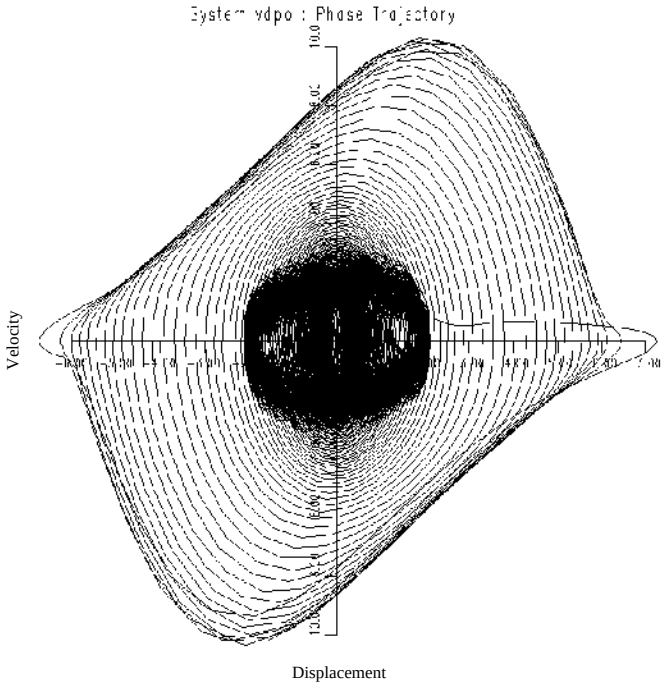


Figure 7.6. Phase trajectory for the Van der Pol oscillator (vdpo) excited by a chirp signal rising in frequency.

shown in figure 7.8 results. The surface is smooth, but not 'sharp' enough in the corners, and a comparison with the exact surface (figure 7.9) gives a MSE of 30.8%. The solution is described earlier, the data for modelling are chosen from a rectangular sub-region (indicated by broken lines in figure 7.7). The resulting interpolated surface is given in figure 7.10. This surface gave a comparison error with the exact surface of 0.04%, which is negligible.

The coefficients for the Chebyshev model and their errors are given in table 7.3.

Some of the results are very good. In fact, the inaccurate coefficients are actually not significant, again this will be clear from the model comparisons. The comparison between the reconstructed force surface and the exact surface is given in figure 7.10. The comparison MSE is 0.13. If data from the system are regenerated from a Runge–Kutta scheme using the Chebyshev model, the initial agreement with the exact data is excellent (figure 7.11—showing the first 1000 points). However, the MSE for the comparison over 10 000 points is 30.6, which is rather poor. The explanation is that the reconstructed data makes the transition to an irregular motion rather earlier than the exact data as shown in figure 7.12 (which shows a later window of 1000 points). There is an important point to be

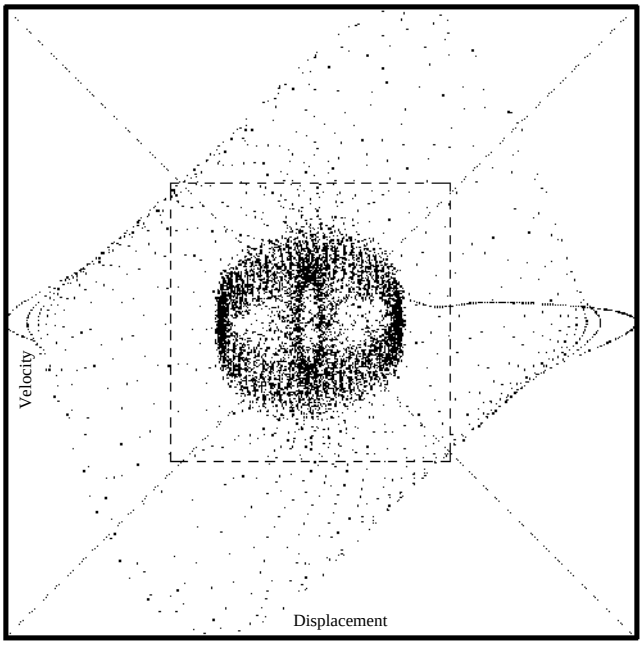


Figure 7.7. Distribution of sample points in the phase plane for figure 7.6.

made here, if the behaviour of the system is very sensitive to initial conditions or coefficient values, it might be impossible to reproduce the time response even though the representation of the internal forces is very good.

7.2.3.3 Piecewise linear systems

This system has the equation of motion

$$\ddot{y} + 20\dot{y} + 10^4 y = x(t) \tag{7.31}$$

in the interval $y \in [-0.001, 0.001]$. Outside this interval, the stiffness is multiplied by a factor of 11. This type of nonlinearity presents problems for parametric approaches, because the position of the discontinuities in the force surface (at $y = \pm 0.001$) do not enter the equations in a sensible way for *linear-in-the-parameters* least-squares estimation. Nonetheless, the restoring force surface (RFS) approach works because it is non-parametric. Working methods are needed for systems of this type because they commonly occur in practice via clearances in systems.

The data were generated by Runge–Kutta integration with a sampling frequency of 10 kHz and 10 000 samples were collected. The excitation was white noise with rms 100.0 band-limited onto the interval $[0, 2000]$ Hz. After

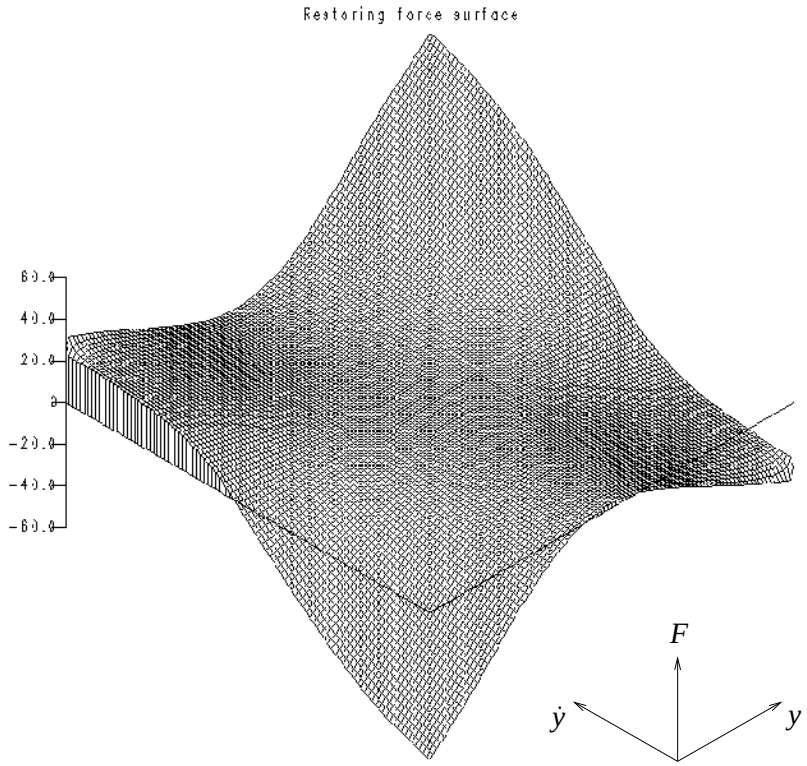


Figure 7.8. Interpolated restoring force surface for the Van der Pol oscillator (vdpo) using all the data.

concentrating on a region of the phase plane covered well by data, a force surface of the form shown in figure 7.13 is obtained. The piecewise linear nature is very clear. Comparison with the true surface gives excellent agreement.

Problems start to occur if one proceeds with the Masri–Caughey procedure and tries to fit a Chebyshev-series model. This is simply because the discontinuities in the surface are very difficult to model using inherently smooth polynomial terms. A ninth-order polynomial fit is shown in figure 7.14 in comparison with the real surface. Despite the high order, the model surface is far from perfect. In fact, when the model was used to predict the displacements using the measured force, the result diverged. The reason for this divergence is simple. The polynomial approximation is not constrained to be physically sensible, i.e. the requirement of a best fit, may fix the higher-order stiffness coefficients *negative*. When the displacements are then estimated on the full data

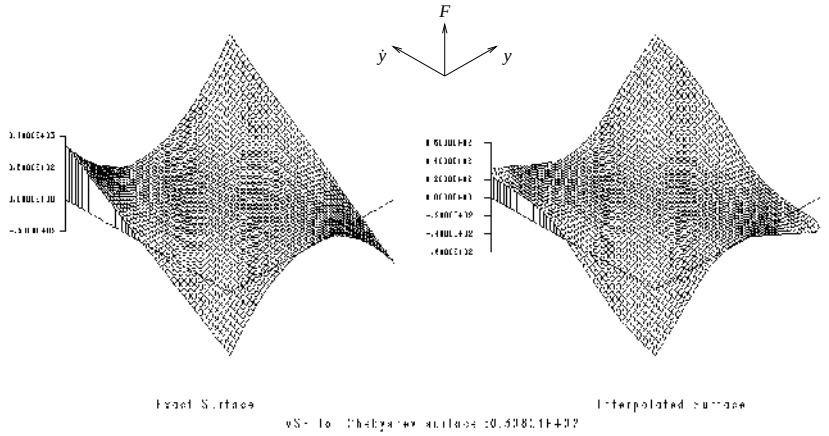


Figure 7.9. Comparison of the restoring force surface in figure 7.8 with the exact surface.

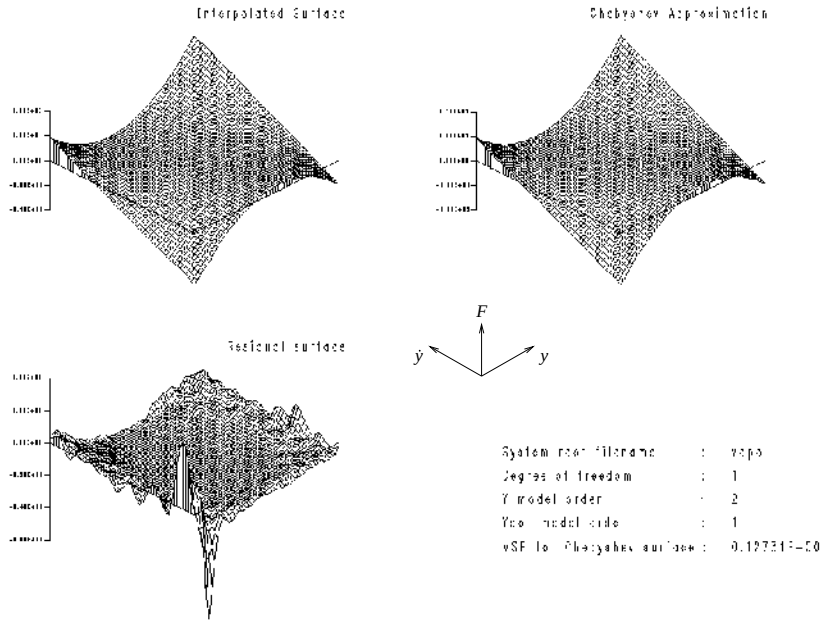


Figure 7.10. Chebyshev model for the Van der Pol oscillator (vdpo) based on a restoring force surface constructed over a restricted data set.

set rather than the reduced data set, it is possible to obtain negative *stiffness forces* and instability results. This is an important issue: if non-polynomial systems are approximated by polynomials, they are only valid over the data used for

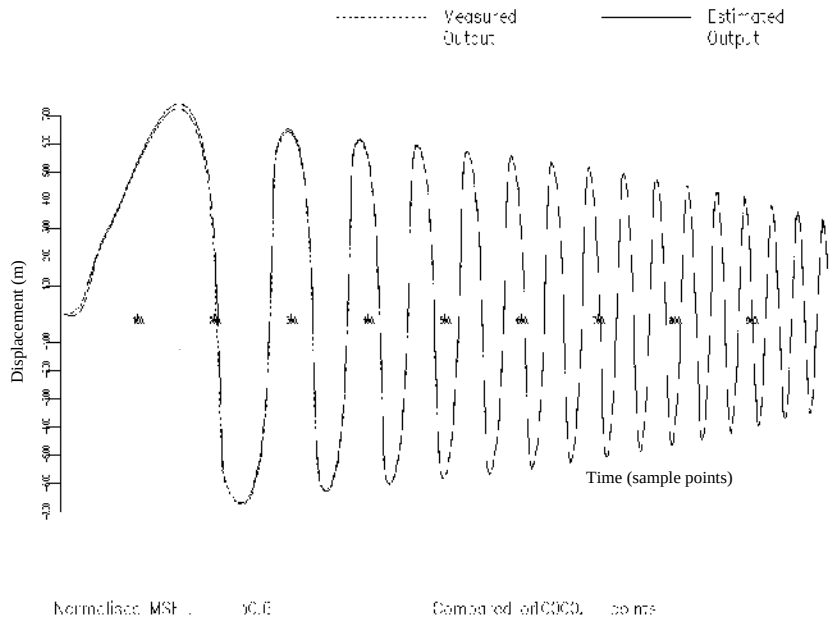


Figure 7.11. Comparison of the measured Van der Pol oscillator (vdpo) response with predictions from the nonlinear Chebyshev model. The early part of the record.

Table 7.3. Chebyshev coefficients for model of linear system.

Coefficient	Exact	Estimated	% Error
C^{ζ}_{00}	0.003	0.078	1994.7
C^{ζ}_{01}	3.441	3.413	−0.80
C^{ζ}_{10}	3.091	3.067	−0.79
C^{ζ}_{11}	0.043	0.082	88.9
C^{ζ}_{20}	−0.005	−0.050	878.7
c^{ζ}_{21}	4.351	4.289	−1.44

estimation—the *estimation set*; the identification is *input dependent*.

The difficulty in fitting a polynomial model increases with the severity of the discontinuity. The 'clearance' system above has a discontinuity in the *first derivative* of the stiffness force. In the commonly occurring situation where dry friction is present, the discontinuity may be in the force itself. An often used

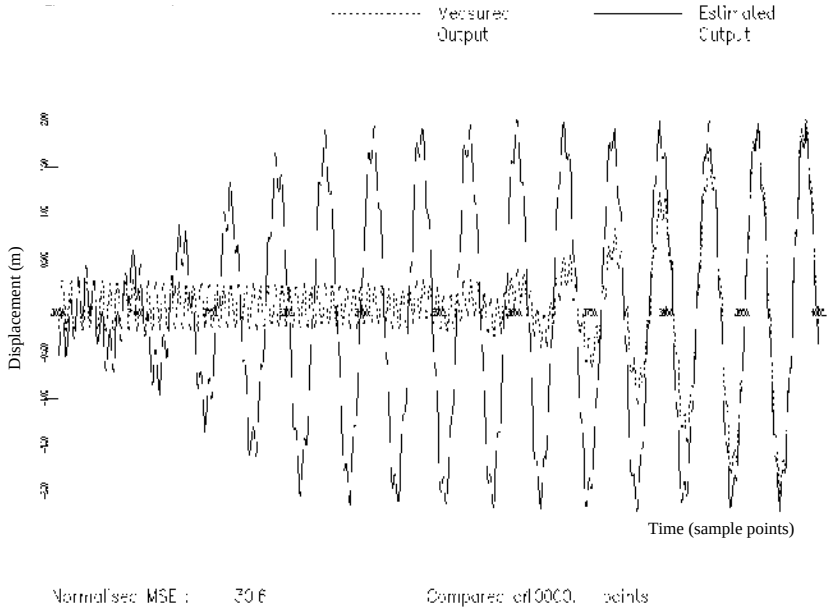


Figure 7.12. Comparison of the measured Van der Pol oscillator (vdpo) response with predictions from the Chebyshev model, a later part of the record.

approximation to dry friction is to add a damping term of the form $\mathrm{sgn}(\dot{y})$ [2]. To illustrate the analysis for such systems, data were simulated from an oscillator with equation of motion

$$\ddot{y} + 20\dot{y} + 10\,\mathrm{sgn}(\dot{y}) + 10^4 y = x(t) \tag{7.32}$$

in more or less the same fashion as before. When the C^1 restoring force surface was computed, the result was as shown in figure 7.15; a number of spikes are visible. These artifacts are the result of the estimation of gradients for the interpolation. Two points on either side of the discontinuity can yield an arbitrarily high estimated gradient depending on their proximity. When the gradient terms (first order in the Taylor expansion) are added to the force estimate, the interpolant can be seriously in error. The way around the problem is to use a C^0 interpolant which does not need gradient information. The lower-order surface for the same data is shown in figure 7.16 and the spikes are absent. If one is concerned about lack of accuracy in regions of low data density, a hybrid

[2] Friction is actually a lot more complicated than this. A brief but good review of real friction forces can be found in [183]. This paper is also interesting for proposing a friction model where the force depends on the acceleration as well as the velocity. Because there are three independent states in such a model, it cannot be visualized using RFS methods.

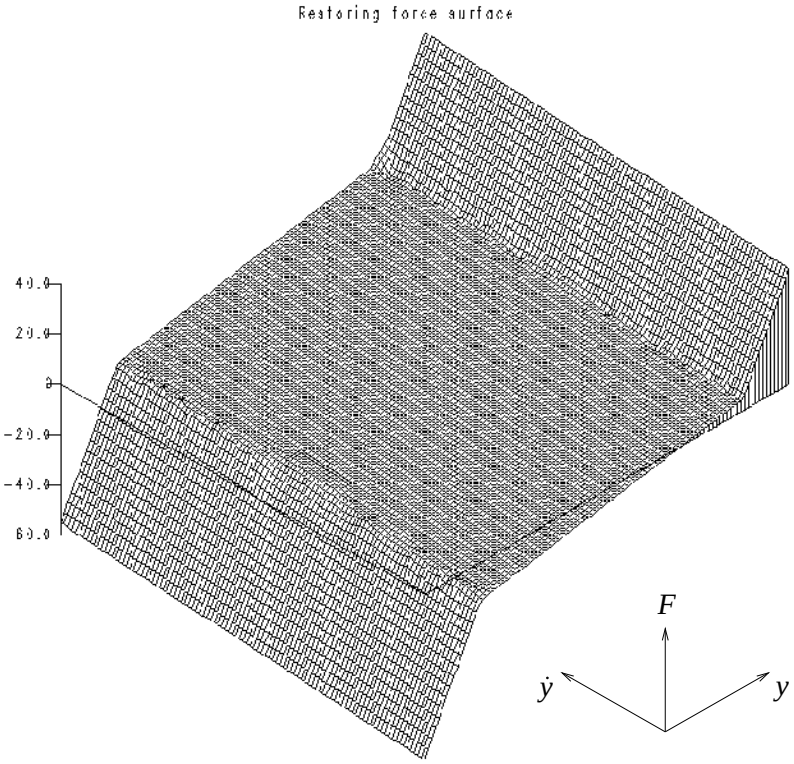


Figure 7.13. Identified restoring force surface for data from a piecewise linear system.

approach can be used where the surface is C^0 in the region of the discontinuity and C^1 elsewhere.

Because the discontinuity is so severe for Coulomb friction, it is even more difficult to produce a polynomial model. The ninth-order model for the surface is shown in figure 7.17. The reproduction of the main feature of the surface is terrible. When the model was used to reconstruct the response to the measured force, the prediction was surprisingly good but diverged in places where badly modelled areas of the phase plane are explored (figure 7.18). These two examples illustrate the fact that polynomial models may or may not work for discontinuous systems, it depends on the leading terms in the polynomial approximations whether the model is stable or not.

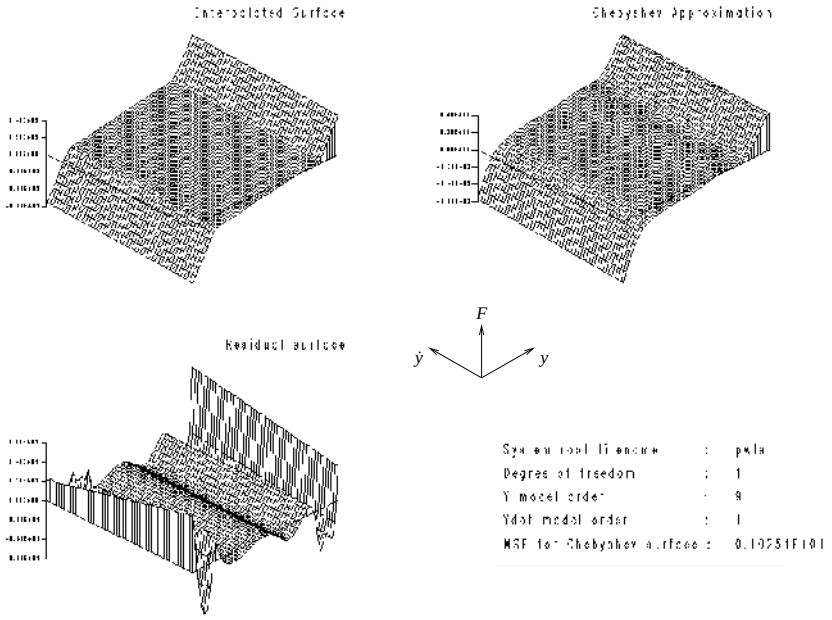


Figure 7.14. Comparison of the Chebyshev model with the interpolated restoring force surface for the piecewise linear system.

7.3 The Masri–Caughey method for MDOF systems

7.3.1 Basic theory

The Masri–Caughey approach would be rather limited if it only applied to SDOF systems. In fact, the extension to MDOF is fairly straightforward and is predominantly a problem of book-keeping. As usual for MDOF analysis, vectors and matrices will prove necessary.

One begins, as before, with Newton's second law

$$[m]\{\ddot{y}\} + \{f(y,\dot{y})\} = \{x(t)\} \tag{7.33}$$

where $[m]$ is the physical-mass matrix and $\{f\}$ is the vector of (possibly) nonlinear restoring forces. It is assumed implicitly, that a lumped-mass model with a finite number of degrees of freedom is appropriate. The number of DOF will be taken as N. The lumped-mass assumption will usually be justified in practice by the fact that band-limited excitations will be used and only a finite number of modes will be excited.

The simplest possible situation is where the system is linear, i.e.

$$[m]\{\ddot{y}\} + [c]\{\dot{y}\} + [k]\{y\} = \{x\} \tag{7.34}$$

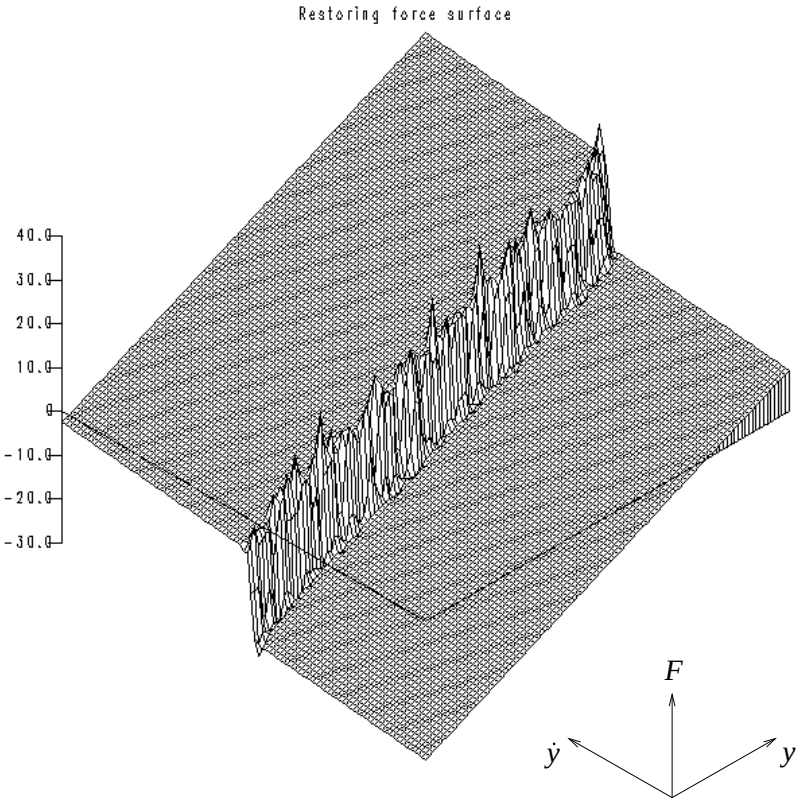


Figure 7.15. The identified restoring force surface for data from a Coulomb friction system: C^1 interpolation.

and the change to normal coordinates

$$\{y\} = [\psi]\{u\} \tag{7.35}$$

decouples the system into N SDOF systems

$$m_i\ddot{u}_i + c_i\dot{u}_i + k_i u_i = p_i, \quad i = 1, \ldots, N \tag{7.36}$$

as described in chapter 1. In this case, each system can be treated by the SDOF Masri–Caughey approach.

The full nonlinear system (7.33) is much more interesting. In general, there is no transformation of variables—linear or nonlinear—which will decouple the system. However, the MDOF Masri–Caughey approach assumes that the transformation to linear normal coordinates (i.e. the normal coordinates of the

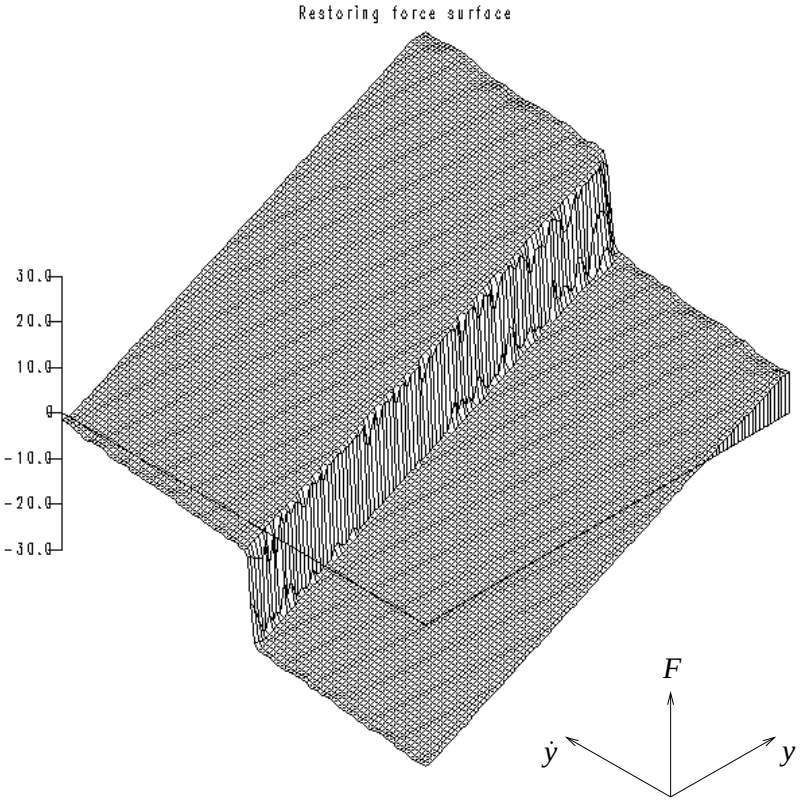


Figure 7.16. The identified restoring force surface for data from a Coulomb friction system: C^0 interpolation.

underlying linear system) will nonetheless yield a worthwhile simplification. Equation (7.33) becomes

$$[M]\{\ddot{u}\} + \{h(u, \dot{u})\} = \{p(t)\} \tag{7.37}$$

where $\{h\} = [\psi]^{\mathrm{T}}\{f\}$. As before, the method assumes that the $\{y\}$, $\{\dot{y}\}$ and $\{\ddot{y}\}$ data are available. However, in the MDOF case, estimates of the mass matrix $[M]$ and modal matrix $[\psi]$ are clearly needed. For the moment assume that this is the case; modal analysis at low excitation can provide $[\psi]$ and there are numerous, well-documented means of estimating $[m]$ [11]. The restoring force vector is obtained from

$$\{h\} = \{p\} - [M]\{\ddot{u}\} = [\psi]^{\mathrm{T}}(\{x\} - [m]\{\ddot{y}\}) \tag{7.38}$$

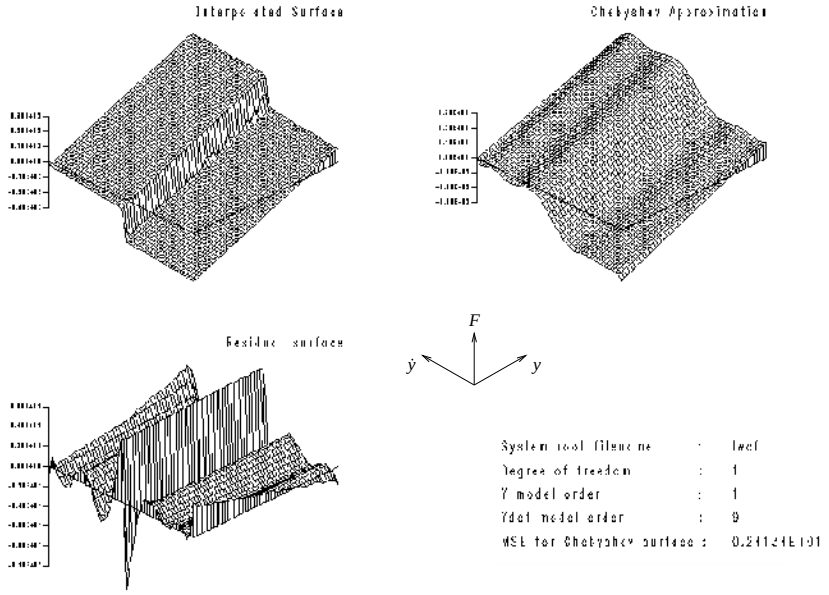


Figure 7.17. Comparison of the Chebyshev model with the interpolated surface for the Coulomb friction system.

and the ith component is simply

$$h_i = p_i - m_i \ddot{u}_i. \tag{7.39}$$

These equations obviously hold at each sampling instant, but as an aid to clarity, time instant labels will be suppressed in the following. Equation (7.39) is formally no more complicated than (7.4) in the SDOF case. Unfortunately, this time h_i is not only a function of u_i and $\dot{u}_i$. In general, h_i can and will depend on *all* u_i and $\dot{u}_i$ for $i = 1, \ldots, N$. This eliminates the possibility of a simple restoring force surface for each modal degree of freedom. However, as a first approximation, it can be assumed that the dominant contribution to h_i is from u_i and $\dot{u}_i$. In exactly the same way as for SDOF systems, one can represent h_i as a surface over the $(u_i, \dot{u}_i)$ plane and fit a Chebyshev model of the form

$$h_i^{(1)}(u_i, \dot{u}_i) = \sum_m \sum_n {}^1C_{mn}^{(i)} T_m(u_i) T_n(\dot{u}_i). \tag{7.40}$$

(For the sake of clarity, the labels for the maps ζ which carry the data onto the squares $[-1, 1] \times [-1, 1]$ have been omitted. However, these transformations are still necessary in order to apply formulae of the form (7.13) to estimate the coefficients.) This expansion will represent dependence of the force on terms such

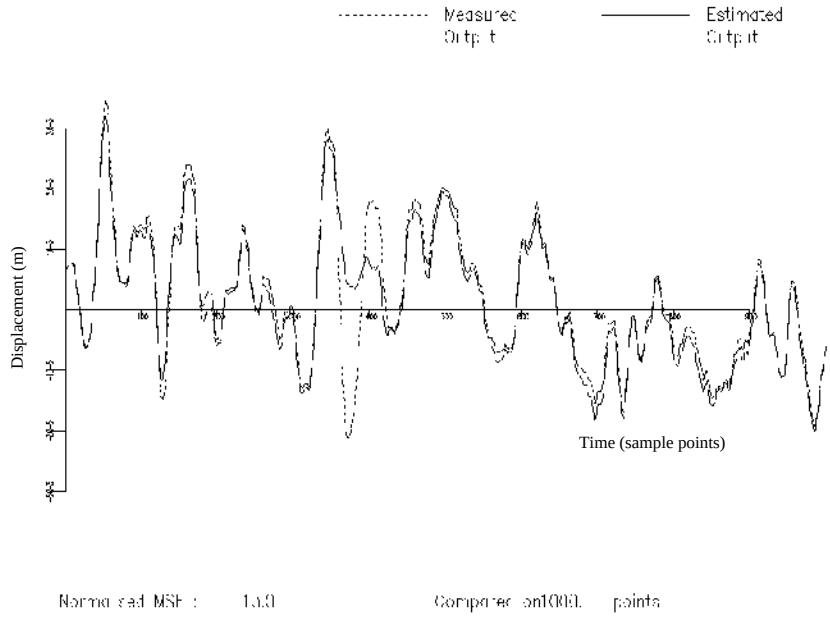


Figure 7.18. Comparison of the measured Coulomb friction system response with predictions from the Chebyshev model.

as $u_i^\alpha \dot{u}_i^\beta$. To include the effects of modal coupling due to the nonlinearity, terms such as $u_i^\alpha u_j^\beta$ are needed with $i \neq j$. Further, if the nonlinearity is in the damping, the model will need terms of the form $\dot{u}_i^\alpha \dot{u}_j^\beta$. Finally, consideration of the Van der Pol oscillator suggests the need for terms such as $u_i^\alpha \dot{u}_j^\beta$. The model for the MDOF restoring force is clearly much more complex than its SDOF counterpart. There are essentially two methods for constructing the required multi-mode model. The first is to fit all terms in the model in one go, but this violates the fundamental property of the Masri–Caughey procedure which allows visualization. The second method, the one adopted by Masri *et al* [175], proceeds as follows.

After fitting the model (7.40), it is necessary to reorganize the data so that the other model components can be obtained. First, the residual term $r_i^{(1)}$ is computed:

$$r_i^{(1)}(\{u\},\{\dot{u}\}) = h_i(\{u\},\{\dot{u}\}) - h_i^{(1)}(u_i,\dot{u}_i). \tag{7.41}$$

This is a time series again, so one can successively order the forces over the (u_i, u_j)-planes and a sequence of models can be formed

$$h_i^{(2)}(\{u\}) = \sum_m \sum_n {}^2C_{mn}^{(i)(j)} T_m(u_i) T_n(u_j) \approx r_i^{(1)}(\{u\},\{\dot{u}\}) \tag{7.42}$$

including only those modes which interact with the ith mode—of course this may be all of them. Velocity–velocity coupling is accounted for in the same way, the residual

$$r_i^{(2)}(\{u\},\{\dot{u}\}) = r_i^{(1)}(\{u\},\{\dot{u}\}) - h_i^{(2)}(\{u\}) \tag{7.43}$$

is formed and yields the model

$$h_i^{(3)}(\{\dot{u}\}) = \sum_m \sum_n {}^3C_{mn}^{(i)(j)} T_m(\dot{u}_i) T_n(\dot{u}_j) \approx r_i^{(2)}(\{u\},\{\dot{u}\}). \tag{7.44}$$

Finally, the displacement–velocity coupling is obtained from the iteration

$$r_i^{(3)}(\{u\},\{\dot{u}\}) = r_i^{(2)}(\{u\},\{\dot{u}\}) - h_i^{(3)}(\{u\}) \tag{7.45}$$

and

$$h_i^{(4)}(\{u\},\{\dot{u}\}) = \sum_m \sum_n {}^4C_{mn}^{(i)(j)} T_m(u_i) T_n(\dot{u}_j). \tag{7.46}$$

A side-effect of this rather complicated process is that one does not require a proportionality constraint on the damping. Depending on the extent of the modal coupling, the approach will require many expansions.

7.3.2 Some examples

The first example of an MDOF system is a 2DOF oscillator with a continuous stiffness nonlinearity, the equations of motion are

$$\begin{pmatrix} \ddot{y}_1 \\ \ddot{y}_2 \end{pmatrix} + 20 \begin{pmatrix} \dot{y}_1 \\ \dot{y}_2 \end{pmatrix} + 10^4 \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + 5 \times 10^9 \begin{pmatrix} y_1^3 \\ 0 \end{pmatrix} = \begin{pmatrix} x \\ 0 \end{pmatrix}. \tag{7.47}$$

As usual, this was simulated with a Runge–Kutta routine and an excitation with rms 150.0 was used. The modal matrix for the underlying linear system is

$$[\psi] = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \tag{7.48}$$

so the equations of motion in modal coordinates are

$$\ddot{u}_1 + c\dot{u}_1 + ku_1 + \frac{1}{4}k_3(u_1+u_2)^3 = \frac{1}{\sqrt{2}}x \tag{7.49}$$

and

$$\ddot{u}_2 + c\dot{u}_2 + ku_2 + \frac{1}{4}k_3(u_1+u_2)^3 = \frac{1}{\sqrt{2}}x \tag{7.50}$$

with $c = 20.0$ Ns m^{-1}, $k = 10^4$ N m^{-1} and $k_3 = 5 \times 10^9$ N m^{-3}. The identification proceeds as follows:

(1) Assemble the data for the $h_1^{(1)}(u_1,\dot{u}_1)$ expansion. The distribution of the data in the $(u_1,\dot{u}_1)$ plane is given in figure 7.19 with the reduced data set in

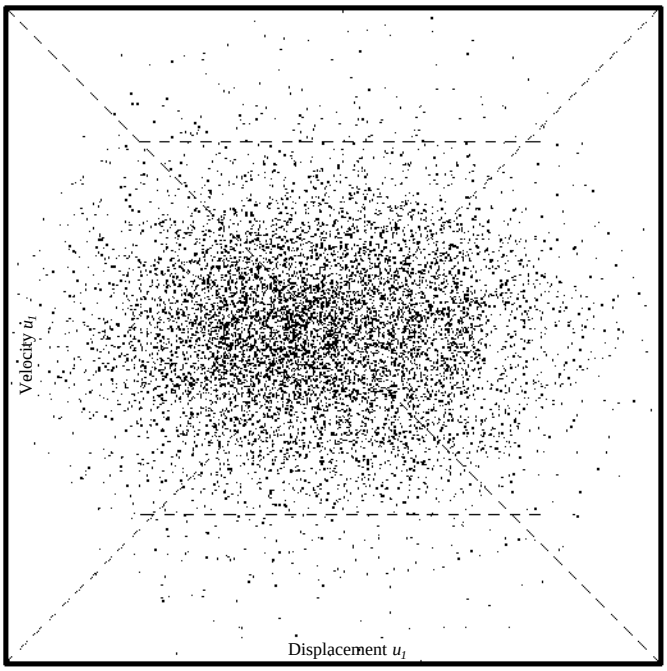


Figure 7.19. Data selected from the $(u_1, \dot{u}_1)$-plane for the interpolation of the force surface $h_1^{(1)}(u_1, \dot{u}_1)$. The system is a 2DOF cubic oscillator.

the rectangle indicated by broken lines. The interpolated surface is shown in figure 7.20 and appears to be very noisy; fortunately, the explanation is quite simple. The force component h_1 actually depends on all four state variables for the system

$$h_1 = c\dot{u}_1 + ku_1 + \tfrac{1}{4}k_3(u_1^3 + 3u_1^2u_2 + u_1u_2^2 + u_2^3). \tag{7.51}$$

However, only u_1 and $\dot{u}_1$ have been ordered to form the surface. Because the excitation is random, the force at a given point $q = (u_{1q}, \dot{u}_{1q})$ is formed from two components: a deterministic part comprising

$$h_{1\mathrm{d}} = c\dot{u}_1 + ku_1 + \tfrac{1}{4}k_3u_1^3 \tag{7.52}$$

and a random part

$$h_{1\mathrm{r}} = \tfrac{1}{4}k_3(3u_1^2X + 3u_1X^2 + X^3) \tag{7.53}$$

where X is a random variable with probability density function $P_q(X) = \alpha P_j(u_{1q}, X)$. P_j is the overall joint probability density function for u_1 and u_2 and α is a normalization constant.

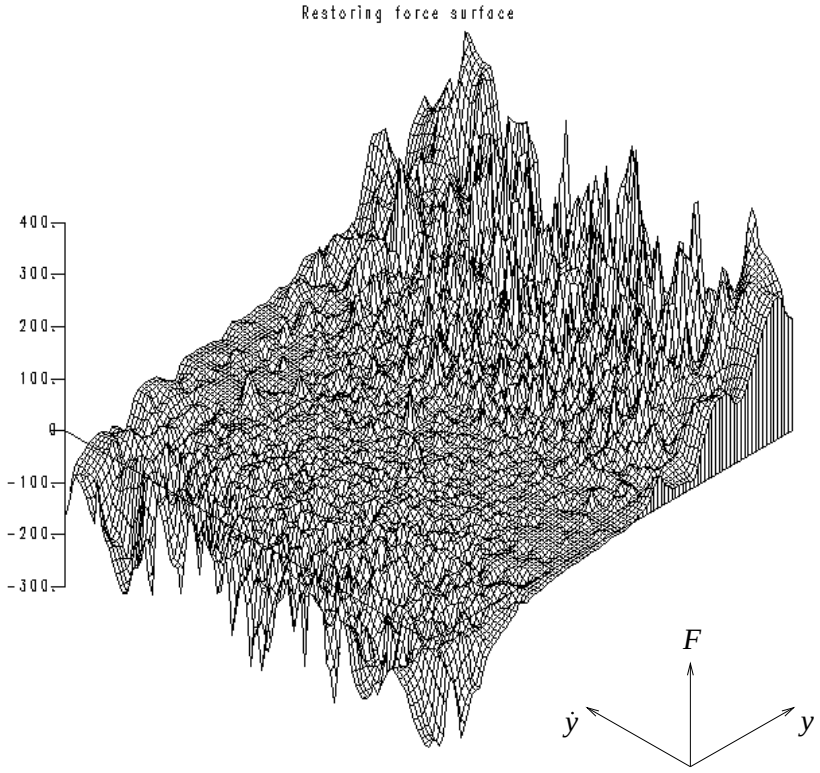


Figure 7.20. Interpolated force surface $h_1^{(1)}(u_1, \dot{u}_1)$ for the 2DOF cubic oscillator.

(2) Fit a Chebyshev series to the interpolated surface (figure 7.21). In this case, the optimum model order was $(3, 1)$ and this was reflected in the model errors. Subtract the model from the time data for h_1 to form the residual time series $r_1^{(1)}$.

(3) Assemble the residual force data over the (u_1, u_2) plane for the $h_1^{(2)}$ expansion. The distribution of the data in this plane is shown in figure 7.22. Note that the variables are strongly correlated. Unfortunately, this means that the model estimated in step 1 will be biased because the first model expansion will include a component dependent on u_2. One can immediately see this from the surface which still appears noisy. However, at this stage one can correct for errors in the u_1 dependence. The interpolated surface is formed as in figure 7.23 and the Chebyshev model coefficients $C2_{mn}^{(i)(j)}$ are identified—in this case the necessary model order is $(3, 3)$ (figure 7.24).

(4) Carry out steps (1) to (3) for the h_2 component.

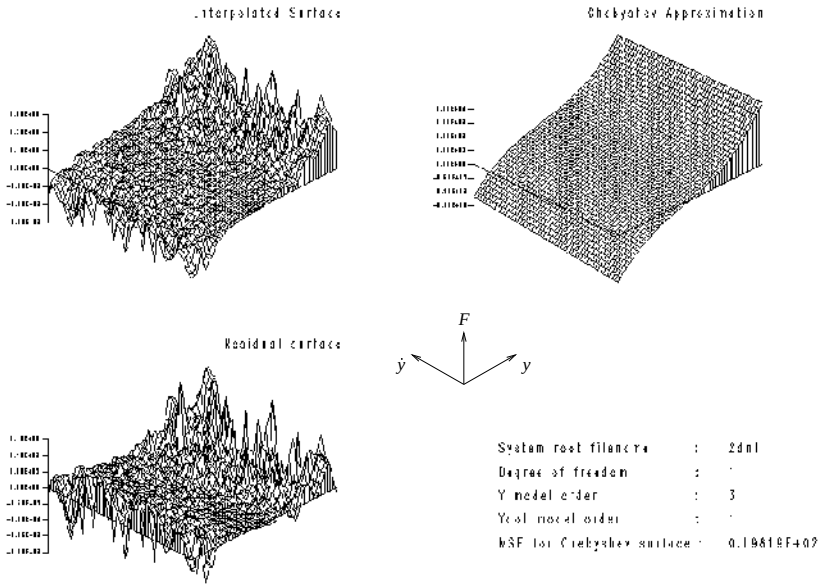


Figure 7.21. Chebyshev model fit of order $(3, 1)$ to the surface in figure 7.20.

If the bias in this procedure is a matter for concern, these steps can be iterated until all dependencies have been properly accounted for. Unfortunately, this renders the process extremely time-consuming.

In order to see how well the procedure works, the displacements u_1 and u_2 can be reconstructed when the Chebyshev model is forced by the measured excitation $x(t)$. The results are shown in figure 7.25. The results are passable; bias has clearly been a problem. The reconstruction from a linear model actually diverges because it has estimated negative damping (figure 7.26).

The second illustration here is for a 3DOF system with a discontinuous nonlinearity as described by the equations of motion:

$$\begin{pmatrix} \ddot{y}_1 \\ \ddot{y}_2 \\ \ddot{y}_3 \end{pmatrix} + 20 \begin{pmatrix} \dot{y}_1 \\ \dot{y}_2 \\ \dot{y}_3 \end{pmatrix} + 10^4 \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} + \begin{pmatrix} 0 \\ f_{nl} \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ x \\ 0 \end{pmatrix}. \tag{7.54}$$

The response was simulated with the same excitation as the 2DOF system. The nonlinear force was piecewise-linear with clearance 0.001 as shown in figure 7.27.

The identification was carried out using the steps described earlier. The formation of the resulting surfaces and expansions is illustrated in figures 7.28–7.35. The restoring force surface for h_2 is flat because the modal matrix for the

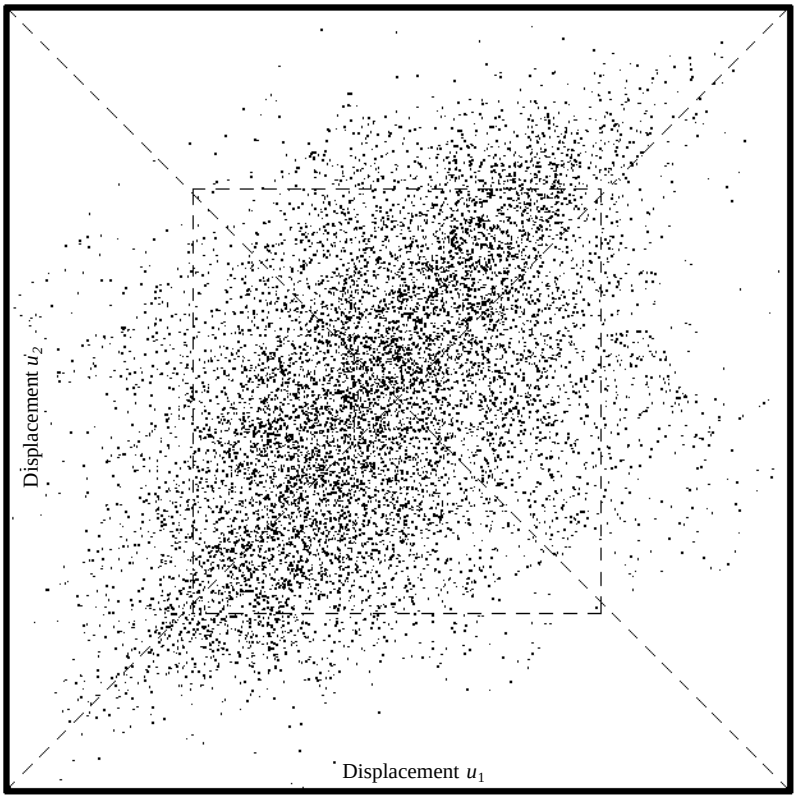


Figure 7.22. Data selected from the (u_1, u_2)-plane for the interpolation of the force surface $h_1^{(2)}(u_1, u_2)$. The system is a 2DOF cubic oscillator.

underlying linear system is

$$[\psi] = \frac{1}{2}\begin{pmatrix} 1 & \sqrt{2} & 1 \\ \sqrt{2} & 0 & \sqrt{2} \\ 1 & -\sqrt{2} & 1 \end{pmatrix} \tag{7.55}$$

and the nonlinear force does not appear in the equation for the second mode. This illustrates nicely one of the drawbacks to moving to a modal coordinate basis; the transformation shuffles the physical coordinates so that one cannot tell from the restoring forces *where* the nonlinearity might be.

Because of the 'noise' in the surfaces caused by interactions with other modes, there is no longer an option of using a C^1 interpolation. This is because two arbitrarily close points in the $(u_1, \dot{u}_1)$-plane might have quite large differences in the force values above them because of contributions from other modes. This means that the gradients will be overestimated as described before

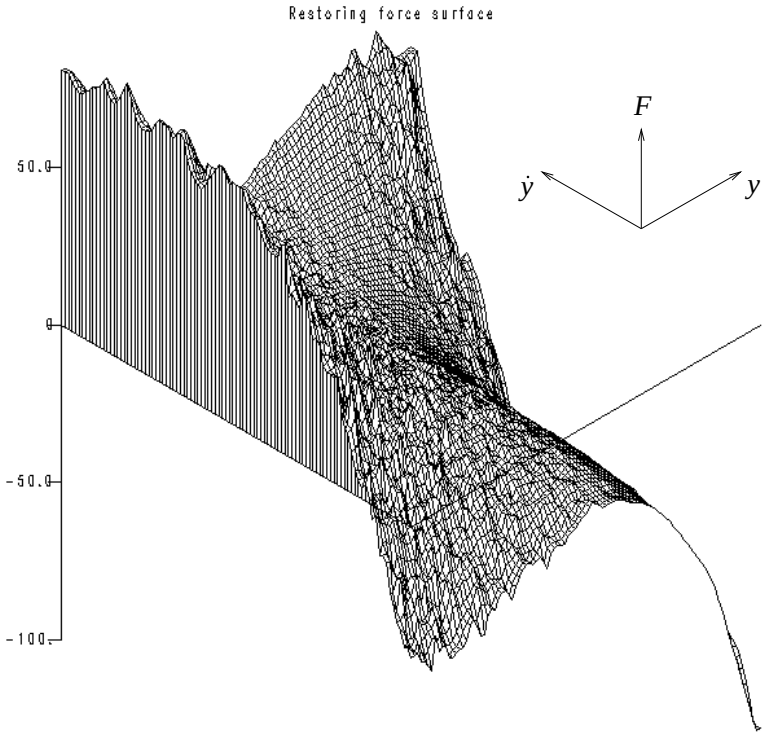


Figure 7.23. Interpolated force surface $h_1^{(2)}(u_1, u_2)$ for the 2DOF cubic oscillator.

and the interpolated surface will contain spurious peaks.

These examples show that the Masri–Caughey method is a potentially powerful means of identifying nearly arbitrary nonlinear systems. In their later work, Masri and Caughey adopted a scheme which made use of direct least-squares estimation to obtain the linear system matrices, while retaining the Chebyshev expansion approach for the nonlinear forces [176, 177]. The following sections discuss an approach based completely on direct least-squares methods which shows some advantages over the hybrid approach.

7.4 Direct parameter estimation for SDOF systems

7.4.1 Basic theory

Certain disadvantages of the Masri–Caughey procedure may already have become apparent: (i) it is time-consuming; (ii) there are many routes by which errors

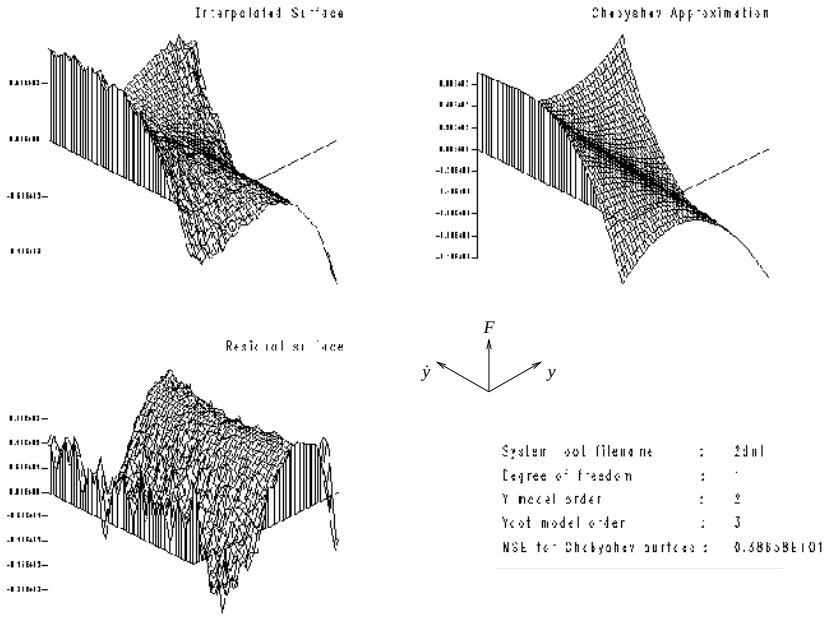


Figure 7.24. Chebyshev model fit of order $(3, 3)$ to the surface in figure 7.23.

accumulate; (iii) the restoring forces are expanded in terms of Chebyshev polynomials which obscures the physical meaning of the coefficients; and (iv) there are no confidence limits for the parameters estimated. The object of this section is to show an alternative approach. This will be termed direct parameter estimation (DPE) and is based on the simple least-squares estimation theory described in the previous chapter. It will be shown that the approach overcomes the problems described earlier.

Consider the SDOF Duffing oscillator

$$m\ddot{y} + c\dot{y} + ky + k_3 y^3 = x(t). \tag{7.56}$$

If the same data are assumed as for the Masri–Caughey procedure, namely samples of displacement y_i, velocity $\dot{y}_i$ and acceleration $\ddot{y}_i$ at N sampling instants i, one can obtain for the matrix least-squares problem:

$$\{Y\} = [A]\{\beta\} + \{\zeta\} \tag{7.57}$$

with $\{Y\} = (x_1, \ldots, x_{\mathrm{N}})^{\mathrm{T}}$, $\beta = (m, c, k, k_3)^{\mathrm{T}}$ and

$$[A] = \begin{pmatrix} \ddot{y}_1 & \dot{y}_1 & y_1 & y_1^3 \\ \vdots & \vdots & \vdots & \vdots \\ \ddot{y}_{\mathrm{N}} & \dot{y}_{\mathrm{N}} & y_{\mathrm{N}} & y_{\mathrm{N}}^3 \end{pmatrix}. \tag{7.58}$$

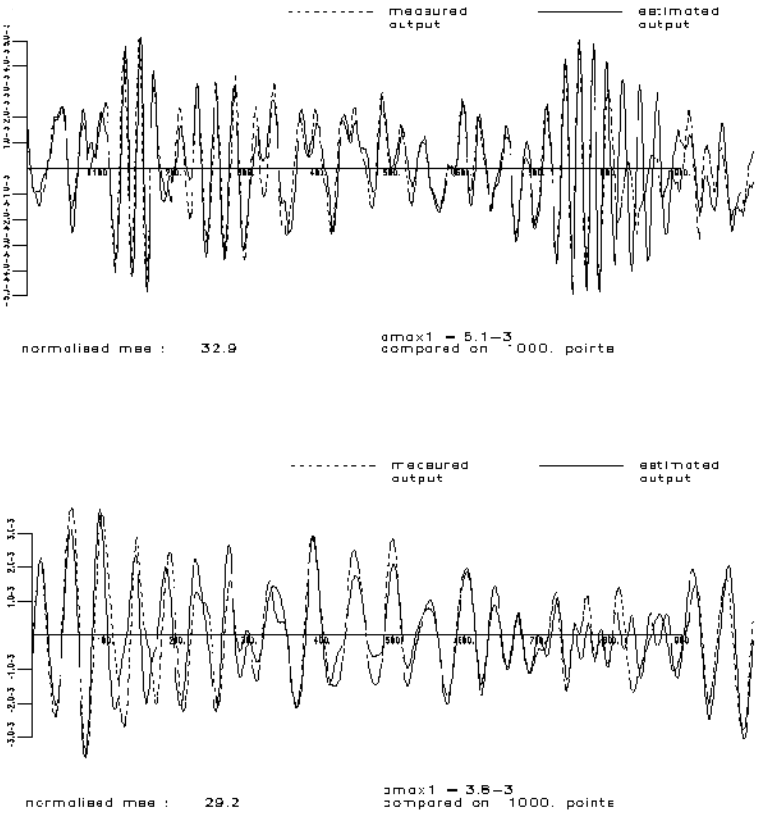


Figure 7.25. Comparison of measured data and that predicted by the Chebyshev model for the 2DOF cubic oscillator: nonlinear model with $h_1^{(1)}$, $h_1^{(2)}$, $h_2^{(1)}$ and $h_2^{(2)}$ used.

This equation (where measurement noise $\{\zeta\}$ has been accounted for) is formally identical to equation (6.14) which set up the estimation problem in discrete time. As a result, all the methods of solution discussed in chapter 6 apply, this time in order to estimate the continuous-time parameters m, c, k and k_3. Furthermore, the standard deviations of the parameter estimates follow directly from (6.30) so the confidence in the parameters is established.

In order to capture all possible dependencies, the general polynomial form

$$m\ddot{y} + \sum_{i=0}^{m}\sum_{j=0}^{n} C_{ij} y^i \dot{y}^j = x(t) \tag{7.59}$$

is adopted. Note that in this formulation, the mass is not singled out; it is estimated in exactly the same way as the other parameters. Significance factors for the model terms can be defined exactly as in (6.31).

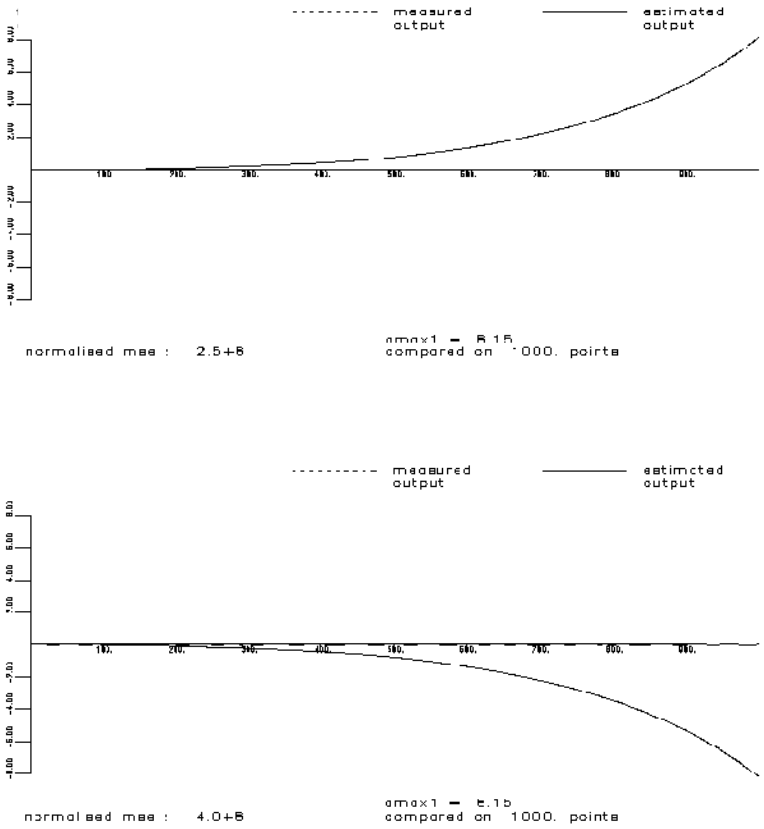


Figure 7.26. Comparison of measured data and that predicted by the Chebyshev model for the 2DOF cubic oscillator: linear model with $h_1^{(1)}$ and $h_2^{(1)}$ used.

If necessary, one can include in the model, basis functions for well-known nonlinearities, i.e. $\mathrm{sgn}(\dot{y})$ for friction. This was first observed in [9].

As an aside, note that there is no reason why a model of the form

$$m\ddot{y} + \sum_{i=0}^{m}\sum_{j=0}^{n} C_{ij} T_i(y) T_j(\dot{y}) = x(t) \tag{7.60}$$

should not be adopted, where T_k is the Chebyshev polynomial of order k. This means that DPE allows the determination of a Masri–Caughey-type model without having to obtain the coefficients from double integrals. In fact, the Chebyshev expansions are obtained much more quickly and with greater accuracy by this method.

To simplify matters, the MSE used for direct least-squares is based on the

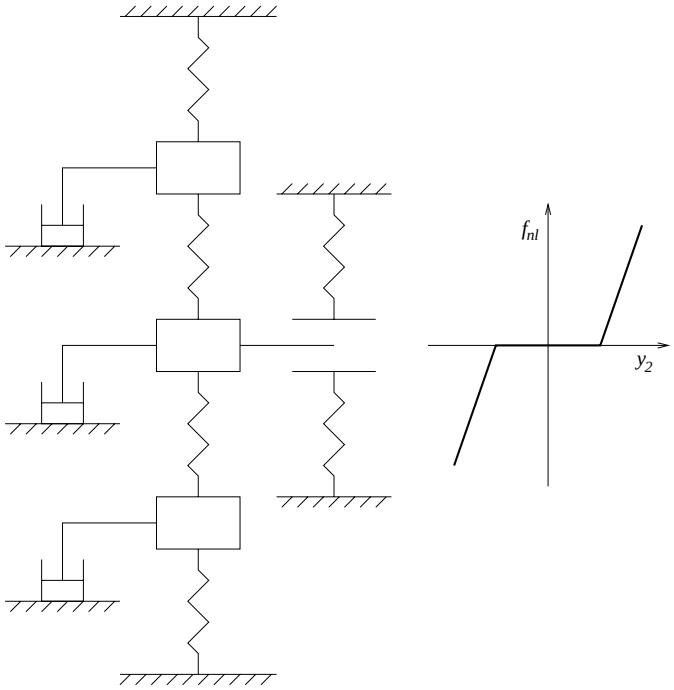


Figure 7.27. A 3DOF simulated piecewise linear system.

excitation force, i.e. for a SDOF linear system, the excitation is estimated from the parameter estimates $\hat{m}$, $\hat{c}$ and $\hat{k}$ as follows:

$$\hat{x}_i = \hat{m}\ddot{y}_i + \hat{c}\dot{y}_i + \hat{k}y_i \tag{7.61}$$

and the MSE is estimated from

$$MSE(\hat{x}) = \frac{100}{N\sigma_x^2}\sum_{i=1}^{N}(x_i - \hat{x}_i)^2. \tag{7.62}$$

When the method is applied to noise-free data from the linear system discussed before, the parameter estimates are $c = 40.000\,000$ and $k = 10\,000.0000$ as compared to $c = 39.96$ and $k = 9994.5$ from the Masri–Caughey procedure. The direct estimate also uses 1000 points as compared to 10 000. Further, the least-squares (LS) estimate is orders of magnitude faster to obtain.

7.4.2 Display without interpolation

The direct least-squares methods described earlier do not produce restoring force surfaces naturally in the course of their use as the Masri–Caughey procedure does.

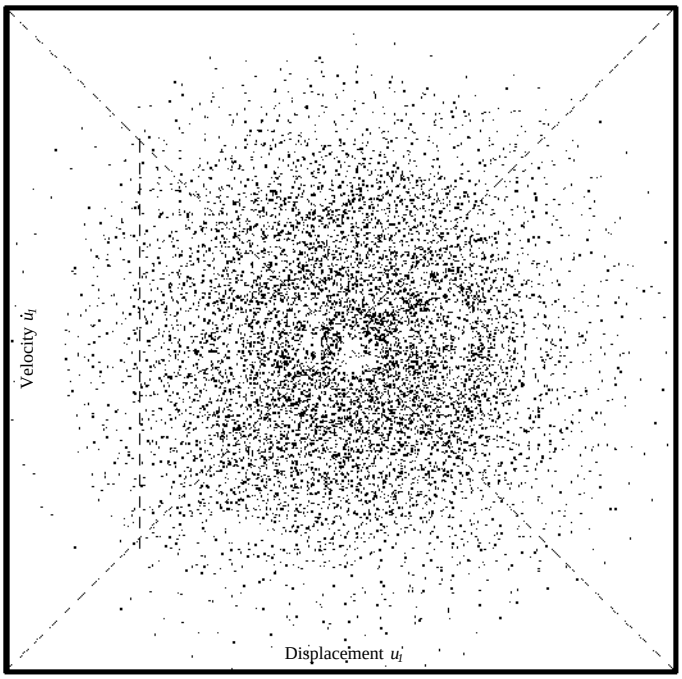


Figure 7.28. Data selected from the $(u_1, \dot{u}_1)$-plane for the interpolation of the force surface $h_1^{(1)}(u_1, \dot{u}_1)$. The system is a 3DOF piecewise linear oscillator.

However, the force surface provides a valuable visual aid to the identification, e.g. the force surface shows directly if a force is piecewise-linear or otherwise, this would not be obvious from a list of polynomial coefficients. Clearly, some means of generating the surfaces is needed which is consistent with the philosophy of direct LS methods. Two methods are available which speedily generate data on a regular grid for plotting.

7.4.2.1 Sections

The idea used here is a modification of the procedure originally used by Masri and Caughey to overcome the extrapolation problem. The stiffness curve or *section* is obtained by choosing a narrow band of width δ through the origin parallel to the y-axis. One then records all pairs of values $(y_i, f(y_i, \dot{y}_i))$ with velocities such that $|\dot{y}_i| < \delta$. The y_i values are saved and placed in increasing order. This gives a $y \to f$ graph which is essentially a slice through the force surface at $\dot{y} = 0$. The procedure is illustrated in figure 7.36. The same procedure can be used to give the damping curve at $y = 0$. If the restoring force separates, i.e.

$$f(y, \dot{y}) = f_{\mathrm{d}}(\dot{y}) + f_s(y) \tag{7.63}$$

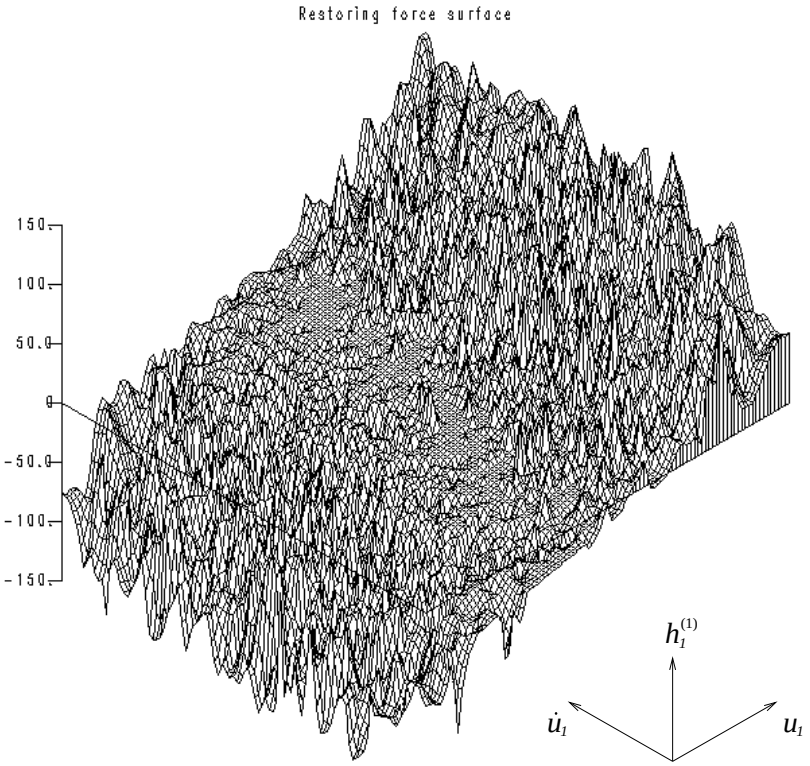


Figure 7.29. Interpolated force surface $h_1^{(1)}(u_1, \dot{u}_1)$ for the 3DOF piecewise linear oscillator.

then identification (i.e. curve-fitting to) of the stiffness and damping sections is sufficient to identify the whole system. Figures 7.37–7.39 show, respectively, the sections for data from a linear system, a Duffing oscillator and a piecewise linear system.

7.4.2.2 Crawley/O'Donnell surfaces

This method of constructing the force surfaces was introduced in [70, 71]. One begins with the triplets obtained from the sampling and processing $(y_i, \dot{y}_i, f_i)$. One then divides the rectangle in the phase plane $[y_{\min}, y_{\max}] \times [\dot{y}_{\min}, \dot{y}_{\max}]$ into small grid squares. If a grid square contains sample points $(y_i, \dot{y}_i)$, the force values above these points are averaged to give an overall force value for the square. This gives a scattering of force values on a regular grid comprising the

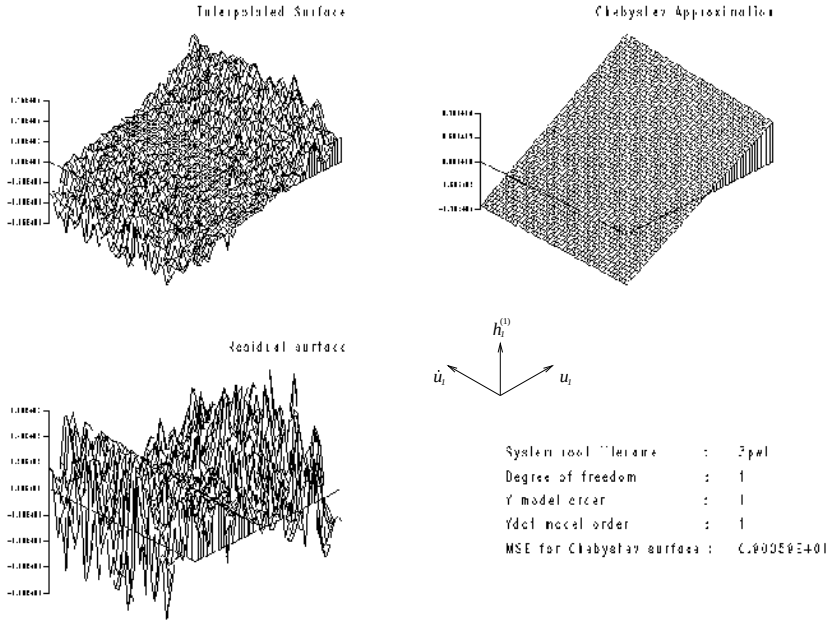


Figure 7.30. Chebyshev model fit of order $(1, 1)$ to the surface in figure 7.29.

centres of the squares. One then checks all the empty squares; if an empty square has four populated neighbours, the relevant force values are averaged to give a value over the formerly empty square. This step is repeated until no new force values are defined. At the next stage, the procedure is repeated for squares with three populated neighbours. As a final optional stage the process can be carried out again for squares with two populated neighbours. The procedure is illustrated in figure 7.40.

The surfaces obtained are not guaranteed to cover the grid and their smoothness properties are generally inferior to those obtained by a more systematic interpolation. In fact, the three-neighbour surface is exact for a linear function in one direction and a constant function in the other at each point. The linear direction will vary randomly from square to square. The surfaces make up for their lack of smoothness with extreme speed of construction. Figures 7.41–7.43 show three-neighbour surfaces for data from a linear system, a Duffing oscillator and a piecewise linear system.

7.4.3 Simple test geometries

The Masri–Caughey procedure was illustrated earlier on simulated data. The direct LS method will be demonstrated a little later on experimental data.

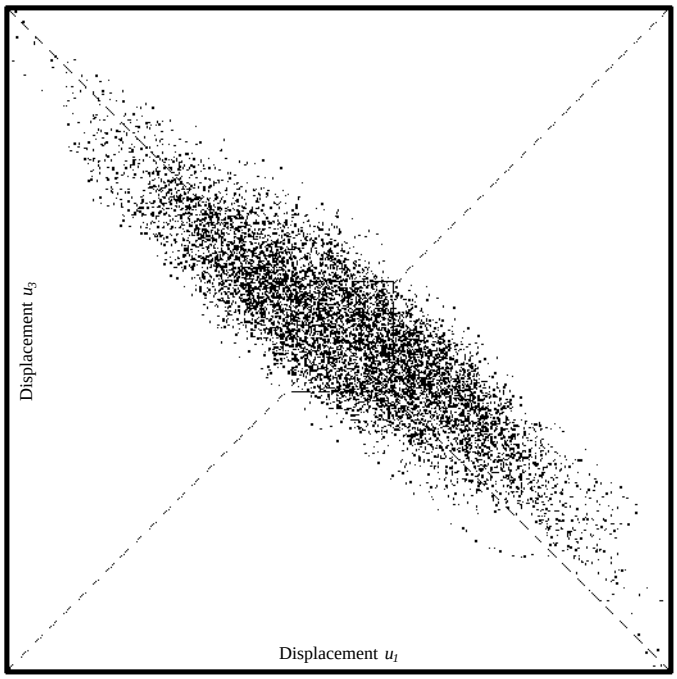


Figure 7.31. Data selected from the (u_1, u_3)-plane for the interpolation of the force surface $h_1^{(2)}(u_1, u_3)$. The system is the 3DOF piecewise linear oscillator.

Before proceeding, it is useful to digress slightly and discuss some useful test configurations. It has been assumed up to now that the force $x(t)$ acts on the mass m with the nonlinear spring grounded and therefore providing a restoring force $f(y, \dot{y})$. This is not always ideal and there are two simple alternatives which each offer advantages.

7.4.3.1 Transmissibility or base excitation

In this geometry (figure 7.44), the base is allowed to move with acceleration $\ddot{y}_{\mathrm{b}}(t)$. This motion is transmitted to the mass through the nonlinear spring and excites the response of the mass $\ddot{y}_m(t)$. The relevant equation of motion is

$$m\ddot{y}_m + f(\delta, \dot{\delta}) = 0 \tag{7.64}$$

where $\delta = y_m - y_{\mathrm{b}}$. In this configuration, the relative acceleration $\ddot{\delta}$ would be computed and integrated to give $\dot{\delta}$ and δ. The advantage is that as the mass only appears as a scaling factor, one can set the mass scale $m = 1$ and form the set of triplets $(\delta_i, \dot{\delta}_i, f_i)$ and produce the force surface. The surface is true up to an overall scale, the type of nonlinearity is represented faithfully. If an estimate of

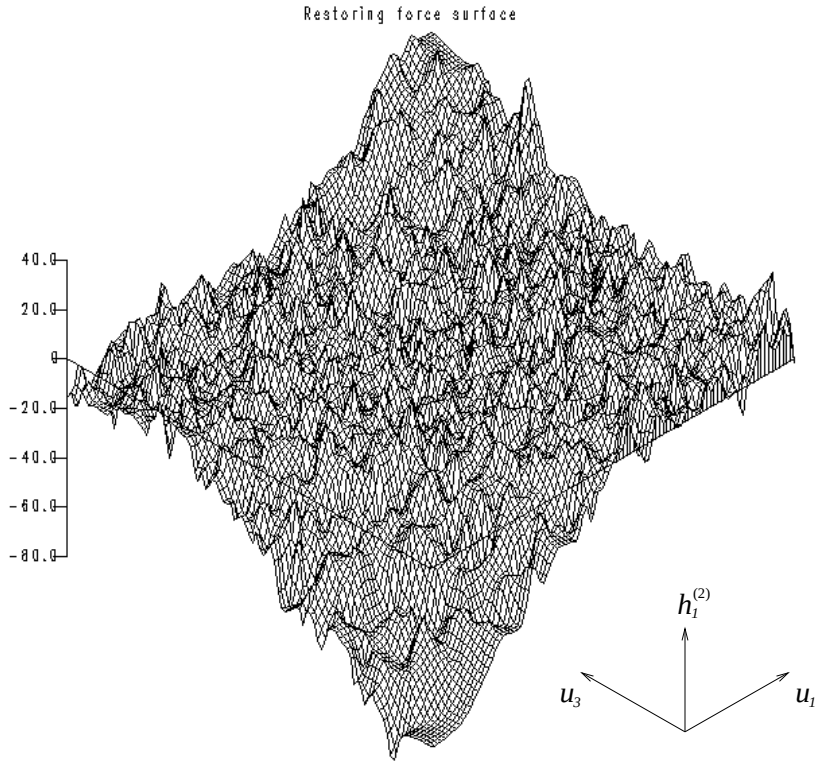


Figure 7.32. Interpolated force surface $h_1^{(2)}(u_1, u_2)$ for the 3DOF piecewise linear oscillator.

the mass becomes available, the force surface can be given the correct scale and the data can be used to fit a model.

7.4.3.2 Mass grounded

Here (figure 7.45), the mass is grounded against a force cell and does not accelerate. Excitation is provided via the base. The equation of motion reduces to

$$f(y_{\mathrm{b}}, \dot{y}_{\mathrm{b}}) = x(t) \tag{7.65}$$

and there is no need to use acceleration. The force triplets can be formed directly using the values measured at the cell. There is no need for an estimate of the mass, yet the overall scale of the force surface is correct.

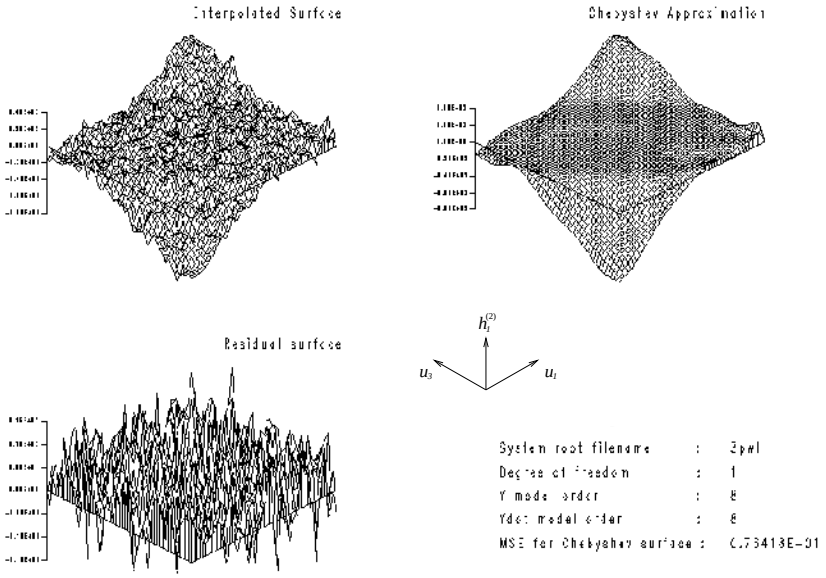


Figure 7.33. Chebyshev model fit of order $(8, 7)$ to the surface in figure 7.32.

7.4.4 Identification of an impacting beam

The system of interest here is a beam made of mild steel, mounted vertically with one encastré end and one free end as shown in figure 4.33. If the amplitude of transverse motion of the beam exceeds a fixed limit, projections fixed on either side of the beam make contact with a steel bush fixed in a steel cylinder surrounding the lower portion of the beam. In the experiments described here, the clearance was set at 0.5 mm. Clearly, when the beam is in contact with the bush, the effective length of the beam is lowered with a consequent rise in stiffness. Overall, for transverse vibrations, the beam has a piecewise linear stiffness. Initial tests showed that the inherent damping of the beam was very light, so this was augmented by the addition of constrained layer damping material to both sides of the beam. Separate tests were carried out at low and high excitation.

7.4.4.1 Low excitation tests

The purpose of this experiment was to study the behaviour of the beam without impacts, when it should behave as a linear system. Because of the linearity, the experiment can be compared with theory. The dimensions and material constants for the beam are given in table 7.4.

According to [42], the first two natural frequencies of a cantilever (fixed-

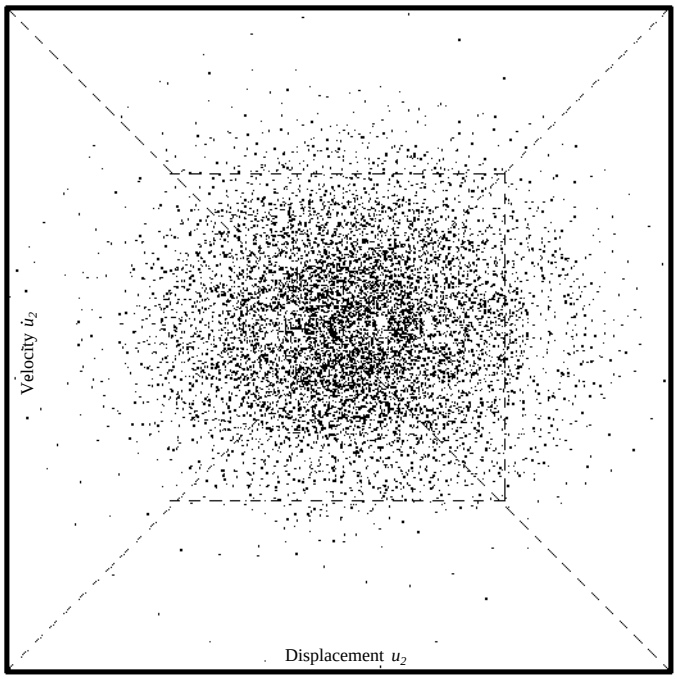


Figure 7.34. Data selected from the $(u_2, \dot{u}_2)$-plane for the interpolation of the force surface $h_2^{(1)}(u_2, \dot{u}_2)$. The system is the 3DOF piecewise linear oscillator.

free) beam are

$$f_i = \frac{1}{2\pi}\left(\frac{\lambda_i}{L}\right)^2 \left(\frac{EI}{m_l}\right)^{\frac{1}{2}} \text{ Hz} \tag{7.66}$$

where $\lambda_1 = 1.8751$ and $\lambda_2 = 4.6941$. This gives theoretical natural frequencies of 16.05 Hz and 100.62 Hz.

A simple impulse test was carried out to confirm these predictions. When an accelerometer was placed at the cross-point (figure 4.33), the frequency response analyser gave peaks at 15.0 Hz and 97.0 Hz (figure 7.46). With the accelerometer at the direct point, the peaks were at 15.5 Hz and 98.5 Hz. These underestimates are primarily due to the additional mass loading of the accelerometer.

One can also estimate the theoretical stiffnesses for the beam using simple theory. If a unit force is applied at a distance a from the root (i.e. the point where the shaker is attached, $a = 0.495$ m), the displacement at a distance d m from the free end is given by

$$y(d) = \frac{1}{6EI}([d-a]^3 - 3(L-a)^2 d + 3(L-a)^2 L - (L-a)^3) \tag{7.67}$$

where $[\ldots]$ is a Macaulay bracket which vanishes if its argument is negative. The

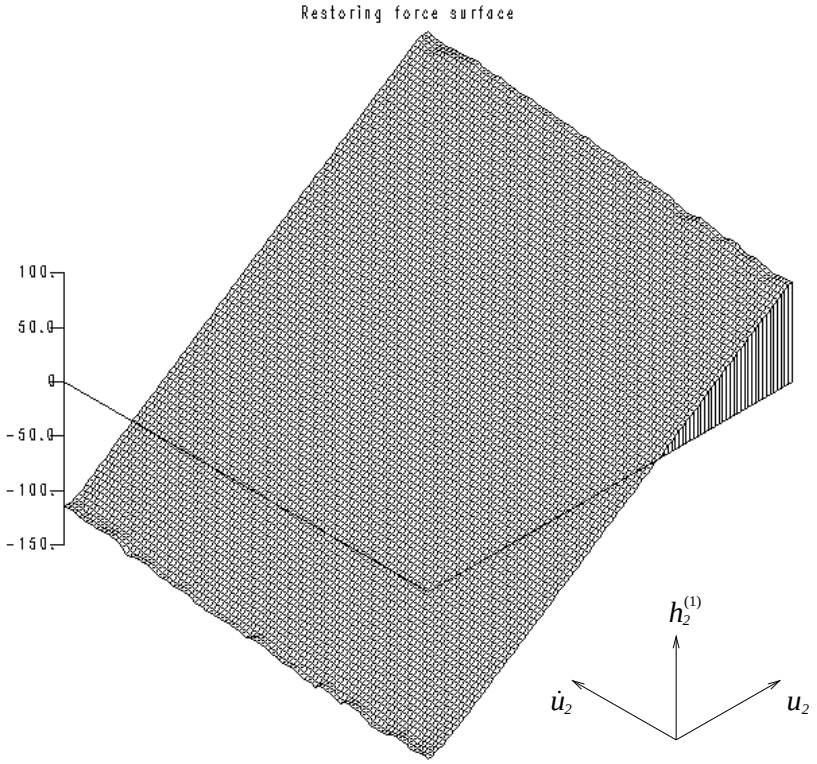


Figure 7.35. Interpolated force surface $h_2^{(1)}(u_2, \dot{u}_2)$ for the 3DOF piecewise linear oscillator.

observable stiffness for the accelerometer at d follows:

$$k(d) = \frac{6EI}{([d-a]^3 - 3(L-a)^2 d + 3(L-a)^2 L - (L-a)^3}. \tag{7.68}$$

When the displacement is measured at the direct point, the *direct stiffness* is estimated as $k_{\mathrm{d}} = 9.654 \times 10^4$ N m^{-1}. At the cross-point, near the free end, the estimated *cross stiffness* is $k_{\mathrm{c}} = 2.769 \times 10^4$ N m^{-1}.

The first two modes of this system are well separated and the first mode is the simple bending mode (which resembles the static deflection curve). It is therefore expected that SDOF methods will suffice if only the first mode is excited, the equation of motion of the system will be, to a good approximation

$$m(d)\ddot{y}_{\mathrm{d}} + c(d)\dot{y}_{\mathrm{d}} + k(d)y_{\mathrm{d}} = x(t) \tag{7.69}$$

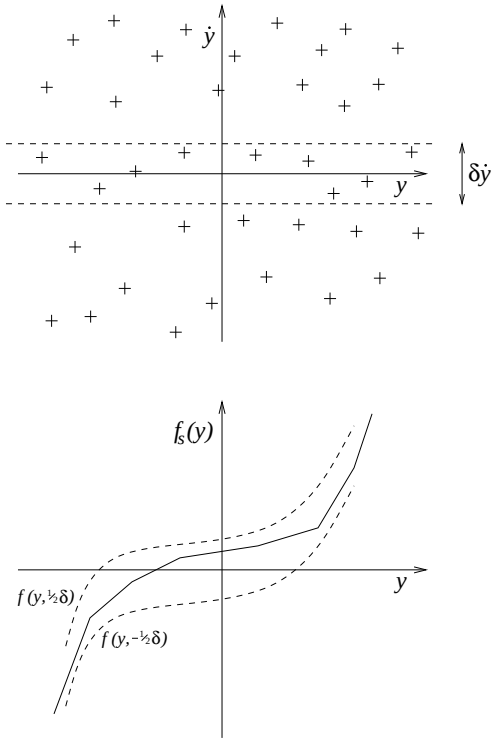


Figure 7.36. Schematic diagram showing the formation of the stiffness section.

where the displacement y_{d} is obtained d m from the free end. The mass $m(d)$ is fixed by the requirement that the natural frequency of the system is given by

$$\omega_{n1} = 2\pi f_1 = \left(\frac{k(d)}{m(d)}\right)^{\frac{1}{2}}. \tag{7.70}$$

Two low level tests were carried out with the accelerometer at the direct-point and cross-point. The instrumentation is shown in figure 7.47. Unfortunately, the CED 1401 sampling instrument was not capable of sampling input and output simultaneously, so the acceleration samples lagged the forces by $\Delta t/2$ with Δt the sampling interval. In order to render the two channels simultaneous, the accelerations were shifted using an interpolation scheme [272].

The first test was carried out with the accelerometer at the cross-point; 5000 points were sampled at 500 Hz. The excitation was white noise band-limited into the interval [10–20] Hz. The accelerations were integrated using the trapezium rule to give velocities and displacements and the estimated signals were band-pass filtered to eliminate spurious components from the integration (the procedures for integration are discussed in some detail in appendix I).

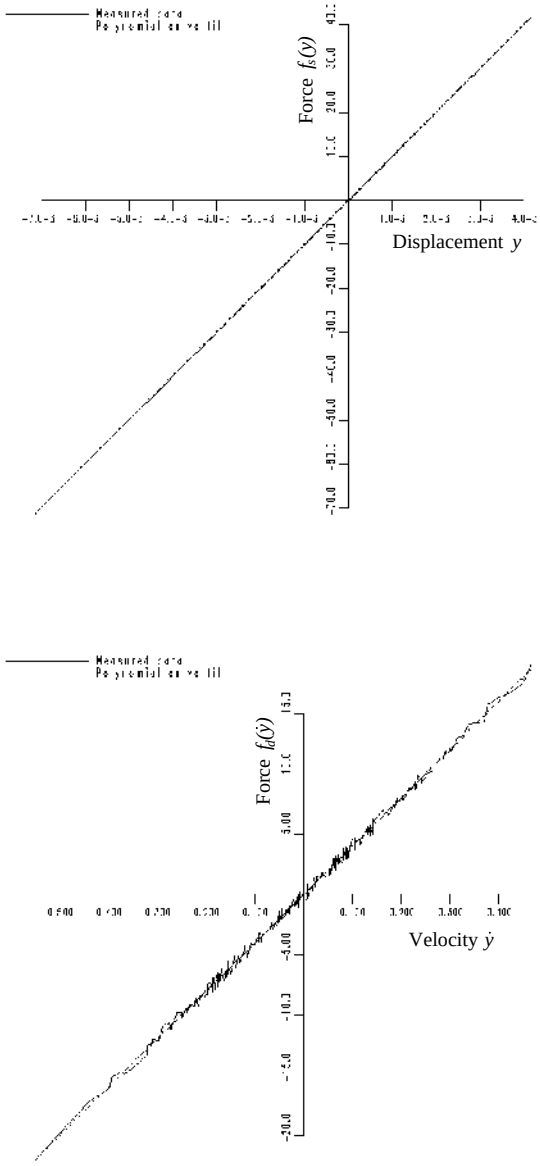


Figure 7.37. Sections from the restoring force surface for a linear system: (*a*) stiffness; (*b*) damping.

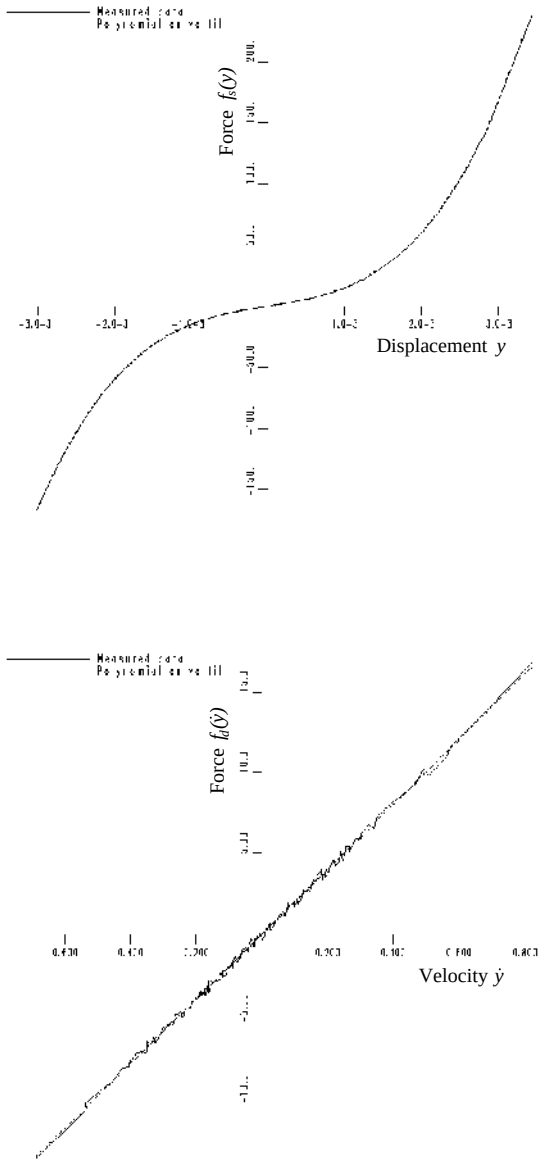


Figure 7.38. Sections from the restoring force surface for a cubic stiffness system: (*a*) stiffness; (*b*) damping.

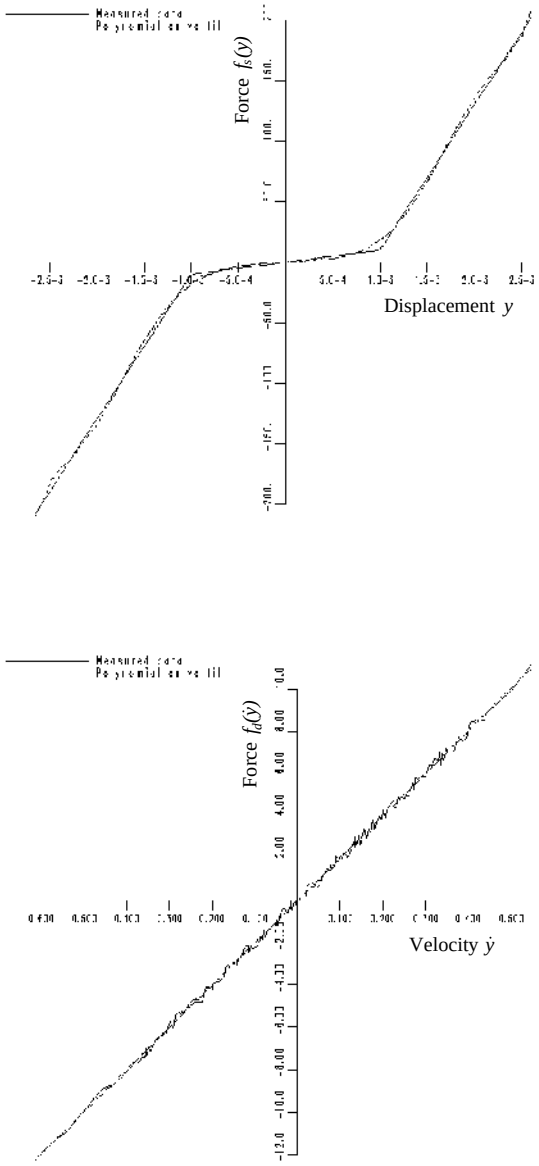


Figure 7.39. Sections from the restoring force surface for a piecewise linear system: (*a*) stiffness; (*b*) damping.

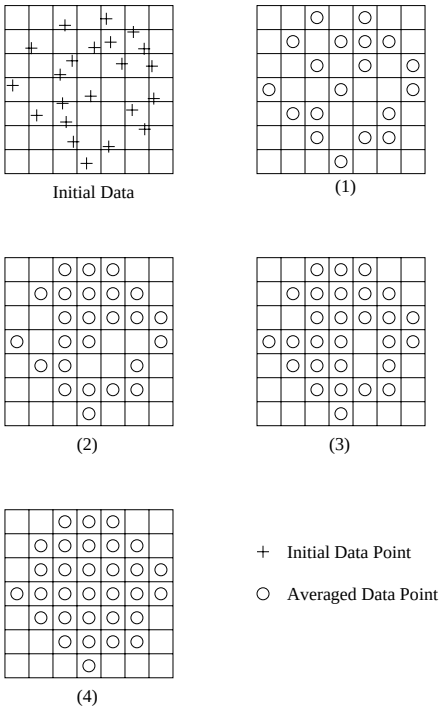


Figure 7.40. Formation of the Crawley–O'Donnell visualization of the restoring force surface.

A direct LS estimation for the model structure (7.64) gave parameters

$$m_{\mathrm{c}} = 3.113\,\mathrm{kg}, \quad c_{\mathrm{c}} = 0.872\ \mathrm{N\,s\,m^{-1}}, \quad k_{\mathrm{c}} = 2.771 \times 10^4\ \mathrm{N\,m^{-1}}.$$

The stiffness shows excellent agreement with the theoretical $k_{\mathrm{c}} = 2.769 \times 10^4$ and the estimated natural frequency of 15.01 Hz compares well with the theoretical 15.00 Hz. Comparing the measured and predicted $x(t)$ data gave an MSE of 0.08%. The estimated restoring force surface is shown in figure 7.48, the linearity of the system is manifest.

The second test used an identical procedure, except data was recorded at the direct point, the LS parameters for the model were

$$m_{\mathrm{d}} = 10.03\,\mathrm{kg}, \quad c_{\mathrm{d}} = 1.389\ \mathrm{N\,s\,m^{-1}}, \quad k_{\mathrm{d}} = 9.69 \times 10^4\ \mathrm{N\,m^{-1}}.$$

Again, the stiffness compares well with the theoretical 9.69×10^4 and the estimated natural frequency $f_1 = 15.66$ Hz compares favourably with the theoretical 15.5 Hz. These tests show that the direct LS approach can accurately identify real systems.

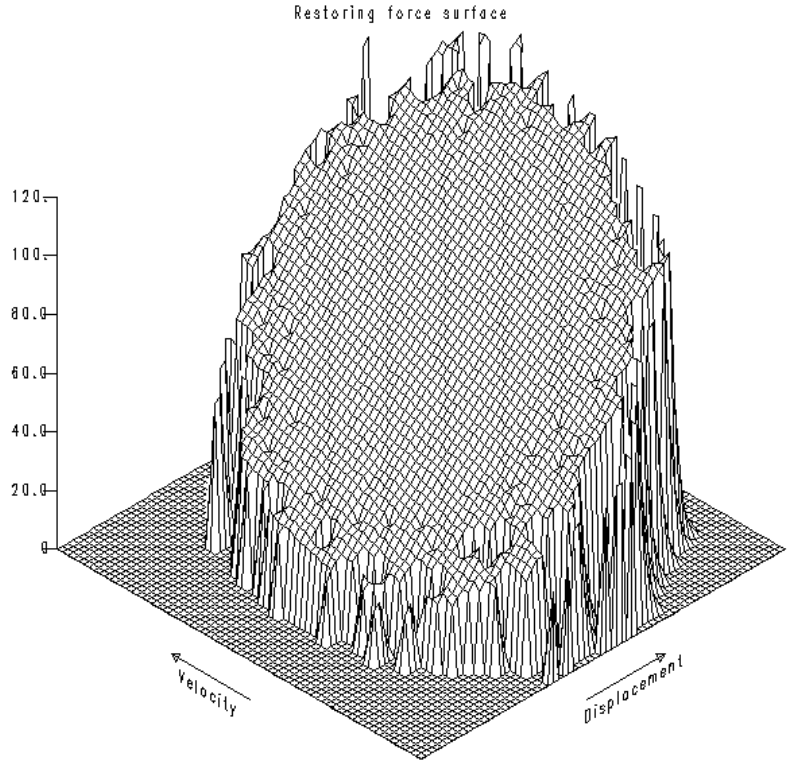


Figure 7.41. Crawley–O'Donnell surface for a linear system.

7.4.4.2 High excitation test

This test was carried out at the cross-point. The level of excitation was increased until the projections on the side of the beam made contact with the bush. As before, the input was band-limited into the range [10–20] Hz. The output spectrum from the test showed a significant component at high frequencies, so the sampling frequency for the test was raised to 2.5 kHz. The high-frequency component made accurate time-shifting difficult, so it was not carried out; the analysis in [272] indicates, in any case, that the main effect would be on the damping, and the stiffness is of interest here. The data were integrated using the trapezium rule and then filtered into the interval $[10, 200]$ in order to include a sufficient number of harmonics in the data. A linear LS fit gave a mass estimate of 2.24 kg which was used to form the restoring force. The stiffness section is given in figure 7.49 (the force surface and damping section are not given as the

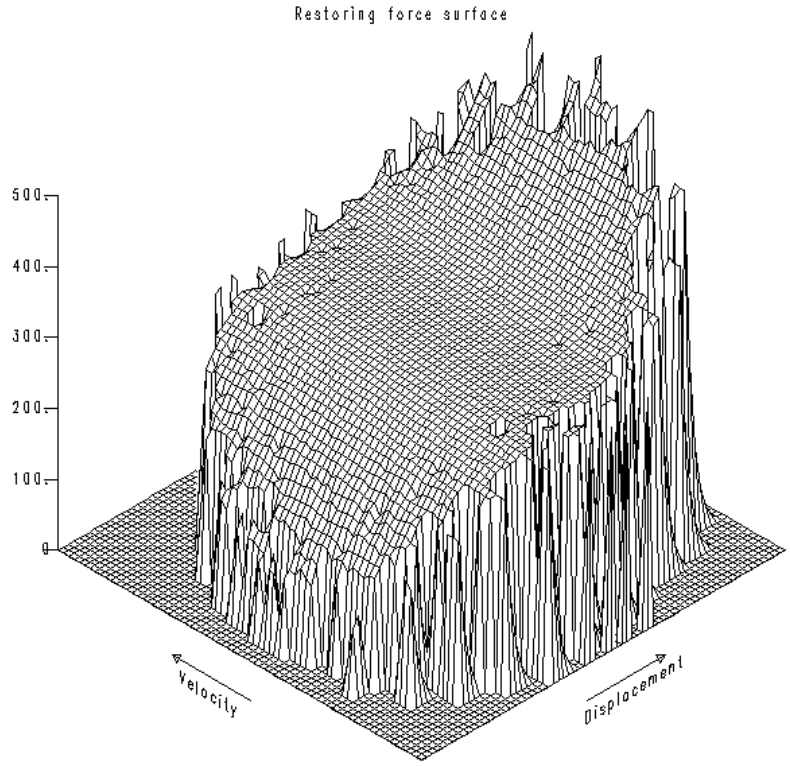


Figure 7.42. Crawley–O'Donnell surface for a cubic stiffness system.

damping behaviour is biased). The section clearly shows the piecewise linear behaviour with discontinuities at ± 0.6 mm. This is acceptably close to the design clearances of ± 0.5 mm.

7.4.5 Application to measured shock absorber data

The automotive shock absorber or damper merits careful study as a fundamental part of the automobile suspension system since the characteristics of the suspension are a major factor in determining the handling properties and ride comfort characteristics of a vehicle.

In vehicle simulations the shock absorber subsystem is usually modelled as a simple linear spring-damper unit. However, experimental work by Lang [157, 223], Hagedorn and Wallaschek [127, 262] and Genta and Campanile [108] on the dynamics of shock absorbers in isolation show that the assumption of

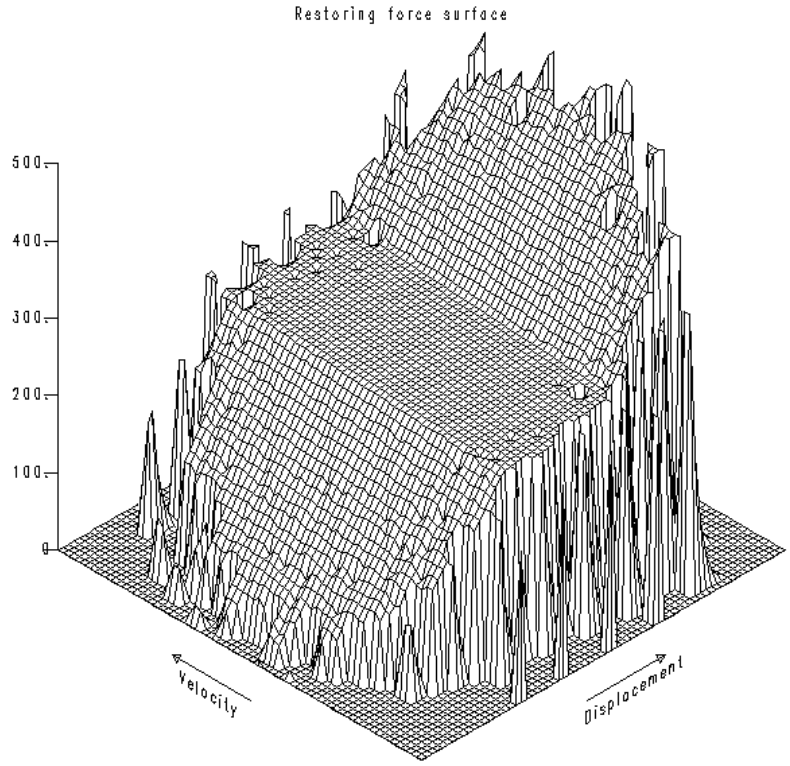


Figure 7.43. Crawley–O'Donnell surface for a piecewise linear system.

linearity is unjustified. This is not a surprising conclusion as automotive dampers are *designed* to have different properties in compression and rebound in order to give balance to the handling and comfort requirements.

On recognizing that the absorber is significantly nonlinear, some means of characterizing this nonlinearity is needed, in order that the behaviour can be correctly represented in simulations.

The most careful theoretical study of an absorber is that of Lang [157]. A physical model was constructed which took properly into account the internal compressive oil/gas flow through the various internal chambers of the absorber; the result was an 87 parameter, highly nonlinear model which was then simulated using an analogue computer; the results showed good agreement with experiment. Unfortunately Lang's model necessarily depends on the detailed construction of a particular absorber and cannot be applied to any other.

Rather than considering the detailed physics, a more straightforward

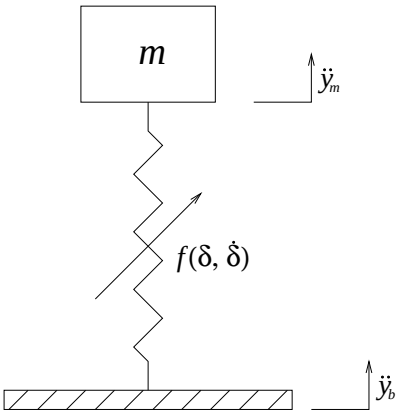


Figure 7.44. Transmissibility configuration for a restoring force surface test.

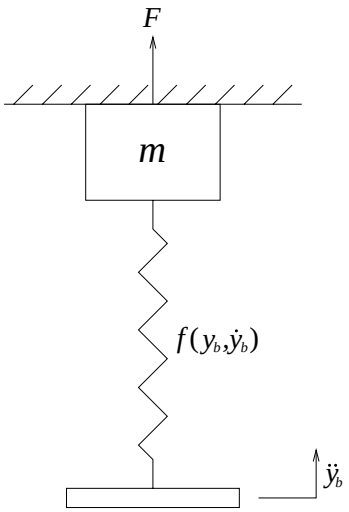


Figure 7.45. Blocked mass configuration for a restoring force surface test.

approach is to obtain an experimental characterization of the absorber. This is usually accomplished by obtaining a force–velocity or *characteristic diagram* (figure 7.50); the force data from a test are simply plotted against the corresponding velocity values. These diagrams show 'hysteresis' loops, i.e. a finite area is enclosed within the curves. This is a consequence of the position dependence of the force. A reduced form of the characteristic diagram is usually produced by testing the absorber several times, each time at the same frequency but with a different amplitude. The maximum and minimum values of the forces and velocities are determined each time and it is these values which are

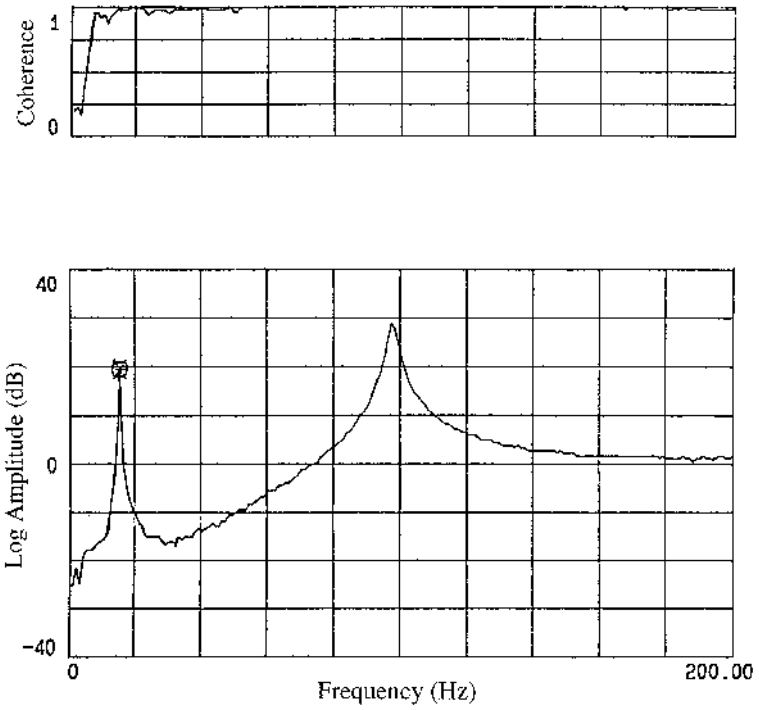


Figure 7.46. FRF for an impacting cantilever experiment at low excitation.

Table 7.4. Dimensions and material constants for cantilever beam.

Length L	0.7 m
Width w	2.525×10^{-2} m
Thickness t	1.25×10^{-2} m
Density ρ	7800 kg m^{-3}
Young's modulus E	2.01×10^{11} N m^{-2}
Second moment of area I	4.1097×10^{-9} m^4
Mass per unit length m_l	2.462 kg m^{-1}

plotted; this procedure actually generates the envelope of the true characteristic diagram and much information is discarded as a consequence. Similar plots of force against displacement—*work diagrams*—can also be produced which convey information about the position dependence of the absorber.

These characterizations of the absorber are too coarse to allow accurate simulation of the absorber dynamics. The approach taken here is to use measured data to construct the restoring force surface for the absorber which simultaneously

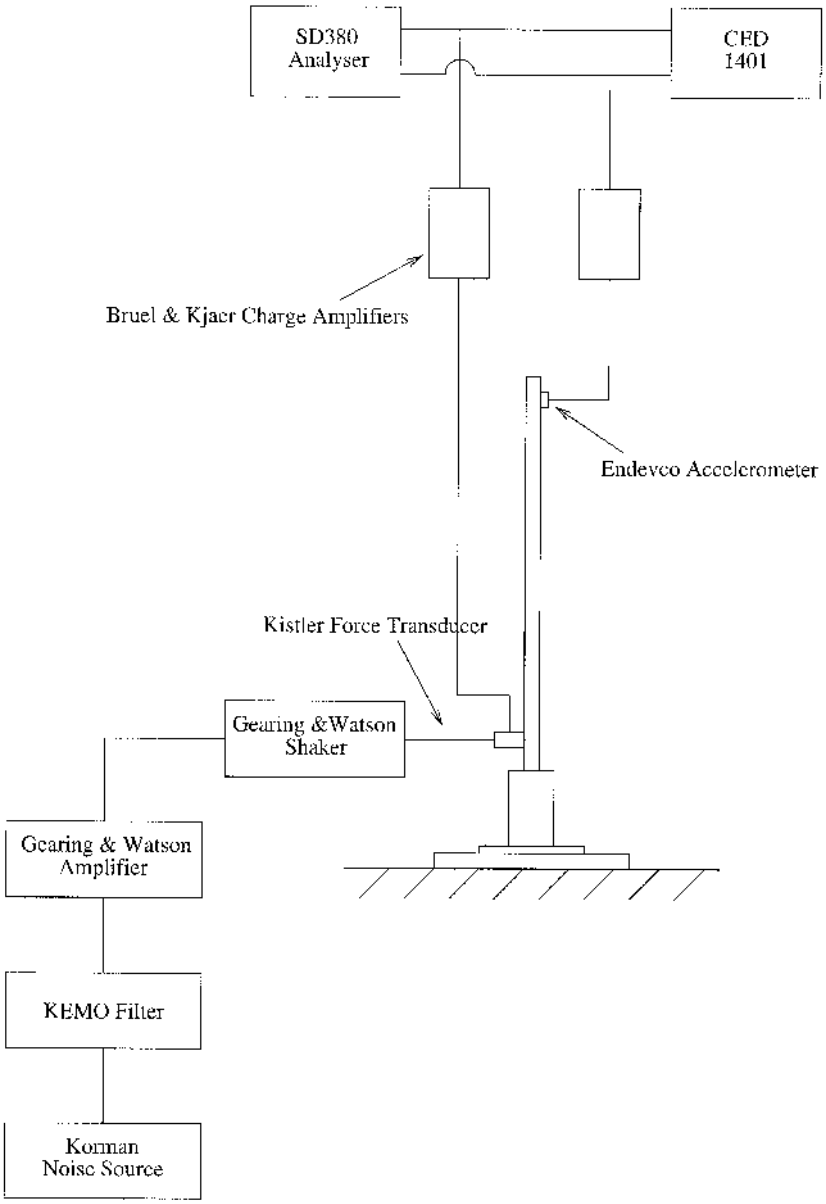


Figure 7.47. Instrumentation for the impacting cantilever identification.

displays the position and velocity dependence of the restoring force in the absorber. This non-parametric representation does not depend on an *a priori*

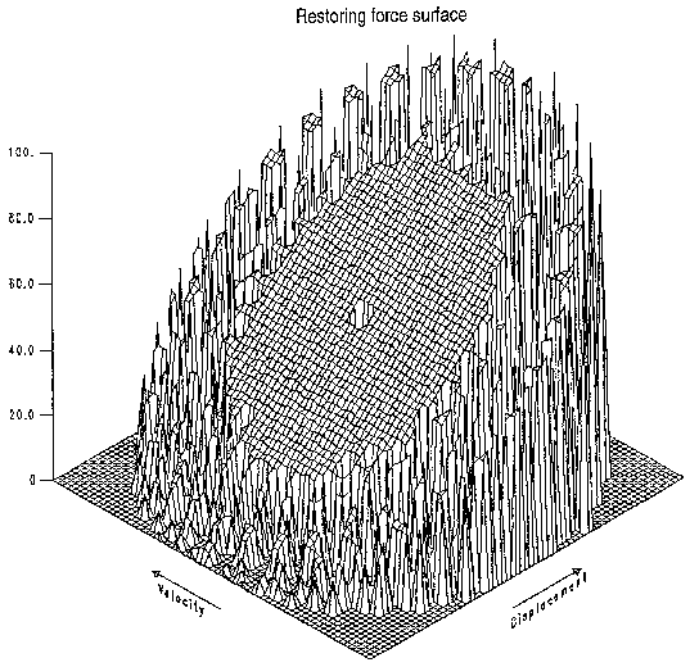


Figure 7.48. Estimated restoring force surface for the impacting cantilever at a low level of excitation.

model of the structure. If necessary, a parametric model can be fitted using the LS methods described earlier or the Masri–Caughey procedure.

The restoring force surface procedure has been applied to the identification of automotive shock absorbers in a number of publications [16, 19, 239]. The most recent work [82] is noteworthy as it also generated fundamental work on restoring force surfaces in general, firstly a new local definition of the surface has been proposed, which fits different models over different sections of the phase plane [83]. Secondly, it has been possible to generate optimal input forces for restoring force surface identification [84].

The results presented here are for a number of sets of test data from a FIAT vehicle shock absorber. The data were obtained by FIAT engineers using the experimental facilities of the vehicle test group at Centro Ricerche FIAT, Orbassano. The apparatus and experimental strategy are shown in figure 7.51 and are described in more detail in [19]; the subsequent data processing and analysis can be found in [239]. Briefly, data were recorded from an absorber which was constrained to move in only one direction in order to justify the assumption of SDOF behaviour. The top of the absorber was fixed to a load cell so that the internal force could be measured directly (it was found that inertial forces were

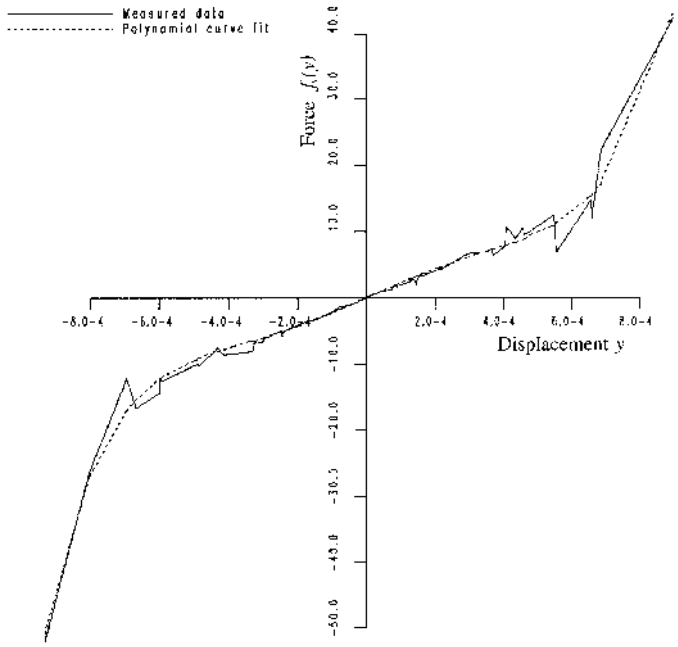


Figure 7.49. Estimated stiffness section for the impacting cantilever at a high level of excitation.

negligible). The base was then excited harmonically using a hydraulic actuator. The absorber was tested at six frequencies, 1, 5, 10, 15, 20, and 30 Hz; the results shown here are for the 10 Hz test showing a range of amplitude levels.

The restoring force surface and the associated contour map are given in figure 7.52, they both show a very clear bilinear characteristic. On the contour map, the contours, which are concentrated in the positive velocity half-plane, are parallel to each other and to the $\dot{y} = 0$ axis showing that the position dependence of the absorber is small. Note that if a parametric representation of the internal force had been obtained, say a LS polynomial, it would have been impossible to infer the bilinear characteristic from the coefficients alone; it is the direct visualization of the nonlinearity which makes the force surface so useful.

The surfaces from the tests at other frequencies showed qualitatively the same characteristics, i.e. a small linear stiffness and a bilinear damping. However, the line of discontinuity in the surface was found to rotate in the phase plane as the test frequency increased. A simple analysis using differenced force surfaces showed that this dependence on frequency was not simply a consequence of disregarding the absorber mass [274]. Force surfaces have also been used to investigate the temperature dependence of shock absorbers [240].

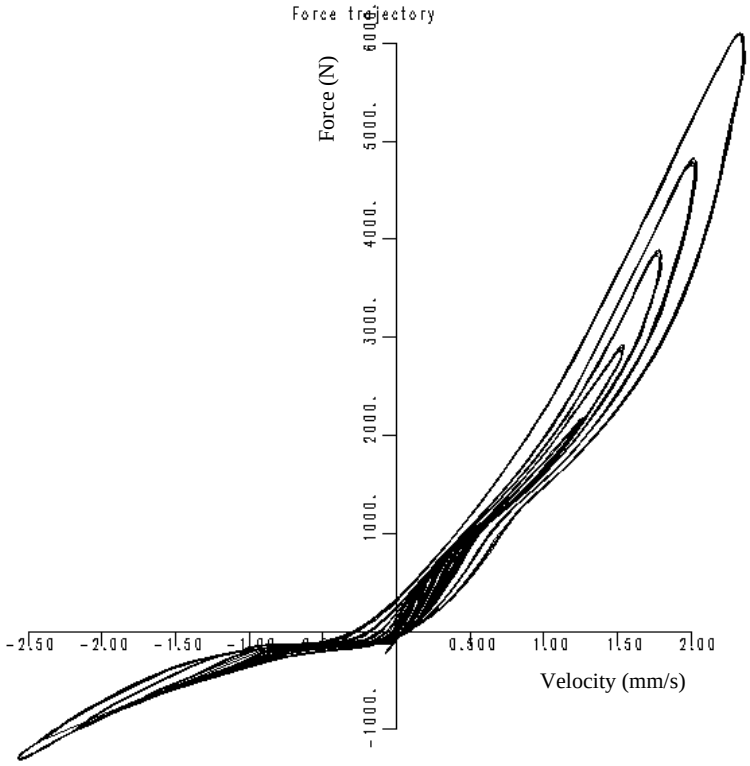


Figure 7.50. Typical shock absorber characteristic diagram.

7.5 Direct parameter estimation for MDOF systems

7.5.1 Basic theory

For a general MDOF system, it is assumed that the mass is concentrated at N measurement points, m_i being the mass at point i. Each point i is then assumed to be connected to each other point j by a link l_{ij}, and to ground by a link l_{ii}. The situation is illustrated in figure 7.53 for a 3DOF system.

If the masses are displaced and released, they are restored to equilibrium by internal forces in the links. These forces are assumed to depend only on the relative displacements and velocities of the masses at each end of the links. If $\delta_{ij} = y_i - y_j$ is the relative displacement of mass m_i relative to mass m_j, and $\dot{\delta}_{ij} = \dot{y}_i - \dot{y}_j$ is the corresponding relative velocity, then

$$\text{force in link } l_{ij} := f_{ij}(\delta_{ij}, \dot{\delta}_{ij}) \tag{7.71}$$

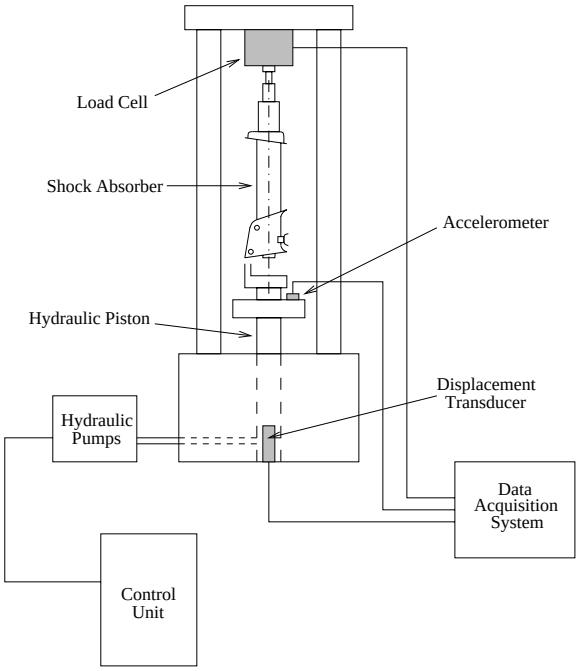


Figure 7.51. Schematic diagram of the shock absorber test bench.

where $\delta_{ii} = y_i$ and $\dot{\delta}_{ii} = \dot{y}_i$ for the link to ground. It will be clear that, as links l_{ij} and l_{ji} are the same,

$$f_{ij}(\delta_{ij}, \dot{\delta}_{ij}) = -f_{ji}(\delta_{ji}, \dot{\delta}_{ji}) = -f_{ji}(-\delta_{ij}, -\dot{\delta}_{ij}). \tag{7.72}$$

If an external force $x_i(t)$ is now applied at each mass, the equations of motion are,

$$m_i\ddot{y}_i + \sum_{j=1}^{N} f_{ij}(\delta_{ij}, \dot{\delta}_{ij}) = x_i(t), \quad i = 1, \ldots, N. \tag{7.73}$$

It is expected that this type of model would be useful for representing a system with a finite number of modes excited. In practice, only the N accelerations and input forces at each point are measured. Differencing yields the relative accelerations $\ddot{\delta}_{ij}$ which can be integrated numerically to give $\dot{\delta}_{ij}$ and δ_{ij}. A polynomial representation is adopted here for f_{ij} giving a model,

$$m_i\ddot{y}_i + \sum_{j=0}^{N}\sum_{k=0}^{p}\sum_{l=0}^{q} a_{(ij)kl}(\delta_{ij})^k(\dot{\delta}_{ij})^l = x_i. \tag{7.74}$$

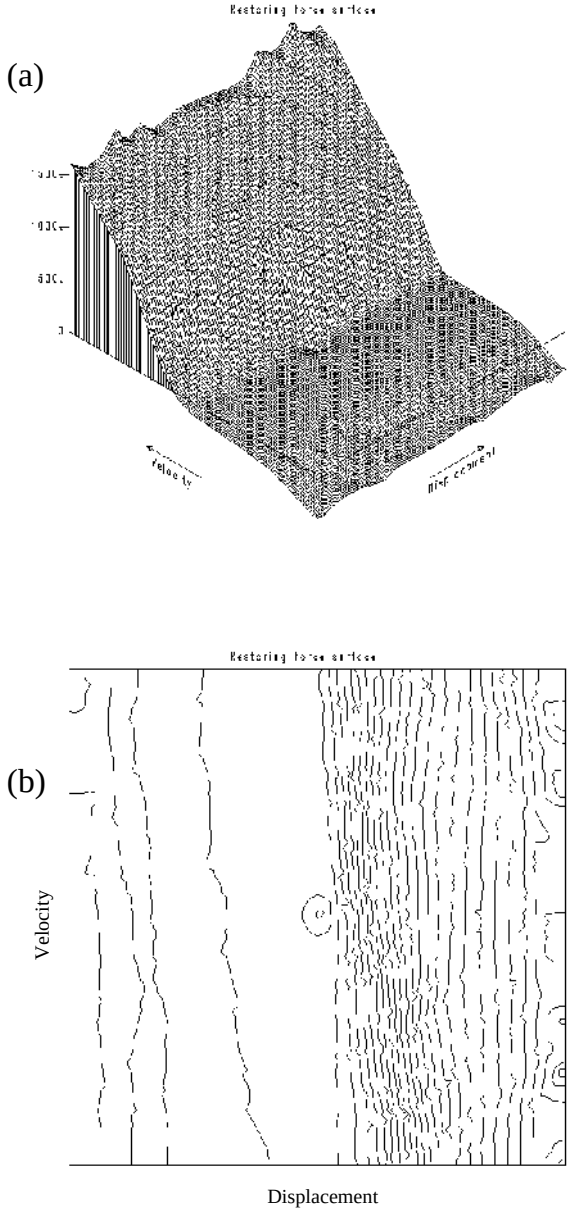


Figure 7.52. Internal restoring force of shock absorber: (*a*) force surface; (*b*) contour map.

LS parameter estimation can be used to obtain the values of the coefficients m_i and $a_{(ij)kl}$ which best fit the measured data. Note that an *a priori* estimate

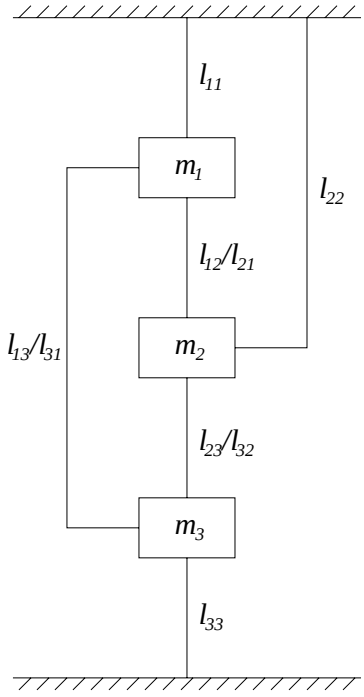


Figure 7.53. Link model of a 3DOF system.

of the mass is *not* required. If there is no excitation at point i, transmissibility arguments yield the appropriate form for the equation of motion of m_i:

$$f'_{ij}(\delta_{ij}, \dot{\delta}_{ij}) = \sum_{j=0}^{N} \sum_{k=0}^{p} \sum_{l=0}^{q} a'_{(ij)kl} (\delta_{ij})^k (\dot{\delta}_{ij})^l = -\ddot{y}_i \tag{7.75}$$

where

$$a'_{(ij)kl} = \frac{1}{m_i} a_{(ij)kl}.$$

Structures of type (7.74) will be referred to as *inhomogeneous* (p, q) models while those of type (7.75) will be termed *homogeneous* (p, q) models. This is in keeping with the terminology of differential equations.

In terms of the expansion coefficients, the symmetry relation (7.72) becomes

$$a_{(ij)kl} = (-1)^{k+l+1} a_{(ji)kl} \tag{7.76}$$

or

$$m_i a'_{(ij)kl} = (-1)^{k+l+1} m_j a'_{(ji)kl}. \tag{7.77}$$

In principle, the inclusion of difference variables allows the model to locate nonlinearity [9]; for example, if a term of the form $(\delta_{23})^3$ appears in the appropriate expansion one can *infer* the presence of a cubic stiffness nonlinearity between points 2 and 3.

Suppose now that only one of the inputs x_i is non-zero. Without loss of generality it can be taken as x_1. The equations of motion become

$$m_1\ddot{y}_i + \sum_{j=1}^{N} f_{ij}(\delta_{ij}, \dot{\delta}_{ij}) = x_1(t) \tag{7.78}$$

$$\ddot{y}_i + \sum_{j=1}^{N} f'_{ij}(\delta_{ij}, \dot{\delta}_{ij}) = 0, \quad i = 2, \ldots, N. \tag{7.79}$$

One can identify all coefficients in the $\ddot{y}_2$ equation up to an overall scale—the unknown m_2 which is embedded in each f'_{2j}. Similarly, all the coefficients in the $\ddot{y}_3$ equation can be known up to the scale m_3. Multiplying the latter coefficients by the ratio m_2/m_3 would therefore scale them with respect to m_2. This means that coefficients for both equations are known up to the same scale m_2. The ratio m_2/m_3 can be obtained straightforwardly; if there is a link l_{23} the two equations will contain terms f'_{23} and f'_{32}. Choosing one particular term, e.g. the linear stiffness term, from each f' expansion gives, via (7.77)

$$\frac{m_2}{m_3} = \frac{a'_{(32)10}}{a'_{(23)01}}. \tag{7.80}$$

The scale m_2 can then be transferred to the $\ddot{y}_4$ equation coefficients by the same method if there is a link l_{24} or l_{34}. In fact, the scale factor can be propagated through all the equations since each mass point must be connected to all other mass points through some sequence of links. If this were not true the system would fall into two or more disjoint pieces.

If the $\ddot{y}_1$ equation has an input, m_1 is estimated and this scale can be transferred to all equations so that the whole MDOF system can be identified using only one input. It was observed in [283] that if the unforced equations of motion are considered, the required overall scale can be fixed by a knowledge of the total system mass, i.e. all system parameters can be obtained from measurements of the free oscillations.

If a restriction is made to linear systems, the equations and notation can be simplified a great deal. Substituting

$$a_{(ij)01} = \gamma_{ij} \tag{7.81}$$

$$a_{(ij)10} = \mu_{ij} \tag{7.82}$$

in the linear versions of the equations of motion (7.78) and (7.79) yields

$$m_1\ddot{y}_i + \sum_{j=1}^{N} \gamma_{ij}\dot{\delta}_{ij} + \sum_{j=1}^{N} \mu_{ij}\delta_{ij} = x_1(t) \tag{7.83}$$

$$\ddot{y}_i + \sum_{j=1}^{N} \gamma'_{ij}\dot{\delta}_{ij} + \sum_{j=1}^{N} \mu'_{ij}\delta_{ij} = 0, \quad i = 2, \ldots, N \tag{7.84}$$

where $\gamma'_{ij} = \gamma_{ij}/m_i$ and $\mu'_{ij} = \mu_{ij}/m_i$.

If estimates for m_i, γ_{ij} and μ_{ij} are obtained, then the usual stiffness and damping matrices $[k]$ and $[c]$ are recovered from the simple relations

$$\begin{aligned} c_{ij} &= -\gamma_{ij}, \quad k_{ij} = -\mu_{ij}, \quad i \neq j \\ c_{ii} &= \sum_{j=1}^{N} \gamma_{ij}, \quad k_{ii} = \sum_{j=1}^{N} \mu_{ij}. \end{aligned} \tag{7.85}$$

The symmetry conditions (7.76) become

$$\gamma_{ij} = \gamma_{ji}, \quad \mu_{ij} = \mu_{ji} \tag{7.86}$$

which imply

$$c_{ij} = c_{ji}, \quad k_{ij} = k_{ji} \tag{7.87}$$

so the model structure forces a symmetry or *reciprocity* condition on the damping and stiffness matrices. By assuming that reciprocity holds at the outset, it is possible to identify all system parameters using one input by an alternative method which is described in [189].

A further advantage of adopting this model is that it allows a natural definition of the restoring force surface for each link. After obtaining the model coefficients the surface f_{ij} can be plotted as a function of δ_{ij} and $\dot{\delta}_{ij}$ for each link l_{ij}. In this case the surfaces are purely a visual aid to the identification, and are more appropriate in the nonlinear case.

7.5.2 Experiment: linear system

The system used for the experiment was a mild steel cantilever (fixed-free) beam mounted so that its motion was confined to the horizontal plane. In order to make the system behave as far as possible like a 3DOF system, three lumped masses of 0.455 kg each, in the form of mild steel cylinders, were attached to the beam at equally spaced points along its length (figure 7.54). The system was described in [111] where a functional-series approach was used in order to identify the characteristics of such systems as discussed in the next chapter. Initial tests showed the damping in the system to be very low; to increase the energy dissipation, constrained layer damping material was fixed to both sides of the beam in between the cylinders.

Details of the various geometrical and material constants for the system are given in [189] in which an alternative approach to DPE is used to analyse data from this system. In order to obtain *theoretical* estimates of the natural frequencies etc, estimates of the mass matrix $[m]$ and the stiffness matrix $[k]$ are

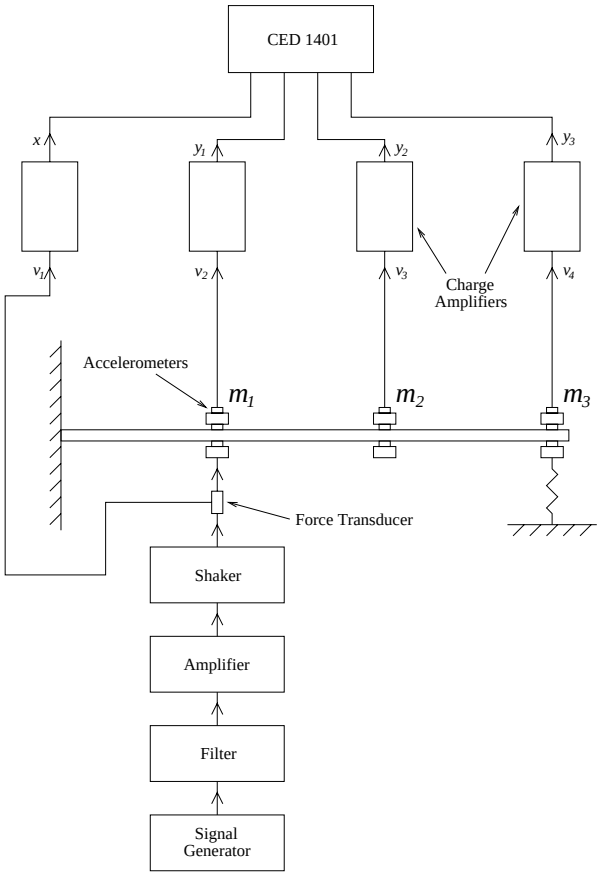


Figure 7.54. Instrumentation for the restoring force surface experiments on the 3DOF experimental nonlinear system.

needed. Assuming that the system can be treated as a 3DOF lumped parameter system, the mass is assumed to be concentrated at the locations of the cylinders. The mass of the portion of beam nearest to each cylinder is transferred to the cylinder. The resulting estimate of the mass matrix was

$$[m] = \begin{vmatrix} 0.7745 & 0.0000 & 0.0000 \\ 0.0000 & 0.7745 & 0.0000 \\ 0.0000 & 0.0000 & 0.6148 \end{vmatrix} \text{[kg]}.$$

Simple beam theory yielded an estimate of the stiffness matrix

$$[k] = 10^5 \times \begin{vmatrix} 1.2579 & -0.7233 & 0.1887 \\ -0.7233 & 0.6919 & -0.2516 \\ 0.1887 & -0.2516 & 0.1101 \end{vmatrix} [\text{N m}^{-1}].$$

Having obtained these estimates, the eigenvalue problem

$$-\omega_i^2[m]\{\psi_i\} = [k]\{\psi_i\} \tag{7.88}$$

was solved, yielding the natural frequencies $f_i = \omega_i/2\pi$ and the modeshapes $\{\psi_i\}$. The predictions for the first three natural frequencies were 4.76, 22.34, and 77.11 Hz. As the integrating procedures used to obtain velocity and displacement data from measured accelerations require a band-limited input to be used, it would have proved difficult to excite the first mode and still have no input at low frequencies. For this reason, a helical compression spring with stiffness 1.106×10^4 N m^{-1} was placed between point 3 and ground as shown in figure 7.54. The added mass of the spring was assumed to be negligible. The modification to the stiffness matrix was minimal, except that k_{33} changed from 1.101×10^4 to 2.207×10^4. However, the first natural frequency changed dramatically, re-solving the eigenvalue problem gave frequencies of 17.2, 32.0 and 77.23 Hz.

The arrangement of the experiment is also shown in figure 7.54. The signals were sampled and digitized using a CED 1401 intelligent interface. A detailed description of the rest of the instrumentation can be found in [267].

The first experiment carried out on the system was a modal analysis to determine accurately the natural frequencies of the system. The FRFs $Y_1(\omega)/X(\omega)$, $Y_2(\omega)/X(\omega)$ and $Y_3(\omega)/X(\omega)$ were obtained; standard curve-fitting to these functions showed that the first three natural frequencies were 16.91, 31.78 and 77.78 Hz in good agreement with the theoretical estimates. The averaged output spectrum for the system when excited by a band-limited input in the range 10–100 Hz is shown in figure 7.55; there seems to be no significant contribution from higher modes than the third and it would therefore be expected that the system could be modelled well by a 3DOF model if the input is band-limited in this way.

An experiment was then carried out with the intention of fitting LS models of the types (7.77) and (7.78) to the data. The excitation used was a noise sequence band-limited in the range 10–100 Hz. The data $x(t)$, $\ddot{y}_1$, $\ddot{y}_2$ and $\ddot{y}_3$ were sampled with frequency 1666.6 Hz, and 3000 points per channel were taken. Equal-interval sampling between channels was performed.

The acceleration signals were integrated using the trapezium rule followed by band-pass filtering in the range 10–100 Hz [274]; the data were passed through the filter in both directions in order to eliminate phase errors introduced by a single pass. To remove any filter transients 500 points of data were deleted from the beginning and end of each channel; this left 2000 points per channel.

An inhomogeneous $(1,1)$ model was fitted to data points 500 to 1500 in order to identify the $\ddot{y}_1$ equation of motion; the result was

$$\begin{aligned} 0.8585\ddot{y}_1 - 4.33\dot{y}_1 + 7.87 \times 10^4 y_1 + 10.1(\dot{y}_1 - \dot{y}_2) \\ + 8.33 \times 10^4(y_1 - y_2) - 2.23 \times 10^4(y_1 - y_3) = x(t). \end{aligned} \tag{7.89}$$

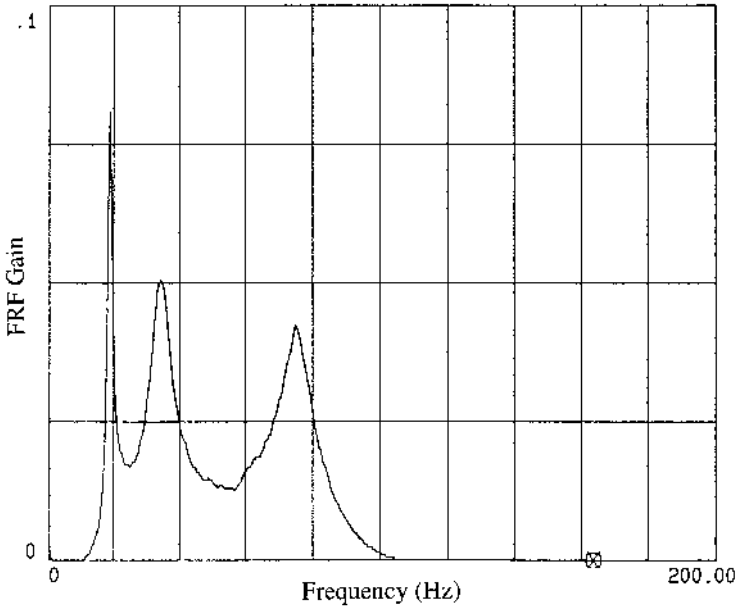


Figure 7.55. Output spectrum for the linear 3DOF system under excitation by a random signal in the range 10–100 Hz.

Comparing the predicted and measured data gave an MSE of 0.035%, indicating excellent agreement. In all models for this system the significance threshold for deleting insignificant terms was set at 0.1%.

A homogeneous $(1,1)$ model was fitted to each of the $\ddot{y}_2$ and $\ddot{y}_3$ equations of motion. The results were

$$\ddot{y}_2 + 9.11 \times 10^4(y_2 - y_1) - 3.55 \times 10^4 y_2 + 3.34 \times 10^4(y_2 - y_3) = 0 \quad (7.90)$$

and

$$\begin{aligned}\ddot{y}_3 + 6.84(\dot{y}_3 - \dot{y}_1) - 7.13\dot{y}_3 - 3.85 \times 10^4(y_3 - y_1) \\ + 4.63 \times 10^4(y_3 - y_2) + 3.00 \times 10^4 y_3 = 0. \quad (7.91)\end{aligned}$$

The comparisons between measured and predicted data gave MSE values of 0.176% and 0.066%, again excellent.

The scale factors were transferred from the first equation of motion to the others as previously described. The final results for the (symmetrized) estimated system matrices were

$$[m] = \left|\begin{array}{ccc} 0.8595 & 0.0000 & 0.0000 \\ 0.0000 & 0.9152 & 0.0000 \\ 0.0000 & 0.0000 & 0.5800 \end{array}\right| \text{ [kg]}$$

Table 7.5. Natural frequencies for linear system.

Mode	Experimental frequency (Hz)	Model frequency (Hz)	Error (%)
1	16.914	17.044	0.77
2	31.781	32.247	1.47
3	77.529	77.614	0.11

$$[k] = 10^5 \times \begin{vmatrix} 1.3969 & -0.8334 & 0.2233 \\ -0.8334 & 0.7949 & -0.2869 \\ 0.2233 & -0.2869 & 0.2379 \end{vmatrix} \ [\mathrm{N\ m^{-1}}]$$

which compare favourably with the theoretical results. In all cases, the damping estimates have low significance factors and large standard deviations, indicating a low level of confidence. This problem is due to the low level of damping in the system, the constrained layer material having little effect. Thus the damping matrix estimates are not given. Using the estimated $[m]$ and $[k]$ matrices, the first three natural frequencies were estimated. The results are shown in table 7.5 and the agreement with the modal test is good. However, the question remains as to whether the parameters correspond to actual *physical* masses and stiffnesses. In order to address this question, another experiment was carried out. An additional 1 kg mass was attached to measurement point 2 and the previous experimental procedure was repeated exactly. The resulting parameter estimates were

$$[m] = \begin{vmatrix} 0.8888 & 0.0000 & 0.0000 \\ 0.0000 & 1.9297 & 0.0000 \\ 0.0000 & 0.0000 & 0.7097 \end{vmatrix} \ [\mathrm{kg}]$$

$$[k] = 10^5 \times \begin{vmatrix} 1.3709 & -0.8099 & 0.2245 \\ -0.8099 & 0.7841 & -0.3014 \\ 0.2245 & -0.3014 & 0.2646 \end{vmatrix} \ [\mathrm{N\ m^{-1}}]$$

and the results have changed very little from the previous experiment, the only exception being that m_{22} has increased by 1.01 kg. The results give confidence that the parameters are physical for this highly discretized system with very small effects from out-of-range modes. The natural frequencies were estimated and compared with those obtained by curve-fitting to transfer functions. The results are shown in table 7.6, again with good agreement.

7.5.3 Experiment: nonlinear system

The final experimental system was based on that in [111]. The same experimental arrangement as in the previous subsection was used with a number

Table 7.6. Natural frequencies for linear system with 1 kg added mass.

Mode	Experimental frequency (Hz)	Model frequency (Hz)	Error (%)
1	13.624	13.252	2.73
2	29.124	29.846	2.48
3	69.500	69.365	0.19

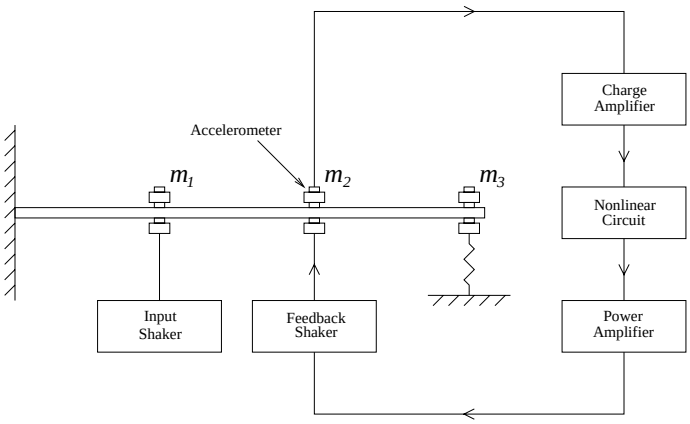


Figure 7.56. Feedback loop for the introduction of a nonlinear force into the 3DOF system.

of modifications. An additional accelerometer was placed at measurement point 2, the signal obtained was then passed to a charge amplifier which was used to integrate the signal giving an output proportional to the velocity $\dot{y}_2$. The velocity signal was then passed through a nonlinear electronic circuit which produced an output proportional to $\dot{y}_2^3$. The cubed signal was then amplified and used to drive an electrodynamic shaker which was attached to measurement point 2 via a rigid link. The overall effect of this feedback loop is to introduce a restoring force at measurement point 2 proportional to the cube of the velocity at point 2. The layout of the feedback loop is shown in figure 7.56.

The experimental procedure was the same as in the linear case. The excitation used was a noise sequence in the range 10–100 Hz. Consideration of the FRFs for the system showed that the damping in the system was clearly increased by the presence of the shaker. The natural frequencies for the system with the shaker attached (but passive) were approximately 19, 32 and 74.9 Hz; the shaker also introduces additional mass and stiffness. The cubic circuit was then switched in and the amplitude of the feedback signal increased until a noticeable

increase in damping and loss of coherence were obtained in the FRF.

Using the CED interface 4000 points of sampled data were obtained for each channel $x(t)$, $\ddot{y}_1$, $\ddot{y}_2$ and $\ddot{y}_3$. After passing the data to the computer, each channel was shifted forward in time as described earlier. The acceleration signals were then integrated using the trapezium rule followed by filtering. In this case the pass-band was 10–300 Hz, the high cut-off being chosen so that any third harmonic content in the data would be retained. As before, 500 points were removed from the beginning and end of each channel in order to eliminate transients.

The $\ddot{y}_1$ equation of motion was obtained by fitting an inhomogeneous $(1,1)$ model to 1000 points of the remaining data. The estimated equation was

$$0.872\ddot{y}_1 - 22.4\dot{y}_1 + 8.59 \times 10^4 y_1 + 20.7(\dot{y}_1 - \dot{y}_2) + 7.96 \times 10^4(y_1 - y_2) - 2.31 \times 10^4(y_1 - y_3) = x(t). \tag{7.92}$$

The comparison between measured and predicted data gave an MSE of 0.056%. The very low MSE indicates that the equation is adequately described by a $(1,1)$ model, i.e. it has no significant nonlinear terms. As a check, a $(3,3)$ model was fitted to the same data. All but the linear terms were discarded as insignificant. The mass and stiffness values did not change but the damping values did alter slightly, further evidence that the damping estimates are not to be trusted.

The second equation of motion was obtained by fitting a inhomogeneous $(1,3)$ model to 2500 points of data. The estimation yielded the equation

$$\ddot{y}_2 - 16.7(\dot{y}_2 - \dot{y}_1) + 154.3\dot{y}_2 + 8.45 \times 10^4(y_2 - y_1) - 2.93 \times 10^4 y_2 + 3.07 \times 10^4(y_2 - y_3) + 228.0(\dot{y}_2 - \dot{y}_1)^3 - 183.0\dot{y}_2^2 + 5.63 \times 10^3 \dot{y}_2^3 = 0. \tag{7.93}$$

The MSE for the comparison between measured and predicted output shown in figure 7.57 was 0.901%. The MSE obtained when a $(1,1)$ model was tried was 1.77%; this increase indicates that the equation truly requires a nonlinear model. The force surfaces for links l_{21}, l_{22} and l_{23} are shown in figures 7.58–7.60. It can be seen that the surface for link l_{21} is almost flat as expected, even though a cubic term is present. In fact, the significance/confidence levels for the $(\dot{y}_1 - \dot{y}_2)^3$ and $\dot{y}_2^2$ terms were so low that the standard errors for the parameters were greater than the parameters themselves. The $\dot{y}_2^3$ term must be retained as the estimate is 5630 ± 4882 for the coefficient; also the significance factor for this term was 2.6. Finally, it can be seen from the force surface in figure 7.59 that the cubic term is significant. It can be concluded that the procedure has identified a cubic velocity term in the link connecting point 2 to ground.

The $\ddot{y}_3$ equation was obtained by fitting a homogeneous $(1,1)$ model to 1000 points of data. The estimated equation was

$$\ddot{y}_3 + 8.37(\dot{y}_3 - \dot{y}_1) + 27.1(\dot{y}_3 - \dot{y}_2) - 36.4\dot{y}_3 - 3.98 \times 10^4(y_3 - y_1) + 4.47 \times 10^4(y_3 - y_2) + 3.35 \times 10^4 y_3 = 0. \tag{7.94}$$

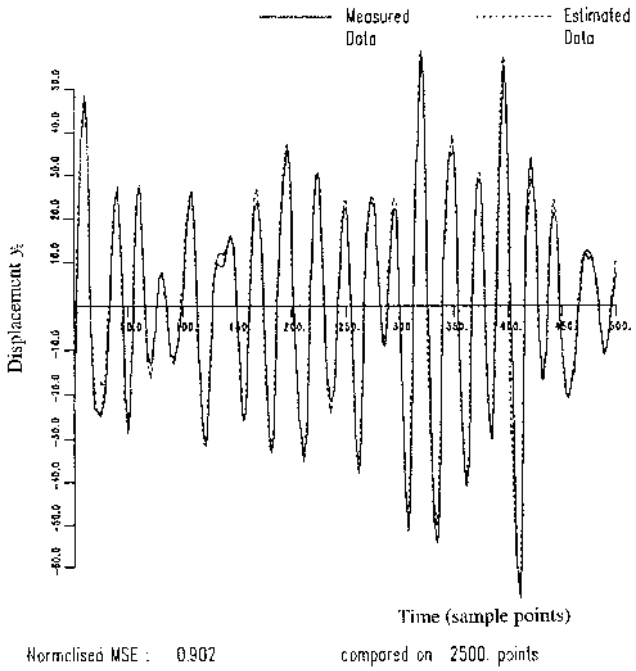


Figure 7.57. Comparison of measured data and that predicted by the nonlinear model for the second equation of motion for the nonlinear 3DOF experimental system.

A comparison between measured and predicted output gave an MSE of 0.31%, indicating that a linear model is adequate. To check, a $(3,3)$ model was fitted and all but the linear terms were discarded as insignificant.

After transferring scales from the $\ddot{y}_1$ equation to the other two, the system matrices could be constructed from the previous estimates. The symmetricized results were

$$[m] = \begin{vmatrix} 0.8720 & 0.0000 & 0.0000 \\ 0.0000 & 0.9648 & 0.0000 \\ 0.0000 & 0.0000 & 0.5804 \end{vmatrix} \text{[kg]}$$

$$[k] = 10^5 \times \begin{vmatrix} 1.4240 & -0.7960 & 0.2310 \\ -0.7960 & 0.7950 & -0.2711 \\ 0.2310 & -0.2711 & 0.2345 \end{vmatrix} [\text{N m}^{-1}].$$

These parameters show good agreement with those for the linear experiment. This time, a significant damping coefficient c_{22} was obtained; this is due to the linear damping introduced by the shaker.

All that remained to be done now was to determine the true cubic coefficient

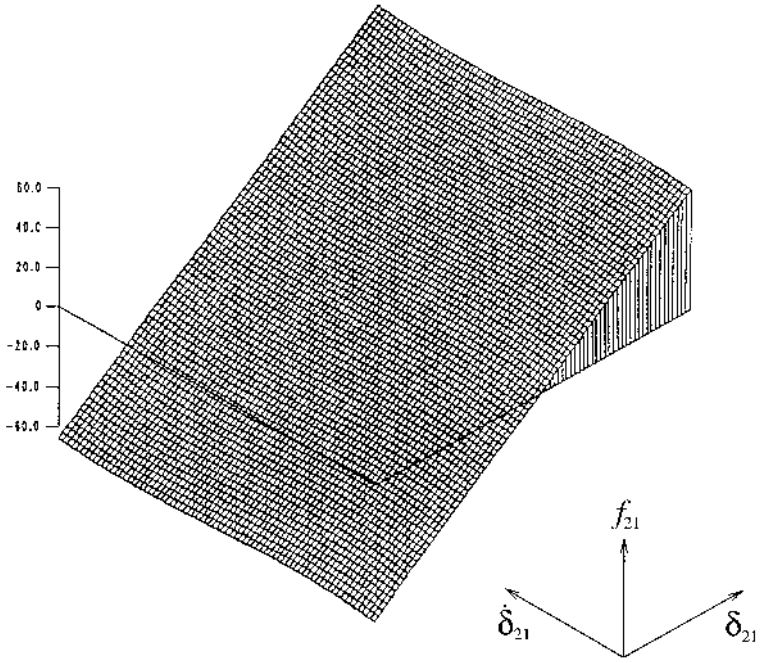


Figure 7.58. Restoring force surface for link l_{21} in the nonlinear 3DOF experimental system.

in the experiment. The details of this calibration experiment are given in [267]. The result was

$$F = -3220.0\dot{y}_2^3. \tag{7.95}$$

The coefficient value estimated by the identification procedure was 5431 ± 4710. The percentage error is therefore 69%; while this is a little high, the estimate has the right order of magnitude and the error interval of the estimate encloses the 'true' value.

The DPE scheme has also been implemented for distributed systems in [165].

It is clear that restoring force methods allow the identification of MDOF nonlinear experimental systems. It should be stressed that high-quality instrumentation for data acquisition is required. In particular, poor phase-matching between sampled data channels can result in inaccurate modelling of damping behaviour. The two approaches presented here can be thought of as complementary. The Masri–Caughey modal coordinate approach allows the construction of restoring force surfaces without specifying an *a priori* model. The main disadvantage is that the surfaces are distorted by nonlinear interference terms from other coordinates unless modes are well separated. The DPE approach produces force

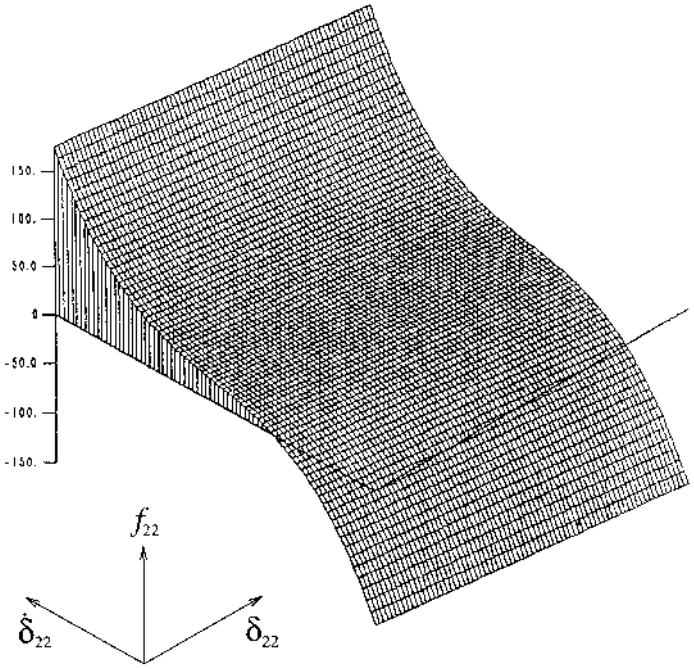


Figure 7.59. Restoring force surface for link l_{22} in the nonlinear 3DOF experimental system.

surfaces only after a parametric model has been specified and fitted, but offers the advantage that systems with close modes present no particular difficulties.

7.6 System identification using optimization

The system identification methods discussed earlier in this chapter and the previous one are only appropriate for *linear-in-the-parameters* system models and, although these form a large class of models, they by no means exhaust the possibilities. Problems begin to arise when the system nonlinearity is not linear-in-the-parameters, e.g. for piecewise-linear systems (which include clearance, deadband and backlash systems) or if the equations of motion contain states which cannot be measured directly, e.g. in the Bouc–Wen hysteresis model discussed later. If the objective function for optimization, e.g. squared-error, depends differentiably on the parameters, traditional minimization techniques like gradient descent or Gauss–Newton [99] can be used. If not, newly developed (or rather, newly exploited) techniques like genetic algorithms (GAs) [117] or downhill simplex [209] can be employed. In [241], a GA with simulated annealing was used to identify linear discrete-time systems. In [100], the GA was used to find the

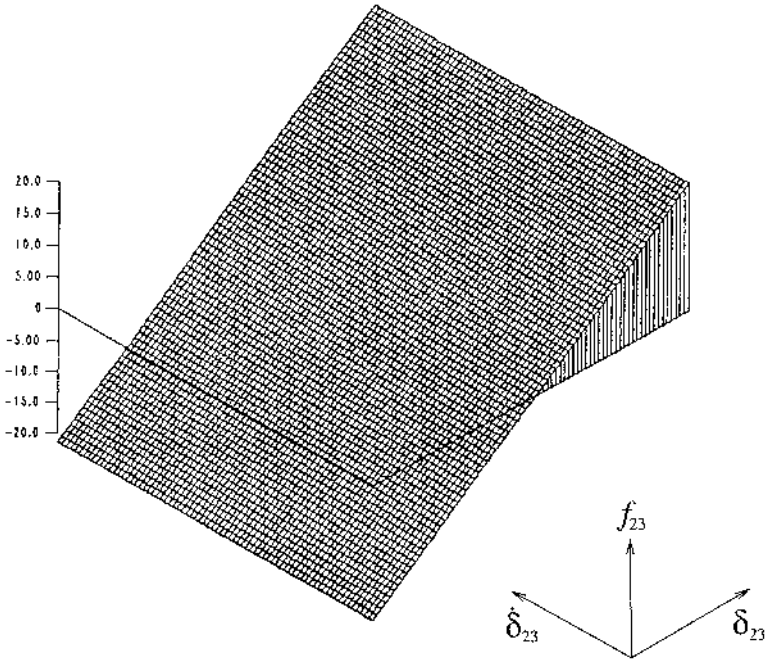


Figure 7.60. Restoring force surface for link l_{23} in the nonlinear 3DOF experimental system.

structure for an NARMAX model. This section demonstrates how optimization methods, GAs and gradient descent, in particular, can be used to solve continuous-time parameter estimation problems.

7.6.1 Application of genetic algorithms to piecewise linear and hysteretic system identification

7.6.1.1 Genetic algorithms

For the sake of completeness, a brief discussion of genetic algorithms (GAs) will be given here, for more detail the reader is referred to the standard introduction to the subject [117].

GAs are optimization algorithms developed by Holland [132], which evolve solutions in a manner based on the Darwinian principle of natural selection. They differ from more conventional optimization techniques in that they work on whole populations of encoded solutions. Each possible solution, in this case each set of possible model parameters, is encoded as a gene. The most usual form for this gene is a binary string, e.g. 0001101010 gives a 10-bit (i.e. accurate to one part in 1024) representation of a parameter. In this illustration, two codes were used:

the first which will be called the *interval code*, is obtained by multiplying a small number Δp_i by the integer obtained from the bit-string, for each parameter p_i. The second code, the *range code*, is obtained by mapping the expected range of the parameter onto $[0, 1023]$ for example.

Having decided on a representation, the next step is to generate, at random, an initial population of possible solutions. The number of genes in a population depends on several factors, including the size of each individual gene, which itself depends on the size of the solution space.

Having generated a population of random genes, it is necessary to decide which of them are fittest in the sense of producing the best solutions to the problem. To do this, a fitness function is required which operates on the encoded genes and returns a single number which provides a measure of the suitability of the solution. These fitter genes will be used for mating to create the next generation of genes which will hopefully provide better solutions to the problem. Genes are picked for mating based on their fitnesses. The probability of a particular gene being chosen is equal to its fitness divided by the sum of the fitnesses of all the genes in the population. Once sufficient genes have been selected for mating, they are paired up at random and their genes combined to produce two new genes. The most common method of combination used is called *crossover*. Here, a position along the genes is chosen at random and the substrings from each gene after the chosen point are switched. This is one-point crossover. In two-point crossover a second position is chosen and the gene substrings switched again. There is a natural fitness measure for identification problem, namely the inverse of the comparison error between the reference data and the model data (see later).

The basic problem addressed here is to construct a mathematical model of an input–output system given a sampled-data record of the input time series $x(t)$ and the corresponding output series $y(t)$ (for displacement say). The 'optimum' model is obtained by minimizing the error between the reference data $y(t)$, and that produced by the model $\hat{y}(t)$ when presented with the sequence $x(t)$. The error function used here is the MSE defined in (6.108), the fitness for the GA is obtained simply by inverting the MSE.

If a gene in a particular generation is extremely fit, i.e. is very close to the required solution, it is almost certain to be selected several times for mating. Each of these matings, however, involves combining the gene with a less fit gene so the maximum fitness of the population may be lower in the next generation. To avoid this, a number of the most fit genes can be carried through unchanged to the next generation. These very fit genes are called the *elite*.

To prevent a population from stagnating, it can be useful to introduce perturbations into the population. New entirely random genes may be added at each generation. Such genes are referred to as *new blood*. Also, by analogy with the biological process of the same name, genes may be *mutated* by randomly switching one of their binary digits with a small probability. The mutation used here considers each bit of each gene separately for switching.

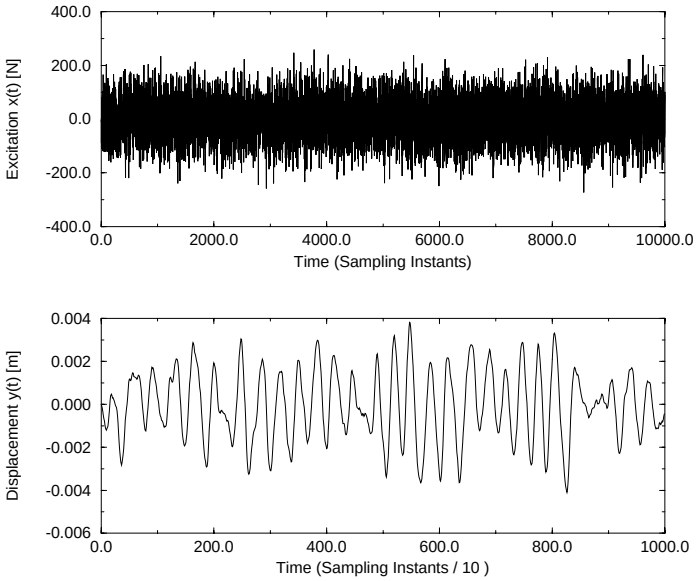


Figure 7.61. Force and displacement reference data for genetic algorithm (GA) identification of a linear system.

With genetic methods it is not always possible to say what the fitness of a perfect gene will be. Thus the iterative process is usually continued until the population is dominated by a few relatively fit genes. One or more of these genes will generally be acceptable as solutions.

7.6.1.2 A linear system

Before proceeding to nonlinear systems, it is important to establish a benchmark, so the algorithm is applied to data from a linear system. For simplicity, the systems considered here are all single-degree-of-freedom (SDOF); this does not represent a limitation of the method. Input and output data were obtained for the system given by

$$m\ddot{y} + c\dot{y} + ky = x(t) \tag{7.96}$$

with $m = 1$ kg, $c = 20$ N s m^{-1} and $k = 10^4$ N m^{-1}, using a fourth-order Runge–Kutta routine. $x(t)$ was a sequence of 10 000 points of Gaussian white noise with rms 75.0 and time step 0.0002. The resulting $y(t)$ was decimated by a factor of 10 giving 1000 points of reference data with sampling frequency 500 Hz. The data are shown in figure 7.61.

The methods of identifying this system shown previously in this chapter would require the availability of displacement, velocity and acceleration data. An advantage of the current method is the need for only one response variable.

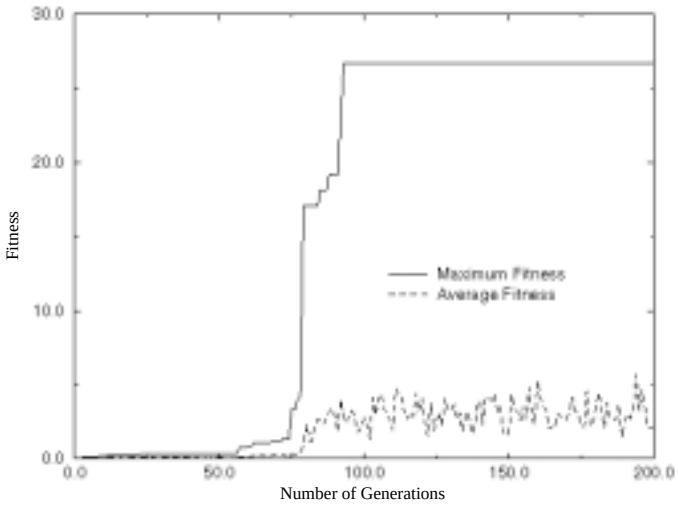


Figure 7.62. Evolution of fitness for GA identification of the linear system.

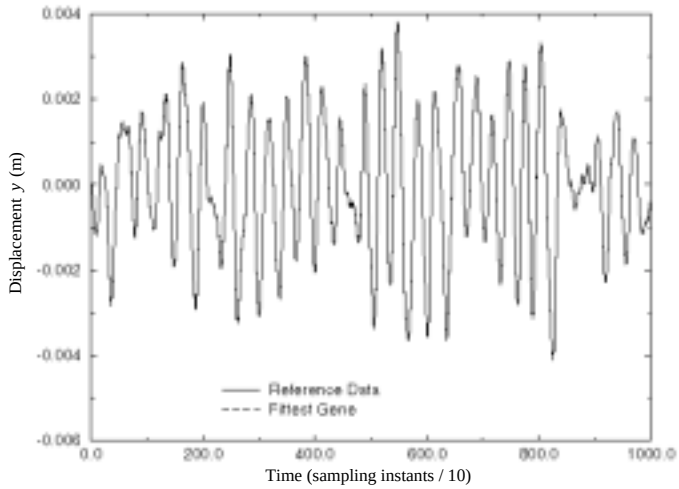


Figure 7.63. Comparison of measured and predicted displacements from GA identification of the linear system.

For the GA, each parameter m, c and k was coded as a 10-bit segment using the interval code with $\Delta m = 0.01$, $\Delta c = 0.1$ and $\Delta k = 20$. This gave a 30-bit gene. The fitness was evaluated by decoding the gene and running the Runge–Kutta routine with the estimated parameters and $x(t)$. The MSE for the model data $\hat{y}$ was obtained and inverted. The GA ran with a population of 50 for 200

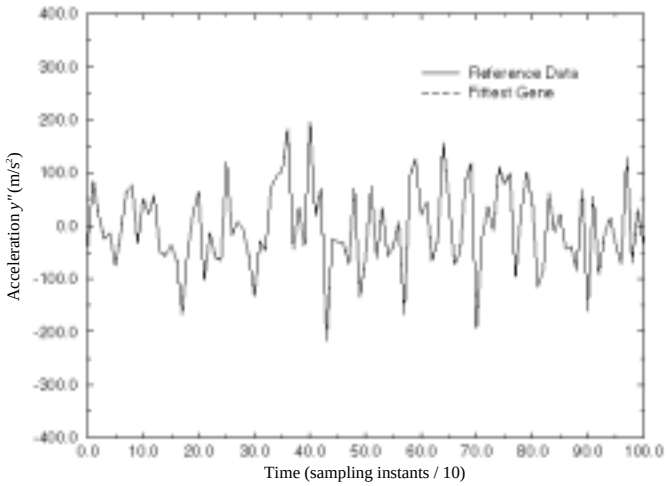


Figure 7.64. Comparison of measured and predicted accelerations from GA identification of the linear system.

generations. It used a single-member elite and introduced five new blood at each generation. The crossover probability was 0.6 and two-point crossover was used. The mutation probability was 0.08. The evolution of the maximum fitness and average fitness is given in figure 7.62. The optimum solution was found at about generation 100 and gave parameters $m = 1.03$, $c = 19.9$ and $k = 10\,280.0$ with a comparison error of 0.04. Figure 7.63 shows the resulting comparison of reference data and model data, the two traces are essentially indistinguishable. Processing for each generation was observed to take approximately 16 s. As the main overhead is fitness evaluation, this could have been been speeded up by a factor of about 10 by using a 1000 point input record with the same time step as the response.

In practice, the response most often measured is acceleration. It is a trivial matter to adapt the GA accordingly. One simply takes the acceleration data from the Runge–Kutta routine for reference and model data. A simulation was carried out using force and acceleration data (the same statistics for $x(t)$ and the same time step as before was used). Using the same GA parameters as before produced parameters $m = 1.01$, $c = 20.0$ and $k = 10\,240.0$ after 25 generations. The MSE for this solution is 0.02. A comparison of model and reference data is given in figure 7.64.

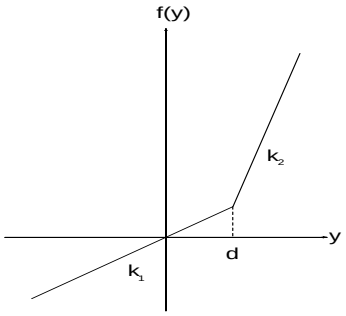


Figure 7.65. Simulated bilinear stiffness under investigation using the GA.

7.6.1.3 A piecewise linear system

The first nonlinear system considered here is a bilinear system with equation of motion

$$m\ddot{y} + c\dot{y} + f(y) = x(t) \tag{7.97}$$

with m and c as before. $f(y)$ has the form (figure 7.65)

$$f(y) = \left.\begin{array}{ll} k_1 y & \\ k_1 d + k_2(y-d) & \end{array}\right\} \begin{array}{l} y < d \\ y \geq d \end{array} \tag{7.98}$$

with $k_1 = 1000.0$, $k_2 = 10\,000.0$ and $d = 0.001$. This system is only physically sensible, i.e. $f(y)$ goes through the origin, if d is positive. This is not a restriction, in the general case, one could allow several negative and positive break points. It is the complicated dependence on d which makes $f(y)$ a problem for standard parameter estimation routines. However, there is essentially nothing to distinguish this system from the linear one from the point of view of the GA. The only difference is that five parameters are needed, so a 50-bit gene is required if the same precision is retained. In this experiment, the same GA parameters as before were used but the code was the range code and linear fitness scaling was used [117]. The ranges were $[0, 100]$ for m and c, $[0, 20\,000]$ for k_1 and k_2 and $[-0.025, 0.025]$ for d. Displacement was used for this run as the bilinear stiffness produces a significant mean level in the displacement which might provide a useful feature for the identification. The GA obtained a solution after 250 generations with $m = 0.952$, $c = 19.91$, $k_1 = 935.1$, $k_2 = 10\,025.6$ and $d = 0.001\,06$ and then failed to refine this further. The resulting comparison of model and reference data gave an MSE of 0.19.

To improve the solution, the GA was run again with the ranges taken from the final population of the previous run. The ranges were $[0.928, 1.147]$ for m, $[19.91, 21.47]$ for c, $[930.3, 11\,566.4]$ for k_1, $[9344.4, 11\,251.4]$ for k_2 and $[0.000\,85, 0.013\,55]$ for d. All other GA parameters were retained. The GA attained a fitness of 205.8 after 400 generations, corresponding to an MSE of

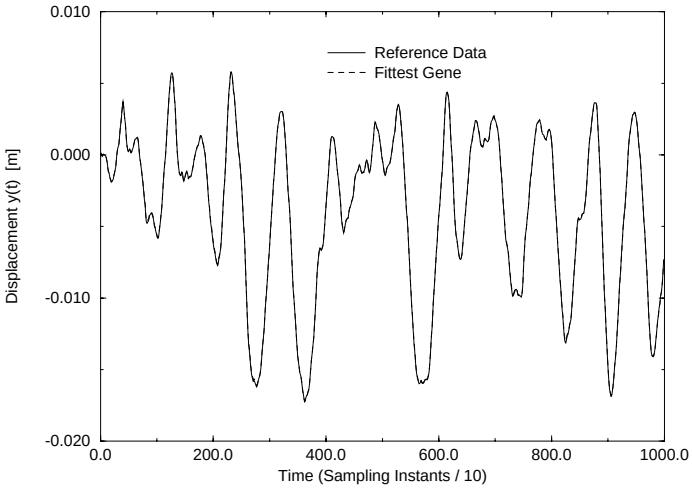


Figure 7.66. Comparison of measured and predicted displacements from GA identification of the bilinear system.

0.005. The final parameters were $m = 1.015$, $c = 20.03$, $k_1 = 1008.0$, $k_2 = 10\,300.0$ and $d = 0.001$. A comparison between model and reference data is given in figure 7.66, the two traces are indistinguishable.

7.6.1.4 A hysteretic system

The Bouc–Wen model [44, 263], briefly discussed before, is a general nonlinear hysteretic restoring force model where the total restoring force $Q(y, \dot{y})$ is composed of a polynomial non-hysteretic and a hysteretic component based on the displacement time history. A general hysteretic system described by Wen is represented next, where $g(y, \dot{y})$ is the polynomial part of the restoring force and $z(y)$ the hysteretic:

$$x(t) = m\ddot{y} + z(y, \dot{y}) \tag{7.99}$$

$$Q(y, \dot{y}) = g(y, \dot{y}) + z(y) \tag{7.100}$$

$$g(y, \dot{y}) = f(y) + h(\dot{y}). \tag{7.101}$$

Where the polynomial function $g(y, \dot{y})$ is separated into its displacement and velocity components

$$f(y) = b_0 \operatorname{sign}(y) + b_1 y + b_2 |y| y + b_3 y^3 + \cdots \tag{7.102}$$

$$h(\dot{y}) = a_0 \operatorname{sign}(\dot{y}) + a_1 \dot{y} + a_2 |\dot{y}| \dot{y} + a_3 \dot{y}^3 + \cdots. \tag{7.103}$$

The system under test here is an SDOF system based on this model with $g(y, \dot{y}) = 0$ for simplicity as studied in [178] $x(t)$ is a random force, with the

hysteretic component $z(y)$ defined in [263] by

$$\dot{z} = \begin{cases} -\alpha|\dot{y}|z^n - \beta\dot{y}|z^n| + A\dot{y}, & \text{for } n \text{ odd} \\ -\alpha|\dot{y}|z^{n-1}|z| - \beta\dot{y}z^n + A\dot{y}, & \text{for } n \text{ even.} \end{cases} \tag{7.104}$$

This may be reduced to

$$\frac{\mathrm{d}z}{\mathrm{d}x} = A \pm (\alpha \pm \beta)z^n. \tag{7.105}$$

Equation (7.105) may be integrated in closed form to show the hysteretic relationship between z and y where A, α, β and n are the constants that govern the scale, shape and smoothness of the hysteresis loop.

A series of experiments were performed using the GA to identify the parameters of a Bouc–Wen based SDOF system as presented earlier. The parameters of the reference system were $m = 1.0$ kg, $n = 2$, $\alpha = 1.5\ \mathrm{N}^{1-n}\ \mathrm{m}^{-1}$, $\beta = -1.5\ \mathrm{N}^{1-n}\ \mathrm{m}^{-1}$ and $A = 6680.0\ \mathrm{N\,m}^{-1}$. The reference system (figure 7.67) and the systems generated from the GA runs were driven with the same Gaussian noise input signal with an rms value of 10.0 N. Reference and experimental data were compared on 2000 data points using the MSE of the displacement data.

The fitness score for each gene was obtained by the reciprocal of the MSE as before. Genes whose parameters matched exactly those of the reference system resulted in a 'divide by zero' which was one test condition for termination of the program. Genes that described systems that produced outputs either falling to zero or exploding to infinity were assigned a zero-fitness score.

The data were sampled at 250 Hz giving a step size of $h = 0.004$. The results included in figures 7.68–7.70 are from the same GA test run with the following parameters.

Population size	500
Gene length	62
Number of generation	100
Crossover type	two-point
Crossover probability	80%
Mutation probability	10%
Number of elite genes	1
Number of new blood	50.

The peak fitness achieved by a gene in the test run was 4.53 giving a corresponding MSE of 0.22%. Figure 7.68 is given to compare the displacement signals of the reference and test system, the plots overlay. Figure 7.69 shows a comparison of the hysteresis loops.

The results shown are the best of many GA runs made while good simulation parameters were being determined. The average GA run achieved a lower fitness than this, but with results still being in the region of only 1% error.

The parameters decoded from the fittest gene in this case were as in table 7.7.

The peak fitness was achieved near the 35th generation of the test. Table 7.7 shows the final output from the run after 100 generations. Figure 7.70 shows the

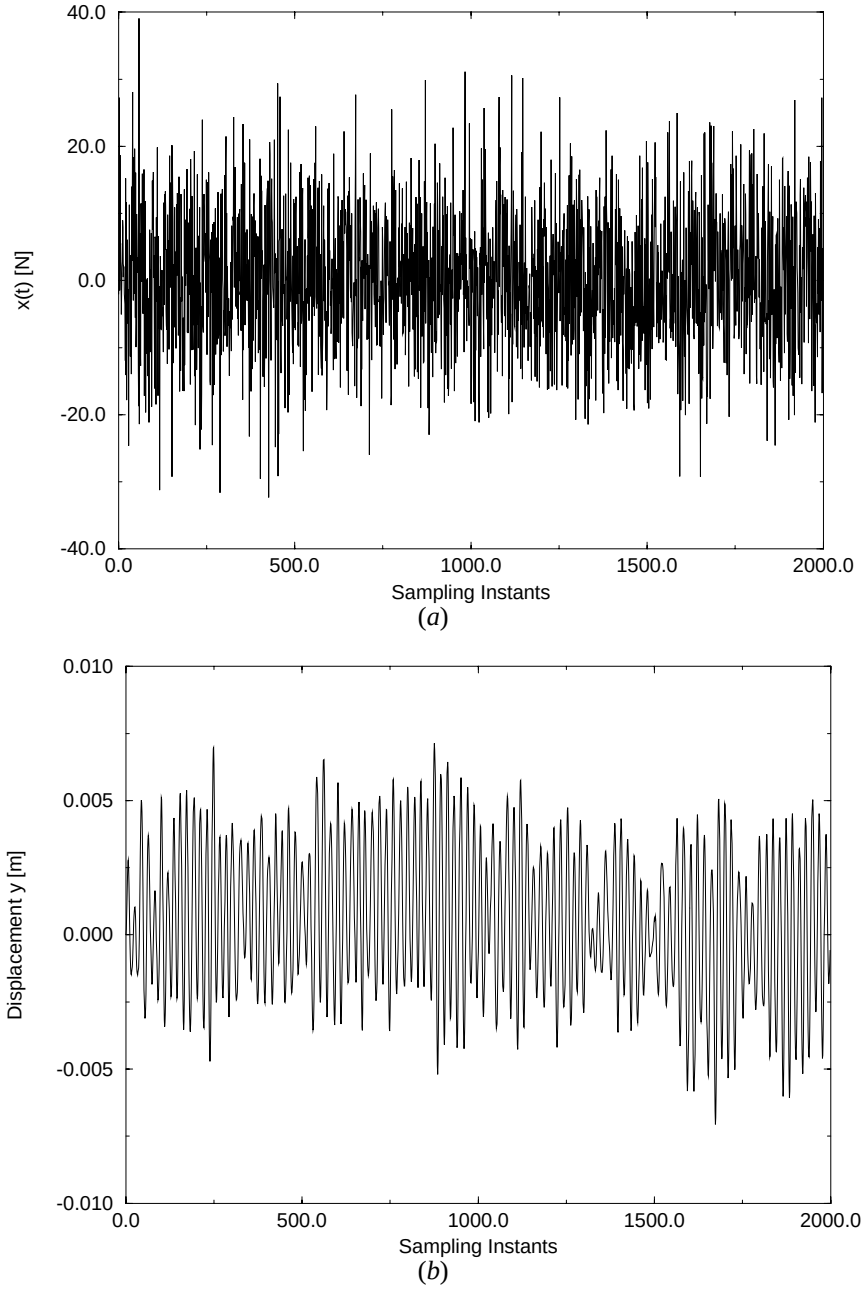


Figure 7.67. Force and displacement reference data for GA identification of the hysteretic system.

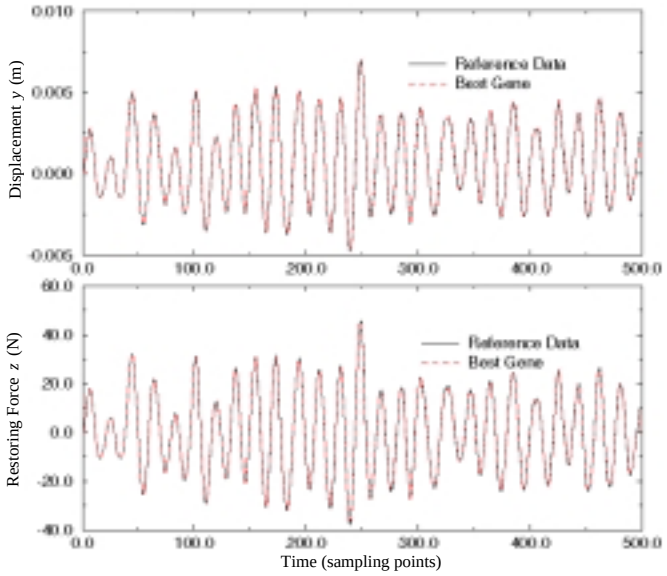


Figure 7.68. Comparison of measured and predicted displacements and internal states z from GA identification of the hysteretic system.

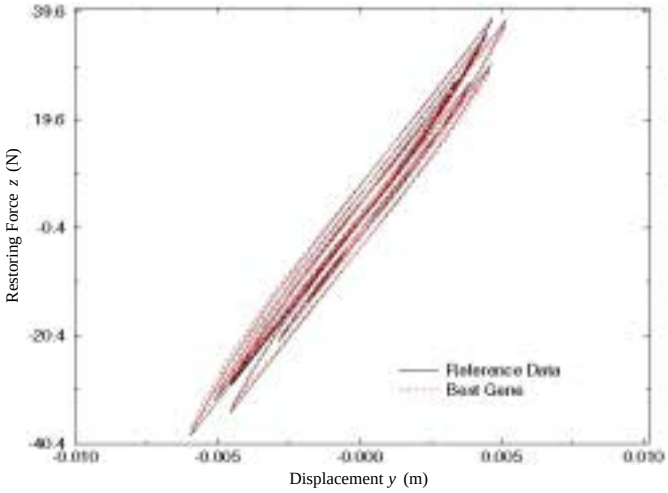


Figure 7.69. Comparison of measured and predicted hysteresis loops from GA identification of the hysteretic system.

growth of both the maximum fitness in the population and the overall average fitness. The average fitness is almost an order of magnitude lower than the

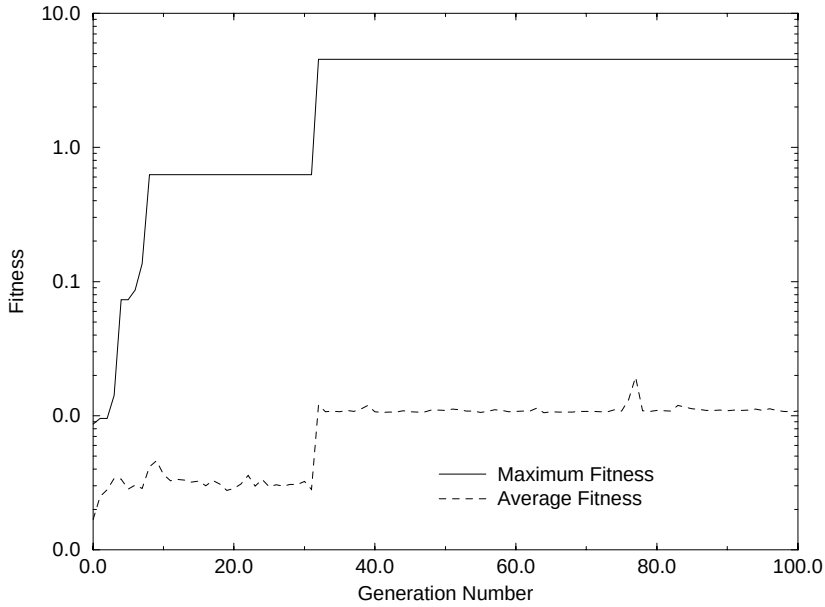


Figure 7.70. Evolution of fitness for GA identification of the linear system.

Table 7.7. Best parameter from GA for hysteretic system.

Parameter	Reference	Best gene	Error (%)
m	1.00	0.97	3
n	2	2	0
α	1.50	1.64	9
β	−1.50	−1.75	17
A	6680	6450	4

maximum; this is a result of the high mutation rates that were used to prevent premature population stagnation.

The method is an improvement on that of [178], in which it was assumed that m was known. The advantage of this assumption is that it allows the separation of the z-variable and reduces the problem to a linear-in-the-parameters estimation; this, however, is unrealistic.

Optimization, and GAs in particular, provide an attractive means of estimating parameters for otherwise troublesome systems. Physical parameters are obtained without going through a discrete model or using costly instrumentation or signal processing to generate displacement, velocity and acceleration data. It is a simple matter to use the algorithm with any of these

response types. The GA is not unique in this respect, any optimization engine could be used here which uses a single scalar objective function and does not require gradients or Hessians. Downhill simplex or simulated annealing should produce similar results. The method could be used for discrete-time systems and would allow the minimization of more effective metrics than the one-step-ahead prediction error. If gradients are available, they may be used with profit as discussed in the next section.

7.6.2 Identification of a shock absorber model using gradient descent

7.6.2.1 The hyperbolic tangent model

The background to shock absorber modelling is given in section 7.4.5, the motivation for the study here was the non-parametric approach to modelling taken in [109], in which a neural network was used to predict the value of the force transmitted by the absorber as a function of lagged displacement and velocity measurements. In the course of this work it was observed that the neural network transfer function—the hyperbolic tangent—bears more than a passing resemblance to the force–velocity characteristics of many shock absorbers (figure 7.50—obtained from a sine test). Many shock absorber force–velocity curves show near-linear behaviour at the higher velocities in the operating range (i.e. the blow-off region), with a smooth transition to high damping centred around zero velocity (i.e. the bleed region). Such functions can be obtained by scaling, translating and rotating a hyperbolic tangent function. The proposed form of the damping force is,

$$f_{\mathrm{d}}(\dot{y}) = c\dot{y} + \alpha[\tanh(\beta\dot{y} + \gamma) - \tanh(\gamma)]. \tag{7.106}$$

For the purposes of testing the absorber, an experimental facility was designed which allowed the possibility of adding mass and a parallel stiffness to the shock absorber (as described a little later). This means that (7.106) should be extended to

$$m\ddot{y} + c\dot{y} + ky + \alpha[\tanh(\beta\dot{y} + \gamma) - \tanh(\gamma)] = x(t) \tag{7.107}$$

which is a simple SDOF nonlinear oscillator (figure 7.71). The usual physical characteristics of the oscillator are represented by m, c, k while α, β, γ characterize the nonlinear damping (figure 7.71). Apart from the additional nonlinear damping, this equation agrees with the minimal model of the suspension system proposed by De Carbon [74], in which case m would be one-quarter of the car-body mass. This minimal model captures much of the essential behaviour of more complex models of the suspension. Note that this model has the structure of a very simple neural network with a linear output neuron (appendix F), as shown in figure 7.72.

There is no physical basis for the new model. The parameters are not related to the structure of the absorber but rather to its behaviour as quantified in the

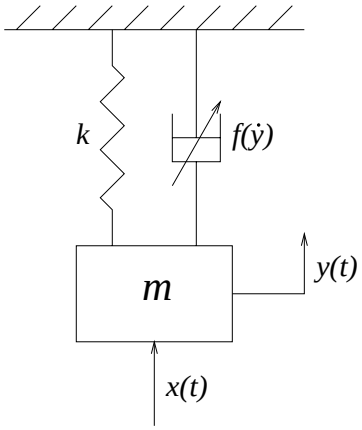


Figure 7.71. Nonlinear De Carbon lumped-mass model of the shock absorber.

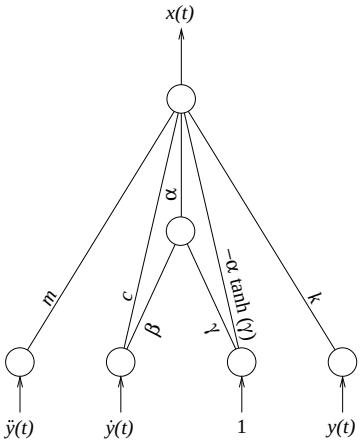


Figure 7.72. Neural network structure of the shock absorber model.

force–velocity curve. This is also the case for polynomial models where,

$$f_{\mathrm{d}}(\dot{y}) = \sum_{i=1}^{N_p} c_i \dot{y}^i \tag{7.108}$$

so it is natural to make a comparison. The De Carbon model corresponding to (7.108) is

$$m\ddot{y} + \sum_{i=1}^{N_p} c_i \dot{y}^i + ky = x(t). \tag{7.109}$$

The advantage of such models is that, with a small number of parameters,

the representation of the suspension system can be improved considerably. The polynomial models can be estimated using the LS methods described earlier in this chapter.

7.6.2.2 Gradient descent parameter estimation

The parameter estimation problem for the model structure (7.107) is a little more complicated as the expression is not linear in the parameters. This means, amongst other things, that it will not always be possible to obtain a global optimum. However, bearing this in mind, numerous methods are available for attacking this type of problem [114]. Given that the model has the structure of a neural network, it seemed appropriate to use a gradient descent or *back-propagation* scheme (appendix F).

The parameter estimate obtained in this case is optimal in the least-squared error sense, i.e. it minimizes $J = \sum_{i=1}^{N} \zeta_i^2$, where

$$\zeta_i = \hat{m}\ddot{y}_i + \hat{c}\dot{y}_i + \hat{k}y_i + \hat{\alpha}[\tanh(\hat{\beta}\dot{y}_i + \hat{\gamma}) - \tanh(\hat{\gamma})] - x_i \tag{7.110}$$

where y_i, $\dot{y}_i$ and $\ddot{y}_i$ are the sampled displacement, velocity and acceleration, and $\hat{m}$ etc are estimates of the parameters. The procedure is iterative; given a current estimate of the parameters, the next estimate is formed by stepping down along the gradient of the error function J; i.e. at step k

$$\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k + \Delta\boldsymbol{\theta}_k = \boldsymbol{\theta}_k - \eta\nabla_{\boldsymbol{\theta}_k} J(\boldsymbol{\theta}_k) \tag{7.111}$$

where the parameters have been ordered in the vector $\boldsymbol{\theta} = (\hat{m}, \hat{c}, \hat{k}, \hat{\alpha}, \hat{\beta}, \hat{\gamma})^{\mathrm{T}}$. The *learning coefficient* η determines the size of the descent step. In order to obtain the parameter update rule, it only remains to obtain the components of the gradient term in (7.110):

$$\nabla_{\boldsymbol{\theta}_k} J(\boldsymbol{\theta}_k) = \left(\frac{\partial J}{\partial m}, \frac{\partial J}{\partial c}, \frac{\partial J}{\partial k}, \frac{\partial J}{\partial \alpha}, \frac{\partial J}{\partial \beta}, \frac{\partial J}{\partial \gamma} \right). \tag{7.112}$$

(As confusion is unlikely to result, the carets denoting estimated quantities will be suppressed in the following discussion.) The update rules are obtained using the definition of J and (7.110). In forming the error-sum J it is not necessary to sum over the residuals for all N points; J can be obtained from a subset of the errors or even the single error which arises from considering one set of measurements $\{x_i, y_i, \dot{y}_i, \ddot{y}_i\}$, i.e.

$$J_i(\theta_i) = \zeta_i^2. \tag{7.113}$$

(In neural network terms, the *epoch* constitutes a single presentation of data.) The

latter course is adopted here and the resulting update rules for the parameters are

$$\begin{aligned}
\Delta m_i &= -\eta\zeta_i\ddot{y}_i \\
\Delta c_i &= -\eta\zeta_i\dot{y}_i \\
\Delta k_i &= -\eta\zeta_i y_i \\
\Delta\alpha_i &= -\eta\zeta_i[\tanh(\beta_i\dot{y}_i+\gamma_i)-\tanh(\gamma_i)] \\
\Delta\beta_i &= -\eta\zeta_i\alpha_i\dot{y}_i\,\mathrm{sech}^2(\beta_i\dot{y}_i+\gamma_i) \\
\Delta\gamma_i &= -\eta\zeta_i\alpha_i[\mathrm{sech}^2(\beta_i\dot{y}_i+\gamma_i)-\mathrm{sech}^2(\gamma_i)]
\end{aligned} \tag{7.114}$$

with ζ_i the resulting error on using the measurements labelled by i at *this iteration* (this will clearly be different at the next presentation of the values labelled by i). In keeping with normal practice in back-propagation, the value of i is chosen randomly between 1 and N at each iteration. Also, a momentum term was added to the iteration to help damp out high-frequency oscillations over the error surface (appendix F). The final update scheme was, therefore,

$$\Delta\boldsymbol{\theta}_k = -\eta\nabla_{\boldsymbol{\theta}_k}J_k(\boldsymbol{\theta}_k)+\mu\Delta\boldsymbol{\theta}_{k-1} \tag{7.115}$$

where μ is the momentum coefficient.

It is well known that nonlinear estimation schemes can be sensitive to the initial estimates; in order to obtain favourable starting values for the iteration, a linear model of the form

$$m_l\ddot{y}+c_l\dot{y}+k_l y = x(t) \tag{7.116}$$

was fitted first, the estimates m_l and k_l were used as starting values for the coefficients m and k in the nonlinear model; the estimate c_l was divided evenly between c and α in the absence of any obvious prescription. The initial values of β and γ were set at 1.0 and 0.0 respectively.

In order to validate the algorithm, data were generated by numerical integration for the system

$$6.3\ddot{y}+75\dot{y}+6300y+2000[\tanh(\dot{y}-0.25)-\tanh(-0.25)] = x(t). \tag{7.117}$$

The coefficient values were motivated by a desire to expose the parameter estimator to the same conditions as might be expected for a real absorber sub-assembly. At low levels of excitation, the effective damping coefficient is $c+\alpha\beta$, in this case 5.2 times critical; at high levels, the effective coefficient is c, giving 0.18 times critical. Data were obtained by taking $x(t)$ to be a Gaussian white noise sequence, initially of rms 6000, band-limited into the interval 0–20 Hz. The equation of motion (7.117) was stepped forward in time using a standard fourth-order Runge–Kutta procedure with a time step of 0.01 s; 10 000 sets of data $\{x_i, y_i, \dot{y}_i, \ddot{y}_i\}$ were obtained.

The algorithm was applied to the simulation data, using learning and momentum coefficients of 0.2 and 0.3 respectively. As the data were noise-free,

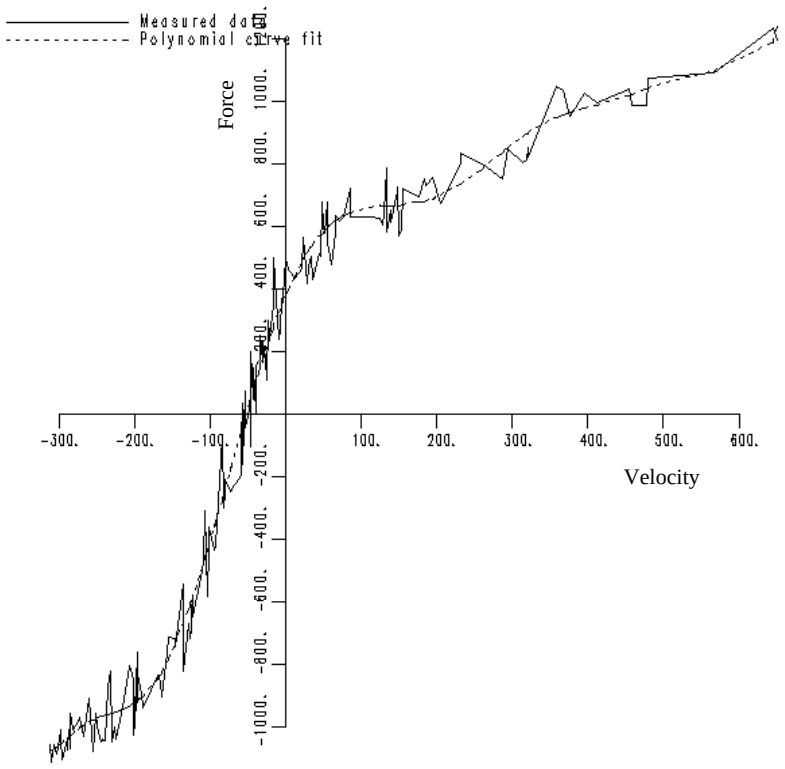


Figure 7.73. Force–velocity curve from shock absorber experiment compared with ninth-order polynomial model fit.

the iteration was required to terminate once the estimates had stabilized to within a fractional tolerance of 10^{-8}. This level of convergence was reached after 15 006 iterations (essentially covering the whole data set twice); the resulting estimates were

$$m = 6.300\,0001$$
$$c = 74.999\,723$$
$$k = 6300.0005$$
$$\alpha = 2000.0012$$
$$\beta = 0.999\,999\,35$$
$$\gamma = -0.249\,999\,84.$$

This gives confidence in the estimator. In practice, the true values will not be

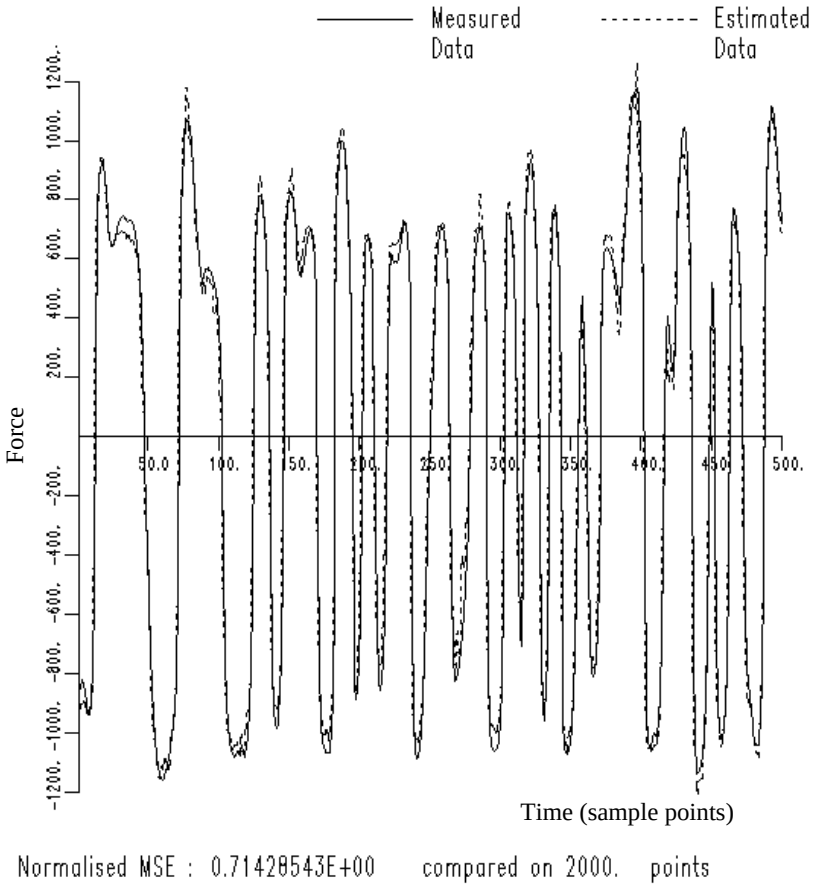


Figure 7.74. Force data from shock absorber experiment compared with ninth-order polynomial model fit.

known and some other objective measure of confidence will be required for the estimates. The measure used here was the normalized mean-square error or $\mathrm{MSE}(\hat{x})$.

7.6.2.3 Results using experimental data

The shock absorber test facility essentially took the form of figure 7.51. Facilities were provided to add a parallel stiffness in the form of a spring of known characteristics and to load the system with an additional mass. This option was not used for the particular test described here. As the shock absorber is essentially

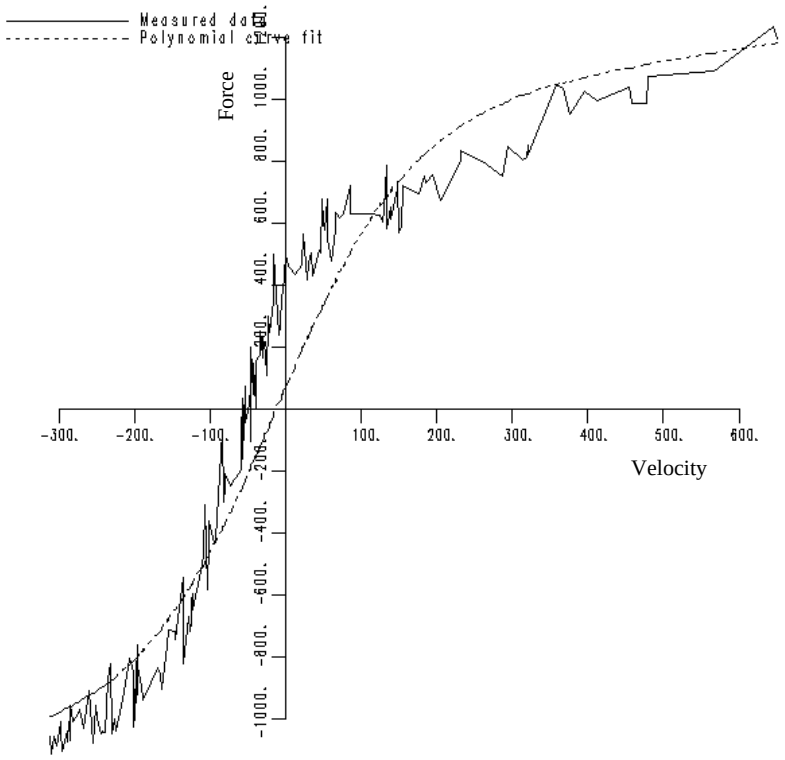


Figure 7.75. Force–velocity curve from shock absorber experiment compared with hyperbolic tangent model.

an SDOF system under vertical excitation in this configuration, the simple model of figure 7.71 applies. The excitation for the system was provided by the random signal generator of a spectrum analyser, amplified and filtered into the interval 2–30 Hz. The band-limited signal facilitates post-processing of measured data, i.e. numerical differentiation or integration (appendix I).

The piezoelectric load cell provided a measurement of $x(t)$. The other signal measured was displacement, the required velocity and acceleration being arrived at by numerical differentiation. This decision was made because the actuator actually incorporates an LVDT (linear voltage displacement transducer) which produces a high quality signal. A detailed account of the test structure and instrumentation can be found in [50].

For the particular test considered here, a displacement of 3.0 mm rms was applied at the base of the absorber and 7000 samples of x_i and y_i were obtained at

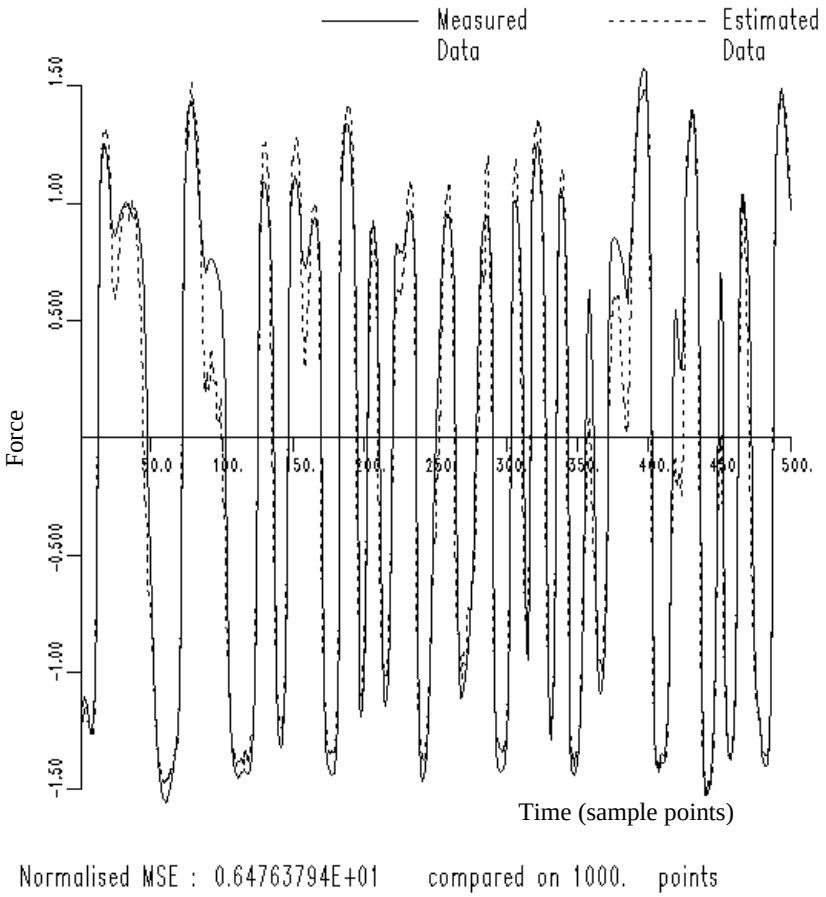


Figure 7.76. Force data from shock absorber experiment compared with hyperbolic tangent model.

a frequency of 500 Hz. A three-point centred difference was used to obtain the $\dot{y}$ and $\ddot{y}$ data. The characteristic force–velocity curve (the full curve in figure 7.73), was obtained using the sectioning method described earlier in this chapter.

Polynomial models were fitted to the data for various model orders, the masses (as expected) could be disregarded as insignificant. In fact, the stiffnesses could also be discarded as their contribution to the total variance of the right-hand side vector $\{x\}$ was small. The resulting models for the damping force f_{d} gave $\mathrm{MSE}(\hat{f}_{\mathrm{d}})$ values:

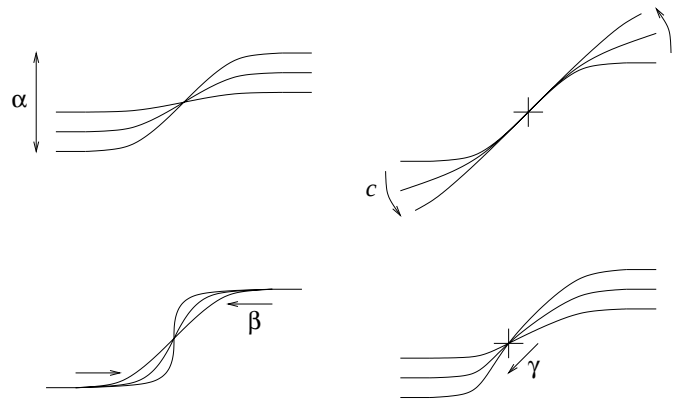


Figure 7.77. Behaviour of the hyperbolic tangent function under variation of the parameters.

Model order	MSE
1	15.5
3	5.8
5	1.9
7	1.1
9	0.9

The curve-fit for the ninth-order polynomial model is given in figure 7.73. The corresponding model-predicted force is given in figure 7.74.

The parameter estimation routine of the last section was applied to 1000 points of data, using a learning coefficient of 0.1 and no momentum. Convergence of the parameters to within a fractional tolerance of 0.000 05 was obtained after 16 287 iterations, the resulting parameters being,

$$m = -0.0007$$
$$c = 0.369$$
$$k = -5.6$$
$$\alpha = 942.8$$
$$\beta = 0.000\,56$$
$$\gamma = 0.0726.$$

The mass and stiffness can be disregarded as insignificant as before, the negative signs can probably be regarded as statistical fluctuations. The MSE value was 6.9 which shows quite good agreement. Figure 7.76 shows a comparison between the measured force time series and that predicted by the model (7.107). Figure 7.75

shows a comparison between the measured force–velocity curve and that of the model. Agreement is quite good.

7.6.2.4 Discussion

The results of the last section show that a better representation of the force–velocity curve could be obtained using high-order polynomials; however, it could be argued that the model (7.107) is preferable for two reasons:

(1) Polynomial models are restricted to the excitation levels at or below the level used for parameter estimation. The reason for this is that a polynomial, on leaving the interval on which the model is defined, will tend to $\pm\infty$ as $\mathrm{O}(x^n)$ depending on the sign and order of the leading term. In many cases this leads to instability because a negative leading term will tend to reinforce rather than oppose the motion at high velocities (see figure 7.73). Alternatively, (7.107) leads asymptotically to linear damping.

(2) The polynomial coefficients will not usually admit a physical interpretation. In the case of the model (7.106) or (7.107), the coefficients have a direct interpretation in terms of the force–velocity characteristics; c generates rotations (shear really) and fixes the asymptotic value of the damping; α governs the overall scale of the central high damping region and β the gradient; variations in γ translate the high damping region along the velocity scale while maintaining a zero force condition at zero velocity (figure 7.77). These characteristics are the main features of interest to designers and have a direct bearing on subjective ride comfort evaluation. The model developed here may also facilitate comparisons between real absorbers.

This concludes the main discussions on system identification. The book now returns to the idea of the FRF and discusses how the concept may be generalized to nonlinear systems.

Chapter 8

The Volterra series and higher-order frequency response functions

8.1 The Volterra series

In the first chapter it was shown that linear systems admit dual time- and frequency-domain characterizations[1]:

$$y(t) = \int_{-\infty}^{\infty} \mathrm{d}\tau \, h(\tau) x(t-\tau) \tag{8.1}$$

and

$$Y(\omega) = H(\omega) X(\omega). \tag{8.2}$$

All information about a single-input–single-output (SISO) system is encoded in either the impulse response function $h(t)$ or the frequency response function (FRF) $H(\omega)$. The representation to be used in a given problem will usually be dictated by the form of the answer required. In vibration problems, the frequency-domain approach is usually adopted; displaying the FRF $H(\omega)$ shows immediately those frequencies at which large outputs can be expected, i.e. peaks in $H(\omega)$ corresponding to the system resonances.

Equations (8.1) and (8.2) are manifestly linear and therefore cannot hold for arbitrary nonlinear systems; however, both admit a generalization. The extended form of equation (8.1) was obtained in the early part of this century by Volterra

[1] There are of course other characterizations. The set $\{m, c, k\}$ fixes the behaviour of a linear SDOF system in just the same way as the functional forms $h(t)$ and $H(\omega)$ do, and arguably provides a more parsimonious means of doing so. However, the $h(t)$ and $H(\omega)$ can provide a *visual* representation that communicates the likely behaviour of the system in a way that the set of numbers does not. A more meaningful combination of the parameters, say $\{m, \zeta, \omega_n\}$ conveys better understanding to the average structural dynamicist. In the case of a SISO (single-input–single-output) continuous system, *all* the representations involve infinite-dimensional sets and the distinction becomes otiose. The authors would like to thank Dr Steve Gifford for discussion on this point.

[261]. It takes the form of an infinite series[2]

$$y(t) = y_1(t) + y_2(t) + y_3(t) + \cdots \tag{8.3}$$

where

$$y_1(t) = \int_{-\infty}^{+\infty} \mathrm{d}\tau\, h_1(\tau) x(t-\tau) \tag{8.4}$$

$$y_2(t) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \mathrm{d}\tau_1\, \mathrm{d}\tau_2\, h_2(\tau_1, \tau_2) x(t-\tau_1) x(t-\tau_2) \tag{8.5}$$

$$y_3(t) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \mathrm{d}\tau_1\, \mathrm{d}\tau_2\, \mathrm{d}\tau_3\, h_2(\tau_1, \tau_2, \tau_3) x(t-\tau_1) x(t-\tau_2) x(t-\tau_3). \tag{8.6}$$

The form of the general term is obvious from the previous statements. The functions $h_1(\tau), h_2(\tau_1, \tau_2), h_3(\tau_1, \tau_2, \tau_3), \ldots h_n(\tau_1, \ldots, \tau_n), \ldots$ are generalizations of the linear impulse response function and are usually referred to as *Volterra kernels*. The use of the Volterra series in dynamics stems from the seminal paper of Barrett [20], in which the series was applied to nonlinear differential equations for the first time. One can think of the series as a generalization of the Taylor series from functions to functionals. The expression (8.1) simply represents the lowest-order truncation which is, of course, exact only for linear systems.

The derivation of the series is beyond the scope of this book, but heuristic arguments can be found in [261, 25, 221]. Note that these kernels are not forced to be symmetric in their arguments. In fact, any non-symmetric kernel can be replaced by a symmetric version with impunity so that $h_2(\tau_1, \tau_2) = h_2(\tau_2, \tau_1)$ etc. A formal proof is fairly straightforward; for simplicity, consider the expression for $y_2(t)$:

$$y_2(t) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \mathrm{d}\tau_1\, \mathrm{d}\tau_2\, h_2(\tau_1, \tau_2) \Pi_2(\tau_1, \tau_2; t) \tag{8.7}$$

with the newly-defined

$$\Pi_2(\tau_1, \tau_2; t) = x(t-\tau_1) x(t-\tau_2) \tag{8.8}$$

and note that Π_2 is manifestly symmetric in its arguments τ_1 and τ_2.

[2] The term *weak* nonlinearity has occasionally appeared in this book without a convincing definition. The Volterra series allows at least a mathematically precise characterization if one defines a weak nonlinear system as one that admits a representation in terms of a Volterra expansion. Because the Volterra series is essentially a polynomial representation it cannot describe systems with multi-valued responses. As a result, this definition of weak agrees with the more imprecise idea that *strongly* nonlinear systems are those that exhibit the sort of bifurcations that result in subharmonic or chaotic behaviour.

Assuming that h_2 has no particular symmetries, it still has a canonical decomposition into symmetric and antisymmetric parts:

$$h_2(\tau_1, \tau_2) = h_2^{\mathrm{sym}}(\tau_1, \tau_2) + h_2^{\mathrm{asym}}(\tau_1, \tau_2) \tag{8.9}$$

where

$$\begin{aligned} h_2^{\mathrm{sym}}(\tau_1, \tau_2) &= \tfrac{1}{2}(h_2(\tau_1, \tau_2) + h_2(\tau_2, \tau_1)) \\ h_2^{\mathrm{asym}}(\tau_1, \tau_2) &= \tfrac{1}{2}(h_2(\tau_1, \tau_2) - h_2(\tau_2, \tau_1)). \end{aligned} \tag{8.10}$$

Now, consider the contribution to $y_2(t)$ from the antisymmetric component of the kernel:

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} h_2^{\mathrm{asym}}(\tau_1, \tau_2) \Pi_2(\tau_1, \tau_2; t)\, \mathrm{d}\tau_1\, \mathrm{d}\tau_2. \tag{8.11}$$

Any (infinitesimal) contribution to this 'summation', say at $\tau_1 = v, \tau_2 = w$ will cancel with the corresponding contribution at $\tau_2 = v, \tau_1 = w$, as

$$h_2^{\mathrm{asym}}(v, w)\Pi_2(v, w; t) = -h_2^{\mathrm{asym}}(w, v)\Pi_2(w, v; t) \tag{8.12}$$

and the overall integral will vanish. This is purely because of the 'contraction' or summation against the symmetric quantity $\Pi_2(\tau_1, \tau_2; t)$. Because h_2^{asym} makes no contribution to the quantity $y_2(t)$ it may be disregarded and the kernel h_2 can be assumed to be symmetric. Essentially, the h_2 picks up all the symmetries of the quantity Π_2. This argument may be generalized to the kernel $h_n(\tau_1, \ldots, \tau_n)$.

In general, for a symmetric kernel, h_n^{sym} is obtained by summing all of the possible permutations of the argument, weighted by an inverse factor which counts the terms. The following section describes a method of extracting the kernel transforms directly, which automatically selects the symmetric kernel. This method will be adopted throughout the remainder of the book. For this reason, the identifying label 'sym' will be omitted on the understanding that *all* kernels and kernel transforms are symmetric. For information about other conventions for kernels, mainly the *triangular* form, the reader can consult [217].

As previously stated, there exists a dual frequency-domain representation for nonlinear systems. The *higher-order FRFs* or *Volterra kernel transforms* $H_n(\omega_1, \ldots, \omega_n)$, $n = 1, \ldots, \infty$ are defined as the multi-dimensional Fourier transforms of the kernels, i.e.

$$\begin{aligned} H_n(\omega_1, \ldots, \omega_n) = &\int_{-\infty}^{+\infty} \ldots \int_{-\infty}^{+\infty} \mathrm{d}\tau_1 \ldots \mathrm{d}\tau_n\, h_n(\tau_1, \ldots, \tau_n) \\ &\times \mathrm{e}^{-\mathrm{i}(\omega_1\tau_1 + \cdots + \omega_n\tau_n)} \end{aligned} \tag{8.13}$$

$$\begin{aligned} h_n(\tau_1, \ldots, \tau_n) = &\frac{1}{(2\pi)^n} \int_{-\infty}^{+\infty} \ldots \int_{-\infty}^{+\infty} \mathrm{d}\omega_1 \ldots \mathrm{d}\omega_n\, H_n(\omega_1, \ldots, \omega_n) \\ &\times \mathrm{e}^{+\mathrm{i}(\omega_1\tau_1 + \cdots + \omega_n\tau_n)}. \end{aligned} \tag{8.14}$$

It is a simple matter to show that symmetry of the kernels implies symmetry of the kernel transforms so, for example, $H_2(\omega_1, \omega_2) = H_2(\omega_2, \omega_1)$.

It is then a straightforward matter to obtain the frequency-domain dual of the expression (8.3)

$$Y(\omega) = Y_1(\omega) + Y_2(\omega) + Y_3(\omega) + \cdots \tag{8.15}$$

where

$$Y_1(\omega) = H_1(\omega)X(\omega) \tag{8.16}$$

$$Y_2(\omega) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \mathrm{d}\omega_1\, H_2(\omega_1, \omega - \omega_1)X(\omega_1)X(\omega - \omega_1) \tag{8.17}$$

$$\begin{aligned} Y_3(\omega) = {} & \frac{1}{(2\pi)^2}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \mathrm{d}\omega_1\, \mathrm{d}\omega_2\, H_3(\omega_1, \omega_2, \omega - \omega_1 - \omega_2) \\ & \times X(\omega_1)X(\omega_2)X(\omega - \omega_1 - \omega_2). \end{aligned} \tag{8.18}$$

The fundamental problem associated with the Volterra series is the determination of either the kernels or the kernel transforms. This must be done analytically if the equations of motion are known or numerically if time series are given for the input and output processes. Section 8.3 will consider what happens if the equations of motion are known, but first some motivation for use of the series will be given.

8.2 An illustrative case study: characterization of a shock absorber

Before proceeding to the main body of the theory of functional series, it is useful to pause and consider what sort of problems one might apply them to. This section illustrates their use on a real engineering system, namely a shock absorber. The system considered will be a *Monroe–McPherson strut*; this is simply a coil spring mounted over an automotive damper of the sort briefly discussed in the previous chapter. It is characterized by a linear stiffness and a nonlinear damper. The work described in this section was carried out by Dr Steve Cafferty and a much more detailed discussion can be found in [50].

The experimental arrangement is shown in figure 7.51. The higher-order FRFs are obtained by a harmonic testing approach. First the system is tested without the coil spring and then with.

There are one or two interesting features of this problem: first, the force in the shock absorber without the spring is a function of the velocity, not the displacement, i.e. assuming linear viscous damping

$$f(t) = c_1\dot{y}. \tag{8.19}$$

The first-order FRF of interest is for the process $y \longrightarrow F$ and this is termed the *dynamic stiffness*. A simple calculation yields

$$\frac{F(\omega)}{Y(\omega)} = H_1(\omega) = \mathrm{i}c_1\omega \tag{8.20}$$

and it follows that the dynamic stiffness varies linearly with ω and the gradient is the linear damping coefficient. The presence of the imaginary term simply shows that the displacement and force are in quadrature (90° out of phase).

The first task is to establish $H_1(\omega)$. The experimental procedure is a standard stepped-sine test. The system is excited by a displacement signal, a sinusoid $Y\cos(\omega t)$ at a given frequency and the amplitude and phase of the force response $F\cos(\omega t - \phi)$ are recorded. The gain and phase of $H_1(\omega)$ are simply F/Y and ϕ as discussed in chapter 1.

In reality it is not quite as simple as this because the damper is nonlinear. Assuming a polynomial expansion up to third order gives

$$f(t) = c_1\dot{y} + c_2\dot{y}^2 + c_3\dot{y}^3. \tag{8.21}$$

Just as the first-order FRF is completely specified by c_1, the higher-order coefficients are encoded in the higher-order FRFs. Anticipating equation (8.32) (in the form appropriate for velocity nonlinearity) gives, for a harmonic input $y(t) = \mathrm{e}^{\mathrm{i}\Omega t}$, $\dot{y} = \mathrm{i}\Omega\mathrm{e}^{\mathrm{i}\Omega t}$

$$f(t) = H_1(\Omega)\mathrm{i}\Omega\mathrm{e}^{\mathrm{i}\Omega t} - H_2(\Omega,\Omega)\Omega^2\mathrm{e}^{\mathrm{i}2\Omega t} - H_3(\Omega,\Omega,\Omega)\mathrm{i}\Omega^3\mathrm{e}^{\mathrm{i}3\Omega t} + \cdots \tag{8.22}$$

and the higher-order FRFs are read off from (8.21):

$$H_2(\Omega,\Omega) = -c_2\Omega^2 \tag{8.23}$$

$$H_3(\Omega,\Omega,\Omega) = -\mathrm{i}c_3\Omega^3. \tag{8.24}$$

The necessary experimental testing program follows from these formulae. In order to find $H_1(\omega)$ the standard linear stepped-sine procedure is used. In order to find $H_2(\omega,\omega)$, the amplitude and phase of the second harmonic is extracted, i.e. the amplitude and phase of the component at 2ω, to find $H_3(\omega,\omega,\omega)$, take the amplitude and phase of the component at 3ω. Note that this procedure only gives values on the *diagonal* line in the frequency space where $\omega_1 = \omega_2 = \cdots = \Omega$. For this reason, the quantities are called the *diagonal HFRFs*.

The second subtlety alluded to earlier comes into play here. The earlier argument assumes that the excitation is a pure harmonic $\mathrm{e}^{\mathrm{i}\Omega t}$ and this is impossible in practice as it is a complex quantity. In reality, a cosinusoid is used which is the sum of two harmonics, $\cos(\Omega) = (\mathrm{e}^{\mathrm{i}\Omega t} + \mathrm{e}^{-\mathrm{i}\Omega t})/2$. It will be shown later that this means that the quantities measured in the test are not the *pure* FRFs $H_i(\Omega,\ldots,\Omega)$. For example, the amplitude and phase of the component at 2Ω is equal to $H_2(\Omega,\Omega)+$ higher-order terms involving H_4, H_6 etc. Fortunately, it can be shown that the contamination of H_2 by the higher-order terms can be ignored

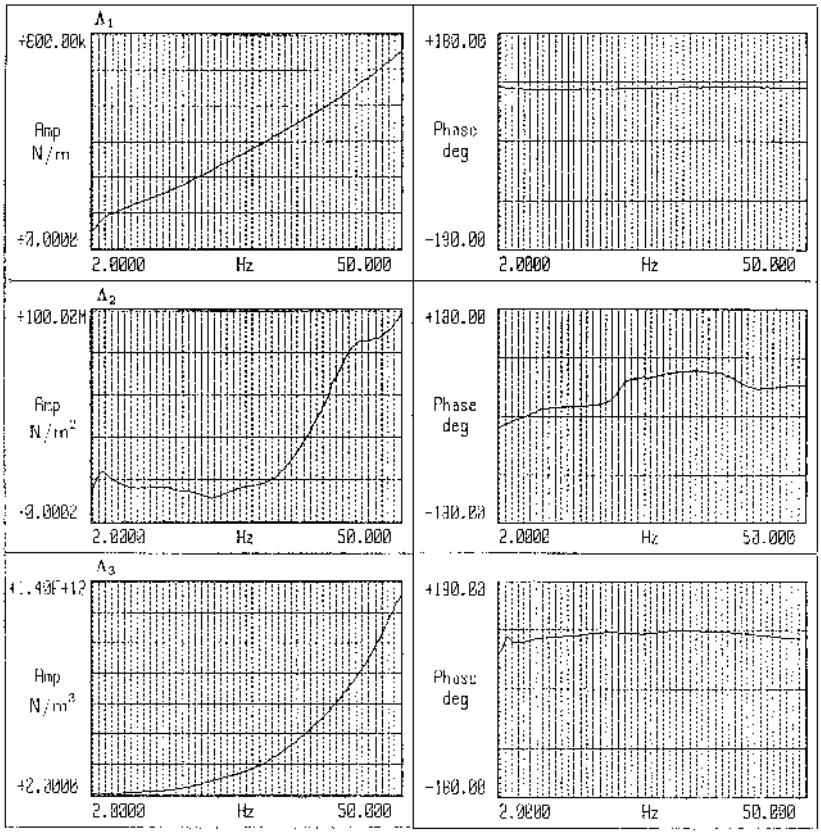


Figure 8.1. Principal diagonals of the first-, second- and third-order composite HFRFs for the shock absorber.

if the amplitude of excitation is small enough. However, in order to eliminate confusion, the *measured* FRFs will be termed *composite* FRFs and will be denoted by $\Lambda_{s1}(\Omega)$, $\Lambda_{s2}(\Omega, \Omega)$ etc. The s-subscript denotes that the FRFs are the result of a sine test.

Figure 8.1 shows the first three measured diagonal HFRFs in terms of amplitude and phase over the testing range 2–50 Hz for a low displacement amplitude. The assumption of linear growth of the Λ_{s1} appears well-justified, also the Λ_{s2} and Λ_{s3} curves have the required polynomial forms. Dividing Λ_{s1} by ω, Λ_{s2} by ω^2 etc should yield constant values by the previous arguments and figure 8.2 shows the results of these operations. At higher frequencies, the HFRFs tend to the required constants; however, there is some distortion at lower frequencies. The estimated coefficients are given in table 8.1. They show the 'softening' behaviour in the damping which might well be expected from

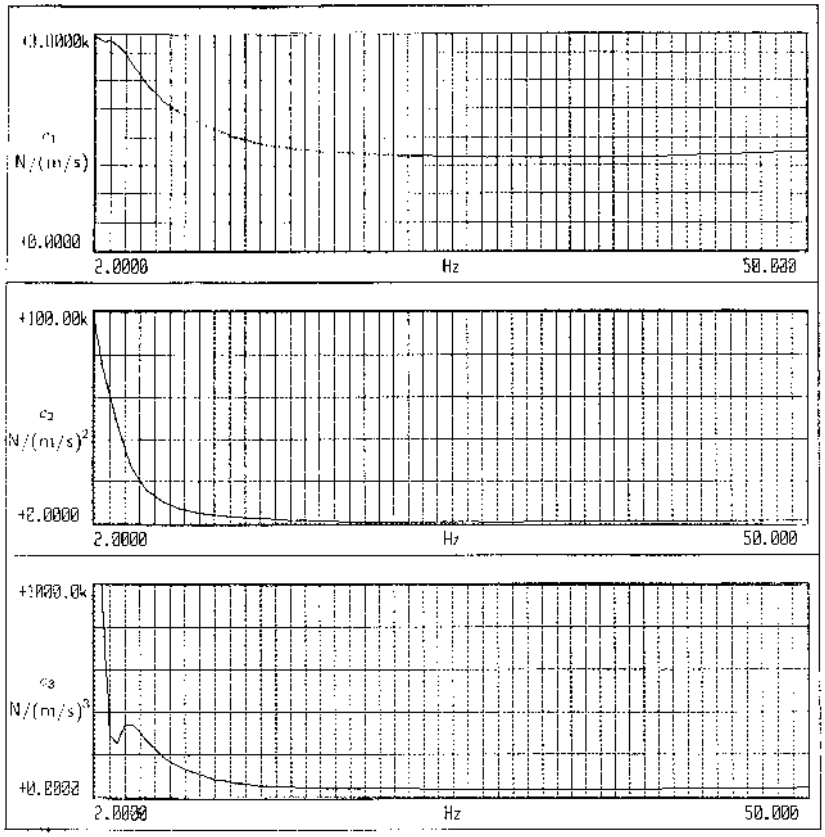


Figure 8.2. Nonlinear damping values for the shock absorber estimated from the principal diagonals.

Table 8.1. Parameter estimates for damping coefficients.

Coefficient	Estimate	Units
c_1	1 600.0	N s m^{-1}
c_2	−832.0	N s^2 m^{-2}
c_3	−38 500.0	N s^3 m^{-3}

characteristic diagrams of the absorber like that in figure 7.50.

The testing procedure is not restricted to producing the diagonal elements of the HFRF. For example, if a two-tone signal is used for the excitation by combining frequencies Ω_1 and Ω_2, then the amplitude and phase of the output

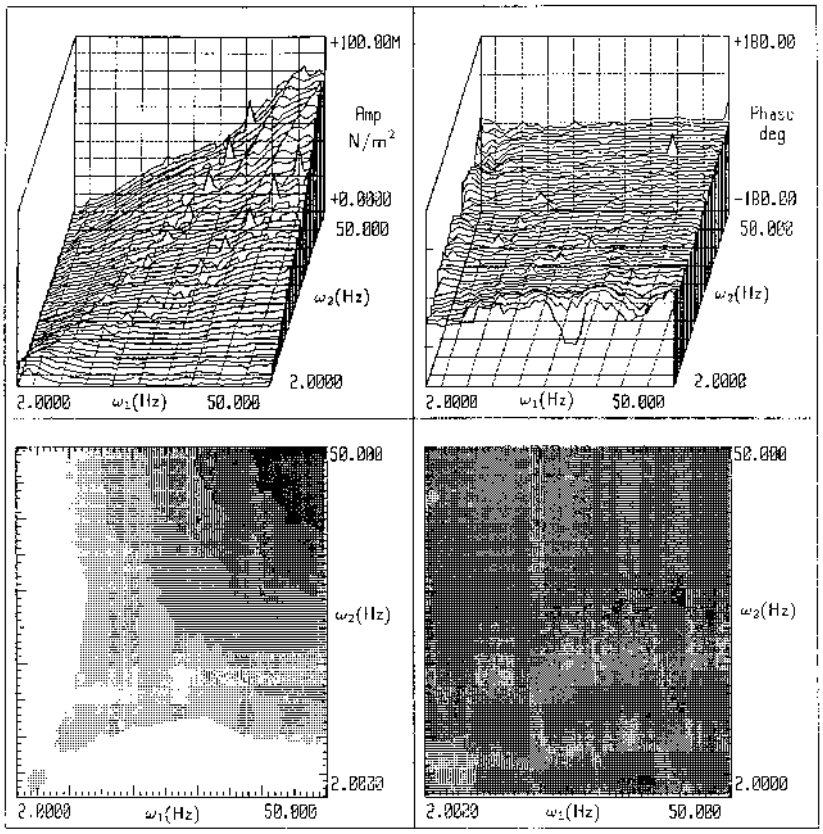


Figure 8.3. Principal quadrant of the second-order composite HFRF $\Lambda_2(\omega_1, \omega_2)$ for the shock absorber.

component at frequency $\Omega_1 + \Omega_2$ approximates the values for $2H_2(\Omega_1, \Omega_2)$. Again it is assumed that the level of excitation is low enough for contributions from H_4 etc to be ignored. Strictly, the measured quantity is the composite FRF $\Lambda_{s2}(\Omega_1, \Omega_2)$. Similarly, if three frequencies are used to excite the system, the amplitude and phase at the sum frequency approximates H_3. Figures 8.3 and 8.4 show $\Lambda_{s2}(\omega_1, \omega_2)$ and $\Lambda_{s3}(\omega_1, \omega_2, \omega_1)$ over the so-called 'principal quadrants'. (Note that it is not possible to plot Λ_{s3} in its full generality as it would require a representation of four-dimensional space.) There is very little structure in these plots, a very smooth variation of the HFRFs is observed with no resonances; this is to be expected of course as there is no stiffness in the system. The theory developed later in this chapter gives

$$H_2(\omega_1, \omega_2) = -c_2\omega_1\omega_2 \tag{8.25}$$

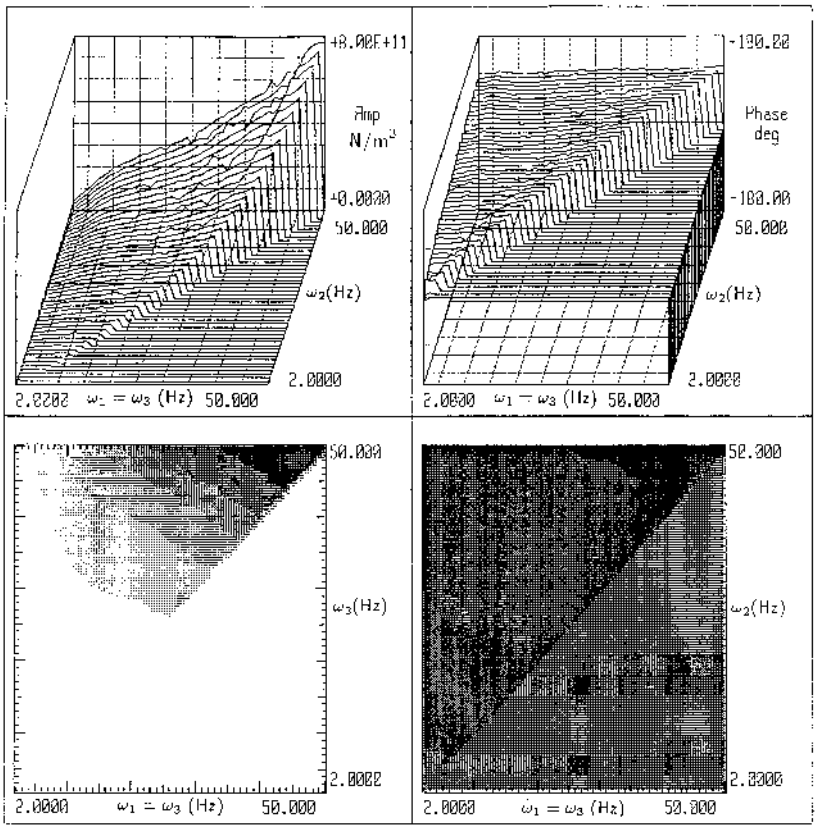


Figure 8.4. Principal quadrant of the third-order composite HFRF $\Lambda_3(\omega_1, \omega_2, \omega_3)$ for the shock absorber.

and

$$H_3(\omega_1, \omega_2, \omega_3) = -c_3\omega_1\omega_2\omega_3. \tag{8.26}$$

The second series of tests were with the coil spring in place. These produced slightly more structured HFRFs due to the internal resonances of the spring. Using basic elasticity theory, a dynamic stiffness FRF for the spring alone was estimated and is shown in figure 8.5, the resonances are clear. A monofrequency test gave the results shown in figure 8.6 for the diagonal composite HFRFs, the polynomial rise from the damping is combined with the spring resonances. A bifrequency test yielded the Λ_{s2} and Λ_{s3} shown in figures 8.7 and 8.8.

This section has shown how the HFRFs can be estimated using sine-testing and also how they allow a parametric identification of the damping characteristics of a shock absorber (although there are easier ways of obtaining estimates of the c_1, c_2 and c_3 as discussed in the previous chapter). The figures showing the

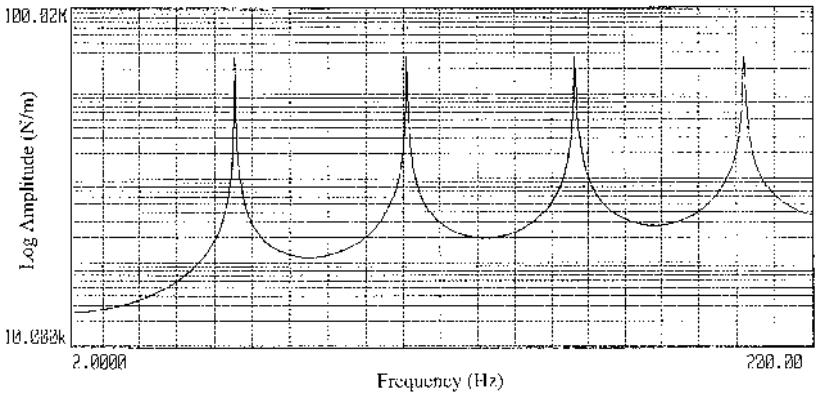


Figure 8.5. Simulated FRF showing the coil spring's first four resonant frequencies calculated from spring theory.

HFRFs themselves actually yield important non-parametric information about the system and the interpretation of the HFRFs is an important subject which will be returned to later. In the meantime, it is important to show how the theoretical HFRFs described earlier were obtained, and this forms the subject of the following section.

8.3 Harmonic probing of the Volterra series

The subject of this section is a direct method of determining the higher-order FRFs for a system given the equations of motion. The method of *harmonic probing* was introduced in [22] specifically for systems with continuous-time equations of motion. The method was extended to discrete-time systems in [30] and [256] An alternative, recursive approach to probing is presented in [205].

In order to explain the harmonic probing procedure, it is necessary to determine how a system responds to a harmonic input in terms of its Volterra series.

First consider a periodic excitation composed of a single harmonic

$$x(t) = \mathrm{e}^{\mathrm{i}\Omega t}. \tag{8.27}$$

The spectral representation of this function follows immediately from the well-known representation of the δ-function (appendix D):

$$\delta(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{d}\omega\, \mathrm{e}^{\mathrm{i}\omega t} \tag{8.28}$$

so that

$$X(\omega) = 2\pi\delta(\omega - \Omega). \tag{8.29}$$

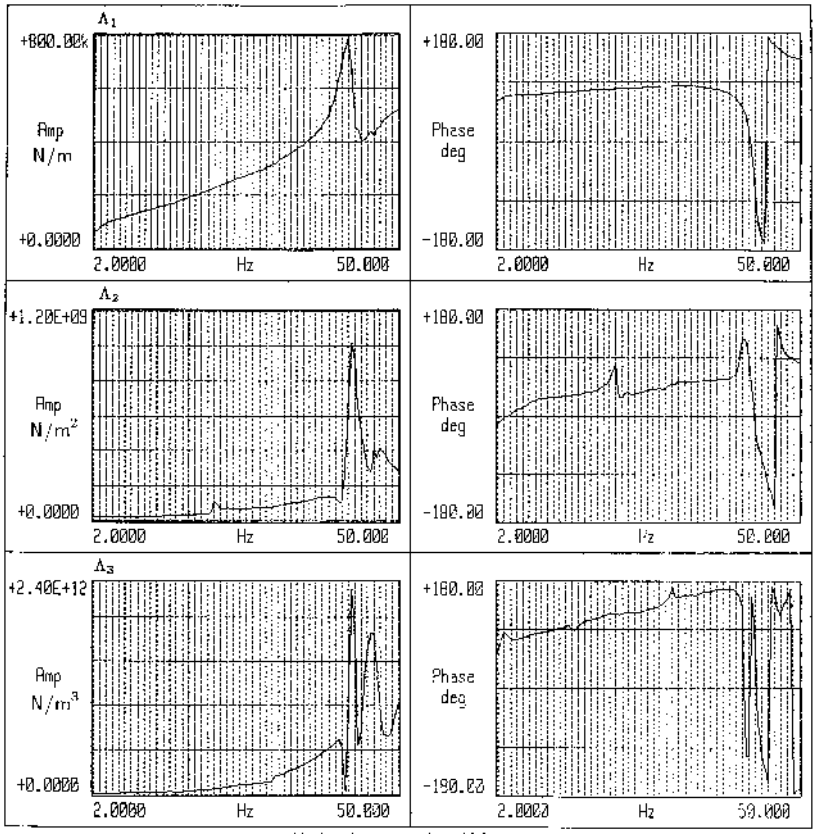


Figure 8.6. Principal diagonal of the first-, second- and third-order composite HFRFs for the shock absorber and coil spring at an input voltage of 0.5 V.

Substituting this expression into equations (8.16)–(8.18) and forming the total response as in (8.15) yields, up to third order,

$$\begin{aligned} Y(\omega) = {} & H_1(\omega)2\pi\delta(\omega-\Omega) + \frac{1}{2\pi}\int_{-\infty}^{+\infty} \mathrm{d}\omega_1\, H_2(\omega_1, \omega-\omega_1) \\ & \times 2\pi\delta(\omega_1-\Omega)2\pi\delta(\omega-\omega_1-\Omega) \\ & + \frac{1}{(2\pi)^2}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \mathrm{d}\omega_1\, \mathrm{d}\omega_2\, H_3(\omega_1, \omega_2, \omega-\omega_1-\omega_2) \\ & \times 2\pi\delta(\omega_1-\Omega)2\pi\delta(\omega_2-\Omega)2\pi\delta(\omega-\omega_1-\omega_2-\Omega) + \cdots \end{aligned} \tag{8.30}$$

using the argument-changing property of the δ-function and carrying out the

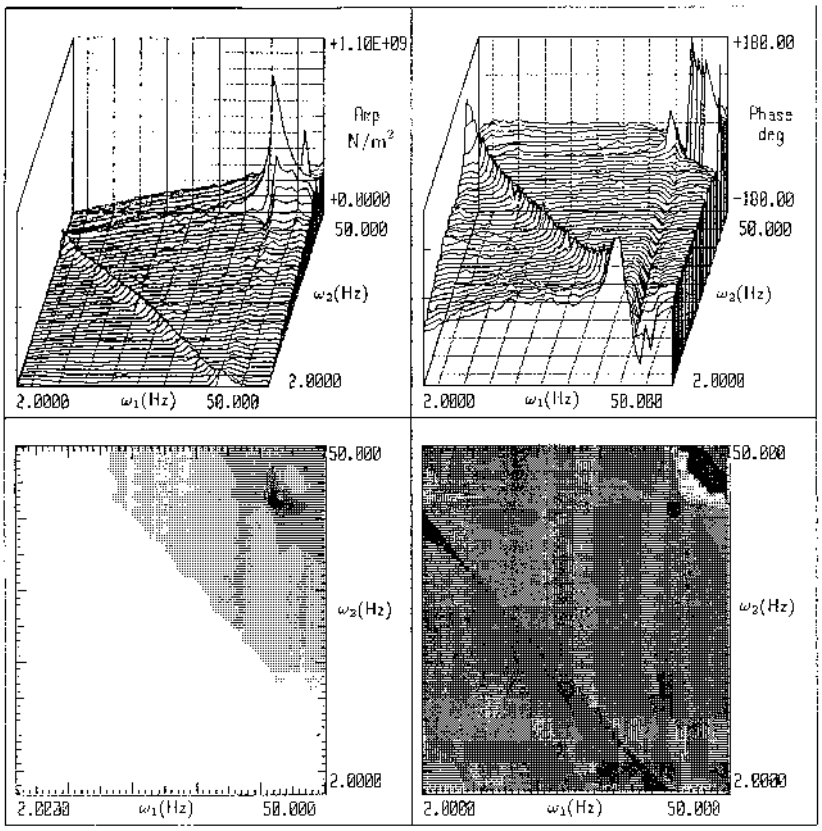


Figure 8.7. Principal quadrant of the second-order composite HFRF $\Lambda_2(\omega_1, \omega_2)$ for the shock absorber and coil spring at an input voltage of 0.5 V.

integrals gives

$$Y(\omega) = 2\pi\{H_1(\Omega)\delta(\omega - \Omega) + H_2(\Omega, \Omega)\delta(\omega - 2\Omega) + H_3(\Omega, \Omega, \Omega)\delta(\omega - 3\Omega) + \cdots\}. \tag{8.31}$$

Taking the inverse Fourier transform yields the required response:

$$y(t) = H_1(\Omega)\mathrm{e}^{\mathrm{i}\Omega t} + H_2(\Omega, \Omega)\mathrm{e}^{\mathrm{i}2\Omega t} + H_3(\Omega, \Omega, \Omega)\mathrm{e}^{\mathrm{i}3\Omega t} + \cdots. \tag{8.32}$$

This shows clearly that components in the output at multiples of the excitation frequency are expected, i.e. harmonics. The important point here is that the component in the output at the forcing frequency is $H_1(\Omega)$.

Probing the system with a single harmonic only yields information about the values of the FRFs on the diagonal line in the frequency spaces. In order to obtain

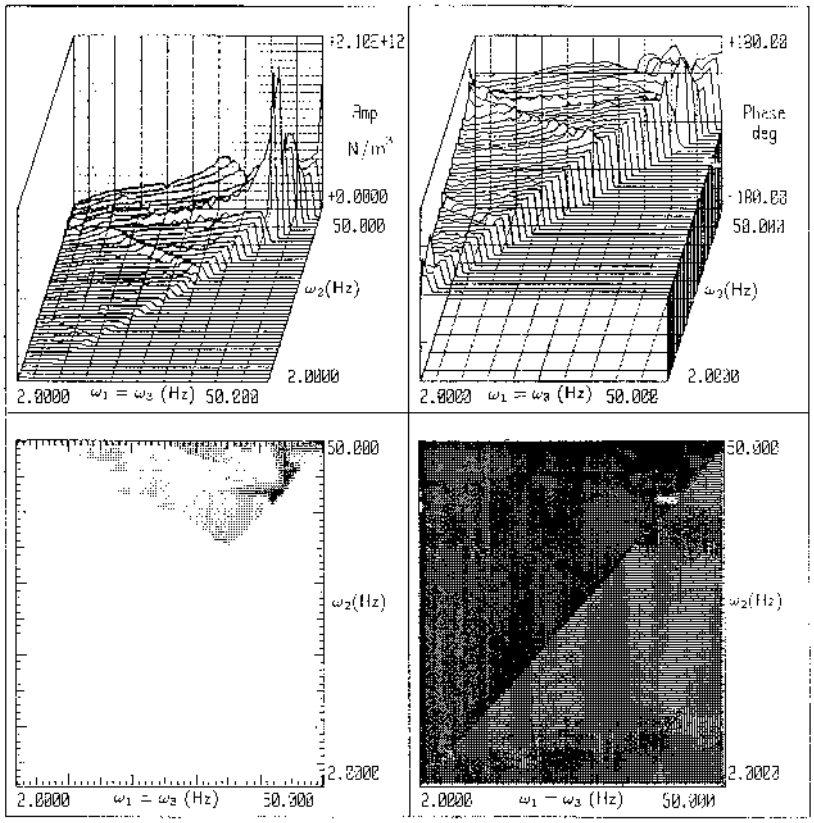


Figure 8.8. Principal quadrant of the third-order composite HFRF $\Lambda_3(\omega_1, \omega_2, \omega_3)$ for the shock absorber and coil spring at an input voltage of 0.5 V.

further information, multi-frequency excitations must be used. With this in mind, consider the 'two-tone' input

$$x(t) = \mathrm{e}^{\mathrm{i}\Omega_1 t} + \mathrm{e}^{\mathrm{i}\Omega_2 t} \tag{8.33}$$

which has spectral representation

$$X(\omega) = 2\pi\delta(\omega - \Omega_1) + 2\pi\delta(\omega - \Omega_2) \tag{8.34}$$

substituting into (8.16)–(8.18) and thence into (8.15) yields

$$\begin{aligned} Y(\omega) &= H_1(\omega)2\pi\delta(\omega - \Omega_1) + H_1(\omega)2\pi\delta(\omega - \Omega_2) \\ &\quad + \frac{1}{2\pi}\int_{-\infty}^{+\infty} \mathrm{d}\omega_1\, H_2(\omega_1, \omega - \omega_1)\{2\pi\delta(\omega_1 - \Omega_1) + 2\pi\delta(\omega_1 - \Omega_2)\} \end{aligned}$$

$$\times \{2\pi\delta(\omega-\omega_1-\Omega_1)+2\pi\delta(\omega-\omega_1-\Omega_2)\}$$
$$+\frac{1}{(2\pi)^2}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\mathrm{d}\omega_1\,\mathrm{d}\omega_2\,H_3(\omega_1,\omega_2,\omega-\omega_1-\omega_2)$$
$$\times \{2\pi\delta(\omega_1-\Omega_1)+2\pi\delta(\omega_1-\Omega_2)\}\{2\pi\delta(\omega_2-\Omega_1)+2\pi\delta(\omega_2-\Omega_2)\}$$
$$\times \{2\pi\delta(\omega-\omega_1-\omega_2-\Omega_1)+2\pi\delta(\omega-\omega_1-\omega_2-\Omega_2)\}+\cdots. \tag{8.35}$$

It is a straightforward but tedious matter to expand this expression and perform the integrals. After making use of the symmetry properties of the higher-order FRFs, namely $H(\omega_1,\omega_2)=H(\omega_2,\omega_1)$ and $H(-\omega_1,-\omega_2)=H^*(\omega_1,\omega_2)$, one obtains

$$\begin{aligned}\frac{Y(\omega)}{2\pi} &= H_1(\Omega_1)\delta(\omega-\Omega_1)+H_1(\Omega_2)\delta(\omega-\Omega_2)+H_2(\Omega_1,\Omega_1)\delta(\omega-2\Omega_1)\\ &\quad+2H_2(\Omega_1,\Omega_2)\delta(\omega-\Omega_1-\Omega_2)+H_2(\Omega_2,\Omega_2)\delta(\omega-2\Omega_2)\\ &\quad+H_3(\Omega_1,\Omega_1,\Omega_1)\delta(\omega-3\Omega_1)+3H_3(\Omega_1,\Omega_1,\Omega_2)\delta(\omega-2\Omega_1-\Omega_2)\\ &\quad+3H_3(\Omega_1,\Omega_2,\Omega_2)\delta(\omega-\Omega_1-2\Omega_2)+H_3(\Omega_2,\Omega_2,\Omega_2)\delta(\omega-3\Omega_2)\\ &\quad+\cdots.\end{aligned} \tag{8.36}$$

On taking the inverse Fourier transform, one obtains the response up to third order:

$$\begin{aligned}y(t) &= H_1(\Omega_1)\mathrm{e}^{\mathrm{i}t\Omega_1}+H_1(\Omega_2)\mathrm{e}^{\mathrm{i}t\Omega_2}\\ &\quad+H_2(\Omega_1,\Omega_1)\mathrm{e}^{\mathrm{i}t2\Omega_1}+2H_2(\Omega_1,\Omega_2)\mathrm{e}^{\mathrm{i}t(\Omega_1+\Omega_2)}+H_2(\Omega_2,\Omega_2)\mathrm{e}^{\mathrm{i}t2\Omega_2}\\ &\quad+H_3(\Omega_1,\Omega_1,\Omega_1)\mathrm{e}^{\mathrm{i}t3\Omega_1}+3H_3(\Omega_1,\Omega_1,\Omega_2)\mathrm{e}^{\mathrm{i}t(2\Omega_1+\Omega_2)}\\ &\quad+3H_3(\Omega_1,\Omega_2,\Omega_2)\mathrm{e}^{\mathrm{i}t(\Omega_1+2\Omega_2)}+H_3(\Omega_2,\Omega_2,\Omega_2)\mathrm{e}^{\mathrm{i}t3\Omega_2}+\cdots.\end{aligned} \tag{8.37}$$

The important thing to note here is that the amplitude of the component at the *sum* frequency for the excitation, i.e. at $\Omega_1+\Omega_2$ is twice the second-order FRF $H_2(\Omega_1,\Omega_2)$. In fact, if a general periodic excitation is used, i.e.

$$x(t)=\mathrm{e}^{\mathrm{i}\Omega t}+\cdots+\mathrm{e}^{\mathrm{i}\Omega_n} \tag{8.38}$$

it can be shown that the amplitude of the output component at the frequency $\Omega_1+\cdots+\Omega_n$ is $n!H_n(\Omega_1,\ldots,\Omega_n)$. This single fact is the basis of the harmonic probing algorithm. In order to find the second-order FRF of a system for example, one substitutes the expressions for the input (8.33) and general output (8.37) into the system equation of motion and extracts the coefficient of $\mathrm{e}^{\,\mathrm{i}(\Omega_1+\Omega_2)t}$; this yields an algebraic expression for H_2.

The procedure is best illustrated by choosing an example. Consider the continuous-time system

$$Dy+y+y^2=x(t) \tag{8.39}$$

where $D = \frac{\mathrm{d}}{\mathrm{d}t}$. In order to find H_1, the probing expressions

$$x(t) = x_1^{\mathrm{p}}(t) = \mathrm{e}^{\mathrm{i}\Omega t} \tag{8.40}$$

and

$$y(t) = y_1^{\mathrm{p}}(t) = H_1(\Omega)\mathrm{e}^{\mathrm{i}\Omega t} \tag{8.41}$$

are substituted into the equation (8.39), the result being

$$(\mathrm{i}\Omega + 1)H_1(\Omega)\mathrm{e}^{\mathrm{i}\Omega t} + H_1(\Omega)^2\mathrm{e}^{\mathrm{i}2\Omega t} = \mathrm{e}^{\mathrm{i}\Omega t} \tag{8.42}$$

equating the coefficients of $\mathrm{e}^{\mathrm{i}\Omega t}$ on each side of this expression yields an equation for H_1

$$(\mathrm{i}\Omega + 1)H_1(\Omega) = 1 \tag{8.43}$$

which is trivially solved, yielding the expression

$$H_1(\Omega) = \frac{1}{\mathrm{i}\Omega + 1}. \tag{8.44}$$

Evaluation of H_2 is only a little more complicated. The probing expressions

$$x(t) = x_2^{\mathrm{p}}(t) = \mathrm{e}^{\mathrm{i}\Omega_1 t} + \mathrm{e}^{\mathrm{i}\Omega_2 t} \tag{8.45}$$

and

$$y(t) = y_2^{\mathrm{p}}(t) = H_1(\Omega_1)\mathrm{e}^{\mathrm{i}\Omega_1 t} + H_1(\Omega_2)\mathrm{e}^{\mathrm{i}\Omega_2 t} + 2H_2(\Omega_1, \Omega_2)\mathrm{e}^{\mathrm{i}(\Omega_1+\Omega_2)t} \tag{8.46}$$

are used. Note that in passing from the general output (8.37) to the probing expression (8.46), all second-order terms except that at the sum frequency have been deleted. This is a very useful simplification and is allowed because no combination of the missing terms can produce a component at the sum frequency and therefore they cannot appear in the final expression for H_2. Substituting (8.45) and (8.46) into (8.39), and extracting the coefficients of $\mathrm{e}^{\mathrm{i}(\Omega_1+\Omega_2)t}$ yields

$$(\mathrm{i}(\Omega_1 + \Omega_2) + 1)H_2(\Omega_1, \Omega_2) + H_1(\Omega_1)H_1(\Omega_2) = 0. \tag{8.47}$$

So that

$$\begin{aligned} H_2(\Omega_1, \Omega_2) &= \frac{-H_1(\Omega_1)H_1(\Omega_2)}{\mathrm{i}(\Omega_1 + \Omega_2) + 1} = -H_1(\Omega_1)H_1(\Omega_2)H_1(\Omega_1 + \Omega_2) \\ &= \frac{-1}{(\mathrm{i}\Omega_1 + 1)(\mathrm{i}\Omega_2 + 1)(\mathrm{i}[\Omega_1 + \Omega_2] + 1)} \end{aligned} \tag{8.48}$$

on using the previously obtained expression for H_1.

The next example is a little more interesting. Consider the asymmetric Duffing equation

$$mD^2y + cDy + ky + k_2y^2 + k_3y^3 = x(t) \tag{8.49}$$

this time with the D notation.

H_1 and H_2 for this system can be evaluated by exactly the same procedure as used on the previous example. The results are

$$H_1(\omega) = \frac{1}{-m\omega^2 + \mathrm{i}c\omega + k} \tag{8.50}$$

$$H_2(\omega_1, \omega_2) = -k_2 H_1(\omega_1) H_1(\omega_2) H_1(\omega_1 + \omega_2). \tag{8.51}$$

Note that the constant k_2 multiplies the whole expression for H_2, so that if the square-law term is absent from the equation of motion, H_2 vanishes. This reflects a quite general property of the Volterra series; if all nonlinear terms in the equation of motion for a system are odd powers of x or y, then the associated Volterra series has no even-order kernels. As a consequence it will possess no even-order kernel transforms.

In order to obtain H_3, the required probing expressions are

$$x(t) = x_3^{\mathrm{p}}(t) = \mathrm{e}^{\mathrm{i}\omega_1 t} + \mathrm{e}^{\mathrm{i}\omega_2 t} + \mathrm{e}^{\mathrm{i}\omega_3 t} \tag{8.52}$$

and

$$\begin{aligned} y(t) = y_3^{\mathrm{p}}(t) = {} & H_1(\omega_1)\mathrm{e}^{\mathrm{i}\omega_1 t} + H_1(\omega_2)\mathrm{e}^{\mathrm{i}\omega_2 t} + H_1(\omega_3)\mathrm{e}^{\mathrm{i}\omega_3 t} \\ & + 2H_2(\omega_1, \omega_2)\mathrm{e}^{\mathrm{i}(\omega_1+\omega_2)t} + 2H_2(\omega_1, \omega_3)\mathrm{e}^{\mathrm{i}(\omega_1+\omega_3)t} \\ & + 2H_2(\omega_2, \omega_3)\mathrm{e}^{\mathrm{i}(\omega_2+\omega_3)t} + 6H_3(\omega_1, \omega_2, \omega_3)\mathrm{e}^{\mathrm{i}(\omega_1+\omega_2+\omega_3)t} \end{aligned} \tag{8.53}$$

which are sufficiently general to obtain H_3 for any system. Substituting into the Duffing equation and extracting the coefficient of $\mathrm{e}^{\mathrm{i}(\omega_1+\omega_2+\omega_3)t}$ yields

$$\begin{aligned} H_3(\omega_1, \omega_2, \omega_3) = {} & -\tfrac{1}{6} H_1(\omega_1 + \omega_2 + \omega_3) \\ & \times \{4k_2(H_1(\omega_1)H_2(\omega_2, \omega_3) + H_1(\omega_2)H_2(\omega_3, \omega_1) \\ & + H_1(\omega_3)H_2(\omega_1, \omega_2)) \\ & + 6k_3 H_1(\omega_1)H_1(\omega_2)H_1(\omega_3)\}. \end{aligned} \tag{8.54}$$

A discussion of the interpretation of these functions is deferred until a little later.

It is property of many systems that all higher-order FRFs can be expressed in terms of H_1 for the system. The exact form of the expression will depend on the particular system.

The harmonic probing algorithm has been established for continuous-time systems, i.e. those whose evolution is governed by differential equations of motion. The NARMAX models discussed in chapter 6 are difference equations so the probing algorithm requires a little modification as in [32] and [256]. Consider the difference equation analogue of equation (8.39):

$$\triangle y + y + y^2 = x(t). \tag{8.55}$$

where $\triangle$ is the backward shift operator, defined by $\triangle y(t) = y(t-1)$. (Throughout this chapter it is assumed, except where indicated, that the sampling

interval for a discrete-time system is scaled to unity. This yields a unit sampling frequency and Nyquist frequency of 0.5.) In the usual notation for difference equations, (8.55) becomes

$$y_{i-1} + y_i + y_i^2 = x_i. \tag{8.56}$$

However, the form containing $\triangle$ allows the most direct comparison with the continuous-time case. It is clear from the previous argument that the only differences for harmonic probing of discrete-time systems will be generated by the fact that the operator $\triangle$ has a different action on functions $\mathrm{e}^{\mathrm{i}\omega t}$ to the operator D. This action is very simple to compute, as shown in chapter 1[3],

$$\triangle \mathrm{e}^{\mathrm{i}\omega t} = \mathrm{e}^{\mathrm{i}\omega(t-1)} = \mathrm{e}^{-\mathrm{i}\omega} \cdot \mathrm{e}^{\mathrm{i}\omega t}. \tag{8.57}$$

It is now clear that one can carry out the harmonic probing algorithm for (8.55) exactly as for the continuous-time (8.39); the only difference will be that the $\triangle$ operator will generate a multiplier $\mathrm{e}^{-\mathrm{i}\omega}$ wherever D generated a factor $\mathrm{i}\omega$. As a consequence H_1 and H_2 for (8.55) are easily computed.

$$H_1(\omega) = \frac{1}{\mathrm{e}^{-\mathrm{i}\omega} + 1} \tag{8.58}$$

$$H_2(\omega_1, \omega_2) = \frac{-H_1(\omega_1)H_1(\omega_2)}{\mathrm{e}^{-\mathrm{i}(\omega_1+\omega_2)} + 1} = -H_1(\omega_1)H_1(\omega_1)H_1(\omega_1 + \omega_2). \tag{8.59}$$

Note that the form of H_2 as a function of H_1 is identical to that for the continuous-time system.

It is possible at this point to make a quite general statement. Given a continuous-time system with linear or nonlinear equation of motion $f(D, y, x) = 0$ and HFRFs $H_n^{\mathrm{c}}(\omega_1, \ldots, \omega_n)$, $n = 1, \ldots, \infty$, the corresponding discrete-time system $f(\triangle, y, x) = 0$ has HFRFs $H_n^{\mathrm{d}}(\omega_1, \ldots, \omega_n) = H_n^{\mathrm{c}}(-\mathrm{i}\mathrm{e}^{-\mathrm{i}\omega_1}, \ldots, -\mathrm{i}\mathrm{e}^{-\mathrm{i}\omega_n})$, $n = 1, \ldots, \infty$. Further the functional relationships between the H_n and H_1 will be identical in both cases.

The system in equation (8.56) is not an NARMAX system as it is a nonlinear function of the most recent sampled value y_i. As discussed in chapter 6, an NARMAX, or more strictly NARX, model has the general form

$$y_i = F(y_{i-1}, \ldots, y_{i_{n_y}}; x_{i-1}, \ldots, x_{i_{n_x}}) \tag{8.60}$$

with appropriate noise modelling if necessary.

The relevant existence theorems obtained in [161, 162] show that this form is general enough to represent almost all input–output systems.

[3] It is amusing to note that this action follows from the fact that $\triangle = \mathrm{e}^{-D}$ as an operator equation; as $\mathrm{e}^{\mathrm{i}\omega t}$ is an eigenfunction of D with eigenvalue $\mathrm{i}\omega$, it is also an eigenfunction of $\triangle$ with eigenvalue $\mathrm{e}^{-\mathrm{i}\omega}$.

8.4 Validation and interpretation of the higher-order FRFs

In order to justify studying the higher-order FRFs it is necessary to show that they contain useful information about whatever system is under examination. In fact, as time- and frequency-domain representations are completely equivalent, the higher-order FRFs contain *all* system information; later in this section it is demonstrated that important facts can be conveyed in a very direct and visible way.

Before discussing matters of interpretation it is important to address the question of uniqueness of the higher-order FRFs as it is critical to any analysis that the non-uniqueness of the time-domain NARMAX representation of a system does not affect the frequency-domain representation.

The first thing which must be established is the correspondence between the FRFs of the continuous system and the FRFs of the discrete approximations. Consider the Duffing oscillator of equation (8.49), a discrete-time representation for this system could be obtained by adopting discrete approximations to the derivatives. The coarsest approximation available is the backward-difference approximation

$$\dot{y}_i \approx \frac{y_i - y_{i-1}}{\Delta t} \tag{8.61}$$

$$\ddot{y}_i \approx \frac{y_{i+1} - 2y_i + y_{i-1}}{\Delta t^2} \tag{8.62}$$

which gives the discrete-time representation

$$\begin{aligned} y_i = & \left\{ \frac{2m + c\Delta t + k\Delta t^2}{m + c\Delta t} \right\} y_{i-1} - \left\{ \frac{m}{m + c\Delta t} \right\} y_{i-2} \\ & - \left\{ \frac{k_2 \Delta t^2}{m + c\Delta t} \right\} y_{i-1}^2 - \left\{ \frac{k_3 \Delta t^2}{m + c\Delta t} \right\} y_{i-1}^3 + \left\{ \frac{\Delta t^2}{m + c\Delta t} \right\} x_{i-1}. \end{aligned} \tag{8.63}$$

In fact, because this is based on the coarse approximations (8.61) and (8.62), it does not yield good representations of the higher-order FRFs. In order to demonstrate accurate FRFs from a NARX model, the following numerical simulation was carried out. A fourth-order Runge–Kutta scheme [209], was used to obtain the response of the system (8.49) under excitation by a Gaussian noise sequence $x(t)$ with rms 10.0 and frequency range 0–90 Hz. The coefficient values adopted were: $m = 1$, $c = 20$, $k = 10^4$, $k_2 = 10^7$, $k_3 = 5 \times 10^9$. This system has a resonant frequency of $\omega_{\mathrm{r}} = 99$ rad s^{-1} or $f_{\mathrm{r}} = \frac{\omega_{\mathrm{r}}}{2\pi} = 15.75$ Hz. The data were generated with a sampling interval of 0.005 s, giving a Nyquist frequency of 100 Hz.

A NARX model was fitted to 1000 points of the resulting discrete x and y data using the estimation and validation methods described in the previous

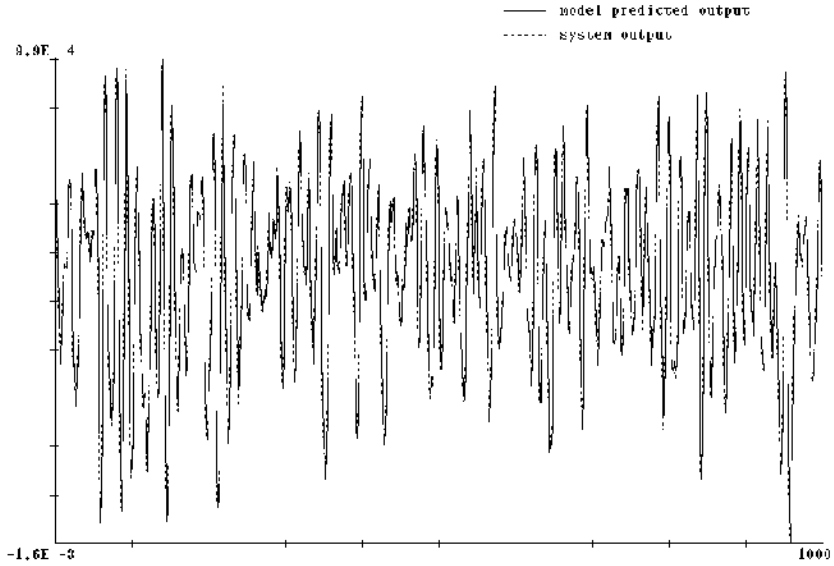


Figure 8.9. Comparison between simulated Duffing oscillator data and the prediction by a NARX model.

section. The result was

$$\begin{aligned} y_i &= 1.6696y_{i-1} - 0.903\,48y_{i-2} \\ &- 2.1830\times 10^{2}y_{i-1}^{2} - 1.0665\times 10^{5}y_{i-1}^{3} \\ &+ 3.0027\times 10^{-6}x_i + 1.8040\times 10^{-5}x_{i-1} \\ &+ 2.7676\times 10^{-6}x_{i-2}. \end{aligned} \tag{8.64}$$

Figure 8.9 shows a comparison between the original y data from the simulation, and that predicted by the NARX model (8.64), when excited by the same input data x; the NARX model clearly gives a good representation of the system in the time domain. The fitted model was then used to generate the higher-order FRFs, H_1, H_2 and H_3, by the method of harmonic probing. As the exact results could also be obtained by harmonic probing of (8.49), direct comparisons could be made. In all cases, the exact FRFs are given with the frequency scale in Hz; the FRFs for the discrete model are given with corresponding normalized frequency scales $f_{\mathrm{n}} = f/f_{\mathrm{s}}$ where f_{s} is the sampling frequency; the Nyquist frequency is 0.5 in these units.

Figure 8.10 shows a comparison between the exact H_1 and that obtained from the model; the agreement looks excellent. However, an important point must be raised here. H_1 for the discrete system is only an approximation to H_1 for the continuous system up to the Nyquist frequency of 0.5 (100 Hz); it is only plotted up to this frequency in figures 8.10(*c*) and 8.10(*d*) because it simply repeats beyond this point and is therefore meaningless.

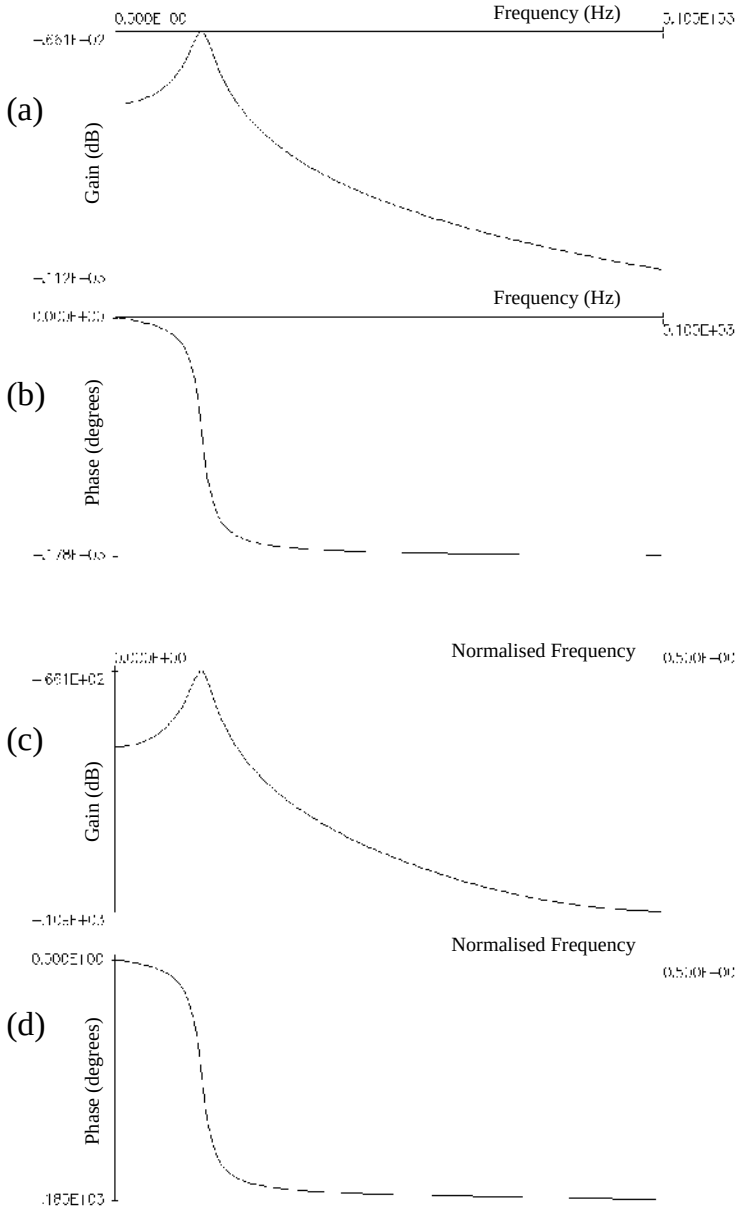


Figure 8.10. $H_1(f)$ for the Duffing oscillator system: (*a*) exact magnitude; (*b*) exact phase; (*c*) NARX model magnitude; (*d*) NARX model phase.

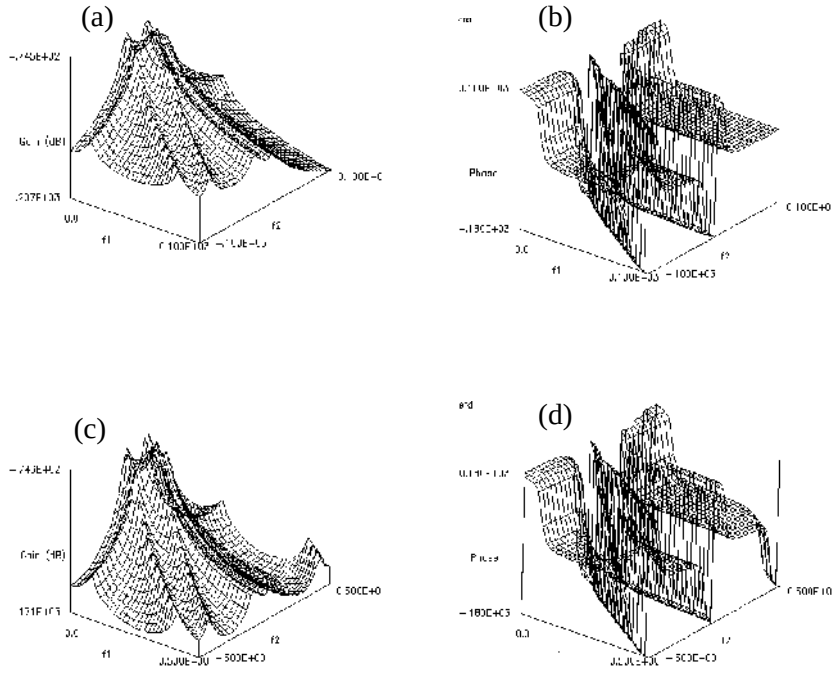


Figure 8.11. $H_2(f_1, f_2)$ surface for the Duffing oscillator system: (*a*) exact magnitude; (*b*) exact phase; (*c*) NARX model magnitude; (*d*) NARX model phase.

The comparison between the exact H_2 and that from the NARMAX model is given in figure 8.11. The same comparison using the contour maps for the functions is shown in figure 8.12; again the agreement is very good. Note that because H_2 contains factors $H_1(2\pi f_1)$ and $H_2(2\pi f_2)$ it would be meaningless to plot it outside the ranges corresponding to $f_1 \leq 100, f_2 \leq 100$. Further, H_2 also contains a factor $H_1(2\pi(f_1 + f_2))$ so that the plots should not extend past the area specified by $f_1 + f_2 \leq 100$. Rather than plot irregularly shaped regions, the H_2 figures presented in this book include information beyond this last bound, which is indicated by the full line in the model contour maps in figure 8.12; information presented outside this region on any H_2 plot should not be regarded as meaningful.

The comparison between the exact H_3 and model H_3 is given in figure 8.13, and in contour map form in figure 8.14. Unfortunately, the whole H_3 surface cannot be plotted as it exists as a three-dimensional manifold embedded in a four-dimensional space over the $(\omega_1, \omega_2, \omega_3)$-'plane'. However, one can plot two-dimensional submanifolds of H_3, and this is the approach which is usually adopted. Figures 8.13 and 8.14 show $H_3(\omega_1, \omega_2, \omega_1)$ plotted over the (ω_1, ω_2)-plane. The region of validity of the H_3 surface is a little more complicated in

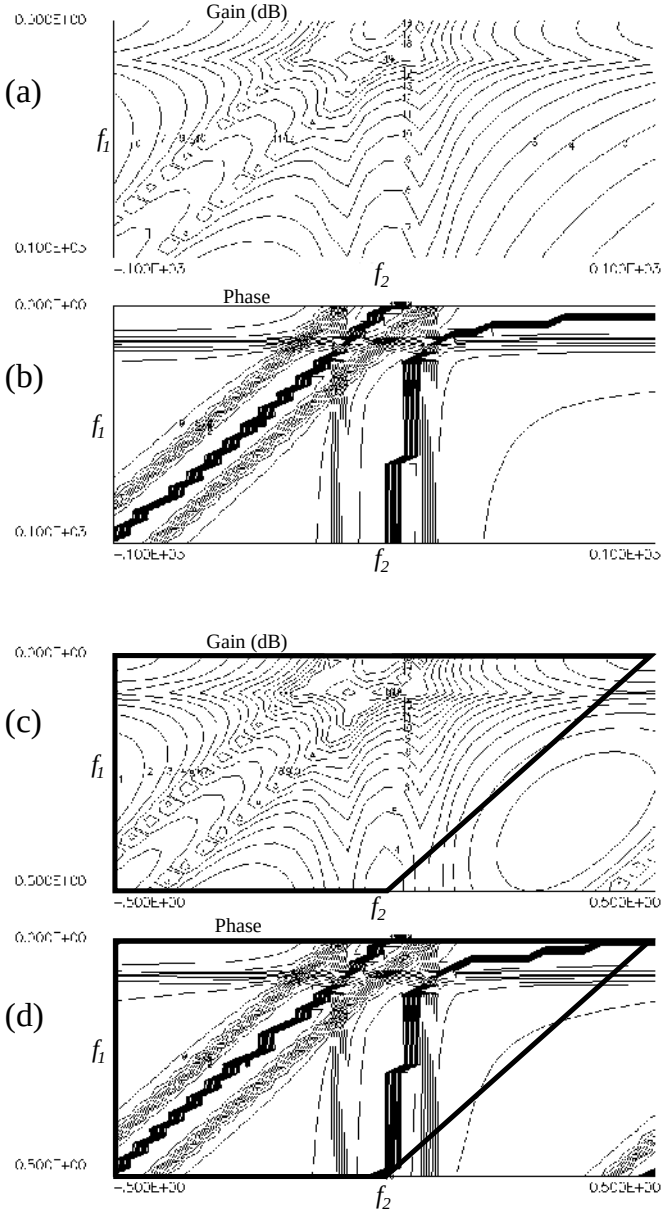


Figure 8.12. $H_2(f_1, f_2)$ contours for the Duffing oscillator system: (*a*) exact magnitude; (*b*) exact phase; (*c*) NARX model magnitude; (*d*) NARX model phase.

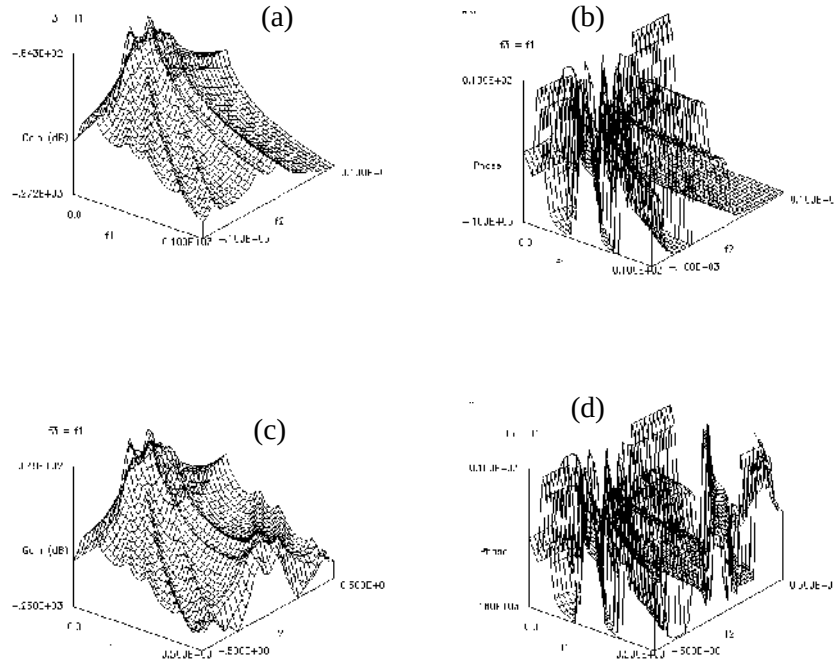


Figure 8.13. $H_3(f_1, f_2, f_1)$ surface for the Duffing oscillator system: (*a*) exact magnitude; (*b*) exact phase; (*c*) NARX model magnitude; (*d*) NARX model phase.

this situation. In all cases, agreement between the exact H_n and those obtained from the NARMAX model appears impressive. For a less passive comparison, figure 8.15 shows the gain and phase of the output components y_1, y_2 and y_3 obtained from the systems defined by the exact and model FRFs when excited by a unit sinusoid at various frequencies. Again, agreement looks excellent. Note that the plot for second harmonic in figure 8.15 contains a peak at $f_r/2$. This is due to the fact that the diagonal HFRF contains a factor $H_1(2\omega)$ as shown by equation (8.51).

Having established that a NARX model can yield good representations of the FRFs from a continuous system, the next question which must be addressed concerns the correspondence between frequency-domain representations of different yet *exactly* equivalent NARX models. (Non-uniqueness is actually a problem with most methods of modelling, it is not specific to NARX). Suppose one has obtained as an accurate discretization of a continuous system, the ARX model,

$$y_i = a_1 y_{i-1} + a_2 y_{i-2} + b_1 x_{i-1}. \tag{8.65}$$

As this expression holds for all values of i (away from the initial points), it

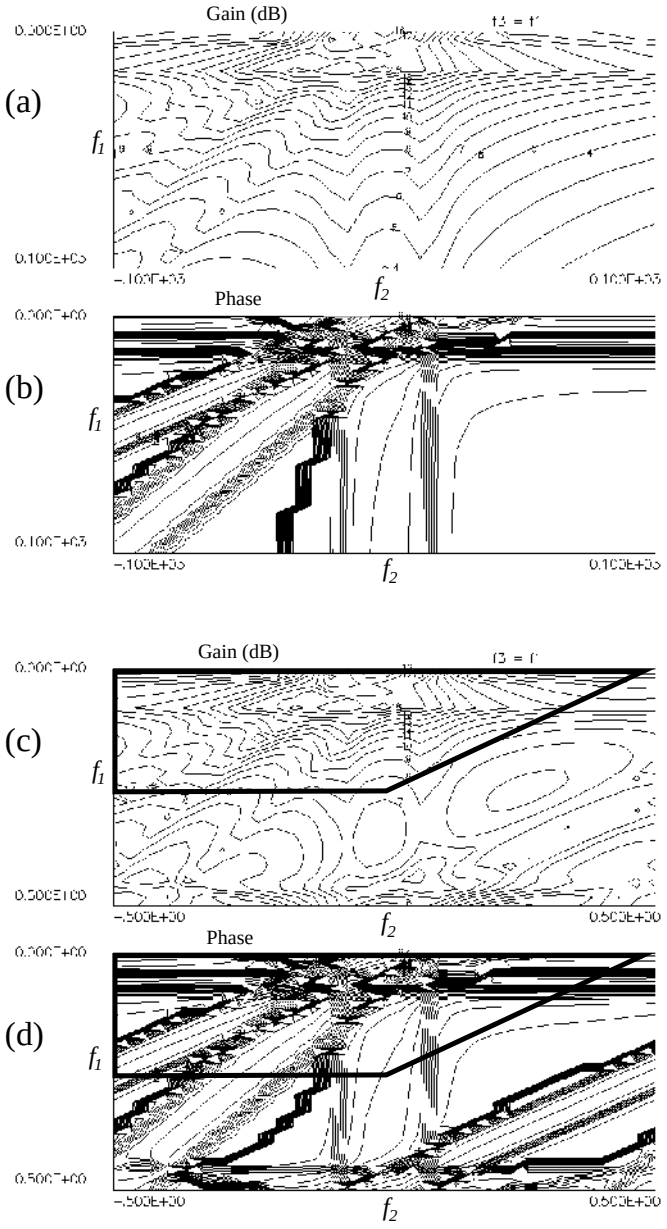


Figure 8.14. $H_3(f_1, f_2, f_1)$ contours for the Duffing oscillator system: (*a*) exact magnitude; (*b*) exact phase; (*c*) NARX model magnitude; (*d*) NARX model phase.

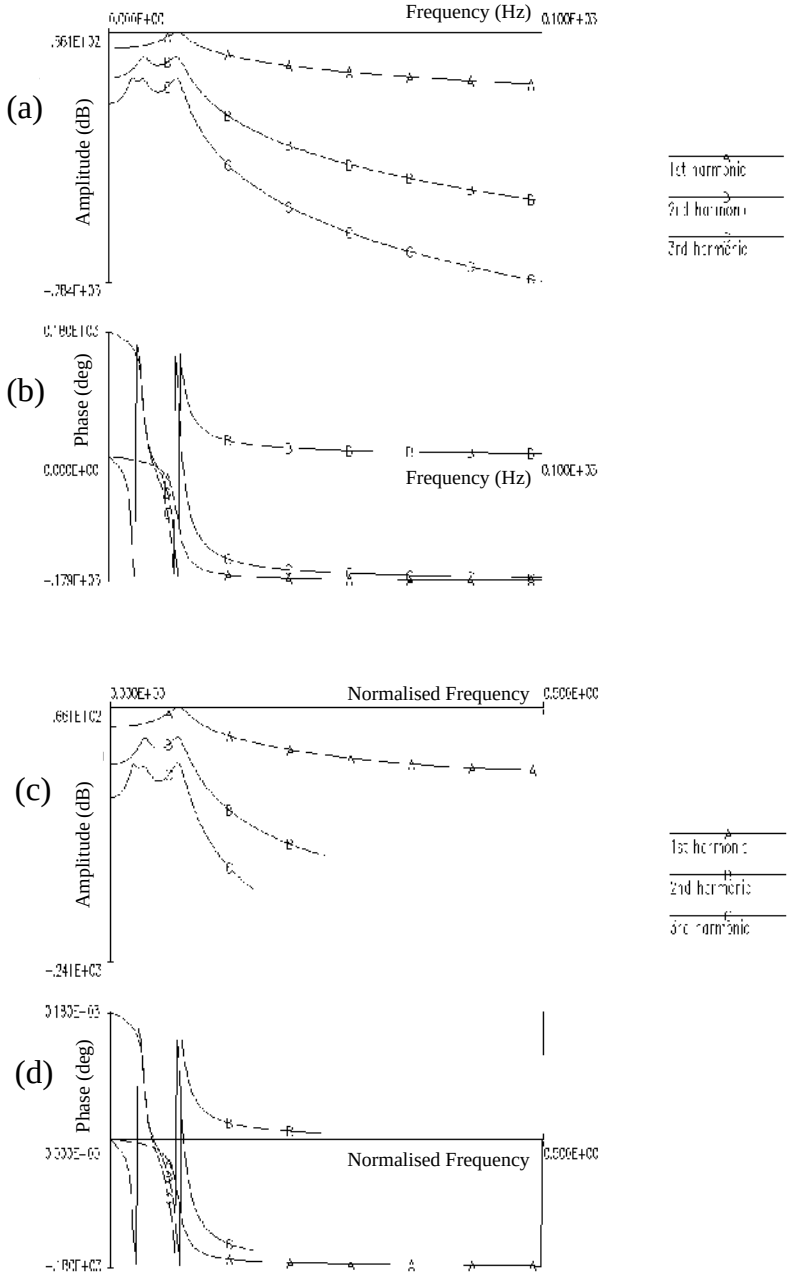


Figure 8.15. H_1, H_2 and H_3 components for the Duffing oscillator response excited by a unit sinusoid: (*a*) exact magnitude; (*b*) exact phase; (*c*) NARX model magnitude; (*d*) NARX model phase.

can just as well be written as

$$y_{i-1} = a_1 y_{i-2} + a_2 y_{i-3} + b_1 x_{i-2} \tag{8.66}$$

and substituting (8.66) into (8.65) yields the ARX model

$$y_i = (a_1^2 + a_2) y_{i-2} + a_1 a_2 y_{i-3} + b_1 x_{i-1} + a_1 b_1 x_{i-2} \tag{8.67}$$

which is exactly equivalent to (8.65) yet contains different terms. This type of ambiguity will occur for any system which regresses the present output onto past values of output. It is a reflection of a type of ambiguity for continuous-time systems; one can always differentiate the equation of motion to obtain a completely equivalent system. The only thing which changes is the set of objects for which initial conditions are required. Harmonic probing of (8.65) yields (in symbolic notation where $\triangle = \mathrm{e}^{-\mathrm{i}\omega}$)

$$H_1^{(8.65)} = \frac{b_1 \triangle}{1 - a_1 \triangle - a_2 \triangle^2} \tag{8.68}$$

while probing of (8.67) gives the superficially different

$$H_1^{(8.67)} = \frac{b_1 \triangle + a_1 b_1 \triangle^2}{1 - (a_1^2 + a_2) \triangle^2 - a_1 a_2 \triangle^3}. \tag{8.69}$$

However, the latter expression factors:

$$H_1^{(8.67)} = \frac{b_1 \triangle (a_1 \triangle + 1)}{(a_1 \triangle + 1)(1 - a_1 \triangle - a_2 \triangle^2)} = \frac{b_1 \triangle}{1 - a_1 \triangle - a_2 \triangle^2} = H_1^{(8.65)}. \tag{8.70}$$

The final type of non-uniqueness is generated by the fact that NARMAX models can be approximately equivalent. As an illustration consider the simple system

$$y_i = \alpha y_{i-1} + x_{i-1}. \tag{8.71}$$

If α is small, a simple application of the binomial theorem gives

$$\begin{aligned} (1 - \alpha\triangle) y_i = x_{i-1} &\Longrightarrow y_i = (1 - \alpha\triangle)^{-1} x_{i-1} \\ &\Longrightarrow y_i = (1 + \alpha\triangle) x_{i-1} + \mathrm{O}(\alpha^2) \end{aligned} \tag{8.72}$$

So the system

$$y_i = x_{i-1} + \alpha x_{i-2} \tag{8.73}$$

is equivalent to the system in (8.71) up to $\mathrm{O}(\alpha^2)$. Now, harmonic probing of system (8.71) yields the FRF

$$H_1^{(8.71)}(\omega) = \frac{1}{1 - \alpha \mathrm{e}^{-\mathrm{i}\omega}} \tag{8.74}$$

and a similar analysis for (8.73) gives

$$H_1^{(8.73)}(\omega) = 1 + \alpha \mathrm{e}^{-\mathrm{i}\omega} = \frac{1}{1 - \alpha \mathrm{e}^{-\mathrm{i}\omega}} + \mathrm{O}(\alpha^2) = H_1^{(8.71)}(\omega) + \mathrm{O}(\alpha^2). \quad (8.75)$$

Note that by retaining n terms in the binomial expansion, the model

$$y_i = x_{i-1} + \alpha x_{i-2} + \cdots + \alpha^{n-1} x_{i-n} \quad (8.76)$$

is obtained which is equivalent to (8.71) up to $\mathrm{O}(\alpha^n)$. As a result, the system (8.71) can be represented with arbitrary accuracy by the binomial expansion if n is large enough. However, note that one representation has only three model terms while the other has n with n possibly large. This serves to illustrate why it is important to correctly detect the model structure or which terms are in the model in order to yield a parsimonious model [32].

One must be careful not to regard these simple arguments as generating a general principle; however, it would seem likely that equivalence of two NARX models up to a given order of accuracy would imply equivalence of the corresponding HFRFs up to the same order of accuracy. This is easy to establish in the case of a general linear system by an extension of the previous argument.

The various cases discussed earlier exhaust all possibilities for obtaining different NARX representations of a given system.

This discussion is simply intended as an argument that all NARX models which are equivalent in the sense that they furnish a discrete approximation to a continuous system will have higher-order FRFs which not only approximate to each other but also to those of the underlying continuous system. It does not constitute a rigorous proof in any sense; however, it is difficult to imagine a situation under which this condition would not hold.

Having established some confidence in their reliability, the interpretation of the higher-order FRFs can be discussed. The Duffing oscillator system (8.49) serves well as an illustration. The magnitude and phase of the expression (8.50) for $H_1(\omega) = H_1(2\pi f)$ is given in figures 8.10(*a*) and (*b*) on the frequency interval 0–100 Hz. The interpretation of these figures, traditionally given together and universally called the Bode plot, has been described in earlier chapters, notably chapter 1. The peak in the magnitude at $f = f_{\mathrm{r}} = 15.75$ Hz shows that for this frequency of excitation the amplitude of the linear part of the response $y_1(t)$ is a maximum. The Bode plot thus allows the immediate identification of those excitation frequencies at which the vibration level of the system is likely to be high.

Interpretation of the second-order FRF is also straightforward. The magnitude and phase of H_2 for the Duffing system given earlier are given in figures 8.11(*a*) and (*b*) as surfaces, or in figures 8.12(*a*) and (*b*) as contour maps, over the $(f_1, f_2) = (\frac{\omega_1}{2\pi}, \frac{\omega_2}{2\pi})$ plane. The frequency ranges for the plot are the same as for H_1 in figure 8.10. A number of ridges are observed. These are in direct correspondence with the peak in H_1 as follows. According to equation (8.51),

H_2 is a constant multiple of $H_1(\omega_1)H_1(\omega_2)H_1(\omega_1+\omega_2)$. As a consequence H_2 possesses local maxima at positions where the H_1 factors have local maxima. Consequently there are two ridges in the H_2 surface corresponding to the lines $\omega_1 = \omega_r = 2\pi f_r$ and $\omega_2 = \omega_r$. These are along lines parallel to the frequency axes. In addition, H_2 has local maxima generated by the $H_1(\omega_1+\omega_2)$ factor along the line $\omega_1+\omega_2 = \omega_r$. This ridge has an important implication; it indicates that one can expect a maximum in the second-order output $y_2(t)$ if the system is excited by two sinusoids whose sum frequency is the linear resonant frequency. This shows clearly why estimation of a transfer function by linear methods is inadequate for nonlinear systems; such a transfer function would usually indicate a maximum in the output for a harmonic excitation close to the linear resonant frequency. However, it would fail to predict that one could excite a large nonlinear component in the output by exciting at $\omega = \frac{\omega_r}{2}$; this is a consequence of the trivial decomposition $2e^{i\frac{\omega_r}{2}t} = e^{i\frac{\omega_r}{2}t} + e^{i\frac{\omega_r}{2}t}$ which means that the signal can be regarded as a 'two-tone' input with a sum frequency at the linear resonance ω_r. The importance of the second-order FRF is now clear. It reveals those pairs of excitation frequencies which will conspire to produce large levels of vibration as a result of second-order nonlinear effects.

The interpretation of H_3 for the system is very similar. Consideration of equation (8.54) shows that for a three-tone input of the form (8.52) one should expect maxima in the third-order output $y_3(t)$ if the following conditions are satisfied: $\omega_1 = \omega_r, \omega_2 = \omega_r, \omega_3 = \omega_r, \omega_1+\omega_2 = \omega_r, \omega_2+\omega_3 = \omega_r, \omega_3+\omega_1 = \omega_r$, $\omega_1+\omega_2+\omega_3 = \omega_r$. The presence of these 'combination resonances' would be indicated by the presence of ridges in the H_3 surface. Although figures 8.13 and 8.14 only show the 'projections' of H_3 over the (ω_1,ω_2)-plane, they are sufficient to indicate the presence of the 'combination resonances' $\omega_1 = \omega_r$, $\omega_2 = \omega_r$, $\omega_1+\omega_2 = \omega_r$, $2\omega_1 = \omega_r$, $2\omega_1+\omega_2 = \omega_r$. It is clear that the local maximum distributions become more and more complex as the order of the HFRF increases.

These arguments show that the higher-order FRFs provide directly visible information about the possible excitation of large nonlinear vibrations through the cooperation of certain frequencies.

8.5 An application to wave forces

The power of the NARX and higher-order FRF approaches can be demonstrated by the following example used in chapter 6 where force and velocity data were obtained from a circular cylinder placed in a planar oscillating fluid flow in a large U-tube [199]. The standard means of predicting forces on cylinders used by the offshore industry is to use Morison's equation (6.121) which expresses the force as a simple nonlinear function of the instantaneous flow velocity and acceleration. For one particular frequency of flow oscillation, Morison's equation gave the force prediction shown in figure 8.16(*a*) compared with the measured force. Morison's equation is inadequate at representing the higher-frequency components of the

force. The model inadequacy is shown clearly by the correlation-based validity tests (section 6.8.3) in figure 8.17(*b*)[4].

A NARX fit to the force–velocity data gave the model prediction shown in figure 8.17(*a*). This model also passes the correlation tests (figure 8.17(*b*)).

A similar analysis has been carried out on fluid-loading data encompassing a broad range of flow conditions ranging from U-tube data to that from a unidirectional wave in a large flume to data from a random directional sea. In all cases, the NARX analysis produced a better model than Morison's equation [276]. Unfortunately the model structures varied. In order to examine the possibility that this was simply due to the non-uniqueness of the NARX representations, the higher-order FRFs were obtained by harmonic probing. The results were very interesting, as an example, H_3 for the U-tube data of figure 8.17 is given in figure 8.18. The pronounced ridges were shown to appear in the third-order FRFs for all of the flow conditions examined; this is in direct contradiction to Morison's equation which forces a constant H_3.

The higher-order FRFs can often throw light onto a problem in this way; the direct visualization of the system properties which they provide is appealing. They have actually been used in wave loading studies for some time now; however, the computational burden imposed by traditional methods of estimation has prohibited the use of functions higher than second order [85]

8.6 FRFs and Hilbert transforms: sine excitation

8.6.1 The FRF

It was shown earlier that the Volterra series provides a convenient means for calculating the nonlinear system response to a single harmonic; this forms the basis of the harmonic probing method. It is only slightly more complicated to calculate the response to multiple harmonics. The benefit is that one can then determine the response to a sinusoid and this, in turn, will allow us to develop an expression for the stepped-sine FRF of the system. Suppose the excitation is a two-tone signal

$$x(t) = A\mathrm{e}^{\mathrm{i}\Omega_{\mathrm{a}}t} + B\mathrm{e}^{\mathrm{i}\Omega_{\mathrm{b}}t} \tag{8.77}$$

which translates into the frequency domain as

$$X(\omega) = 2\pi\{A\delta(\omega - \Omega_{\mathrm{a}}) + B\delta(\omega - \Omega_{\mathrm{b}})\}. \tag{8.78}$$

Substituting this into (8.16)–(8.18) and thence into (8.15) leads, after a long but straightforward calculation, to

$$y(t) = AH_1(\Omega_{\mathrm{a}})\mathrm{e}^{\mathrm{i}\Omega_{\mathrm{a}}t} + BH_1(\Omega_{\mathrm{b}})\mathrm{e}^{\mathrm{i}\Omega_{\mathrm{b}}t} + A^2H_2(\Omega_{\mathrm{a}}, \Omega_{\mathrm{a}})\mathrm{e}^{2\mathrm{i}\Omega_{\mathrm{a}}t}$$

[4] Of course, with enough parameters, one can fit a model to an arbitrary level of accuracy on a given estimation set of data. The modeller should always carry out appropriate levels of model validity testing in order to ensure that the model is genuine and does not simply represent an isolated data set. This is particularly pressing in the situation where one might abandon a *physical* model like Morison's equation in favour of a non-physical model on the grounds of model accuracy.

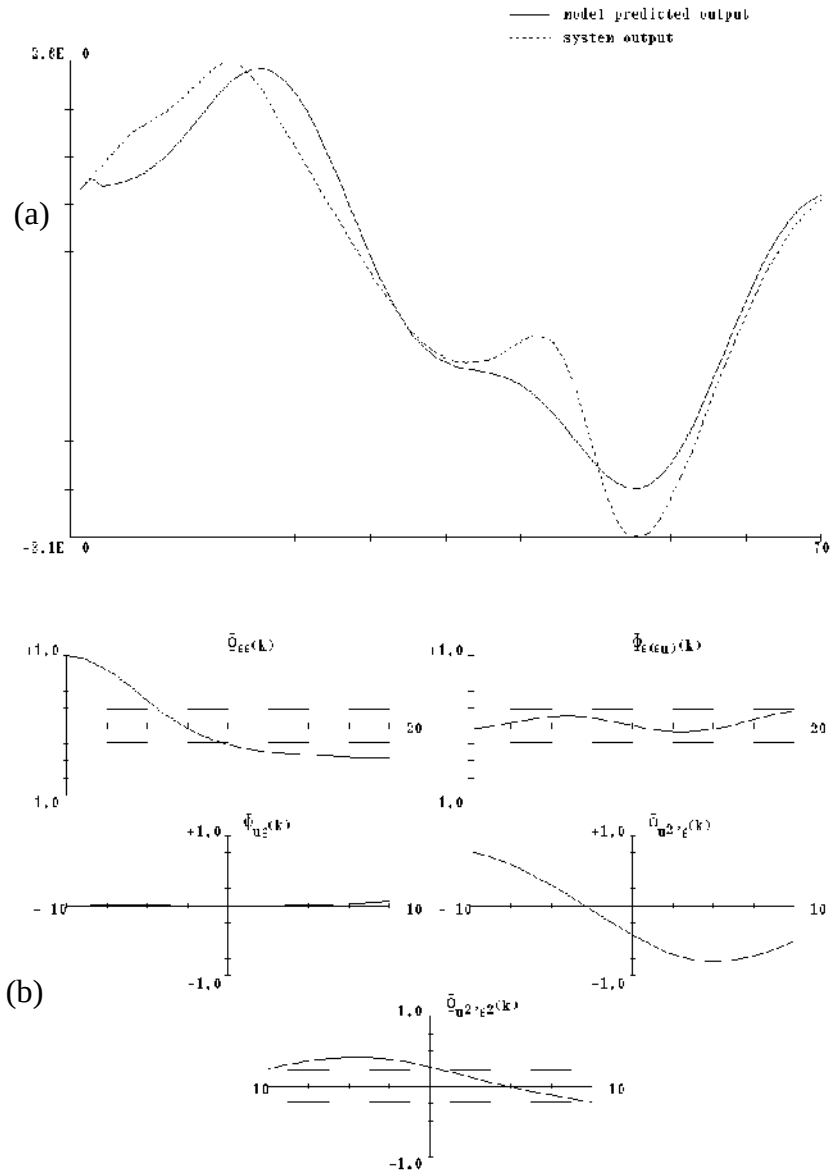


Figure 8.16. Morison equation fit to experimental U-tube data: (*a*) model-predicted output; (*b*) correlation tests.

$$+ 2ABH_2(\Omega_\mathrm{a}, \Omega_\mathrm{b})\mathrm{e}^{\mathrm{i}(\Omega_\mathrm{a}+\Omega_\mathrm{b})t} + B^2H_2(\Omega_\mathrm{b}, \Omega_\mathrm{b})\mathrm{e}^{2\mathrm{i}\Omega_\mathrm{b}t}$$
$$+ A^3H_3(\Omega_\mathrm{a}, \Omega_\mathrm{a}, \Omega_\mathrm{a})\mathrm{e}^{3\mathrm{i}\Omega_\mathrm{a}t} + A^2BH_3(\Omega_\mathrm{a}, \Omega_\mathrm{a}, \Omega_\mathrm{b})\mathrm{e}^{\mathrm{i}(2\Omega_\mathrm{a}+\Omega_\mathrm{b})t}$$

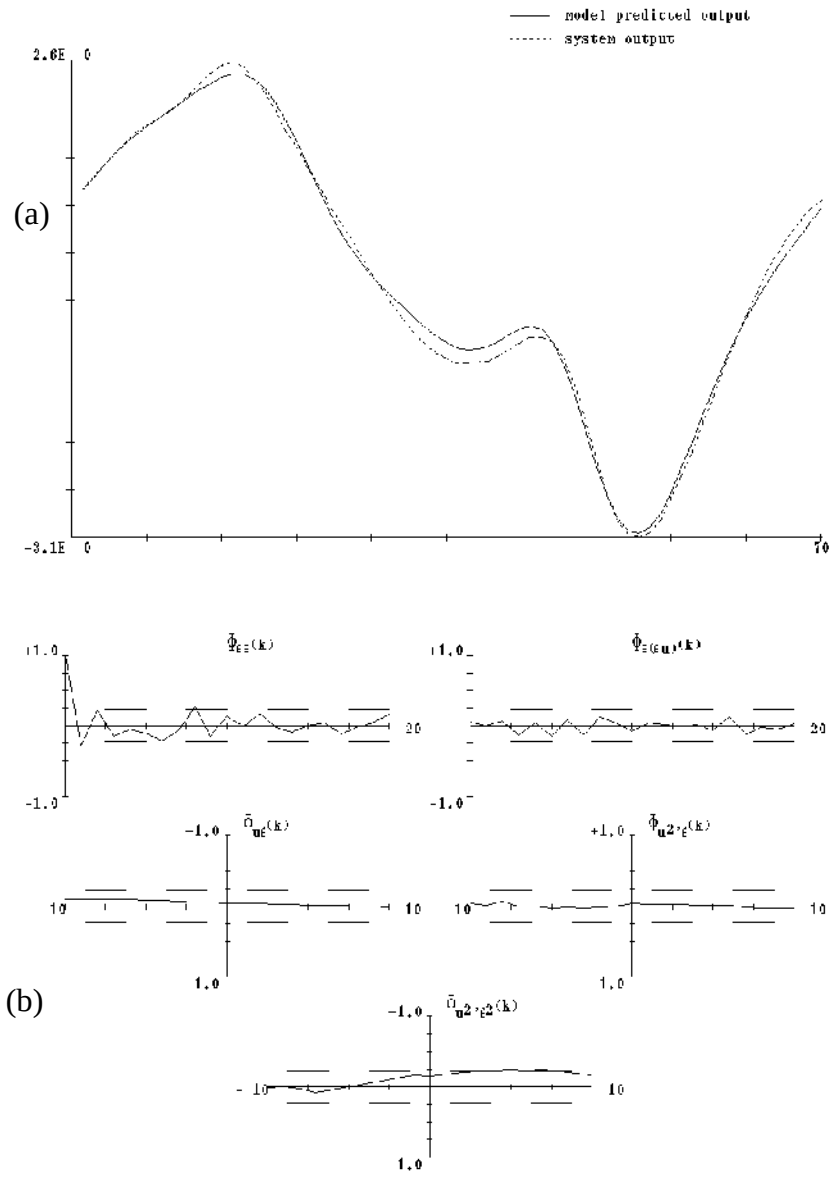


Figure 8.17. NARX model fit to experimental U-tube data: (*a*) model-predicted output; (*b*) correlation tests.

$$+ AB^2 H_3(\Omega_\mathrm{a}, \Omega_\mathrm{b}, \Omega_\mathrm{b}) \mathrm{e}^{\mathrm{i}(\Omega_\mathrm{a}+2\Omega_\mathrm{b})t} + B^3 H_3(\Omega_\mathrm{b}, \Omega_\mathrm{b}, \Omega_\mathrm{b}) \mathrm{e}^{3\mathrm{i}\Omega_\mathrm{b} t} + \cdots \tag{8.79}$$

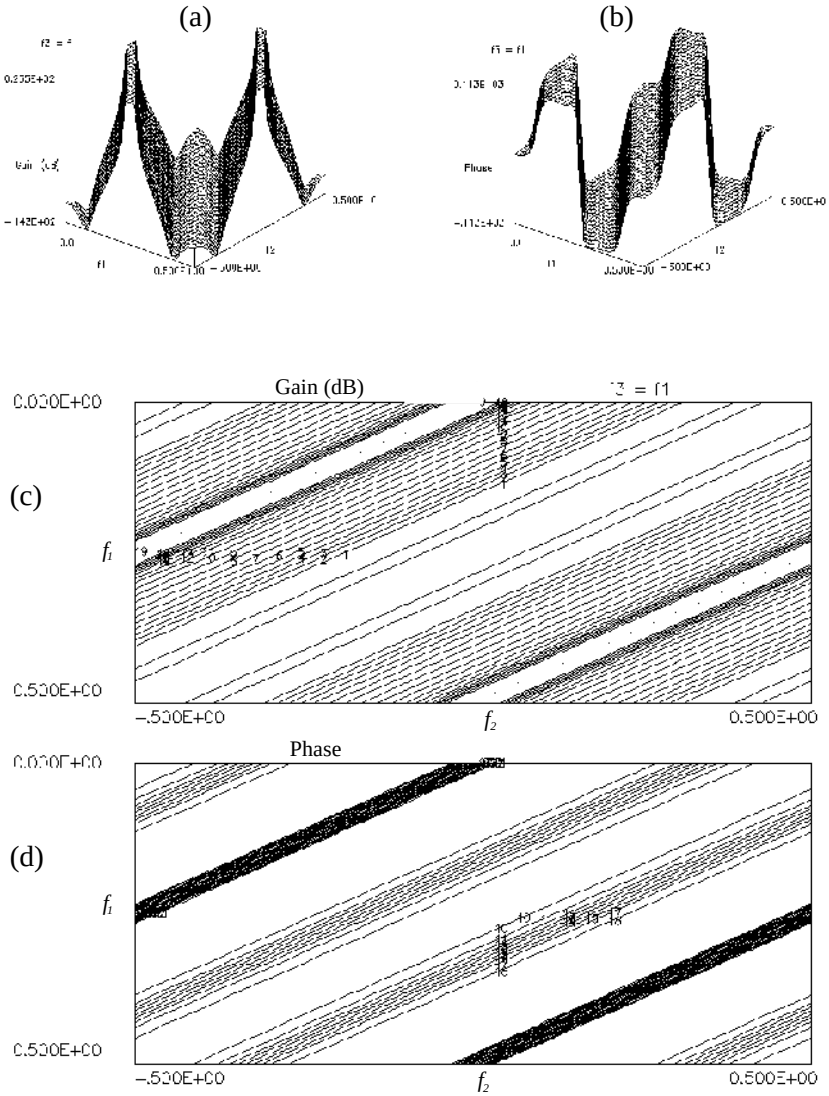


Figure 8.18. $H_3(f_1, f_2, f_1)$ from NARX fit to U-tube data: (*a*) magnitude; (*b*) phase; (*c*) magnitude contours; (*d*) phase contours.

to third order.

Now, for the response to a cosinusoid

$$x(t) = \cos(\Omega t) = \tfrac{1}{2}X(\mathrm{e}^{\mathrm{i}\Omega t} + \mathrm{e}^{-\mathrm{i}\Omega t}) \tag{8.80}$$

one simply substitutes $A = B = X/2$, $\Omega_{\mathrm{a}} = \Omega$ and $\Omega_{\mathrm{b}} = -\Omega$. To third order

again, the result is

$$\begin{aligned} y(t) = {} & \frac{X}{2}H_1(\Omega)\mathrm{e}^{\mathrm{i}\Omega t} + \frac{X}{2}H_1(-\Omega)\mathrm{e}^{-\mathrm{i}\Omega t} + \frac{X^2}{4}H_2(\Omega,\Omega)\mathrm{e}^{2\mathrm{i}\Omega t} \\ & + \frac{X^2}{2}H_2(\Omega,-\Omega) + \frac{X^2}{4}H_2(-\Omega,-\Omega)\mathrm{e}^{-2\mathrm{i}\Omega t} + \frac{X^3}{8}H_3(\Omega,\Omega,\Omega)\mathrm{e}^{3\mathrm{i}\Omega t} \\ & + \frac{3X^3}{8}H_3(\Omega,\Omega,-\Omega)\mathrm{e}^{\mathrm{i}\Omega t} + \frac{3X^3}{8}H_3(\Omega,-\Omega,-\Omega)\mathrm{e}^{-\mathrm{i}\Omega t} \\ & + \frac{X^3}{8}H_3(-\Omega,-\Omega,-\Omega)\mathrm{e}^{-3\mathrm{i}\Omega t} + \cdots . \end{aligned} \tag{8.81}$$

Making use of the reflection properties $H_1(-\Omega) = H^*(\Omega)$ etc, and applying de Moivre's theorem in the form

$$z\mathrm{e}^{\mathrm{i}\Omega t} + z^*\mathrm{e}^{-\mathrm{i}\Omega t} = |z|\mathrm{e}^{\mathrm{i}(\Omega t + \angle z)} + |z|\mathrm{e}^{-\mathrm{i}(\Omega + \angle z)} = 2|z|\cos(\Omega t + \angle z) \tag{8.82}$$

yields

$$\begin{aligned} y(t) = {} & X|H_1(\Omega)|\cos(\Omega t + \angle H_1(\Omega)) \\ & + \frac{X^2}{2}|H_2(\Omega,\Omega)|\cos(2\Omega t + \angle H_2(\Omega,\Omega)) + \frac{X^2}{2}H_2(\Omega,-\Omega) \\ & + \frac{X^3}{4}|H_3(\Omega,\Omega,\Omega)|\cos(3\Omega t + \angle H_3(\Omega,\Omega,\Omega)) \\ & + \frac{3X^3}{4}|H_3(\Omega,\Omega,-\Omega)|\cos(\Omega t + \angle H_3(\Omega,\Omega,-\Omega)) + \cdots \end{aligned} \tag{8.83}$$

which shows again that the response contains all odd and even harmonics. The component of the response at the forcing frequency is

$$\begin{aligned} y(t) = {} & X|H_1(\Omega)|\cos(\Omega t + \angle H_1(\Omega)) \\ & + \frac{3X^3}{4}|H_3(\Omega,\Omega,-\Omega)|\cos(\Omega t + \angle H_3(\Omega,\Omega,-\Omega)) + \cdots \end{aligned} \tag{8.84}$$

and this immediately identifies the composite FRF $\Lambda_{\mathrm{s}}(\Omega)$ as

$$\Lambda_{\mathrm{s}}(\Omega) = H_1(\Omega) + \frac{3X^2}{4}H_3(\Omega,\Omega,-\Omega) + \cdots \tag{8.85}$$

or

$$\Lambda_{\mathrm{s}}(\Omega) = H_1(\Omega) + \frac{3X^2}{4}H_3(\Omega,\Omega,-\Omega) + \frac{5X^4}{8}H_5(\Omega,\Omega,\Omega,-\Omega,-\Omega) + \cdots \tag{8.86}$$

to the next highest order. Again, it is useful to take the Duffing oscillator (8.49) as an example. Equation (8.54) with $k_2 = 0$, gives

$$H_3(\omega_1,\omega_2,\omega_3) = -k_3 H_1(\omega_1)H_1(\omega_2)H_1(\omega_3)H_1(\omega_1+\omega_2+\omega_3) \tag{8.87}$$

(adopting lower-case ω from now on) or

$$H_3(\omega,\omega,-\omega) = -k_3 H_1(\omega)^3 H_1^*(\omega). \tag{8.88}$$

Harmonic balance gives for (8.49)

$$\begin{aligned} H_5(\omega_1,\omega_2,\omega_3,\omega_4,\omega_5) = {} & \frac{1}{5!} H_1(\omega_1+\omega_2+\omega_3+\omega_4+\omega_5) \\ & \times (-3!3!k_3\{H_3(\omega_1,\omega_2,\omega_3)H_1(\omega_4)H_1(\omega_5) \\ & + H_3(\omega_1,\omega_2,\omega_4)H_1(\omega_3)H_1(\omega_5) \\ & + H_3(\omega_1,\omega_3,\omega_4)H_1(\omega_2)H_1(\omega_5) + H_3(\omega_2,\omega_3,\omega_4)H_1(\omega_1)H_1(\omega_5) \\ & + H_3(\omega_1,\omega_2,\omega_5)H_1(\omega_3)H_1(\omega_4) + H_3(\omega_1,\omega_3,\omega_5)H_1(\omega_2)H_1(\omega_4) \\ & + H_3(\omega_2,\omega_3,\omega_5)H_1(\omega_1)H_1(\omega_4) + H_3(\omega_1,\omega_4,\omega_5)H_1(\omega_2)H_1(\omega_3) \\ & + H_3(\omega_2,\omega_4,\omega_5)H_1(\omega_1)H_1(\omega_3) + H_3(\omega_3,\omega_4,\omega_5)H_1(\omega_1)H_1(\omega_2)\}) \end{aligned} \tag{8.89}$$

and thence

$$\begin{aligned} H_5(\omega,\omega,\omega,-\omega,-\omega) = {} & \tfrac{3}{10}k_3^2(3H_1(\omega)^4 H_1^*(\omega)^3 + 6H_1(\omega)^5 H_1^*(\omega)^2 \\ & + H_1(\omega)^4 H_1^*(\omega)^2 H_1(3\omega)). \end{aligned} \tag{8.90}$$

Substituting (8.90) and (8.88) into (8.86) gives the FRF up to $\mathrm{O}(X^4)$

$$\begin{aligned} \Lambda_{\mathrm{s}}(\Omega) = {} & H_1(\Omega) - \frac{3X^2}{4}k_3 H_1(\omega)^3 H_1^*(\omega) + \frac{3X^4}{16}k_3^2(3H_1(\omega)^4 H_1^*(\omega)^3 \\ & + 6H_1(\omega)^5 H_1^*(\omega)^2 + H_1(\omega)^4 H_1^*(\omega)^2 H_1(3\omega)) + \mathrm{O}(X^6). \end{aligned} \tag{8.91}$$

(Amongst other places, this equation has been discussed in [236], where it was used to draw some conclusions regarding the amplitude dependence of the stepped-sine composite FRF.)

In order to illustrate these expressions, the system

$$\ddot{y} + 20\dot{y} + 10^4 y + 5 \times 10^9 y^3 = X\cos(\omega t) \tag{8.92}$$

was chosen. Figure 8.19 shows the FRF magnitude plots obtained from (8.91) for $X = 0.01$ (near linear), $X = 0.5$ and $X = 0.75$. At the higher amplitudes, the expected FRF distortion is obtained, namely the resonant frequency shifts up and the magnitude at resonance falls. Figure 8.20 shows the corresponding Nyquist plots. (Note the unequal scales in the Real and Imaginary axes; the plots are effectively circular.) Figure 8.21 shows the $\mathrm{O}(X^4)$ FRF compared with the 'exact' result from numerical simulation. There is a small degree of error near resonance which is the result of premature truncation of the Volterra series.

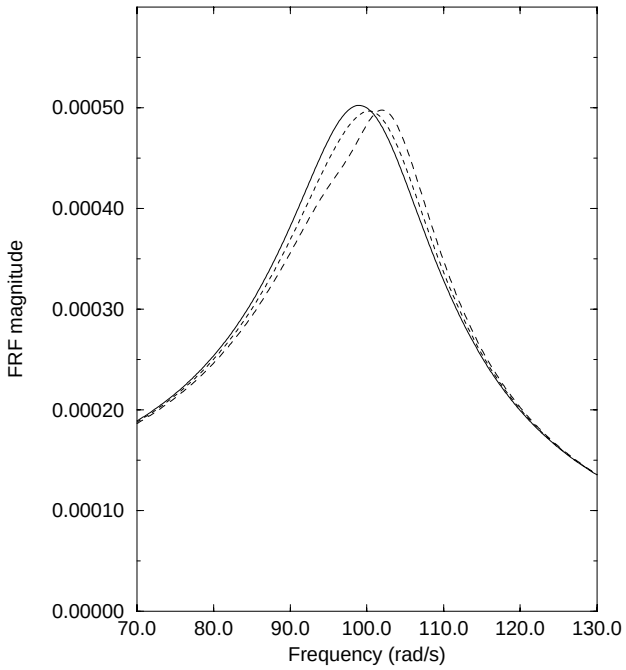


Figure 8.19. Distortion in the magnitude plot of $\lambda_{s1}(\omega)$ computed from the Volterra series for different levels of excitation.

8.6.2 Hilbert transform

Recall from chapters 4 and 5 that the Hilbert transform provides a mean of diagnosing structural nonlinearity on the basis of FRF data. The mapping on the FRF $\Lambda(\omega)$ reduces to the identity on those functions corresponding to linear systems. For nonlinear systems, the Hilbert transform results in a distorted version $\tilde{\Lambda}$, of the original FRF.

From chapter 5, if $\Lambda(\omega)$ is decomposed so

$$\Lambda(\omega) = \Lambda^{+}(\omega) + \Lambda^{-}(\omega) \tag{8.93}$$

where $\Lambda^{+}(\omega)$ (respectively $\Lambda^{-}(\omega)$) has poles only in the upper (respectively lower) half of the complex ω-plane. It is shown in chapter 5 that

$$\mathcal{H}[\Lambda^{\pm}(\omega)] = \pm\Lambda^{\pm}(\omega) \tag{8.94}$$

and the distortion suffered in passing from the FRF to the Hilbert transform is given by the simple relation

$$\Delta\Lambda(\omega) = \mathcal{H}[\Lambda(\omega)] - \Lambda(\omega) = -2\Lambda^{-}(\omega). \tag{8.95}$$

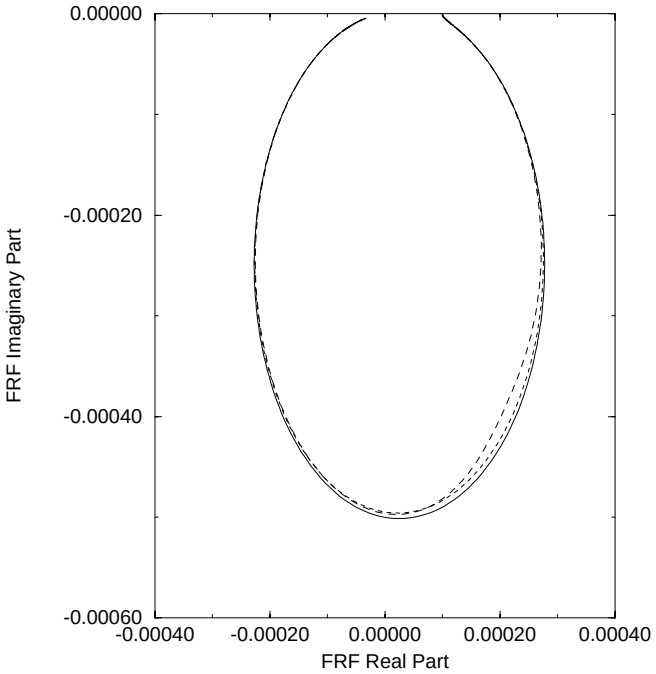


Figure 8.20. Distortion in the Nyquist plot of $\lambda_{s1}(\omega)$ computed from the Volterra series for different levels of excitation. (Note that the Real and Imaginary axes do not have equal scales.)

This section presents a technique which allows the Hilbert transform distortion to be derived term by term from a Volterra series expansion of the system FRF, the expansion parameter being X, the magnitude of the applied sinusoidal excitation. It is illustrated on the Duffing oscillator (8.49), and the basic form of FRF used is the $\mathrm{O}(X^4)$ approximation given in (8.91). If the FRF is known, the Hilbert transform follows from the distortion (8.95). In order to obtain the distortion, the pole–zero form of the FRF is needed.

8.6.2.1 Pole–zero form of the Duffing oscillator FRF

As the approximate nonlinear FRF has been expressed in terms of the linear FRF in (8.91), it is necessary to find the pole–zero form of $H_1(\omega)$; this will then yield the pole–zero form of (8.91). The poles of $H_1(\omega)$ are well known:

$$p_1, p_2 = \pm\omega_{\mathrm{d}} + \mathrm{i}\zeta\omega_n \tag{8.96}$$

where $\omega_{\mathrm{d}} = \omega_n(1-\zeta^2)^{1/2}$ is the damped natural frequency. In terms of these quantities $H_1(\omega)$ may now be expressed as

$$H_1(\omega) = \frac{1}{-m(\omega - p_1)(\omega - p_2)} \tag{8.97}$$

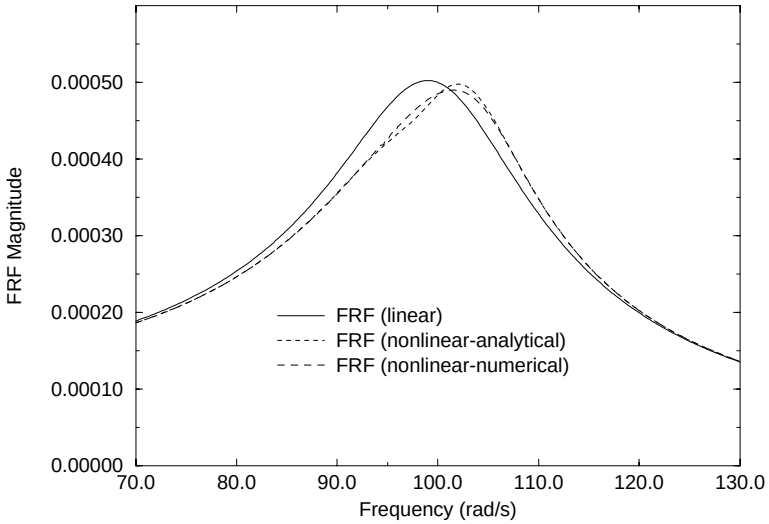


Figure 8.21. Comparison between FRFs $\lambda_{s1}(\omega)$ computed from the Volterra series and from numerical simulation.

and this is the required 'pole–zero' expansion. Note that p_1 and p_2 are both in the upper half-plane so $H_1(\omega) = H_1^+(\omega)$ and the Hilbert transform is therefore the identity on $H_1(\omega)$ as required. However, the expression for $\Lambda_s(\omega)$ in (8.91) contains terms of the form $H_1^*(\omega)$ with poles p_1^* and p_2^*; these are in the lower half-plane and are the cause of the Hilbert transform distortion for $\Lambda_s(\omega)$. In pole–zero form (8.91) becomes

$$\begin{aligned}\Lambda_s(\omega) = {} & \frac{1}{-m(\omega-p_1)(\omega-p_2)} \\ & -\frac{3X^2}{4}k_3\frac{1}{m^4(\omega-p_1)^3(\omega-p_2)^3(\omega-p_1^*)(\omega-p_2^*)} \\ & -\frac{3}{16}k_3^2X^4\left[\frac{3}{m^7(\omega-p_1)^4(\omega-p_2)^4(\omega-p_1^*)^3(\omega-p_2^*)^3}\right. \\ & +\frac{6}{m^7(\omega-p_1)^5(\omega-p_2)^5(\omega-p_1^*)^2(\omega-p_2^*)^2} \\ & \left.+\frac{1}{m^7(\omega-p_1)^4(\omega-p_2)^4(\omega-p_1^*)^2(\omega-p_2^*)^2(3\omega-p_1)(3\omega-p_2)}\right]\end{aligned} \tag{8.98}$$

up to $\mathrm{O}(X^4)$. This is the appropriate form for calculating the distortion.

8.6.2.2 Partial fraction expansion

The method of effecting the decomposition (8.93) for the nonlinear FRF (8.91) is to find the partial fraction expansion. Due to the complexity of the task, this is accomplished using computer algebra. The $\mathrm{O}(X^2)$ and $\mathrm{O}(X^4)$ terms in the transfer function may be considered separately.

The partial fraction expansion of the $\mathrm{O}(X^2)$ is easily found to have the form

$$\frac{A_1}{(\omega - p_1)} + \frac{A_2}{(\omega - p_1)^2} + \frac{A_3}{(\omega - p_1)^3} + \frac{A_4}{(\omega - p_1^*)} + \frac{B_1}{(\omega - p_2)} + \frac{B_2}{(\omega - p_2)^2} + \frac{B_3}{(\omega - p_2)^3} + \frac{B_4}{(\omega - p_2^*)} \tag{8.99}$$

where

$$A_4 = -\frac{1}{(-p_1 + p_1^*)^3(-p_1^* + p_2)^3(-p_1^* + p_2^*)} \tag{8.100}$$

$$A_3 = -\frac{1}{(-p_1 + p_1^*)(-p_1 + p_2)^3(p_1 - p_2^*)} \tag{8.101}$$

$$A_2 = -\frac{-5p_1^2 + 4p_1p_1^* + 2p_1p_2 - p_1^*p_2 + 4p_1p_2^* - 3p_1^*p_2^* - p_2p_2^*}{(-p_1 + p_1^*)^2(-p_1 + p_2)^4(-p_1 + p_2^*)^2} \tag{8.102}$$

and finally

$$A_1 = -\frac{N_1}{D_1} \tag{8.103}$$

with

$$\begin{aligned} N_1 = {} & -15p_1^4 + 24p_1^3p_1^* - 10p_1^2p_1^{*2} + 12p_1^3p_2 - 15p_1^2p_1^*p_2 + 5p_1p_1^{*2}p_2 \\ & - 3p_1^2p_2^2 + 3p_1p_1^*p_2^2 - p_1^{*2}p_2^2 + 24p_1^3p_2^* - 37p_1^2p_1^*p_2^* + 15p_1p_1^{*2}p_2^* \\ & - 15p_1^2p_2p_2^* + 14p_1p_1^*p_2p_2^* - 3p_1^{*2}p_2p_2^* + 3p_1p_2^2p_2^* - p_1^*p_2^2p_2^* - 10p_1^2p_2^{*2} \\ & + 15p_1p_1^*p_2^{*2} - 6p_1^{*2}p_2^{*2} + 5p_1p_2p_2^{*2} - 3p_1^*p_2p_2^{*2} - p_2^2p_2^{*2} \end{aligned} \tag{8.104}$$

$$D_1 = (-p_1 + p_1^*)^3(-p_1 + p_2)^5(-p_1 + p_2^*)^3. \tag{8.105}$$

The B coefficients are obtained simply by interchanging the 1 and 2 subscripts throughout.

Given the formula for the distortion, it is sufficient to consider only those terms in (8.99) with poles in the lower half-plane. Further, it is sufficient to concentrate on the pole at p_1^* as the expression for p_2^* will follow on interchanging the subscripts 1 and 2. Hence

$$\Lambda_{\mathrm{s}}^-(\omega) = \frac{3X^2k_3}{2m^4(p_1 - p_1^*)^3(p_2 - p_1^*)^3(p_1^* - p_2^*)(\omega - p_1^*)} + (p_1 \longleftrightarrow p_2). \tag{8.106}$$

On substituting for p_1 and p_2, in terms of the physical parameters, the $\mathrm{O}(X^2)$ distortion (denoted here by $\Delta^{(2)}\Lambda_{\mathrm{s}}(\omega)$) finally emerges as

$$\Delta^{(2)}\Lambda_{\mathrm{s}}(\omega) = -\left(\frac{3X^2k_3}{2m^4}\right) \times \frac{4\omega_n\zeta(-\omega_{\mathrm{d}}^2+\omega_n^2\zeta^2)+\mathrm{i}\omega(\omega_{\mathrm{d}}^2-3\omega_n^2\zeta^2)}{64\omega_n^3\zeta^3(\omega_{\mathrm{d}}^2+\omega_n^2\zeta^2)^3(\omega-\omega_{\mathrm{d}}+\mathrm{i}\omega_n\zeta)(\omega+\omega_{\mathrm{d}}+\mathrm{i}\omega_n\zeta)}. \tag{8.107}$$

A similar but more involved analysis for the $\mathrm{O}(X^4)$ distortion yields the following six terms which generate $\Delta^{(4)}\Lambda_{\mathrm{s}}(\omega)$—all other terms lie in $\Lambda_{\mathrm{s}}^{+}(\omega)$.

$$\begin{aligned}\Delta^{(4)}\Lambda_{\mathrm{s}}(\omega) = \left(\frac{3X^4k_3^2}{16m^7}\right)\Bigg[&\frac{3}{2048\omega_{\mathrm{d}}^3\omega_n^4\zeta^4(\omega-\omega_{\mathrm{d}}+\mathrm{i}\omega_n\zeta)^3(\mathrm{i}\omega_{\mathrm{d}}+\omega_n\zeta)^4}\\ &+\frac{-3}{2048\omega_{\mathrm{d}}^3\omega_n^4\zeta^4(\omega+\omega_{\mathrm{d}}+\mathrm{i}\omega_n\zeta)^3(-\mathrm{i}\omega_{\mathrm{d}}+\omega_n\zeta)^4}\\ &+\frac{12\mathrm{i}\omega_{\mathrm{d}}^3+89\omega_{\mathrm{d}}^2\omega_n\zeta-150\mathrm{i}\omega_{\mathrm{d}}\omega_n^2\zeta^2-36\omega_n^3\zeta^3}{8192\omega_{\mathrm{d}}^4\omega_n^5\zeta^5(\omega-\omega_{\mathrm{d}}+\mathrm{i}\omega_n\zeta)^2(\mathrm{i}\omega_{\mathrm{d}}+\omega_n\zeta)^5(\mathrm{i}\omega_{\mathrm{d}}+2\omega_n\zeta)}\\ &+\frac{-12\mathrm{i}\omega_{\mathrm{d}}^3+89\omega_{\mathrm{d}}^2\omega_n\zeta+150\mathrm{i}\omega_{\mathrm{d}}\omega_n^2\zeta^2-36\omega_n^3\zeta^3}{8192\omega_{\mathrm{d}}^4\omega_n^5\zeta^5(\omega+\omega_{\mathrm{d}}+\mathrm{i}\omega_n\zeta)^2(-\mathrm{i}\omega_{\mathrm{d}}+\omega_n\zeta)^5(-\mathrm{i}\omega_{\mathrm{d}}+2\omega_n\zeta)}\\ &+T_5+T_6\end{aligned} \tag{8.108}$$

where T_5 is the quotient N_5/D_5 with

$$\begin{aligned}N_5 = {}&280\mathrm{i}\omega_{\mathrm{d}}^5+1995\omega_{\mathrm{d}}^4\omega_n\zeta-5080\mathrm{i}\omega_{\mathrm{d}}^3\omega_n^2\zeta^2-5344\omega_{\mathrm{d}}^2\omega_n^3\zeta^3\\ &+2016\mathrm{i}\omega_{\mathrm{d}}\omega_n^4\zeta^4+288\omega_n^5\zeta^5\end{aligned} \tag{8.109}$$

$$D_5 = 32\,768\omega_{\mathrm{d}}^5\omega_n^5\zeta^5(\omega-\omega_{\mathrm{d}}+\mathrm{i}\omega_n\zeta)(\mathrm{i}\omega_{\mathrm{d}}+\omega_n\zeta)^6(\mathrm{i}\omega_{\mathrm{d}}+2\,\omega_n\zeta)^2 \tag{8.110}$$

and T_6 is given by N_6/D_6 where

$$\begin{aligned}N_6 = {}&280\mathrm{i}\omega_{\mathrm{d}}^5-1995\omega_{\mathrm{d}}^4\omega_n\zeta-5080\mathrm{i}\omega_{\mathrm{d}}^3\omega_n^2\zeta^2+5344\omega_{\mathrm{d}}^2\omega_n^3\zeta^3\\ &+2016\mathrm{i}\omega_{\mathrm{d}}\omega_n^4\zeta^4-288\omega_n^5\zeta^5\end{aligned} \tag{8.111}$$

and

$$D_6 = 32\,768\omega_{\mathrm{d}}^5\omega_n^5\zeta^5(\omega+\omega_{\mathrm{d}}+\mathrm{i}\omega_n\zeta)(-\mathrm{i}\omega_{\mathrm{d}}+\omega_n\zeta)^6(-\mathrm{i}\,\omega_{\mathrm{d}}+2\omega_n\zeta)^2]. \tag{8.112}$$

Using the $\mathrm{O}(X^2)$ and $\mathrm{O}(X^4)$ distortion terms, $\Delta^{(2)}\Lambda_{\mathrm{s}}(\omega)$ and $\Delta^{(4)}\Lambda_{\mathrm{s}}(\omega)$, the Hilbert transform of the Duffing oscillator FRF $\Lambda_{\mathrm{s}}(\omega)$ (represented by a three-term Volterra series) may be expressed as

$$\mathcal{H}[\Lambda_{\mathrm{s}}(\omega)] = \tilde{\Lambda}_{\mathrm{s}}(\omega) = \Lambda_{\mathrm{s}}(\omega)+\Delta^{(2)}\Lambda_{\mathrm{s}}(\omega)+\Delta^{(4)}\Lambda_{\mathrm{s}}(\omega). \tag{8.113}$$

This relationship may be used to calculate numerical values for the Hilbert transform as a function of frequency, forcing and level of nonlinearity.

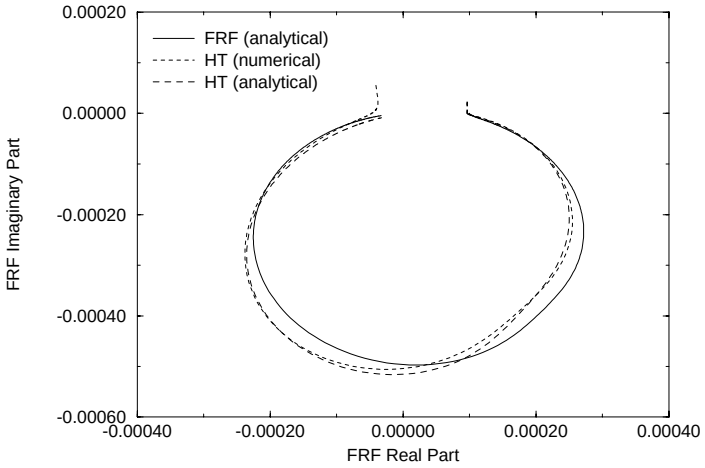


Figure 8.22. Comparison between the numerical estimate of the Hilbert transform and the $\mathrm{O}(X^2)$ Volterra series estimate for the Duffing oscillator under sine excitation. (Note that the Real and Imaginary axes do not have equal scales.)

8.6.2.3 Numerical example

Using the expressions for the $\mathrm{O}(X^2)$ and $\mathrm{O}(X^4)$ contributions to the nonlinear FRF (equation (8.91)), and the $\Delta^{(2)}\Lambda_{\mathrm{s}}(\omega)$ and $\Delta^{(4)}\Lambda_{\mathrm{s}}(\omega)$ distortion terms, a FORTRAN program was used to evaluate the FRF and Hilbert transform numerically for the particular Duffing oscillator given in (8.92). The expressions were obtained for 1024 spectral lines from 0 to 200 rad s^{-1}.

The FRF and HT expressions were evaluated for two levels of excitation, specified by $X = 0.5$ and 0.75.

Figure 8.22 shows an overlay of the Volterra series FRF (full line), and the associated analytical Hilbert transform (broken), as obtained from the $\Delta^{(2)}$ distortion term; this result was obtained using the excitation with $X = 0.75N$. The rotation of the Hilbert transform towards the left and the increase in amplitude over that of the FRF are both established features of the Hilbert transform of a Duffing oscillator (see chapter 4). The broken trace in figure 8.22 shows the Hilbert transform evaluated from the FRF by numerical means. Even using only the $\Delta^{(2)}$ distortion, the theory gives excellent agreement. With the fourth-order distortion included (figure 8.23), agreement is almost perfect. Note that the plots are effectively circular but that in the figures the Real and Imaginary axes are not of equal scales.

8.7 FRFs and Hilbert transforms: random excitation

The object of this section is to derive the composite FRF for a Duffing Oscillator under *random* excitation. Although the FRF mirrors the sine-excitation FRF

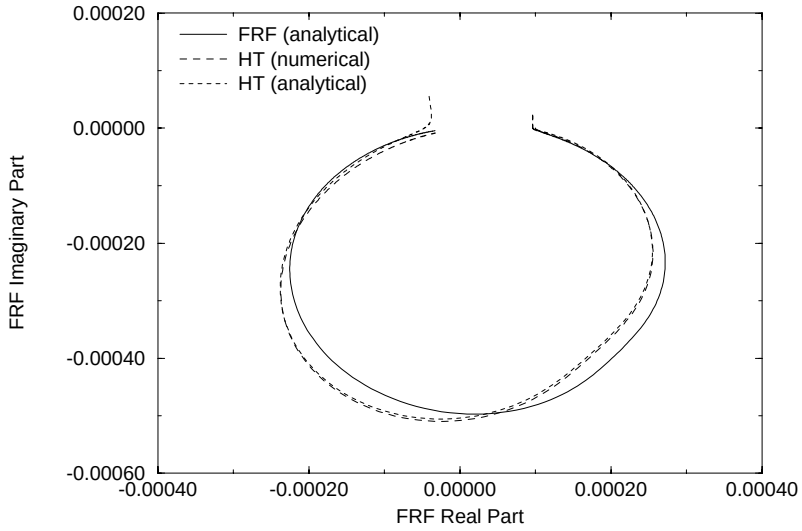


Figure 8.23. Comparison between the numerical estimate of the Hilbert transform and the $\mathrm{O}(X^4)$ Volterra series estimate for the Duffing oscillator under sine excitation. (Note that the Real and Imaginary axes do not have equal scales.)

in many respects, there are important differences. This section is this book's only real foray into the realm of random vibration. If the reader would like to study the subject in more depth, [198] is an excellent example of an introductory textbook. A considerably more advanced treatment can be found in [52], which treats nonlinear random vibration amongst other topics.

There have been a number of related calculations over the years. The simplest method of approximating an FRF for a nonlinear system is based on *equivalent linearization* [54]. This approach estimates the parameters of the linear system which is closest (in a statistical sense) to the original nonlinear system. The FRF of the linearized system is computed. In [75], statistical linearization was combined with perturbation analysis [68], in order to calculate the spectral response of a Duffing oscillator to white noise excitation. (This is equivalent to the FRF calculation up to a multiplicative constant.) It was shown that the FRF exhibits a secondary peak at three times the natural frequency, a result which is unavailable from statistical linearization alone. An approach based on perturbation theory alone is described in [147] and the calculation is carried to first order in the perturbation parameter. A number of studies of spectra have appeared based on the use of the Fokker–Planck–Kolmogorov equation (FPK) [55, 15, 137, 138, 284]. The latter two references actually examine the Duffing oscillator system which is studied in the current work. Good representations of the spectra were obtained; however, to the order of approximation pursued, the approach was unable to explain the presence of the secondary peak described

earlier. An interesting approach to approximating the spectral response of a Duffing oscillator is adopted in [184]. There, the expected response of an equivalent linear system was calculated where the natural frequency of the linear system was a random variable. The results compared favourably with numerical simulation, but the secondary peak could not be obtained. The Volterra series approach given here has been applied in [53] and [250], amongst others; however, the calculation was not carried far enough to allow a description of FRF distortions or the occurrence of the secondary peak. Using a Volterra series approach also allows the definition of higher(polynomial)-order equivalent systems, for example, the method of statistical quadratization is discussed in [78].

8.7.1 Volterra system response to a white Gaussian input

The problem of nonlinear system response to a generic random input is completely intractable. In order to make progress, it is usually assumed that the noise is white Gaussian. The power spectrum of such an input is constant over all frequencies and, as a consequence, Gaussian white noise is a physically unrealizable signal since it has infinite power. In practice, Gaussian white noise is approximated by Gaussian random processes that have sufficiently broad frequency bandwidth for the application of interest.

The definition of the FRF of a linear system based on the input/output cross-spectrum, $S_{yx}(\omega)$, and input auto-spectrum, $S_{xx}(\omega)$, is well known (and is repeated here for convenience)

$$H(\omega) = H_1(\omega) = \frac{S_{yx}(\omega)}{S_{xx}(\omega)}. \tag{8.114}$$

The *composite FRF*, $\Lambda_{\mathrm{r}}(\omega)$, of a nonlinear system under random excitation is defined similarly

$$\Lambda_{\mathrm{r}}(\omega) = \frac{S_{yx}(\omega)}{S_{xx}(\omega)}. \tag{8.115}$$

The term composite FRF is used again because $\Lambda_{\mathrm{r}}(\omega)$, for a nonlinear system, will not in general be equal to $H_1(\omega)$ but will receive contributions from all H_n. It will be shown that random excitation leads to a different composite FRF than sine excitation, hence the identifying subscript. The FRF also depends on the power spectral density of the input. However, $\Lambda_{\mathrm{r}}(\omega)$ tends to the linear FRF as the power spectral density of the excitation tends to zero.

In order to obtain a more detailed expression for $\Lambda_{\mathrm{r}}(\omega)$, an expression for $S_{yx}(\omega)$ must be derived. Using the Volterra series representation given in (8.3) results in the expression

$$\Lambda_{\mathrm{r}}(\omega) = \frac{S_{y_1x}(\omega) + S_{y_2x}(\omega) + \cdots + S_{y_nx}(\omega) + \cdots}{S_{xx}(\omega)}. \tag{8.116}$$

$\Lambda_{\mathrm{r}}(\omega)$ will be approximated here by obtaining expressions for the various cross-spectra between the input and the individual output components. First, consider

the cross-correlation function $\phi_{y_1x}(\tau)$; this is defined by $\phi_{y_1x}(\tau) = E[y_1(t)x(t-\tau)]$ where $E[\,]$ is the expected value operator. Substituting in the expression for the first-order component of the Volterra series response from (8.4) gives

$$\phi_{y_1x}(\tau) = E\left[\int_{-\infty}^{+\infty} \mathrm{d}\tau_1\, h_1(\tau_1)x(t-\tau_1)x(t-\tau)\right]. \tag{8.117}$$

It is known that the operations of taking the expected value and integrating commute, thus

$$\begin{aligned}\phi_{y_1x}(\tau) &= \int_{-\infty}^{+\infty} \mathrm{d}\tau_1\, h_1(\tau_1)E[x(t-\tau_1)x(t-\tau)] \\ &= \int_{-\infty}^{+\infty} \mathrm{d}\tau_1\, h_1(\tau_1)\phi_{xx}(\tau-\tau_1)\end{aligned} \tag{8.118}$$

where $\phi_{xx}(\tau)$ is the input autocorrelation function defined by $\phi_{xx}(\tau) = E[x(t)x(t-\tau)]$.

Taking Fourier transforms of both sides of this equation gives

$$S_{y_1x}(\omega) = \int_{-\infty}^{+\infty} \mathrm{d}\tau\, \mathrm{e}^{-\mathrm{i}\omega\tau} \int_{-\infty}^{+\infty} \mathrm{d}\tau_1\, h_1(\tau_1)\phi_{xx}(\tau-\tau_1) \tag{8.119}$$

and, changing the order of integration, gives

$$S_{y_1x}(\omega) = \int_{-\infty}^{+\infty} \mathrm{d}\tau_1\, h_1(\tau_1) \int_{-\infty}^{+\infty} \mathrm{d}\tau\, \mathrm{e}^{-\mathrm{i}\omega\tau}\phi_{xx}(\tau-\tau_1). \tag{8.120}$$

Using the Fourier transform shift theorem yields

$$S_{y_1x}(\omega) = \int_{-\infty}^{+\infty} \mathrm{d}\tau_1\, h_1(\tau_1)\mathrm{e}^{-\mathrm{i}\omega\tau_1}S_{xx}(\omega) = H_1(\omega)S_{xx}(\omega). \tag{8.121}$$

The result is no more than the expression for the linear FRF as stated in (8.2). However, the example serves to illustrate the methods used to obtain expressions for the cross-spectra between the input and higher-order output components.

To obtain the $S_{y_2x}(\omega)$ term the expression for the second-order component of the Volterra series response is substituted into the equation $\phi_{y_2x}(\tau) = E[y_2(t)x(t-\tau)]$. Following a similar procedure as before gives

$$\begin{aligned}\phi_{y_2x}(\tau) &= E\left[\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \mathrm{d}\tau_1\, \mathrm{d}\tau_2\, h_2(\tau_1,\tau_2)x(t-\tau_1)x(t-\tau_2)x(t-\tau)\right] \\ &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \mathrm{d}\tau_1\, \mathrm{d}\tau_2\, h_2(\tau_1,\tau_2)E[x(t-\tau_1)x(t-\tau_2)x(t-\tau)].\end{aligned} \tag{8.122}$$

It can be shown [158] that for zero-mean Gaussian variables $x_1, x_2, \ldots, x_n, \ldots$

$$E[x_1 x_2 \ldots x_n] = 0 \tag{8.123}$$

if n is odd and if n is even

$$E[x_1 x_2 \ldots x_n] = \sum \prod E[x_i x_j] \tag{8.124}$$

where $\sum \prod$ means the sum of the products of $E[x_i x_j]$, the pairs $x_i x_j$ being taken from $x_1, x_2, \ldots, x_n$ in all the possible distinct ways.

It follows from (8.123) that all cross-correlation functions, and hence all cross-spectra, between the input and the even-order output components will be zero, i.e. $S_{y_{2n}x}(\omega) = \phi_{y_{2n}x}(\tau) = 0, \forall n$.

Moving on to the $S_{y_3x}(\omega)$ term, this is given by

$$\begin{aligned}\phi_{y_3x}(\tau) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \mathrm{d}\tau_1\,\mathrm{d}\tau_2\,\mathrm{d}\tau_3\, h_3(\tau_1,\tau_2,\tau_3) \\ \times E[x(t-\tau_1)x(t-\tau_2)x(t-\tau_3)x(t-\tau)].\end{aligned} \tag{8.125}$$

From (8.124) the expected value of the product of inputs, i.e. the fourth-order moment of the input, reduces to the following product of second-order moments,

$$\begin{aligned}&E[x(t-\tau_1)x(t-\tau_2)x(t-\tau_3)x(t-\tau)] \\ &\quad = E[x(t-\tau_1)x(t-\tau_2)]E[x(t-\tau_3)x(t-\tau)] \\ &\qquad + E[x(t-\tau_1)x(t-\tau_3)]E[x(t-\tau_2)x(t-\tau)] \\ &\qquad + E[x(t-\tau_1)x(t-\tau)]E[x(t-\tau_2)x(t-\tau_3)].\end{aligned} \tag{8.126}$$

Using this equation and taking advantage of the symmetry of the Volterra kernels leads to

$$\begin{aligned}\phi_{y_3x}(\tau) &= 3\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \mathrm{d}\tau_1\,\mathrm{d}\tau_2\,\mathrm{d}\tau_3\, h_3(\tau_1,\tau_2,\tau_3) \\ &\quad \times E[x(t-\tau_1)x(t-\tau_2)]E[x(t-\tau_3)x(t-\tau)] \\ &= 3\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \mathrm{d}\tau_1\,\mathrm{d}\tau_2\,\mathrm{d}\tau_3\, h_3(\tau_1,\tau_2,\tau_3) \\ &\quad \times \phi_{xx}(\tau_2-\tau_1)\phi_{xx}(\tau-\tau_3).\end{aligned} \tag{8.127}$$

Fourier transforming this equation and manipulating the result eventually yields

$$S_{y_3x}(\omega) = \frac{3S_{xx}(\omega)}{2\pi}\int_{-\infty}^{+\infty} \mathrm{d}\omega_1\, H_3(\omega_1, -\omega_1, \omega) S_{xx}(\omega_1). \tag{8.128}$$

This result is already available in the literature [25]. Its presence here is justified by the fact that the derivation of the general term is a simple modification.

The general term is

$$S_{y_{2n-1}x}(\omega) = \frac{(2n)!S_{xx}(\omega)}{n!\,2^n(2\pi)^{n-1}}\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}\mathrm{d}\omega_1\ldots\mathrm{d}\omega_{n-1} \times H_{2n-1}(\omega_1,-\omega_1,\ldots,\omega_{n-1},-\omega_{n-1},\omega) \times S_{xx}(\omega_1)\ldots S_{xx}(\omega_{n-1}). \tag{8.129}$$

Now, given that the input autospectrum is constant over all frequencies for a Gaussian white noise input (i.e. $S_{xx}(\omega) = P$), the composite FRF for random excitation follows. Substituting (8.129) into (8.116) gives

$$\Lambda_{\mathrm{r}}(\omega) = \sum_{n=1}^{n=\infty}\frac{(2n)!P^{n-1}}{n!\,2^n(2\pi)^{n-1}}\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}\mathrm{d}\omega_1\ldots\mathrm{d}\omega_{n-1} \times H_{2n-1}(\omega_1,-\omega_1,\ldots,\omega_{n-1},-\omega_{n-1},\omega). \tag{8.130}$$

This equation will be used to analyse the effect of a Gaussian white noise input on the SDOF Duffing oscillator system.

8.7.2 Random excitation of a classical Duffing oscillator

Using the theory developed in the last section, an expression for $\Lambda_{\mathrm{r}}(\omega)$ up to $\mathrm{O}(P^2)$ will be calculated for the standard system (8.49) with $k_2 = 0$. From (8.130) the first three terms are given by

$$\begin{aligned}\frac{S_{y_1x}(\omega)}{S_{xx}(\omega)} &= H_1(\omega)\\ \frac{S_{y_3x}(\omega)}{S_{xx}(\omega)} &= \frac{3P}{2\pi}\int_{-\infty}^{+\infty}\mathrm{d}\omega_1\, H_3(\omega_1,-\omega_1,\omega)\\ \frac{S_{y_5x}(\omega)}{S_{xx}(\omega)} &= \frac{15P^2}{(2\pi)^2}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\mathrm{d}\omega_1\,\mathrm{d}\omega_2\, H_5(\omega_1,-\omega_1,\omega_2,-\omega_2,\omega).\end{aligned} \tag{8.131}$$

The first term of this equation needs no further work but the others require expressions for the HFRF terms as functions of the H_1s and k_3. The results for H_3 and H_5 are given in (8.87) and (8.89) respectively, the specific forms needed for (8.131) are

$$H_3(\omega_1,-\omega_1,\omega) = -k_3H_1(\omega)^2H_1(\omega_1)H_1(-\omega_1) = -k_3H_1(\omega)^2|H_1(\omega_1)|^2 \tag{8.132}$$

and

$$\begin{aligned}H_5(\omega_1,-\omega_1,\omega_2,-\omega_2,\omega) = \frac{3k_3^2}{10}&H_1(\omega)^2H_1(\omega_1)H_1(-\omega_1)H_1(\omega_2)H_1(-\omega_2)\\ &\times\{2H_1(\omega)+H_1(\omega_1)+H_1(-\omega_1)+H_1(\omega_2)+H_1(-\omega_2)\\ &+H_1(\omega_1+\omega_2+\omega)+H_1(\omega_1-\omega_2+\omega)\end{aligned}$$

$$
\begin{aligned}
&+ H_1(-\omega_1+\omega_2+\omega) + H_1(-\omega_1-\omega_2+\omega)\} \\
&= \frac{3k_3^2}{10} H_1(\omega)^2 |H_1(\omega_1)|^2 |H_1(\omega_2)|^2 \{2H_1(\omega) + H_1(\omega_1) + H_1(-\omega_1) \\
&+ H_1(\omega_2) + H_1(-\omega_2) + H_1(\omega_1+\omega_2+\omega) + H_1(\omega_1-\omega_2+\omega) \\
&+ H_1(-\omega_1+\omega_2+\omega) + H_1(-\omega_1-\omega_2+\omega)\} \qquad (8.133)
\end{aligned}
$$

So only one integral needs to be evaluated for $\frac{S_{y_3x}(\omega)}{S_{xx}(\omega)}$ compared to nine for $\frac{S_{y_5x}(\omega)}{S_{xx}(\omega)}$.

Substituting (8.132) into the $\frac{S_{y_3x}(\omega)}{S_{xx}(\omega)}$ term of (8.131) gives

$$
\frac{S_{y_3x}(\omega)}{S_{xx}(\omega)} = -\frac{3Pk_3H_1(\omega)^2}{2\pi}\int_{-\infty}^{+\infty} \mathrm{d}\omega_1 \, |H_1(\omega_1)|^2. \qquad (8.134)
$$

This integral may be found in standard tables of integrals used for the calculation of mean-square response, e.g. [198]. However, the analysis is instructive and it will allow the definition of notation for the integrals which follow.

Consider the common expression for the linear FRF (8.97). In terms of this, the integral in (8.134), $|H_1(\omega_1)|^2$ may be written

$$
|H_1(\omega_1)|^2 = \frac{1}{m^2(\omega_1-p_1)(\omega_1-p_2)(\omega_1-p_1^*)(\omega_1-p_2^*)} \qquad (8.135)
$$

where p_1^* (respectively p_2^*) is the complex conjugate of p_1 (respectively p_2). The integral is straightforwardly evaluated using the calculus of residues[5]. The approach is well known and details can be found in numerous textbooks, e.g. [6]. The appropriate contour is given in figure 8.24. (In the calculation, the radius of the semicircle is allowed to go to infinity.) The result of the calculation is the standard

$$
\int_{-\infty}^{+\infty} \mathrm{d}\omega \, |H(\omega)|^2 = \frac{\pi}{2m^2\zeta\omega_n(\omega_\mathrm{d}^2+\zeta^2\omega_n^2)} = \frac{\pi}{ck_1}. \qquad (8.136)
$$

Substituting this expression into (8.134) gives

$$
\frac{S_{y_3x}(\omega)}{S_{xx}(\omega)} = -\frac{3Pk_3H_1(\omega)^2}{2ck_1} \qquad (8.137)
$$

for this system. It can be seen that the third-order component of the response does not change the position of the poles of $\Lambda_\mathrm{r}(\omega)$ from those of the linear FRF. In fact it creates double poles at the positions where simple poles were previously found. The effect of the next non-zero term will now be analysed.

[5] Although the *approach* is straightforward, the algebraic manipulations involve rather complicated expressions. In order to facilitate the analysis, a computer algebra package was used throughout this work.

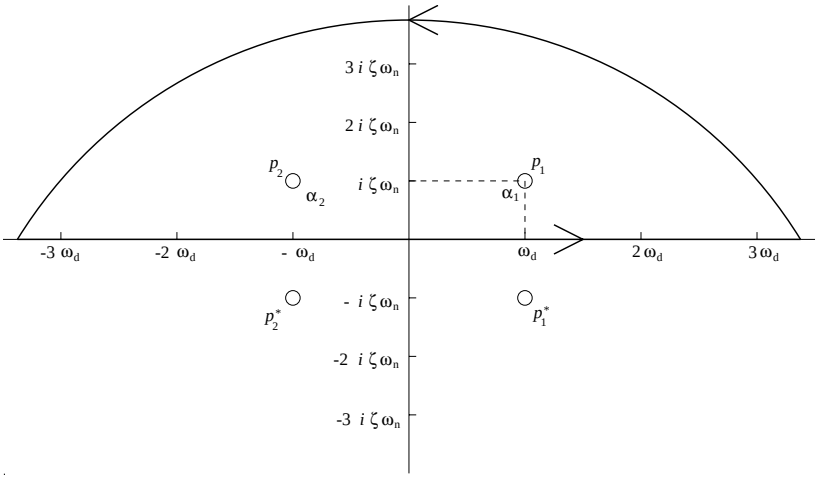


Figure 8.24. The pole structure of $|H_1(\omega_1)|^2$ with the integration contour shown closed in the upper-half of the ω_1-plane.

Substituting (8.133) into the $\frac{S_{y_5x}(\omega)}{S_{xx}(\omega)}$ term of (8.131) gives the integrals

$$\begin{aligned}\frac{S_{y_5x}(\omega)}{S_{xx}(\omega)} = \frac{9P^2k_3^2H_1(\omega)^2}{8\pi^2}\Bigg\{&2H_1(\omega)\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\mathrm{d}\omega_1\,\mathrm{d}\omega_2\,|H_1(\omega_1)|^2|H_1(\omega_2)|^2\\
&+\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\mathrm{d}\omega_1\,\mathrm{d}\omega_2\,H_1(\omega_1)|H_1(\omega_1)|^2|H_1(\omega_2)|^2\\
&+\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\mathrm{d}\omega_1\,\mathrm{d}\omega_2\,H_1(-\omega_1)|H_1(\omega_1)|^2|H_1(\omega_2)|^2\\
&+\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\mathrm{d}\omega_1\,\mathrm{d}\omega_2\,H_1(\omega_2)|H_1(\omega_1)|^2|H_1(\omega_2)|^2\\
&+\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\mathrm{d}\omega_1\,\mathrm{d}\omega_2\,H_1(-\omega_2)|H_1(\omega_1)|^2|H_1(\omega_2)|^2\\
&+\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\mathrm{d}\omega_1\,\mathrm{d}\omega_2\,H_1(\omega_1+\omega_2+\omega)|H_1(\omega_1)|^2|H_1(\omega_2)|^2\\
&+\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\mathrm{d}\omega_1\,\mathrm{d}\omega_2\,H_1(\omega_1-\omega_2+\omega)|H_1(\omega_1)|^2|H_1(\omega_2)|^2\\
&+\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\mathrm{d}\omega_1\,\mathrm{d}\omega_2\,H_1(-\omega_1+\omega_2+\omega)|H_1(\omega_1)|^2|H_1(\omega_2)|^2\\
&+\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\mathrm{d}\omega_1\,\mathrm{d}\omega_2\,H_1(-\omega_1-\omega_2+\omega)|H_1(\omega_1)|^2|H_1(\omega_2)|^2\Bigg\}.\end{aligned} \tag{8.138}$$

However, this is not as daunting as it first appears; exchanging ω_1 for $-\omega_1$ in the second integral results in the expression

$$\int_{+\infty}^{-\infty}\int_{-\infty}^{+\infty}(-\mathrm{d}\omega_1)\,\mathrm{d}\omega_2\,H_1(-\omega_1)|H_1(-\omega_1)|^2|H_1(\omega_2)|^2. \tag{8.139}$$

Noting that $|H_1(-\omega_1)|^2 = |H_1(\omega_1)|^2$ results in the expression

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\mathrm{d}\omega_1\,\mathrm{d}\omega_2\,H_1(-\omega_1)|H_1(\omega_1)|^2|H_1(\omega_2)|^2 \tag{8.140}$$

which is identical to the third integral. A similar argument shows that the fourth and fifth integrals are identical to the second and third. It is also straightforward to show that the last four integrals are identical to each other. This means that (8.138) reduces to only three integrals, i.e.

$$\begin{aligned}\frac{S_{y_5x}(\omega)}{S_{xx}(\omega)} &= \frac{9P^2k_3^2H_1(\omega)^3}{4\pi^2}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\mathrm{d}\omega_1\,\mathrm{d}\omega_2\,|H_1(\omega_1)|^2|H_1(\omega_2)|^2\\ &\quad+\frac{9P^2k_3^2H_1(\omega)^2}{2\pi^2}\\ &\quad\times\left\{\mathrm{Re}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\mathrm{d}\omega_1\,\mathrm{d}\omega_2\,H_1(\omega_1)|H_1(\omega_1)|^2|H_1(\omega_2)|^2\right.\\ &\quad\left.+\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\mathrm{d}\omega_1\,\mathrm{d}\omega_2\,H_1(\omega_1+\omega_2+\omega)|H_1(\omega_1)|^2|H_1(\omega_2)|^2\right\}.\end{aligned} \tag{8.141}$$

The first integral may be evaluated straightforwardly as

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\mathrm{d}\omega_1\,\mathrm{d}\omega_2\,|H_1(\omega_1)|^2|H_1(\omega_2)|^2 = \left\{\int_{-\infty}^{+\infty}\mathrm{d}\omega_1\,|H_1(\omega_1)|^2\right\}^2 = \frac{\pi^2}{c^2k_1^2} \tag{8.142}$$

on making use of (8.136).

The second integral can also be factorized

$$\begin{aligned}&\mathrm{Re}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\mathrm{d}\omega_1\,\mathrm{d}\omega_2\,H_1(\omega_1)|H_1(\omega_1)|^2|H_1(\omega_2)|^2\\ &\quad=\mathrm{Re}\int_{-\infty}^{+\infty}\mathrm{d}\omega_1\,H_1(\omega_1)|H_1(\omega_1)|^2\int_{-\infty}^{+\infty}\mathrm{d}\omega_2\,|H_1(\omega_2)|^2\end{aligned} \tag{8.143}$$

and the second integral is given by (8.136). Contour integration can again be used to evaluate the first part. The integrand may be written

$$H_1(\omega_1)|H_1(\omega_1)|^2 = \frac{1}{-m^3(\omega_1-p_1)^2(\omega_1-p_2)^2(\omega_1-p_1^*)(\omega_1-p_2^*)} \tag{8.144}$$

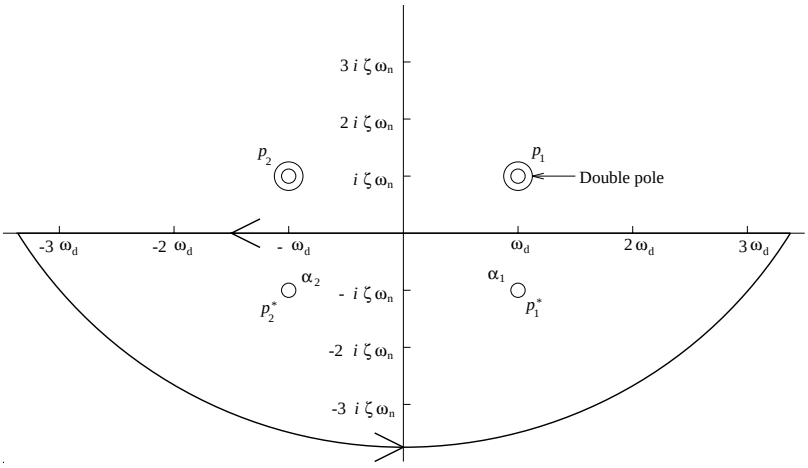


Figure 8.25. The pole structure of $H_1(\omega_1)|H_1(\omega_1)|^2$ with the integration contour shown closed in the lower-half of the ω_1-plane.

and this has a similar pole structure to the $\frac{S_{y_3x}(\omega)}{S_{xx}(\omega)}$ integrand, except the poles at p_1 and p_2 are now double poles (figure 8.25). The contour should be closed in the lower half of the ω_1-plane as there are only simple poles there. The result is

$$\mathrm{Re}\int_{-\infty}^{+\infty} \mathrm{d}\omega_1\, H_1(\omega_1)|H_1(\omega_1)|^2 = \frac{\pi}{4m^3\zeta\omega_n(\omega_\mathrm{d}^2+\zeta^2\omega_n^2)^2} = \frac{\pi}{2ck_1^2} \tag{8.145}$$

and this combined with (8.136) yields

$$\mathrm{Re}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \mathrm{d}\omega_1\,\mathrm{d}\omega_2\, H_1(\omega_1)|H_1(\omega_1)|^2|H_1(\omega_2)|^2 = \frac{\pi^2}{2c^2k_1^3}. \tag{8.146}$$

The third and final integral of the $\frac{S_{y_5x}(\omega)}{S_{xx}(\omega)}$ expression is slightly more complicated as the integrand does not factorize.

$$\begin{aligned} I(\omega) &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \mathrm{d}\omega_1\,\mathrm{d}\omega_2\, H_1(\omega_1+\omega_2+\omega)|H_1(\omega_1)|^2|H_1(\omega_2)|^2 \\ &= \int_{-\infty}^{+\infty} \mathrm{d}\omega_2\,|H_1(\omega_2)|^2 \int_{-\infty}^{+\infty} \mathrm{d}\omega_1\, H_1(\omega_1+\omega_2+\omega)|H_1(\omega_1)|^2. \end{aligned} \tag{8.147}$$

The second integral must be solved first. In terms of poles,

$$\begin{aligned} &H_1(\omega_1+\omega_2+\omega)|H_1(\omega_1)|^2 \\ &\quad = \frac{1}{m^3(\omega_1-p_1)(\omega_1-p_2)(\omega_1-p_1^*)(\omega_1-p_2^*)(\omega_1-q_1)(\omega_1-q_2)} \end{aligned} \tag{8.148}$$

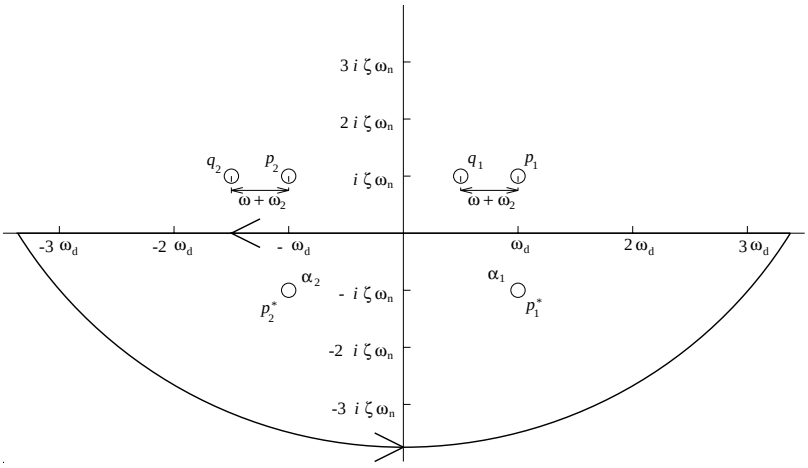


Figure 8.26. The pole structure of $H_1(\omega_1+\omega_2+\omega_3)|H_1(\omega_1)|^2$ with the integration contour shown closed in the lower-half of the ω_1-plane.

where p_1, p_2, p_1^* and p_2^* are the same as before and q_1 and q_2 are the poles of $H_1(\omega_1+\omega_2+\omega)$

$$\begin{aligned} q_1 &= p_1-\omega-\omega_2 = (\omega_{\rm d}-\omega-\omega_2)+{\rm i}\zeta\omega_n \\ q_2 &= p_2-\omega-\omega_2 = (-\omega_{\rm d}-\omega-\omega_2)+{\rm i}\zeta\omega_n. \end{aligned} \tag{8.149}$$

For simplicity, the contour is again closed on the lower half of the ω_1-plane (figure 8.26). The result is

$$\begin{aligned} &\int_{-\infty}^{+\infty} {\rm d}\omega_1\, H_1(\omega_1+\omega_2+\omega)|H_1(\omega_1)|^2 \\ &\quad = [-\pi(\omega+\omega_2-4{\rm i}\zeta\omega_n)][2m^3\zeta\omega_n(\omega_{\rm d}^2+\zeta^2\omega_n^2)(\omega+\omega_2-2{\rm i}\zeta\omega_n) \\ &\qquad \times(\omega+\omega_2+2\omega_{\rm d}-2{\rm i}\zeta\omega_n)(\omega+\omega_2-2\omega_{\rm d}-2{\rm i}\zeta\omega_n)]^{-1} \\ &\quad = [-\pi(\omega+\omega_2-4{\rm i}\zeta\omega_n)][mck_1(\omega+\omega_2-2{\rm i}\zeta\omega_n) \\ &\qquad \times(\omega+\omega_2+2\omega_{\rm d}-2{\rm i}\zeta\omega_n)(\omega+\omega_2-2\omega_{\rm d}-2{\rm i}\zeta\omega_n)]^{-1}. \end{aligned} \tag{8.150}$$

The expression above is then substituted into (8.147) and the integral over ω_2 evaluated. The integrand is this expression multiplied by $|H_1(\omega_2)|^2$. In terms of poles it is

$$\frac{-\pi(\omega+\omega_2-4{\rm i}\zeta\omega_n)}{m^3ck_1(\omega_2-p_1)(\omega_2-p_2)(\omega_2-p_1^*)(\omega_2-p_2^*)(\omega_2-q_1)(\omega_2-q_2)(\omega_2-r)} \tag{8.151}$$

where p_1, p_2, p_1^* and p_2^* are as before and

$$q_1 = -\omega+2\omega_{\rm d}+2{\rm i}\zeta\omega_n, \quad q_2 = -\omega-2\omega_{\rm d}+2{\rm i}\zeta\omega_n$$

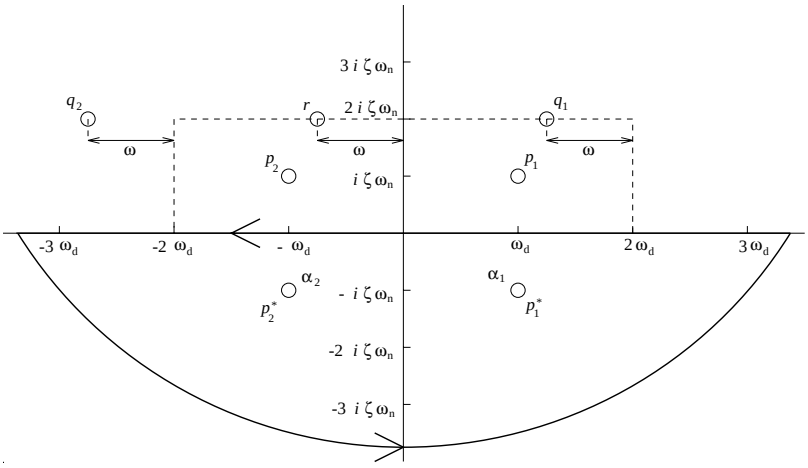


Figure 8.27. The pole structure of equation (8.151) with the integration contour shown closed in the lower-half of the ω_1-plane.

$$r = -\omega + 2\mathrm{i}\zeta\omega_n. \tag{8.152}$$

The contour is again closed on the lower half of the ω_2-plane (figure 8.27). Finally

$$\begin{aligned} I(\omega) = {} & [-\pi^2(\omega^2 - 3\omega_\mathrm{d}^2 - 10\mathrm{i}\zeta\omega_n\omega - 27\zeta^2\omega_n^2)] \\ & \times [mc^2k_1^2(\omega - \omega_\mathrm{d} - 3\mathrm{i}\zeta\omega_n)(\omega + \omega_\mathrm{d} - 3\mathrm{i}\zeta\omega_n) \\ & \times (\omega - 3\omega_\mathrm{d} - 3\mathrm{i}\zeta\omega_n)(\omega + 3\omega_\mathrm{d} - 3\mathrm{i}\zeta\omega_n)]^{-1}. \end{aligned} \tag{8.153}$$

Substituting (8.153), (8.146) and (8.142) into (8.141) gives the overall expression for $\frac{S_{y_5x}(\omega)}{S_{xx}(\omega)}$:

$$\frac{S_{y_5x}(\omega)}{S_{xx}(\omega)} = \frac{9P^2k_3^2H_1(\omega)^3}{4c^2k_1^2} + \frac{9P^2k_3^2H_1(\omega)^2}{4c^2k_1^3} + \frac{9P^2k_3^2H_1(\omega)^2}{2\pi^2}I(\omega) \tag{8.154}$$

for the classical Duffing oscillator. This equation shows that the first two terms do not affect the position of the poles of the linear system. The first term does, however, introduce triple poles at the position of the linear system poles. The term of greatest interest, though, is the final one which has introduced four new poles at

$$\omega_\mathrm{d} + 3\mathrm{i}\zeta\omega_n; \quad -\omega_\mathrm{d} + 3\mathrm{i}\zeta\omega_n; \quad 3\omega_\mathrm{d} + 3\mathrm{i}\zeta\omega_n; \quad -3\omega_\mathrm{d} + 3\mathrm{i}\zeta\omega_n. \tag{8.155}$$

The pole structure to this order is shown in figure 8.28. The poles at $\pm 3\omega_\mathrm{d} + 3\mathrm{i}\zeta\omega_n$ explain the secondary observed peak at three times the resonant frequency in the output spectra of nonlinear oscillators [285].

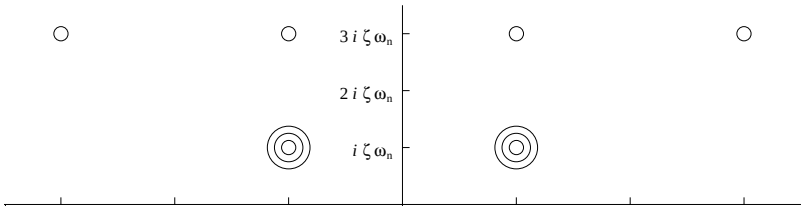


Figure 8.28. The pole structure of the first three terms of $\Lambda_{\mathrm{r}}(\omega)$ for the classical Duffing oscillator.

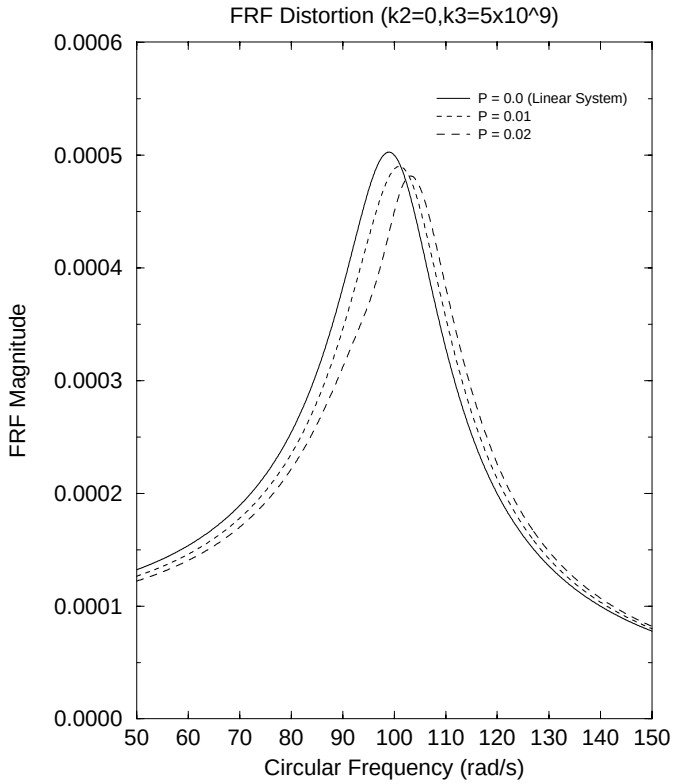


Figure 8.29. Composite FRF $\Lambda_{\mathrm{r}}(\omega)$ to order $\mathrm{O}(P^2)$ for the classical Duffing oscillator.

Combining (8.154) and (8.137) into (8.131) yields an expression for the composite FRF $\Lambda_{\mathrm{r}}(\omega)$ up to $\mathrm{O}(P^2)$. The magnitude of the composite FRF is plotted in figure 8.29 for values of P equal to 0 (linear system), 0.01 and 0.02. The Duffing oscillator parameters are $m = 1$, $c = 20$, $k_1 = 10^4$ and $k_3 = 5 \times 10^9$.

As observed in practice, the resonant frequency shifts upwards with increasing P, while the peak magnitude reduces. This is an encouraging level of agreement with experiment given that only three terms are taken in the expansion (8.116). Although the frequency shifts, there is no real evidence of the sort of distortions observed during stepped-sine testing. This lends support to the view that random excitation produces a 'linearized' FRF.

It is significant that all the poles are located in the upper half of the ω-plane. It is known (e.g. see chapter 5) that applying the Hilbert transform test to a system with all its poles in the upper half of the complex plane results in the system being labelled linear. If this behaviour continues for higher terms this then shows agreement with apparent linearization of FRFs obtained under random excitation.

In order to determine whether or not the inclusion of further terms in the $\Lambda_{\mathrm{r}}(\omega)$ approximation results in further poles arising at new locations, the fourth non-zero term (i.e. $\frac{S_{y_7x}(\omega)}{S_{xx}(\omega)}$) for this system was considered. The $H_7(\omega_1, -\omega_1, \omega_2, -\omega_2, \omega_3, -\omega_3, \omega)$ expression consists of 280 integrals when the problem is expressed in H_1 terms. However repeating the procedure of combining terms which yield identical integrals results in 13 integrals. The expression for $\frac{S_{y_7x}(\omega)}{S_{xx}(\omega)}$ is

$$
\begin{aligned}
\frac{S_{y_7x}(\omega)}{S_{xx}(\omega)} = & -\frac{27P^3k_3^3H_1(\omega)^4}{8\pi^3}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\mathrm{d}\omega_1\,\mathrm{d}\omega_2\,\mathrm{d}\omega_3 \\
& \times |H_1(\omega_1)|^2|H_1(\omega_2)|^2|H_1(\omega_3)|^2 \\
& -\frac{27P^3k_3^3H_1(\omega)^3}{2\pi^3}\left\{\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\mathrm{d}\omega_1\,\mathrm{d}\omega_2\,\mathrm{d}\omega_3\right. \\
& \times H_1(\omega_1)|H_1(\omega_1)|^2|H_1(\omega_2)|^2|H_1(\omega_3)|^2 \\
& +\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\mathrm{d}\omega_1\,\mathrm{d}\omega_2\,\mathrm{d}\omega_3 \\
& \left.\times H_1(\omega+\omega_1+\omega_2)|H_1(\omega_1)|^2|H_1(\omega_2)|^2|H_1(\omega_3)|^2\right\} \\
& -\frac{9P^3k_3^3H_1(\omega)^2}{8\pi^3}\left\{6\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\mathrm{d}\omega_1\,\mathrm{d}\omega_2\,\mathrm{d}\omega_3\right. \\
& \times H_1(\omega_1)^2|H_1(\omega_1)|^2|H_1(\omega_2)|^2|H_1(\omega_3)|^2 \\
& +3\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\mathrm{d}\omega_1\,\mathrm{d}\omega_2\,\mathrm{d}\omega_3\,|H_1(\omega_1)|^4|H_1(\omega_2)|^2|H_1(\omega_3)|^2 \\
& +12\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\mathrm{d}\omega_1\,\mathrm{d}\omega_2\,\mathrm{d}\omega_3\,H_1(\omega_1)H_1(\omega_2) \\
& \times |H_1(\omega_1)|^2|H_1(\omega_2)|^2|H_1(\omega_3)|^2 \\
& +24\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\mathrm{d}\omega_1\,\mathrm{d}\omega_2\,\mathrm{d}\omega_3\,H_1(\omega_1)H_1(\omega+\omega_1+\omega_2)
\end{aligned}
$$

$$
\begin{aligned}
&\times |H_1(\omega_1)|^2|H_1(\omega_2)|^2|H_1(\omega_3)|^2 \\
&+ 12\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \mathrm{d}\omega_1\,\mathrm{d}\omega_2\,\mathrm{d}\omega_3\, H_1(\omega_1)H_1(\omega_1+\omega_2+\omega_3) \\
&\times |H_1(\omega_1)|^2|H_1(\omega_2)|^2|H_1(\omega_3)|^2 \\
&+ 6\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \mathrm{d}\omega_1\,\mathrm{d}\omega_2\,\mathrm{d}\omega_3\, H_1(\omega+\omega_1+\omega_2)^2 \\
&\times |H_1(\omega_1)|^2|H_1(\omega_2)|^2|H_1(\omega_3)|^2 \\
&+ 24\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \mathrm{d}\omega_1\,\mathrm{d}\omega_2\,\mathrm{d}\omega_3\, H_1(\omega+\omega_1+\omega_2) \\
&\times H_1(\omega+\omega_1+\omega_3)|H_1(\omega_1)|^2|H_1(\omega_2)|^2|H_1(\omega_3)|^2 \\
&+ 12\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \mathrm{d}\omega_1\,\mathrm{d}\omega_2\,\mathrm{d}\omega_3\, H_1(\omega+\omega_1+\omega_2) \\
&\times H_1(\omega_1+\omega_2+\omega_3)|H_1(\omega_1)|^2|H_1(\omega_2)|^2|H_1(\omega_3)|^2 \\
&+ 12\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \mathrm{d}\omega_1\,\mathrm{d}\omega_2\,\mathrm{d}\omega_3\, H_1(\omega+\omega_1+\omega_2) \\
&\times H_1(-\omega_1-\omega_2+\omega_3)|H_1(\omega_1)|^2|H_1(\omega_2)|^2|H_1(\omega_3)|^2 \\
&+ 2\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \mathrm{d}\omega_1\,\mathrm{d}\omega_2\,\mathrm{d}\omega_3\, |H_1(\omega_1+\omega_2+\omega_3)|^2 \\
&\times |H_1(\omega_1)|^2|H_1(\omega_2)|^2|H_1(\omega_3)|^2\Bigg\}. \qquad (8.156)
\end{aligned}
$$

The evaluation of these integrals was carried out as before. However, the results will not be included here. The important point is whether the calculation introduces new poles. By extrapolating from the poles in the first three terms of $\Lambda_{\mathrm{r}}(\omega)$, a fourth term might be expected to introduce poles at

$$
\begin{aligned}
&\omega_{\mathrm{d}}+5\mathrm{i}\zeta\omega_n; \quad -\omega_{\mathrm{d}}+5\mathrm{i}\zeta\omega_n; \quad 3\omega_{\mathrm{d}}+5\mathrm{i}\zeta\omega_n; \\
&-3\omega_{\mathrm{d}}+5\mathrm{i}\zeta\omega_n; \quad 5\omega_{\mathrm{d}}+5\mathrm{i}\zeta\omega_n; \quad -5\omega_{\mathrm{d}}+5\mathrm{i}\zeta\omega_n. \qquad (8.157)
\end{aligned}
$$

However, after evaluating the integrals in (8.156), again by contour integration, it was found that no new poles arose. Instead, three of the integrals resulted in simple poles at the locations given in equation (8.155), whilst another three integrals resulted in double poles at these locations.

Due to the rapidly increasing level of difficulty associated with the addition of further terms to $\Lambda_{\mathrm{r}}(\omega)$ it was not possible to completely examine the $\frac{S_{y_9x}(\omega)}{S_{xx}(\omega)}$ term. It was possible, however, to consider one integral which would be included in the overall expression. The expression for $\frac{S_{y_9x}(\omega)}{S_{xx}(\omega)}$ is given by

$$
\begin{aligned}
\frac{S_{y_9x}(\omega)}{S_{xx}(\omega)} = \frac{945P^4}{(2\pi)^4}&\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \mathrm{d}\omega_1\,\mathrm{d}\omega_2\,\mathrm{d}\omega_3\,\mathrm{d}\omega_4 \\
&\times H_9(\omega_1,-\omega_1,\omega_2,-\omega_2,\omega_3,-\omega_3,\omega_4,-\omega_4,\omega). \qquad (8.158)
\end{aligned}
$$

One of the possible HFRF products which make up $H_9(\omega_1, -\omega_1, \omega_2, -\omega_2, \omega_3, -\omega_3, \omega_4, -\omega_4, \omega)$ is $H_5(\omega_1, \omega_2, \omega_3, \omega_4, \omega)H_3(-\omega_1, -\omega_2, -\omega_3)H_1(-\omega_4)$. This in turn results in several integrals, one of which is given by

$$\frac{27P^4k_3^4H_1(\omega)^2}{128\pi^4}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\mathrm{d}\omega_1\,\mathrm{d}\omega_2\,\mathrm{d}\omega_3\,\mathrm{d}\omega_4 \times H_1(\omega+\omega_1+\omega_2+\omega_3+\omega_4)H_1(\omega+\omega_1+\omega_2)H_1(-\omega_1-\omega_2-\omega_3) \times |H_1(\omega_1)|^2|H_1(\omega_2)|^2|H_1(\omega_3)|^2|H_1(\omega_4)|^2. \tag{8.159}$$

This integral was evaluated as before and was found to have triple poles at the locations given in equation (8.155), simple poles at the locations given in equation (8.157) and also at the following locations

$$\begin{aligned}&\omega_\mathrm{d}+7\mathrm{i}\zeta\omega_n; \quad -\omega_\mathrm{d}+7\mathrm{i}\zeta\omega_n; \quad 3\omega_\mathrm{d}+7\mathrm{i}\zeta\omega_n; \quad -3\omega_\mathrm{d}+7\mathrm{i}\zeta\omega_n;\\ &5\omega_\mathrm{d}+7\mathrm{i}\zeta\omega_n; \quad -5\omega_\mathrm{d}+7\mathrm{i}\zeta\omega_n; \quad 7\omega_\mathrm{d}+7\mathrm{i}\zeta\omega_n; \quad -7\omega_\mathrm{d}+7\mathrm{i}\zeta\omega_n.\end{aligned} \tag{8.160}$$

Although it is possible that these contributions cancel when combined with other integrals from $\frac{S_{yyx}(\omega)}{S_{xx}(\omega)}$, it can be conjectured that including all further terms would result in FRF poles for this system being witnessed at all locations $a\omega_\mathrm{d}+b\mathrm{i}\zeta\omega_n$ where $a \leq b$ are both odd integers. Note that this implies the existence of an infinite sequence of FRF peaks, each one associated with an odd multiple of the natural frequency.

The restriction of a and b to the odd integers might be expected from the fact that only odd-order HFRFs exist for systems which only contain odd polynomial nonlinearities. In that case the introduction of an even nonlinearity results in both odd- and even-order HFRFs. This is discussed in appendix K where a Duffing oscillator with additional quadratic spring stiffness is considered. Appendix K also shows an extension of the analysis to a simple MDOF system.

An interesting feature of this analysis is that the multiplicity of the poles increases with order P. The implication is that the poles will become isolated essential singularities in the limit. This implies rather interesting behaviour in the ω-plane near the poles as Picard's theorem [6] asserts that the FRF will take *all* complex values in any neighbourhood of the poles regardless of the size of the neighbourhood. This behaviour need not, of course, be visible from the real line.

8.8 Validity of the Volterra series

The question of validity is often ignored when the Volterra series is used. Not only should the series exist, but it should converge. Both of these requirements have been the subject of intense study over the years.

The first condition of importance for the Volterra series is the *existence* condition. When does an input–output functional S, as in $y(t) = S[x(t)]$, admit a Volterra representation of the form (8.3)? The definitive work on this question can

be found in [200], and requires fairly advanced techniques of functional analysis which are beyond the scope of this book. The important conclusion from [200] is that *'the class of Volterra-like representable systems is only restricted by a kind of smoothness condition'*. 'Smoothness' here means *sufficiently differentiable*. Now strictly speaking, the derivatives referred to are Fréchet derivatives of the functional S, but this translates easily onto a smoothness condition on the nonlinear *functions* in the equation of motion. For the purposes of this book, which is almost exclusively concerned with polynomial nonlinearities (which are, of course, infinitely differentiable), the Volterra series will always exist. The review [201] is a less rigorous but more readable introduction to questions of existence.

As the Volterra series is an infinite series, establishing existence immediately raises the vexed question of convergence. Namely, for what range of inputs $x(t)$ does the series converge. A concrete example for illustration is available now: the calculation of the composite FRF in section 8.8 assumed values for the input spectral density P. The convergence of the series for these P will be examined here.

General convergence results are few and far between. Ku and Wolf [156] established some useful results. For a bounded-input deterministic system, i.e. if

$$|x(t)| \leq K, \quad \text{for all } t \tag{8.161}$$

they showed that

$$\left| \sum_{i=1}^{\infty} y_i(t) \right| \leq \sum_{i=1}^{\infty} a_n K^n \tag{8.162}$$

where

$$a_n = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \mathrm{d}\tau_1 \dots \mathrm{d}\tau_n \, h(\tau_1, \dots, \tau_n) \tag{8.163}$$

and the radius of convergence of the Volterra series is given by

$$R = \left(\lim_{i \to \infty} |a_i|^{\frac{1}{i}} \right)^{-1}. \tag{8.164}$$

For the random case, they established that if $x(t)$ is stationary with bounded statistical moments then $\sum_{i=1}^{\infty} y_i$ converges in the mean if

$$\lim_{n \to \infty} \sum_{k=1}^{n} a_k < \infty \tag{8.165}$$

and

$$\lim_{k \to \infty} a_k = 0. \tag{8.166}$$

Fortunately, in order to validate the results of section 8.8, it is not necessary to use the general theory as results can be obtained for the classical Duffing

oscillator using the criteria developed by Barrett [21][6]. The first step is to convert the equation of motion (8.49) with $k_2 = 0$, to the normalized form

$$\ddot{y}' + 2\zeta\dot{y}' + y' + \epsilon y'^3 = x'(t') \tag{8.167}$$

and this is accomplished by the transformation

$$y' = \omega_n^2 y; \quad x' = \frac{x}{m}; \quad t' = \omega_n t \tag{8.168}$$

so that

$$\epsilon = \frac{k_3}{m\omega_n^6} \tag{8.169}$$

and ζ has the usual definition. Once in this coordinate system, convergence of the Volterra series is ensured as long as [21]

$$\|y'\| < y'_{\mathrm{b}} = \frac{1}{\sqrt{3\epsilon H}} \tag{8.170}$$

where

$$H = \coth\frac{\pi\zeta}{\sqrt{1-\zeta^2}}. \tag{8.171}$$

The norm $\|y'\|$ on an interval of time is simply the maximum value of y' over that interval. Using the values of section 8.4, $m = 1$, $c = 20$ and $k_1 = 10^4$, $k_3 = 5 \times 10^9$. The value of y'_{b} obtained is 4.514. This translates into a physical bound $y_{\mathrm{b}} = 4.514 \times 10^{-4}$.

Now, the mean-square response, σ_{y_l}, of the underlying linear system ($k_3 = 0$) is given by the standard formula

$$\sigma_{y_l}^2 = \int_{-\infty}^{\infty} \mathrm{d}\omega\, |H(\omega)|^2 S_{xx}(\omega) = P\int_{-\infty}^{\infty} \mathrm{d}\omega\, |H(\omega)|^2 = \frac{\pi P}{ck_1} \tag{8.172}$$

and in this case, $\sigma_{y_l} = 3.963\,32 \times 10^{-4}$ if $P = 0.01$ and $\sigma_{y_l} = 5.604\,99 \times 10^{-4}$ if $P = 0.02$. However, these results will be conservative if a non-zero k_3 is assumed. In fact for the nonlinear system [54, 68]

$$\sigma_{y_{nl}}^2 = \sigma_{y_l}^2 - 3\alpha\sigma_{y_l}^2 \tag{8.173}$$

to first order in $\alpha = k_3/k_1$. If this result is assumed valid (α is by no means small), the mean-square response of the cubic system can be found. It is $\sigma_{y_{nl}} = 3.465 \times 10^{-4}$ when $P = 0.01$ and $\sigma_{y_{nl}} = 4.075 \times 10^{-4}$ when $P = 0.02$. In the first case, the Barrett bound is 1.3 standard deviations and in the second case it is 1.11 standard deviations. Thus, using standard tables for Gaussian statistics [140], it is found that the Volterra series is valid with 80.6%

[6] Barrett derives his convergence criterion using a recursion relation; an alternative proof using functional analysis can be found in [63].

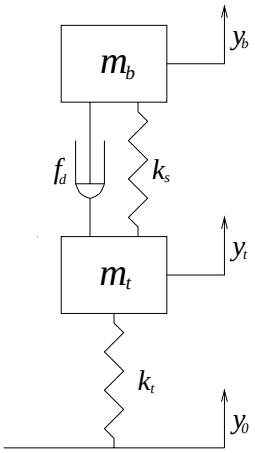


Figure 8.30. 2DOF automotive suspension lumped-mass model.

confidence if $P = 0.01$ and with 73.3 confidence if $P = 0.02$. As the Barrett bound is known to be conservative [173], these results were considered to lend support to the assumption of validity for the Volterra series.

For practical purposes, establishing convergence may not actually help, as the important question is *how many terms of the series need to be included in order to obtain a required accuracy in the representation?*. Not surprisingly, there has been little success in establishing an answer to this question. Some discussion can be found in [282] and [254].

8.9 Harmonic probing for a MDOF system

Having established the harmonic probing techniques needed for SDOF systems, the study can now proceed to more realistic MDOF models. A lumped-mass model of an automotive suspension (figure 8.30) [115] will be used to illustrate the techniques. Note that when a MDOF system is considered, there are two possible positions for the nonlinearities, between two of the masses and between a mass and ground; both cases will be discussed here.

The equations of motion of the system are

$$\begin{aligned} m_{\mathrm{t}}\ddot{y}_{\mathrm{t}} + f_{\mathrm{d}}(\dot{y}_{\mathrm{t}} - \dot{y}_{\mathrm{b}}) + (k_{\mathrm{t}} + k_{\mathrm{s}})y_{\mathrm{t}} - k_{\mathrm{s}}y_{\mathrm{b}} &= k_{\mathrm{t}}y_0 \\ m_{\mathrm{b}}\ddot{y}_{\mathrm{b}} - f_{\mathrm{d}}(\dot{y}_{\mathrm{t}} - \dot{y}_{\mathrm{b}}) + k_{\mathrm{s}}(y_{\mathrm{b}} - y_{\mathrm{t}}) &= 0 \end{aligned} \tag{8.174}$$

where m_{t}, k_{t} and y_{t} are the mass, stiffness and displacement of the tyre—the *unsprung mass*. m_{b} and y_{b} are the mass and displacement of the body or the *sprung mass* (this is usually taken as one-quarter of the total car body mass). f_{d} is the characteristic of the nonlinear damper and k_{s} is the (linear) stiffness

of the suspension. A cubic characteristic will be assumed for the damper, i.e. $f_{\mathrm{d}}(z) = c_1 z + c_2 z^2 + c_3 z^3$. y_0 is the displacement at the road surface and acts here as the excitation.

Each of the processes $y_0 \longrightarrow y_{\mathrm{t}}$ and $y_0 \longrightarrow y_{\mathrm{b}}$ has its own Volterra series and its own HFRFs and it will be shown later that the responses will depend on the kernel transforms from both series. The notation is defined by

$$\begin{aligned} Y_{\mathrm{t}}(\omega) = {} & H_1^{\mathrm{t}}(\omega)Y_0(\omega) + \frac{1}{(2\pi)}\int_{-\infty}^{\infty} \mathrm{d}\omega_1\, H_2^{\mathrm{t}}(\omega_1, \omega-\omega_1)Y_0(\omega_1)Y_0(\omega-\omega_1) \\ & + \frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{d}\omega_1\, \mathrm{d}\omega_2\, H_3^{\mathrm{t}}(\omega_1, \omega_2, \omega-\omega_1-\omega_2) \\ & \times Y_0(\omega_1)Y_0(\omega_2)Y_0(\omega-\omega_1-\omega_2) + \cdots \end{aligned} \tag{8.175}$$

$$\begin{aligned} Y_{\mathrm{b}}(\omega) = {} & H_1^{\mathrm{b}}(\omega)Y_0(\omega) + \frac{1}{(2\pi)}\int_{-\infty}^{\infty} \mathrm{d}\omega_1\, H_2^{\mathrm{b}}(\omega_1, \omega-\omega_1)Y_0(\omega_1)Y_0(\omega-\omega_1) \\ & + \frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{d}\omega_1\, \mathrm{d}\omega_2\, H_3^{\mathrm{b}}(\omega_1, \omega_2, \omega-\omega_1-\omega_2) \\ & \times Y_0(\omega_1)Y_0(\omega_2)Y_0(\omega-\omega_1-\omega_2) + \cdots . \end{aligned} \tag{8.176}$$

The harmonic probing procedure for MDOF systems is a straightforward extension of that for SDOF systems. In order to obtain the first-order (linear) kernel transforms, the probing expressions

$$y_0^{\mathrm{p}} = \mathrm{e}^{\mathrm{i}\Omega t} \tag{8.177}$$

$$y_{\mathrm{t}}^{\mathrm{p}} = H_1^{\mathrm{t}}(\Omega)\mathrm{e}^{\mathrm{i}\Omega t} + \cdots \tag{8.178}$$

and

$$y_{\mathrm{b}}^{\mathrm{p}} = H_1^{\mathrm{b}}(\Omega)\mathrm{e}^{\mathrm{i}\Omega t} + \cdots \tag{8.179}$$

are used. These expressions are substituted into the equations of motion (8.174), and the coefficients of $\mathrm{e}^{\mathrm{i}\Omega t}$ are extracted as before. The resulting equations are, in matrix form

$$\begin{pmatrix} -m_{\mathrm{t}}\Omega^2 + \mathrm{i}c_1\Omega + k_{\mathrm{t}} + k_{\mathrm{s}} & -\mathrm{i}c_1\Omega - k_{\mathrm{s}} \\ -\mathrm{i}c_1\Omega - k_{\mathrm{s}} & -m_{\mathrm{b}}\Omega^2 + \mathrm{i}c_1\Omega + k_{\mathrm{s}} \end{pmatrix} \begin{pmatrix} H_1^{\mathrm{t}}(\Omega) \\ H_1^{\mathrm{b}}(\Omega) \end{pmatrix} = \begin{pmatrix} k_{\mathrm{t}} \\ 0 \end{pmatrix}. \tag{8.180}$$

This 2×2 system has a straightforward solution:

$$H_1^{\mathrm{t}}(\omega) = \frac{k_{\mathrm{t}}(-m_{\mathrm{b}}\omega^2 + \mathrm{i}c_1\omega + k_{\mathrm{s}})}{(-m_{\mathrm{t}}\omega^2 + \mathrm{i}c_1\omega + k_{\mathrm{t}} + k_{\mathrm{s}})(-m_{\mathrm{b}}\omega^2 + \mathrm{i}c_1\omega + k_{\mathrm{s}}) - (\mathrm{i}c_1\omega + k_{\mathrm{s}})^2} \tag{8.181}$$

$$H_1^{\mathrm{b}}(\omega) = \frac{k_{\mathrm{t}}(\mathrm{i}c_1\omega + k_{\mathrm{s}})}{(-m_{\mathrm{t}}\omega^2 + \mathrm{i}c_1\omega + k_{\mathrm{t}} + k_{\mathrm{s}})(-m_{\mathrm{b}}\omega^2 + \mathrm{i}c_1\omega + k_{\mathrm{s}}) - (\mathrm{i}c_1\omega + k_{\mathrm{s}})^2}. \tag{8.182}$$

It will prove useful later to establish a little notation:

$$\Phi(\omega) = \begin{pmatrix} -m_{\mathrm{t}}\omega^2 + \mathrm{i}c_1\omega + k_{\mathrm{t}} + k_{\mathrm{s}} & -\mathrm{i}c_1\omega - k_{\mathrm{s}} \\ -\mathrm{i}c_1\omega - k_{\mathrm{s}} & -m_{\mathrm{b}}\omega^2 + \mathrm{i}c_1\omega + k_{\mathrm{s}} \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ -1 \end{pmatrix}. \tag{8.183}$$

The second-order kernel transforms are obtained using the probing expressions

$$y_0^{\mathrm{p}} = \mathrm{e}^{\mathrm{i}\Omega_1 t} + \mathrm{e}^{\mathrm{i}\Omega_2 t} \tag{8.184}$$

$$y_{\mathrm{t}}^{\mathrm{p}} = H_1^{\mathrm{t}}(\Omega_1)\mathrm{e}^{\mathrm{i}\Omega_1 t} + H_1^{\mathrm{t}}(\Omega_2)\mathrm{e}^{\mathrm{i}\Omega_2 t} + 2H_2^{\mathrm{t}}(\Omega_1,\Omega_2)\mathrm{e}^{\mathrm{i}(\Omega_1+\Omega_2)t} + \cdots \tag{8.185}$$

and

$$y_{\mathrm{b}}^{\mathrm{p}} = H_1^{\mathrm{b}}(\Omega_1)\mathrm{e}^{\mathrm{i}\Omega_1 t} + H_1^{\mathrm{b}}(\Omega_2)\mathrm{e}^{\mathrm{i}\Omega_2 t} + 2H_2^{\mathrm{b}}(\Omega_1,\Omega_2)\mathrm{e}^{\mathrm{i}(\Omega_1+\Omega_2)t} + \cdots. \tag{8.186}$$

These expressions are substituted into the equations of motion (8.174) and the coefficients of $\mathrm{e}^{\mathrm{i}(\Omega_1+\Omega_2)t}$ are extracted. The resulting equations are, in matrix form

$$\begin{aligned} &\begin{pmatrix} -m_{\mathrm{t}}(\Omega_1+\Omega_2)^2 + \mathrm{i}c_1(\Omega_1+\Omega_2) + k_{\mathrm{t}} + k_{\mathrm{s}} & -\mathrm{i}c_1(\Omega_1+\Omega_2) - k_{\mathrm{s}} \\ -\mathrm{i}c_1(\Omega_1+\Omega_2) - k_{\mathrm{s}} & -m_{\mathrm{b}}(\Omega_1+\Omega_2)^2 + \mathrm{i}c_1(\Omega_1+\Omega_2) + k_{\mathrm{s}} \end{pmatrix} \\ &\quad \times \begin{pmatrix} H_2^{\mathrm{t}}(\Omega_1,\Omega_2) \\ H_1^{\mathrm{b}}(\Omega_1,\Omega_2) \end{pmatrix} \\ &\quad = \begin{pmatrix} 1 \\ -1 \end{pmatrix} c_2\Omega_1\Omega_2[H_1^{\mathrm{t}}(\Omega_1) - H_1^{\mathrm{b}}(\Omega_1)][H_1^{\mathrm{t}}(\Omega_2) - H_1^{\mathrm{b}}(\Omega_2)] \end{aligned} \tag{8.187}$$

so

$$\begin{pmatrix} H_2^{\mathrm{t}}(\omega_1,\omega_2) \\ H_1^{\mathrm{b}}(\omega_1,\omega_2) \end{pmatrix} = \Phi(\omega_1+\omega_2)c_2\omega_1\omega_2[H_1^{\mathrm{t}}(\omega_1) - H_1^{\mathrm{b}}(\omega_1)][H_1^{\mathrm{t}}(\omega_2) - H_1^{\mathrm{b}}(\omega_2)]. \tag{8.188}$$

The calculation of the third-order kernel transforms proceeds as before, except a three-tone probing expression is used:

$$y_0^{\mathrm{p}} = \mathrm{e}^{\mathrm{i}\Omega_1 t} + \mathrm{e}^{\mathrm{i}\Omega_2 t} + \mathrm{e}^{\mathrm{i}\Omega_3 t}. \tag{8.189}$$

The result of the computation is

$$\begin{pmatrix} H_3^{\mathrm{t}}(\omega_1,\omega_2,\omega_3) \\ H_3^{\mathrm{b}}(\omega_1,\omega_2,\omega_3) \end{pmatrix} = \Phi(\omega_1+\omega_2+\omega_3)(F(\omega_1,\omega_2,\omega_3) + G(\omega_1,\omega_2,\omega_3)) \tag{8.190}$$

where

$$\begin{aligned} F(\omega_1,\omega_2,\omega_3) = \frac{2}{3}c_2 \sum_{\mathcal{C}} &\omega_1(\omega_2+\omega_3)[H_1^{\mathrm{t}}(\omega_1) - H_1^{\mathrm{b}}(\omega_1)] \\ &\times [H_2^{\mathrm{t}}(\omega_2,\omega_3) - H_2^{\mathrm{b}}(\omega_2,\omega_3)] \end{aligned} \tag{8.191}$$

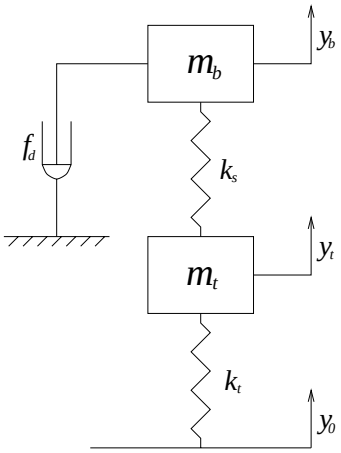


Figure 8.31. 2DOF skyhook model of the automotive suspension.

where $\sum_{\mathcal{C}}$ denotes a sum over cyclic permutations of ω_1, ω_2 and ω_3. Also

$$G(\omega_1, \omega_2, \omega_3) = \mathrm{i}c_3 \prod_{n=1}^{3} \omega_n [H_1^{\mathrm{t}}(\omega_n) - H_1^{\mathrm{b}}(\omega_n)]. \tag{8.192}$$

Note that to obtain $Y_{\mathrm{t}}(\omega)$ (respectively $Y_{\mathrm{b}}(\omega)$) requires a specification of H_3^{t} (respectively H_3^{b}) and H_1^{t}, H_1^{b}, H_2^{t} and H_2^{b}.

This theory is easily adapted to the case of the 'sky-hook' suspension model (figure 8.31) [128]. The equations of motion are

$$\begin{aligned} m_{\mathrm{t}}\ddot{y}_{\mathrm{t}} + (k_{\mathrm{t}} + k_{\mathrm{s}})y_{\mathrm{t}} - k_{\mathrm{s}}y_{\mathrm{b}} &= k_{\mathrm{t}}y_0 \\ m_{\mathrm{b}}\ddot{y}_{\mathrm{b}} + f_{\mathrm{d}}(\dot{y}_{\mathrm{b}}) + k_{\mathrm{s}}(y_{\mathrm{b}} - y_{\mathrm{t}}) &= 0. \end{aligned} \tag{8.193}$$

The H_1 functions for this system are given by

$$H_1^{\mathrm{t}}(\omega) = \frac{k_{\mathrm{t}}(-m_{\mathrm{b}}\omega^2 + \mathrm{i}c_1\omega + k_{\mathrm{s}})}{(-m_{\mathrm{t}}\omega^2 + k_{\mathrm{t}} + k_{\mathrm{s}})(-m_{\mathrm{b}}\omega^2 + ic_1\omega + k_{\mathrm{s}}) - k_{\mathrm{s}}^2} \tag{8.194}$$

$$H_1^{\mathrm{b}}(\omega) = \frac{k_{\mathrm{t}} + k_{\mathrm{s}}}{(-m_{\mathrm{t}}\omega^2 + k_{\mathrm{t}} + k_{\mathrm{s}})(-m_{\mathrm{b}}\omega^2 + \mathrm{i}c_1\omega + k_{\mathrm{s}}) - k_{\mathrm{s}}^2}. \tag{8.195}$$

The H_2 functions are given by

$$\begin{pmatrix} H_2^{\mathrm{t}}(\omega_1, \omega_2) \\ H_1^{\mathrm{b}}(\omega_1, \omega_2) \end{pmatrix} = \Psi(\omega_1 + \omega_2)c_2\omega_1\omega_2 H_1^{\mathrm{b}}(\omega_1)H_1^{\mathrm{b}}(\omega_2) \tag{8.196}$$

where

$$\Psi(\omega) = \begin{pmatrix} -m_{\mathrm{t}}\omega^2 + k_{\mathrm{t}} + k_{\mathrm{s}} & -k_{\mathrm{s}} \\ -k_{\mathrm{s}} & -m_{\mathrm{b}}\omega^2 + \mathrm{i}c_1\omega + k_{\mathrm{s}} \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{8.197}$$

Finally, the H_3s are given by

$$\begin{pmatrix} H_3^{\mathrm{t}}(\omega_1,\omega_2,\omega_3) \\ H_3^{\mathrm{b}}(\omega_1,\omega_2,\omega_3) \end{pmatrix} = \Psi(\omega_1+\omega_2+\omega_3)(J(\omega_1,\omega_2,\omega_3)+K(\omega_1,\omega_2,\omega_3)) \tag{8.198}$$

where

$$J(\omega_1,\omega_2,\omega_3) = \frac{2}{3}c_2\sum_{\mathcal{C}}\omega_1(\omega_2+\omega_3)H_1^{\mathrm{b}}(\omega_1)H_2^{\mathrm{b}}(\omega_2,\omega_3) \tag{8.199}$$

and

$$K(\omega_1,\omega_2,\omega_3) = \mathrm{i}c_3\prod_{n=1}^{3}\omega_n H_1^{\mathrm{b}}(\omega_n). \tag{8.200}$$

8.10 Higher-order modal analysis: hypercurve fitting

Linear modal analysis was discussed briefly in chapter 1. The philosophy of the approach was to extract the modal parameters for a given linear system by curve-fitting in the time or frequency domain. For a given linear system, this is an exercise in nonlinear optimization as the parameters of the system do not enter into the FRF expression in a linear manner, i.e.

$$H_1(\omega) = \sum_{i=1}^{N}\frac{A_i}{-(\omega^2-\omega_{ni}^2)+2i\zeta_i\omega_{ni}\omega+k}. \tag{8.201}$$

A remarkable observation due to Gifford [110] is that for a nonlinear system, the nonlinear parameters are much easier to obtain from the HFRFs than the linear parameters are from the H_1. To illustrate this, consider an asymmetrical Duffing oscillator as in (8.49), and assume for the moment that the FRFs are measurable. The linear parameters m, c and k can be obtained from the H_1:

$$H_1(\omega) = \frac{1}{-m\omega^2+\mathrm{i}c\omega+k} \tag{8.202}$$

using whatever curve-fitter is available. Now, if one considers H_2:

$$H_2(\omega_1,\omega_2) = -k_2H_1(\omega_1)H_1(\omega_2)H_1(\omega_1+\omega_2). \tag{8.203}$$

The quadratic stiffness coefficient k_2 enters as a *linear multiplier* for the product of H_1s. Most importantly, H_1 is now known as the linear parameters have been identified at the first stage. This useful property also holds for the H_3, which has the form

$$\begin{aligned} H_3(\omega_1,\omega_2,\omega_3) = & -\tfrac{1}{6}H_1(\omega_1+\omega_2+\omega_3)\{4k_2(H_1(\omega_1)H_2(\omega_2,\omega_3) \\ & + H_1(\omega_2)H_2(\omega_3,\omega_1)+H_1(\omega_3)H_2(\omega_1,\omega_2)) \\ & + 6k_3H_1(\omega_1)H_1(\omega_2)H_1(\omega_3)\} \end{aligned} \tag{8.204}$$

or

$$H_3(\omega_1,\omega_2,\omega_3) = k_2 F_1[H_1,H_2] + k_3 F_2[H_1] \tag{8.205}$$

so the H_3 function is also linear in the parameters k_2 and k_3. The importance of these observations is clear. Once H_1 and the linear parameters have been obtained from a nonlinear optimization step, the nonlinear parameters for the system can be obtained by linear least-squares analysis of the HFRFs.

The basis for Gifford's identification procedure is the general formulae for H_2 and H_3 for a general NDOF lumped-mass system. These are stated here without proof, the derivation can be found in [110].

$$\begin{aligned} H_2^{rs}(\omega_1,\omega_2) = & \sum_{m=1}^{N} (\omega_1\omega_2 c_{2mm} - k_{2mm}) H_1^{sm}(\omega_1+\omega_2) H_1^{rm}(\omega_1) H_1^{rm}(\omega_2) \\ & - \sum_{m=1}^{N-1}\sum_{n=m+1}^{N} (\omega_1\omega_2 c_{2mn} - k_{2mn})[H_1^{sm}(\omega_1+\omega_2) - H_1^{sn}(\omega_1+\omega_2)] \\ & \times [H_1^{rm}(\omega_1) - H_1^{rn}(\omega_1)][H_1^{rm}(\omega_2) - H_1^{rn}(\omega_2)] \end{aligned} \tag{8.206}$$

where c_{2mn} is the quadratic velocity coefficient for the connection between mass m and mass n (c_{2mm} is the connection to ground). k_{2mn} is the corresponding stiffness coefficient. H_2^{rs} etc. refer to the FRFs between DOF r and DOF s. Note that despite the considerable increase in complexity, the expressions are still linear in the parameters of interest. H_3 is more complex still:

$$\begin{aligned} H_3^{rs}(\omega_1,\omega_2,\omega_3) = & \frac{2}{3}\sum_{m=1}^{N} H_1^{sm}(\omega_1+\omega_2+\omega_3) \\ & \times [[\omega_1(\omega_2+\omega_3)c_{2mm} - k_{2mm}] H_1^{rm}(\omega_1) H_2^{rm}(\omega_2,\omega_3) \\ & + [\omega_2(\omega_3+\omega_1)c_{2mm} - k_{2mm}] H_1^{rm}(\omega_2) H_2^{rm}(\omega_3,\omega_1) \\ & + [\omega_3(\omega_1+\omega_2)c_{2mm} - k_{2mm}] H_1^{rm}(\omega_3) H_2^{rm}(\omega_1,\omega_2)] \\ & - \frac{2}{3}\sum_{m=1}^{N-1}\sum_{n=m+1}^{N} [H_1^{sm}(\omega_1+\omega_2+\omega_3) - H_1^{sn}(\omega_1+\omega_2+\omega_3)] \\ & \times [[\omega_1(\omega_2+\omega_3)c_{2mn} - k_{2mn}][H_1^{rm}(\omega_1) - H_1^{rn}(\omega_1)] \\ & \times [H_2^{rm}(\omega_2,\omega_3) - H_2^{rn}(\omega_2,\omega_3)][\omega_2(\omega_3+\omega_1)c_{2mn} - k_{2mn}] \\ & \times [H_1^{rm}(\omega_2) - H_1^{rn}(\omega_2)][H_2^{rm}(\omega_3,\omega_1) - H_2^{rn}(\omega_3,\omega_1)] \\ & \times [\omega_3(\omega_1+\omega_2)c_{2mn} - k_{2mn}][H_1^{rm}(\omega_3) - H_1^{rn}(\omega_3)] \\ & \times [H_2^{rm}(\omega_1,\omega_2) - H_2^{rn}(\omega_1,\omega_2)]] \\ & + \sum_{m=1}^{N} (i\omega_1\omega_2\omega_3 c_{3mm} - k_{3mm}) \\ & \times H_1^{sm}(\omega_1+\omega_2+\omega_3) H_1^{rm}(\omega_1) H_1^{rm}(\omega_2) H_1^{rm}(\omega_3) \end{aligned}$$

$$
\begin{aligned}
&- \sum_{m=1}^{N-1} \sum_{n=m+1}^{N} (i\omega_1\omega_2\omega_3 c_{3mn} - k_{3mn}) \\
&\times [H_1^{sm}(\omega_1+\omega_2+\omega_3) - H_1^{sn}(\omega_1+\omega_2+\omega_3)] \\
&\times [H_1^{rm}(\omega_1) - H_1^{rn}(\omega_1)][H_1^{rm}(\omega_2) - H_1^{rn}(\omega_2)][H_1^{rm}(\omega_3) - H_1^{rn}(\omega_3)]
\end{aligned}
\tag{8.207}
$$

and again, despite the complexity of the expression, the coefficients c_{2mn}, k_{2mn}, c_{3mn} and k_{3mn} enter the equation in a linear manner.

In principle then, the availability of measurements of the HFRFs up to order n would allow the identification of the parameters in the differential equations of motion up to the nth-power terms. The problem is to obtain the HFRFs accurately and without bias. Gifford used a random excitation test which is described in the next section along with an illustrative example.

8.10.1 Random excitation

The basic premise of Gifford's testing procedure is that at low excitations, the HFRFs can be approximated by certain correlation functions[7]. For the first-order FRF, this is confirmed by (8.130) where

$$
\frac{S_{yx}(\omega)}{S_{xx}(\omega)} = \Lambda_{\mathrm{r}1}(\omega) = H_1(\omega) + \mathrm{O}(P). \tag{8.208}
$$

This type of relation holds true for the higher-order FRFs also. It can be shown by the methods of section 8.7 that

$$
\begin{aligned}
\frac{S_{y'xx}(\omega_1,\omega_2)}{S_{xx}(\omega_1)S_{xx}(\omega_2)} &= \Lambda_{\mathrm{r}2}(\omega_1,\omega_2) \\
&= \Gamma_n \left(\frac{P}{2\pi}\right)^{\frac{n}{2}-1} \int_{-\infty}^{\infty} \mathrm{d}\omega^{(1)} \ldots \mathrm{d}\omega^{(\frac{n}{2}-1)} \\
&\quad \times H_n(\omega_1,\omega_2,\omega^{(1)},-\omega^{(1)},\ldots,\omega^{(\frac{n}{2}-1)},-\omega^{(\frac{n}{2}-1)})
\end{aligned}
\tag{8.209}
$$

where y' is the output signal with the mean removed and

$$
\Gamma_n = \begin{cases} n(n-1)(n-3)\ldots 1, & n \text{ even} \\ 0, & n \text{ odd.} \end{cases} \tag{8.210}
$$

Recall, that (8.208) assumes that the input $x(t)$ is white Gaussian. This is also an assumption used in the derivation of (8.209). The most important

[7] The arguments of this section do not directly follow [110]. The analysis there makes extensive use of the Wiener series, which is essentially an orthogonalized version of the Volterra series. The important facts which are needed here are simply that the correlations used approach the HFRFs in the limit of vanishing P and these can be demonstrated without recourse to the Wiener series.

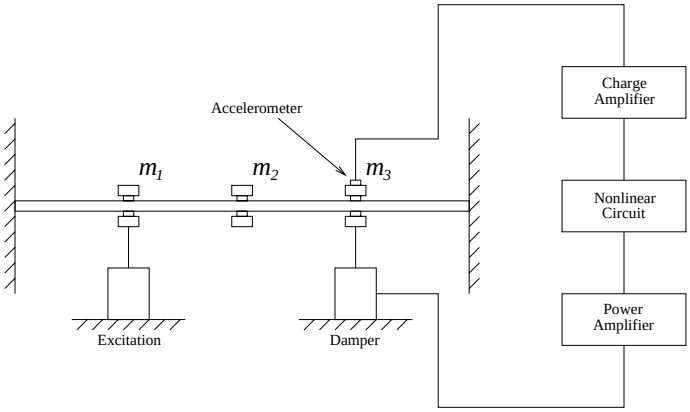


Figure 8.32. General arrangement of experimental nonlinear beam rig.

consequence of (8.209) is that

$$\frac{S_{y'xx}(\omega_1,\omega_2)}{S_{xx}(\omega_1)S_{xx}(\omega_2)} = H_2(\omega_1,\omega_2) + \mathrm{O}(P) \tag{8.211}$$

so as $P \longrightarrow 0$, the correlation function tends towards H_2.

The theory described here was validated by experiment in [110]. The structure used for the experiment was a forerunner of the beam used in section 7.5.3 and is shown in figure 8.32. The main difference between this and the rig used to validate the restoring force surface approach is that Gifford's beam was fixed–fixed and the nonlinear damping force was applied to the third lumped mass instead of the second. Figure 8.33 shows a first-order FRF for the system (the feedback circuit which provides the nonlinear damping force is switched off), the fourth mode is sufficiently far removed from the first three to make this a credible 3DOF system as the excitation force is band-limited appropriately.

Figure 8.34 shows the form of the feedback circuit used to provide the cubic velocity force on mass m_3. After collecting input v_1, and output voltage v_2, data from the nonlinear circuit alone, a polynomial curve-fit established the circuit characteristics

$$v_2 = 1.34v_1 + 1.25v_1^2 + 0.713v_1^3. \tag{8.212}$$

The overall gain of the feedback loop was obtained by measuring the FRF between input v_1 and output v_4 when the circuit was in linear mode; this is shown in figure 8.35. This FRF is very flat at all the frequencies of interest. Using the gain from the plot and all the appropriate calibration factors from the instrumentation, the linear force–velocity characteristics of the nonlinear feedback loop were obtained as

$$F_{\mathrm{f}} = -120\dot{y}_3 \tag{8.213}$$

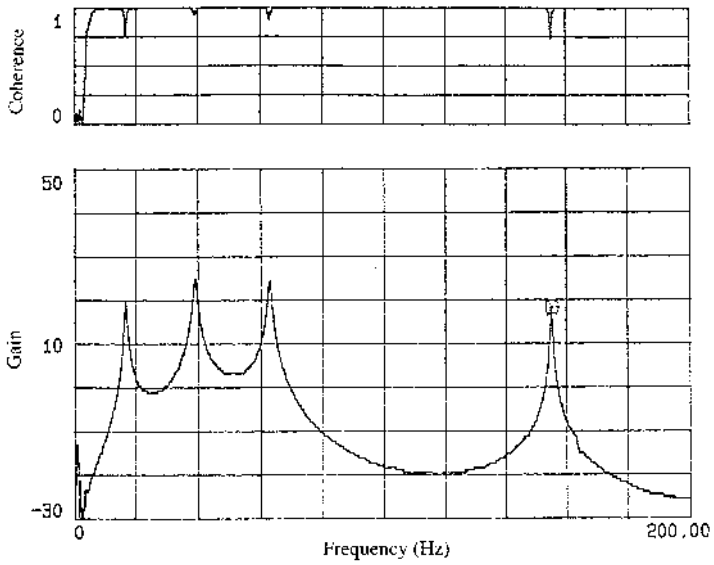


Figure 8.33. First-order frequency response function for experimental nonlinear beam structure under low-level random excitation.

so from (8.213), it follows that the nonlinear characteristic of the loop is

$$F_{\mathrm{f}} = -120\dot{y}_3 - 3540\dot{y}_3^2 - 638\,000\dot{y}_3^3. \tag{8.214}$$

When the experiment was carried out, the level of excitation was set low enough for distortion effects on the H_1 and H_2 to be minimal. Figure 8.36 shows a driving-point FRF for the system with and without the nonlinear part of the circuit switched in; the effect of the nonlinearity is apparently invisible.

In order to estimate the H_2 functions, the autocorrelation function

$$\phi_{y'xx}(\tau_1, \tau_2) = E[y'(t)x(t+\tau_1)x(t+\tau_2)] \tag{8.215}$$

was estimated by an averaging process and then Fourier transformed. The resulting HFRFs are shown in figure 8.37 for the processes $x \longrightarrow y_1$, $x \longrightarrow y_2$ and $x \longrightarrow y_3$. Altogether, 286 000 time samples were used to form these estimates. This is because a substantial number of averages are needed to adequately smooth the HFRFs. Having said this it may be that 'smoothness' from a visual point of view may not be critical; it depends on the noise tolerance of the curve-fitting procedure. Also if a sufficient number of H_2 points are sampled, it may be possible to tolerate a higher degree of noise.

Before the LS procedure for fitting the quadratic terms can be carried out, the H_1 functions for the system must be determined by the nonlinear curve-fitting procedure. Figure 8.38 shows the three first-order FRFs (in real and imaginary

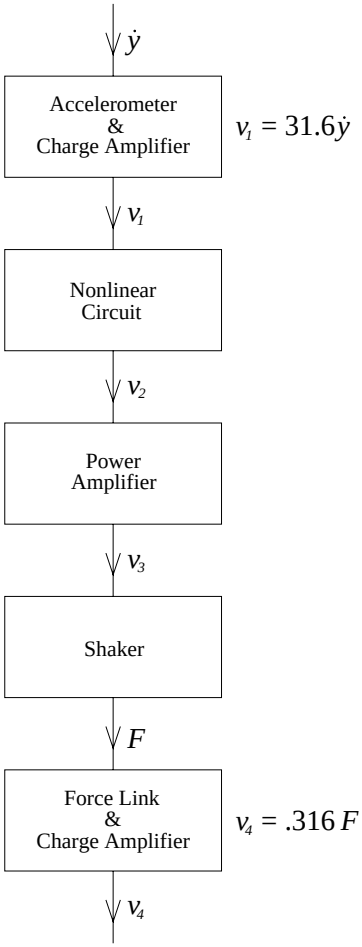


Figure 8.34. Block diagram of all the elements that make up the linear damping system with all their calibration factors in the configuration in which the results were measured.

form), together with the results of a standard modal analysis global curve-fitter; the results are very close indeed, even by linear system standards.

In order to carry out the LS step which fits the model (8.206), it is necessary to select the data from the (ω_1, ω_2)-plane. By analogy with common H_1 curve-fitters, the data are chosen from around the peaks in order to maximize the signal-to-noise ratio. Figure 8.39 shows how the data is selected to fit H_2.

Gifford fitted three models to the H_2 measurements: one assuming a quadratic stiffness characteristic, one assuming quadratic damping and one assuming both. Figure 8.40 shows the H_2 reconstruction from the two 'pure'

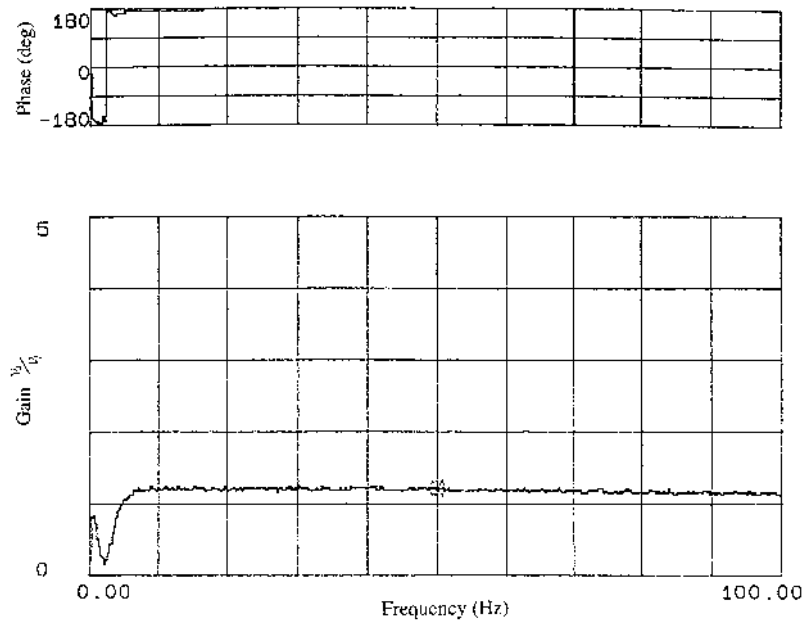


Figure 8.35. A measured FRF of v_4 over v_1, see figure 8.34.

models compared to the measured data. It is clear that the damping model (on the right) is the appropriate one. The coefficients obtained from a local curve-fit to the second FRF H_2^{12} were as given in table 8.2.

By far the dominant contribution comes from the c_{233} term as expected. The estimated value compares well with the calibration value of -3540 (15.8% error). When a global LS fit was carried out using data from all three H_2 curves at once, the coefficient estimate was refined to 3807, giving a percentage error of 10.3%.

The results of this exercise validate Gifford's method and offer a potentially powerful means of using the Volterra series and HFRFs for *parametric* identification.

8.10.2 Sine excitation

Gifford's thesis inspired Storer [237] to make a number of useful observations. First and foremost, he noted that all the structural parameters of interest appear in the *diagonal* FRFs and it is therefore sufficient to curve-fit to them[8]. To illustrate, consider the Duffing oscillator again, the diagonal forms of (8.203) and (8.204)

[8] There is an implicit assumption throughout here that the FRFs contain enough modal information to completely characterize the system, given an assumed model order. The methods may fail if the FRFs are too severely truncated.

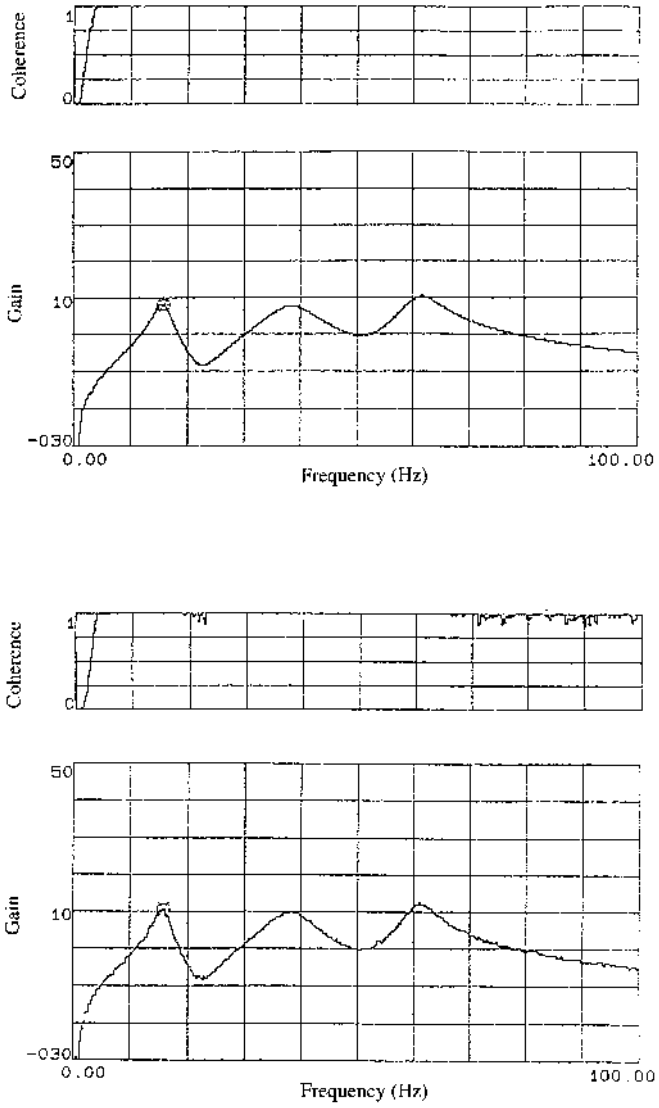


Figure 8.36. First-order FRF from point 1 of the nonlinear rig in both its linear and nonlinear modes.

can easily be found:

$$H_2(\omega,\omega) = -k_2 H_1(\omega)^2 H_1(2\omega) \tag{8.216}$$

$$H_3(\omega,\omega,\omega) = -\tfrac{1}{6} H_1(3\omega)(12k_2 H_1(\omega) H_2(\omega,\omega) + 6k_3 H_1(\omega)^3). \tag{8.217}$$

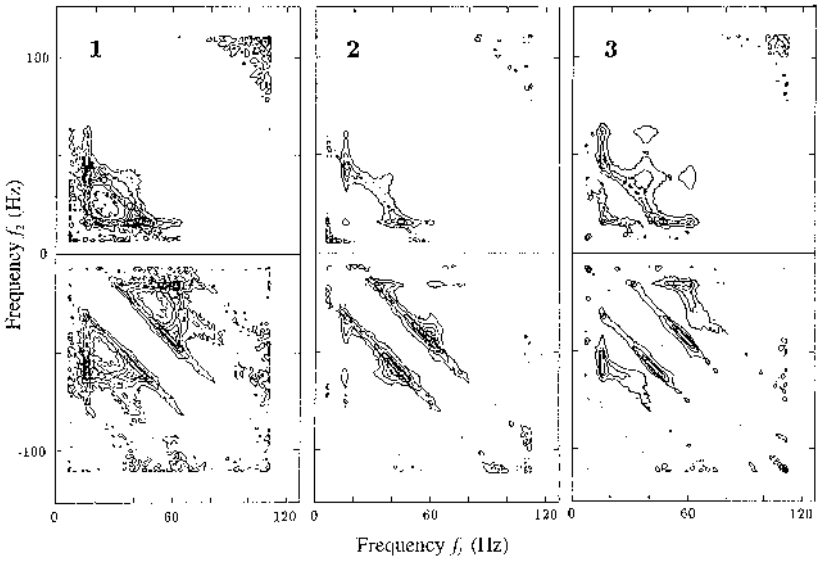


Figure 8.37. Contour plots of the moduli of the three second-order HFRFs of the nonlinear rig operating in its nonlinear mode.

The equivalent forms of (8.206) and (8.207) for a general NDOF system form the basis of Storer's approach to identification discussed in detail in [237, 238, 253][9].

Storer's experimental procedures were based on sinusoidal testing and he made essentially the same assumptions as Gifford regarding the need for low excitation levels. In the case of a SDOF system, equation (8.86) confirms that

$$\Lambda_{\mathrm{s}}(\omega) = H_1(\omega) + \mathrm{O}(X^2) \tag{8.218}$$

where X is the amplitude of excitation.

The sine-testing techniques were exploited in [50, 51] in order to identify automotive shock absorbers as described in an earlier section here.

One way of overcoming the distortion problems on the measured FRFs is given by the interpolation method [90, 45]. The approach is as follows: if sine excitation is used, the response components at multiples of the forcing frequency are:

$$Y(\Omega) = XH_1(\Omega) + \tfrac{3}{4}X^3H_3(\Omega,\Omega,-\Omega) + \tfrac{5}{8}X^5H_5(\Omega,\Omega,\Omega,-\Omega,-\Omega) + \cdots \tag{8.219}$$

[9] In Gifford's later work, he observed that it was only necessary to measure the FRF at a limited number of points and that this number was only a linear function of the spectral densities and not the quadratic function expected if the whole second-order FRF were to be measured. This allowed him to reduce the size of the data sets involved to the same order as those considered by Storer. See [113].

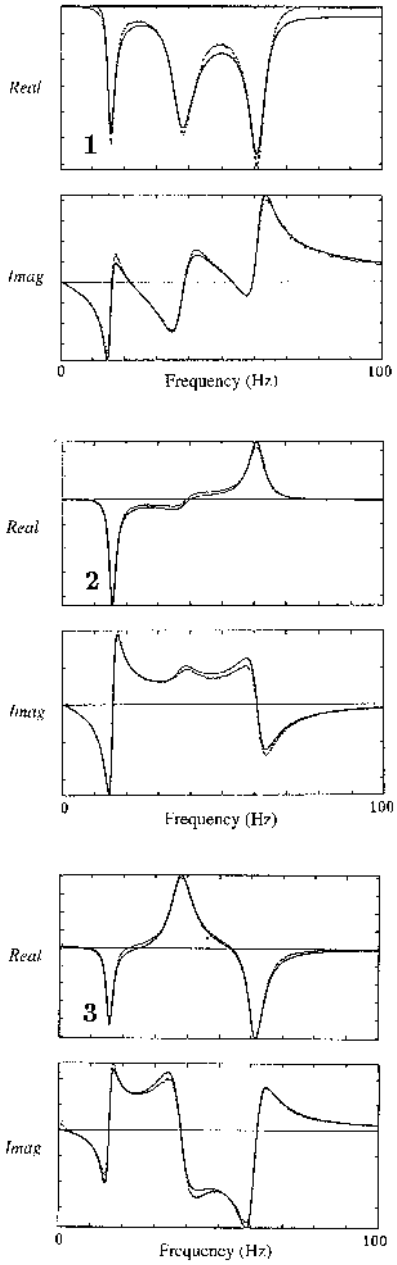


Figure 8.38. The three first-order FRFs of the system formed from time-domain correlation measurements (dotted).

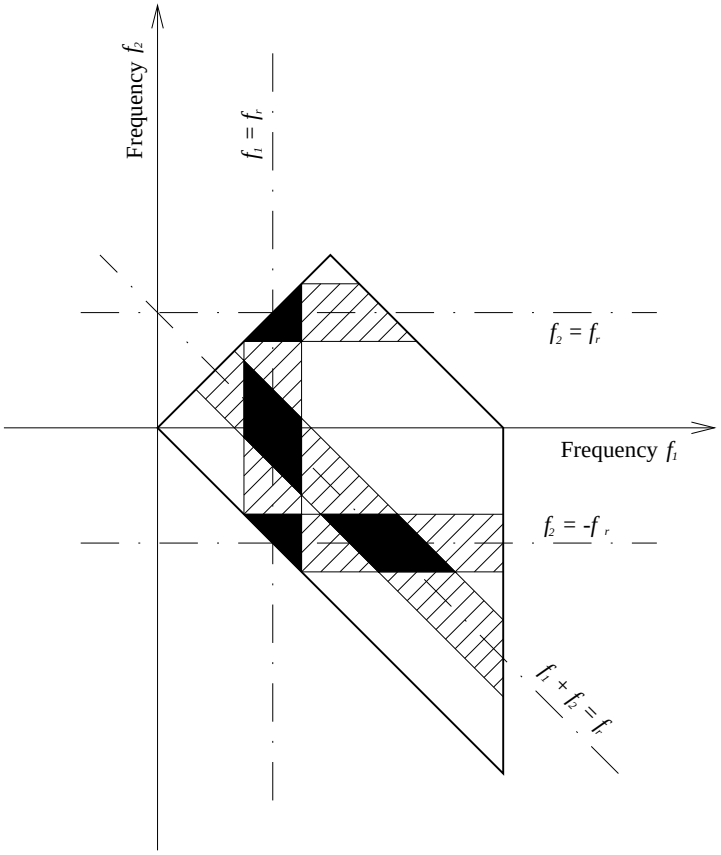


Figure 8.39. Curve-fitting to $\Lambda_2(\omega_1, \omega_2)$ takes place in the black area where the bands containing the poles overlap.

$$Y(2\Omega) = \tfrac{1}{2}X^2 H_2(\Omega,\Omega) + \tfrac{1}{2}X^4 H_4(\Omega,\Omega,\Omega,-\Omega) + \tfrac{15}{32}X^6 H_6(\Omega,\Omega,\Omega,\Omega,-\Omega,-\Omega) + \cdots \tag{8.220}$$

$$Y(3\Omega) = \tfrac{1}{4}X^3 H_3(\Omega,\Omega,\Omega) + \tfrac{5}{16}X^5 H_5(\Omega,\Omega,\Omega,\Omega,-\Omega) + \cdots. \tag{8.221}$$

If the system is harmonically excited at a series of amplitude levels $X_1, X_2, \ldots, X_{\mathrm{N}}$, an overdetermined system can be constructed (in this case up to sixth order) as follows:

$$\begin{pmatrix} Y_1(\Omega) \\ Y_2(\Omega) \\ \vdots \\ Y_{\mathrm{N}}(\Omega) \end{pmatrix} = \begin{pmatrix} X_1 & \frac{3}{4}X_1^2 & \frac{5}{8}X_1^3 \\ X_2 & \frac{3}{4}X_2^2 & \frac{5}{8}X_2^3 \\ \vdots & \vdots & \vdots \\ X_{\mathrm{N}} & \frac{3}{4}X_{\mathrm{N}}^2 & \frac{5}{8}X_{\mathrm{N}}^3 \end{pmatrix} \begin{pmatrix} H_1(\Omega) \\ H_3(\Omega,\Omega,-\Omega) \\ H_5(\Omega,\Omega,\Omega,-\Omega,-\Omega) \end{pmatrix} \tag{8.222}$$

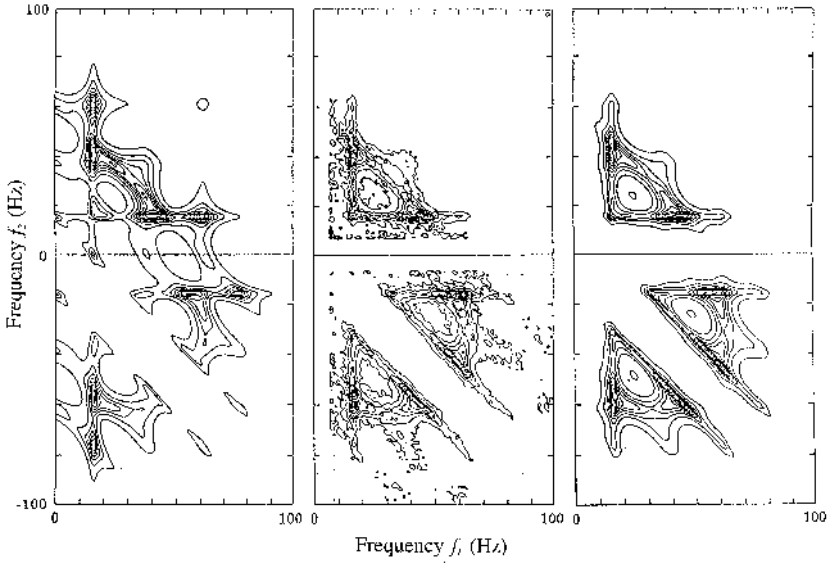


Figure 8.40. Comparison of contour plots of the modulus of the second-order HFRF for point 1 and both stiffness (left) and damping (right) curve-fits.

Table 8.2. Parameter estimates for quadratic damping model.

Coefficient	Estimate
c_{211}	-246.24 ± 37.53
c_{222}	84.24 ± 44.28
c_{233}	-4099.68 ± 84.78
c_{212}	62.1 ± 44.28
c_{213}	-188.19 ± 84.78
c_{223}	67.23 ± 14.31

$$\begin{pmatrix} Y_1(2\Omega) \\ Y_2(2\Omega) \\ \vdots \\ Y_{\mathrm{N}}(2\Omega) \end{pmatrix} = \begin{pmatrix} \frac{1}{2}X_1^2 & \frac{1}{2}X_1^4 & \frac{15}{32}X_1^6 \\ \frac{1}{2}X_2^2 & \frac{1}{2}X_2^4 & \frac{15}{32}X_2^6 \\ \vdots & \vdots & \vdots \\ \frac{1}{2}X_{\mathrm{N}}^2 & \frac{1}{2}X_{\mathrm{N}}^4 & \frac{15}{32}X_{\mathrm{N}}^6 \end{pmatrix} \begin{pmatrix} H_2(\Omega,\Omega) \\ H_4(\Omega,\Omega,\Omega,-\Omega) \\ H_6(\Omega,\Omega,\Omega,\Omega,-\Omega,-\Omega) \end{pmatrix} \tag{8.223}$$

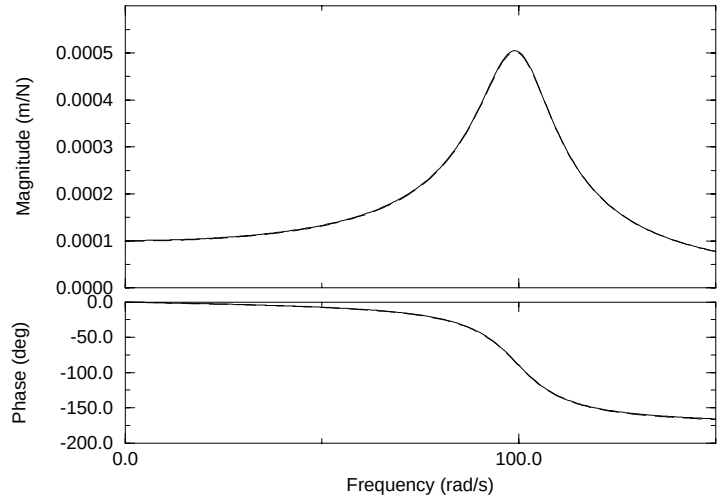


Figure 8.41. Comparison of the theoretical (dotted) against calculated (full) $H_1(\omega)$ from the Duffing oscillator system sinusoidally excited with amplitudes of 0.1, 0.15 and 0.2 N.

$$\begin{pmatrix} Y_1(3\Omega) \\ Y_2(3\Omega) \\ \vdots \\ Y_{\mathrm{N}}(3\Omega) \end{pmatrix} = \begin{pmatrix} \frac{1}{4}X_1^3 & \frac{5}{16}X_1^5 \\ \frac{1}{4}X_2^3 & \frac{5}{16}X_2^5 \\ \vdots & \vdots \\ \frac{1}{4}X_{\mathrm{N}}^3 & \frac{5}{16}X_{\mathrm{N}}^5 \end{pmatrix} \begin{pmatrix} H_3(\Omega,\Omega,\Omega) \\ H_5(\Omega,\Omega,\Omega,\Omega,-\Omega) \end{pmatrix}. \quad (8.224)$$

These systems of equations can be solved by standard LS methods.

The illustration for this method will be taken from [56]. A Duffing oscillator (8.49) was chosen with coefficients $m = 1$, $c = 20$, $k = 10^4$, $k_2 = 0$ and $k_3 = 5 \times 10^9$. The response to a stepped-sine excitation was computed using fourth-order Runge–Kutta. The magnitudes and phases of the harmonic components were extracted using an FFT after the transients had died down at each frequency. Data for equations (8.222) to (8.224) were assembled for up to H_4 at amplitudes 0.1, 0.15 and 0.2 N and the first three diagonal FRFs were calculated. The results are shown in figures 8.41–8.43 compared with the analytical results; the agreement is impressive.

8.11 Higher-order FRFs from neural network models

The NARX class of neural networks was introduced in chapter 6 as a useful non-parametric model structure for system identification. It is not immediately obvious how the models relate to the Volterra series and to higher-order FRFs; however, Wray and Green [280] have shown that there is a rather close connection. That study demonstrated that the Volterra kernels of a given time-

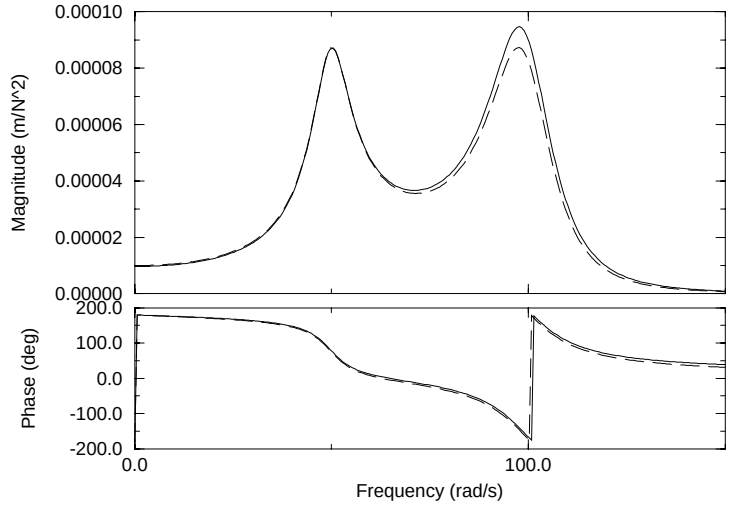


Figure 8.42. Comparison of the theoretical (dotted) against calculated (full) principal diagonal of $H_2(\omega_1, \omega_2)$ from the Duffing oscillator system sinusoidally excited with amplitudes of 0.1, 0.15 and 0.2 N.

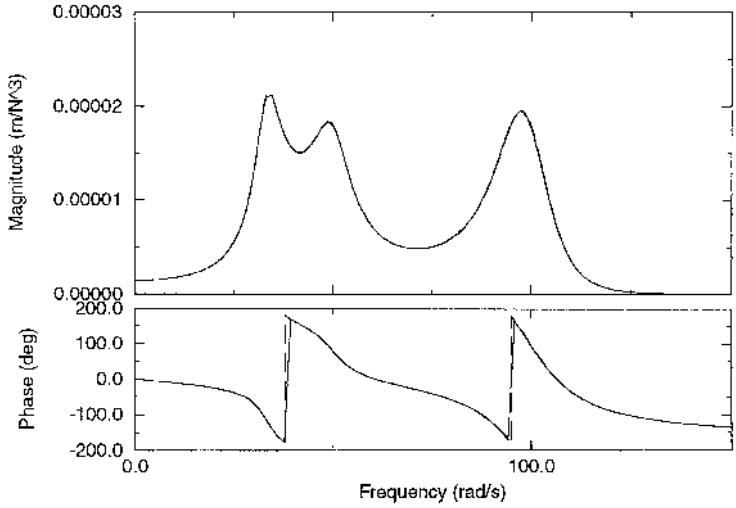


Figure 8.43. Comparison of the theoretical (dotted) against calculated (full) principal diagonal of $H_2(\omega_1, \omega_2, \omega_3)$ from the Duffing oscillator system sinusoidally excited with amplitudes of 0.1, 0.15 and 0.2 N.

delay neural network (TDNN) are rather simple functions of the network weights. The work is discussed here together with the extension to NARX networks and

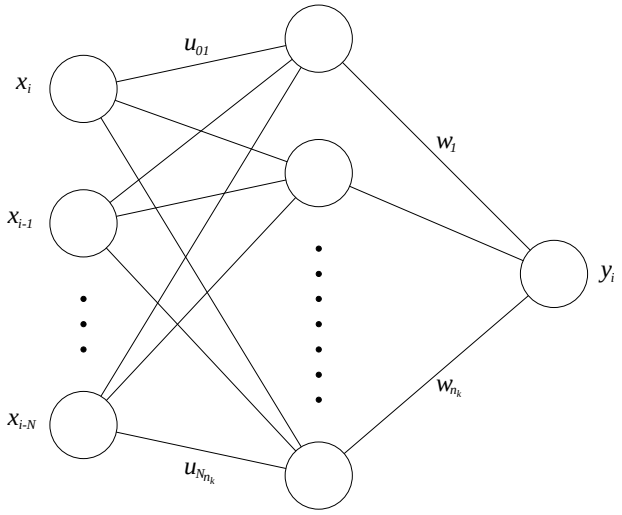


Figure 8.44. Example of a time-delay neural network (TDNN).

loosely follows the more detailed discussion in [56].

As in the case of NARMAX or NARX models generally, the structure and parameters of the neural network NARX model will not necessarily be unique. However, the HFRFs completely characterize the network at each order of nonlinearity and therefore offer a means of validating neural network models used in identification and control.

8.11.1 The Wray–Green method

Recently a method of directly calculating a system's Volterra kernels was presented by Wray and Green [280]. As physiologists, Wray and Green were primarily interested in the time-domain Volterra kernels or higher-order impulse responses. They established that if a TDNN could be trained to model a given system, the Volterra kernels of the network could be calculated directly from its weights. Consequently, if the network was an accurate model of the system, its Volterra kernels would approximate closely to those of the system for an appropriate set of input functions.

The basic TDNN is very similar to the NARX network, the main difference being that only lagged inputs are used to form the model (figure 8.44). The mathematical form is

$$y_i = s + \sum_{j=1}^{n_{\mathrm{h}}} w_j \tanh\left(\sum_{m=0}^{n_x-1} u_{jm} x_{i-m} + b_j\right) \tag{8.225}$$

where w_j is the weight from the jth hidden unit to the output unit and n_{h} is the

number of hidden layer units.

The method is based around the fact that the equation of a TDNN can be shown equivalent to the discrete form of the Volterra series which is given by

$$\begin{aligned} y(t) = h_0 &+ \sum_{-\infty}^{\infty} \Delta\tau_1 h_1(\tau_1) x(t-\tau_1) \\ &+ \sum_{-\infty}^{\infty}\sum_{-\infty}^{\infty} \Delta\tau_1 \Delta\tau_2 h_2(\tau_1,\tau_2) x(t-\tau_1) x(t-\tau_2) + \cdots \\ &+ \sum_{-\infty}^{\infty} \cdots \sum_{-\infty}^{\infty} \Delta\tau_1 \ldots \Delta\tau_n \, h_n(\tau_1,\ldots,\tau_n) x(t-\tau_1)\ldots x(t-\tau_n) + \cdots \end{aligned} \tag{8.226}$$

where h_n is the usual nth-order Volterra kernel and the $\Delta\tau_i$ are the sampling intervals which can all be taken equal if desired. The requirement of causality allows the lower index to be replaced by zero and the effect of damping is to impose a finite memory T on the system, so the discrete series becomes

$$\begin{aligned} y(t) = h_0 &+ \sum_{\tau_1=0}^{\mathrm{T}} \Delta\tau h_1(\tau_1) x(t-\tau_1) \\ &+ \sum_{\tau_1=0}^{\mathrm{T}}\sum_{\tau_2=0}^{\mathrm{T}} \Delta\tau^2 h_2(\tau_1,\tau_2) x(t-\tau_1) x(t-\tau_2) + \cdots \\ &+ \sum_{\tau_1=0}^{\mathrm{T}} \cdots \sum_{\tau_n=0}^{\mathrm{T}} \Delta\tau^n h_n(\tau_1,\ldots,\tau_n) x(t-\tau_1)\ldots x(t-\tau_n) + \cdots . \end{aligned} \tag{8.227}$$

The first step of the Wray–Green method is to expand the activation function of the neural network—the hyperbolic tangent as a Taylor series, then

$$\tanh(z) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1} B_n (2^{4n} - 2^{2n})}{(2n)!} z^{2n-1} \tag{8.228}$$

where the B_n are the Bernoulli numbers defined by

$$B_n = \frac{2(2n)!}{(2\pi)^{2n}} \sum_{h=1}^{\infty} \frac{1}{h^{2n}}. \tag{8.229}$$

The first three non-zero coefficients for this series are $\alpha_1 = 1$, $\alpha_3 = -1/3$ and $\alpha_5 = 2/15$. The function is odd so all the even coefficients vanish. In practice, in order to deal with the bias b_j for each of the hidden nodes, Wray and Green expand the activation function not around zero, but around b_j of each

individual hidden node j. This effectively yields a different function $f_j(d_j)$ at each node where

$$d_j = \sum_{m=0}^{n_x-1} u_{jm} x_{i-m}. \tag{8.230}$$

The coefficients in each expansion are labelled a_{pj} following the definition

$$f_j(d_j) = \sum_{p=0}^{\infty} a_{pj} d_j^{\mathrm{p}} \tag{8.231}$$

and for the hyperbolic tangent function

$$a_{pj} = \frac{1}{p} \tanh^{(p)}(b_j) \tag{8.232}$$

where $\tanh^{(p)}(b_j)$ is the pth derivative of $\tanh(b_j)$.

Adopting the definitions in (8.225), (8.231) and (8.232), the equation of the TDNN network may be written:

$$\begin{aligned} y_i = s &+ w_1(a_{01} + a_{11}d_1 + a_{21}d_1^2 + \cdots) \\ &+ w_2(a_{02} + a_{12}d_2 + a_{22}d_2^2 + \cdots) + \cdots \\ &+ w_{n_\mathrm{h}}(a_{0n_\mathrm{h}} + a_{1n_\mathrm{h}}d_{n_\mathrm{h}} + a_{2n_\mathrm{h}}d_{n_\mathrm{h}}^2 + \cdots) + \cdots. \end{aligned} \tag{8.233}$$

Now, collecting the coefficients of each power of x_i gives, after a little algebra

$$\begin{aligned} y_i = {} & w_1 a_{01} + w_2 a_{02} + \cdots + w_{n_\mathrm{h}} a_{0n_\mathrm{h}} \\ & + \sum_{m=0}^{N} (w_1 a_{11} u_{m1} + w_2 a_{12} u_{m2} + \cdots + w_{n_\mathrm{h}} a_{1n_\mathrm{h}} u_{mn_\mathrm{h}}) x_{i-m} \\ & + \sum_{m=0}^{N}\sum_{k=0}^{N} (w_1 a_{21} u_{m1} u_{k1} + w_2 a_{22} u_{m2} u_{k2} + \cdots + w_{n_\mathrm{h}} a_{2n_\mathrm{h}} u_{mn_\mathrm{h}} u_{kn_\mathrm{h}}) \\ & \times x_{i-m} x_{i-k} + \cdots \end{aligned} \tag{8.234}$$

and equating the coefficients of (8.234) and (8.226) yields a series of expressions for the Volterra kernels

$$h_0 = \sum_{j=1}^{n_\mathrm{h}} w_j a_{0j} \tag{8.235}$$

$$h_1(m) = \sum_{j=1}^{n_\mathrm{h}} \Delta\tau^{-1} w_j a_{1j} u_{mj} \tag{8.236}$$

$$h_2(m,l) = \sum_{j=1}^{n_\mathrm{h}} \Delta\tau^{-2} w_j a_{2j} u_{mj} u_{lj} \tag{8.237}$$

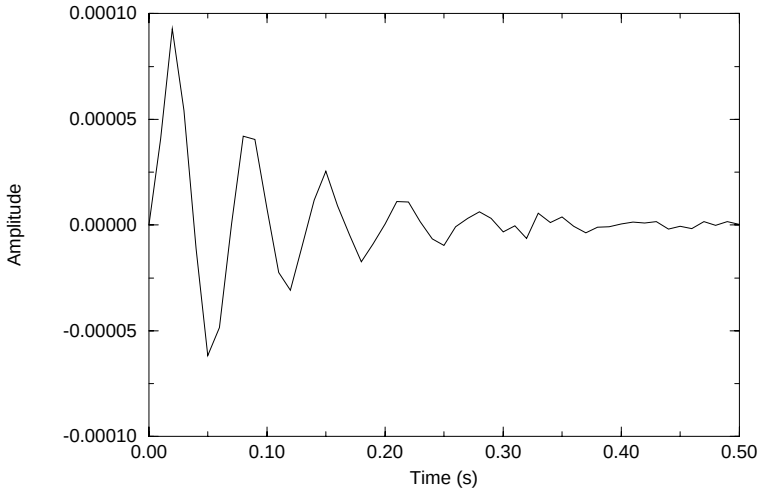


Figure 8.45. First-order time-domain Volterra kernel from the Duffing oscillator obtained from a neural network model by the Wray–Green method.

and, in general,

$$h_n(m_1 \dots m_n) = \sum_{j=1}^{n_\mathrm{h}} \Delta\tau^{-n} w_j a_{nj} u_{m_1 j} u_{m_2 j} \dots u_{m_n j}. \tag{8.238}$$

In order to illustrate this approach, time data for an asymmetric Duffing oscillator (8.49) were simulated with the usual coefficients $m = 1$, $c = 20$, $k = 10^4$, $k_2 = 10^7$ and $k_3 = 5 \times 10^9$. A TDNN with 20 hidden units was fitted using an MLP package and the resulting predictions of the first- and second-order impulse responses are shown in figures 8.45 and 8.46. Wray and Green consider that this method produces considerably better kernel estimates than the Toepliz matrix inversion method which they cite as the previous best method of kernel estimation [150] and their results are confirmed by an independent study by Marmarelis and Zhao [172].

8.11.2 Harmonic probing of NARX models: the multi-layer perceptron

The object of this section is to extend the method of Wray and Green to NARX neural networks with the form

$$y_i = s + \sum_{j=1}^{n_\mathrm{h}} w_j \tanh\left(\sum_{k=1}^{n_y} v_{jk} y_{i-k} + \sum_{m=0}^{n_x - 1} u_{jm} x_{i-m} + b_j \right). \tag{8.239}$$

(In practice, the situation is slightly more complicated than equation (8.239)—and for that matter (8.253)—implies. During training of the network,

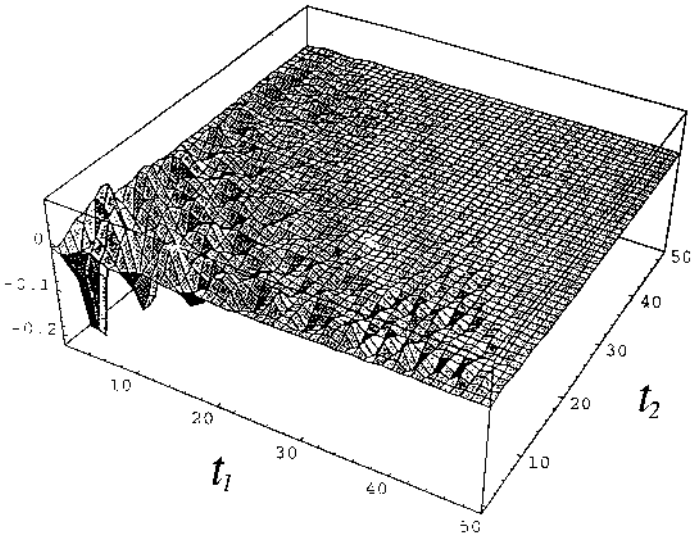


Figure 8.46. Second-order time-domain Volterra kernel from the Duffing oscillator obtained from a neural network model by the Wray–Green method.

all *network* inputs are normalized to lie in the interval $[-1, 1]$ and the output is normalized onto the interval $[-0.8, 0.8]$. Equation (8.239) actually holds between the normalized quantities. The transformation back to the physical quantities, once the HFRFs have been calculated, is derived a little later.)

The NARX models considered here arise from models of dynamical systems and a further simplification can be made. It is assumed that the effects of all the bias terms cancel overall, as the systems being modelled will not contain constant terms in their equations of motion. In dynamical systems this can always be accomplished with an appropriate choice of origin for y (i.e. the equilibrium position of the motion) if the excitation x is also adjusted to remove its d.c. term.

Because of the presence of the lagged outputs in (8.239), the correspondence between the network and Volterra series does not hold as it does for the TDNN, and an alternative approach is needed. It will be shown here that the Volterra kernels are no longer accessible; however, the HFRFs are.

The method of deriving the HFRFs is to apply the standard method of harmonic probing. Because of the complex structure of the model however, the algebra is not trivial. In order to identify $H_1(\omega)$, for example, the system is probed with a single harmonic

$$x_i^{\mathrm{p}} = \mathrm{e}^{\mathrm{i}\Omega t} \tag{8.240}$$

and the expected response is as usual

$$y_i^{\mathrm{p}} = H_1(\Omega)\mathrm{e}^{\mathrm{i}\Omega t} + H_2(\Omega,\Omega)\mathrm{e}^{2\mathrm{i}\Omega t} + H_3(\Omega,\Omega,\Omega)\mathrm{e}^{3\mathrm{i}\Omega t} + \cdots. \tag{8.241}$$

Before proceeding, note that (8.239) is in an inappropriate form for this operation as it stands. The reason is that the term of order n in the expansion of the $\tanh$ function will contain harmonics of all orders up to n, so extracting the coefficient of the fundamental requires the summation of an infinite series. The way around this problem is to use the same trick as Wray and Green and expand the $\tanh$ around the bias; this yields

$$y_i = s + \sum_{j=1}^{n_{\mathrm{h}}} w_j \sum_{\mathrm{t}=0}^{\infty} \left\{ \frac{\tanh^{(t)}(b_j)}{t!} \left(\sum_{k=1}^{n_y} v_{jk} y_{i-k} + \sum_{m=0}^{n_x-1} u_{jm} x_{i-m} \right)^{\mathrm{t}} \right\} \quad (8.242)$$

so each term in the expansion is now a homogeneous polynomial in the lagged x's and y's. $\tanh^{(t)}$ is the tth derivative of tanh at b_j.

The only term in this expansion which can affect the coefficient of the fundamental harmonic is the linear one; therefore take

$$y_i = \sum_{j=1}^{n_{\mathrm{h}}} w_j \frac{\tanh^{(1)}(b_j)}{1} \left(\sum_{k=1}^{n_y} v_{jk} y_{i-k} + \sum_{m=0}^{n_x-1} u_{jm} x_{i-m} \right). \quad (8.243)$$

Following the discrete-time harmonic probing algorithm, substituting x^{p} and y^{p} and extracting the coefficient of $\mathrm{e}^{\mathrm{i}\Omega t}$, yields

$$H_1(\Omega) = \sum_{j=1}^{n_{\mathrm{h}}} w_j \tanh^{(1)}(b_j) \sum_{k=1}^{n_y} \Delta^k H_1(\Omega) + \sum_{j=1}^{n_{\mathrm{h}}} w_j \tanh^{(1)}(b_j) \sum_{m=0}^{n_x-1} u_{jm} \Delta^m \quad (8.244)$$

which can be rearranged to give

$$H_1(\Omega) = \frac{\sum_{j=1}^{n_{\mathrm{h}}} w_j \tanh^{(1)}(b_j) \sum_{m=0}^{n_x-1} u_{jm} \mathrm{e}^{-\mathrm{i}\Omega m \delta t}}{1 - \sum_{j=1}^{n_{\mathrm{h}}} w_j \tanh^{(1)}(b_j) \sum_{k=1}^{n_y} v_{jk} \mathrm{e}^{-\mathrm{i}\Omega k \delta t}}. \quad (8.245)$$

Extraction of H_2 requires a probe with two independent harmonics:

$$x_i^{\mathrm{p}} = \mathrm{e}^{\mathrm{i}\Omega_1 t} + \mathrm{e}^{\mathrm{i}\Omega_1 t} \quad (8.246)$$

and

$$y_i^{\mathrm{p}} = H_1(\Omega_1)\mathrm{e}^{\mathrm{i}\Omega_1 t} + H_1(\Omega_2)\mathrm{e}^{\mathrm{i}\Omega_2 t} + 2H_2(\Omega_1, \Omega_2)\mathrm{e}^{\mathrm{i}(\Omega_1+\Omega_2)t} + \cdots. \quad (8.247)$$

The argument proceeds as for H_1; if these expressions are substituted into the network function (8.242), the only HFRFs to appear in the coefficient of the sum harmonic $\mathrm{e}^{\mathrm{i}(\Omega_1+\Omega_2)t}$ are H_1 and H_2, where H_1 is already known from equation (8.245). So as before, the coefficient can be rearranged to give an expression for H_2 in terms of the network weights and H_1. The only terms in (8.242) which are relevant for the calculation are those at first and second order. The calculation is straightforward but lengthy:

$$H_2(\Omega_1, \Omega_2) = \frac{1}{2!D} \sum_{j=1}^{n_{\mathrm{h}}} w_j \frac{\tanh^{(2)}(b_j)}{2!} \{A_j + B_j + C_j\} \quad (8.248)$$

where

$$A_j = \sum_{k=1}^{n_y}\sum_{l=1}^{n_y} v_{jk}\, v_{jl} H_1(\Omega_1) H_1(\Omega_2)(\mathrm{e}^{-\mathrm{i}\Omega_1 k\delta t}\mathrm{e}^{-\mathrm{i}\Omega_2 l\delta t} + \mathrm{e}^{-\mathrm{i}\Omega_2 k\delta t}\mathrm{e}^{-\mathrm{i}\Omega_1 l\delta t}) \tag{8.249}$$

$$B_j = \sum_{k=0}^{n_x-1}\sum_{l=0}^{n_x-1} u_{jk}\, u_{jl}(\mathrm{e}^{-\mathrm{i}\Omega_1 k\delta t}\mathrm{e}^{-\mathrm{i}\Omega_2 l\delta t} + \mathrm{e}^{-\mathrm{i}\Omega_2 k\delta t}\mathrm{e}^{-\mathrm{i}\Omega_1 l\delta t}) \tag{8.250}$$

$$C_j = 2\sum_{k=1}^{n_y}\sum_{l=0}^{n_x-1} v_{jk}\, u_{jl}(H_1(\Omega_1)\mathrm{e}^{-\mathrm{i}\Omega_1 k\delta t}\mathrm{e}^{-\mathrm{i}\Omega_2 l\delta t} + H_1(\Omega_2)\mathrm{e}^{-\mathrm{i}\Omega_2 k\delta t}\mathrm{e}^{-\mathrm{i}\Omega_1 l\delta t}) \tag{8.251}$$

and

$$D = 1 - \sum_{j=1}^{n_\mathrm{h}} w_j \tanh^{(1)}(b_j)\sum_{k=1}^{n_y} v_{jk}\mathrm{e}^{-\mathrm{i}(\Omega_1+\Omega_2)k\delta t}. \tag{8.252}$$

Derivation of H_3 is considerably more lengthy and requires probing with three harmonics. The expression can be found in [57].

8.11.3 Radial basis function networks

Much of the recent work on system identification has abandoned the MLP structure in favour of the radial basis function networks introduced by Broomhead and Lowe [47]. The essential differences between the two approaches are in the computation of the hidden node activation and in the form of the nonlinear activation function. At each hidden node in the MLP network, the activation z is obtained as a weighted sum of incoming signals from the input layer:

$$z_i = \sum_j w_{ij} x_j. \tag{8.253}$$

This is then passed through a nonlinear activation function which is sigmoidal in shape, the important features of the function are its continuity, its monotonicity and its asymptotic approach to constant values. The resulting hidden node response is *global* in the sense that it can take non-zero values at all points in the space spanned by the network input vectors.

In contrast, the RBF network has *local* hidden nodes. The activation is obtained by taking the Euclidean distance squared from the input vector to a point defined independently for each hidden node—its centre c_i (which is of course a vector of the same dimension as the input layer):

$$z_i = \|x_i - c_i\|. \tag{8.254}$$

This is then passed through a *basis function* which decays rapidly with its argument, i.e. it is significantly non-zero only for inputs close to c_i. The overall

output of the RBF network is therefore the summed response from several locally-tuned units. It is this ability to cover selectively connected regions of the input space which makes the RBF so effective for pattern recognition and classification problems. The RBF structure also allows an effective means of implementing the NARX model for control and identification [61, 62].

For the calculation given here, a Gaussian basis function is assumed as this is by far the most commonly used to date. Also, following Poggio and Girosi [206], the network is modified by the inclusion of direct linear connections from the input layer to the output. The resulting NARX model is summarized by

$$y_i = s + \sum_{j=1}^{n_{\mathrm{h}}} w_j \exp\left\{ -\frac{1}{2\sigma_j^2}\left[\sum_{k=1}^{n_y}(y_{i-k} - v_{jk})^2 + \sum_{m=0}^{n_x-1}(x_{i-m} - u_{jm})^2\right]\right\} + \underbrace{\sum_{j=1}^{n_y} a_j y_{i-j} + \sum_{j=0}^{n_x-1} b_j x_{i-j}}_{\text{from linear connections}} \tag{8.255}$$

where the quantities v_{jk} and u_{jm} are the hidden node centres and the σ_i is the standard deviation or *radius* of the Gaussian at hidden node i. The first part of this expression is the standard RBF network.

As with the MLP network the appearance of constant terms in the exponent will lead to difficulties when this is expanded as a Taylor series. A trivial rearrangement yields the more useful form

$$\begin{aligned} y_i = s &+ \sum_{j=1}^{n_{\mathrm{h}}} w_j \gamma_j \exp\left\{ -\frac{1}{2\sigma_j^2}\left[\sum_{k=1}^{n_y}(y_{i-k}^2 - 2v_{jk}y_{i-k})\right.\right. \\ &\left.\left. + \sum_{m=0}^{n_x-1}(x_{i-m}^2 - 2u_{jm}x_{i-m})\right]\right\} \\ &+ \sum_{j=1}^{n_y} a_j y_{i-j} + \sum_{j=0}^{n_x-1} b_j x_{i-j} \end{aligned} \tag{8.256}$$

where

$$\gamma_j = \exp\left\{ -\frac{1}{2\sigma_j^2}\left[\sum_{k=1}^{n_y} v_{jk}^2 + \sum_{m=0}^{n_x-1} u_{jm}^2\right]\right\}. \tag{8.257}$$

Now, expanding the exponential and retaining only the linear terms leads to the required expression for obtaining H_1:

$$y_i = \sum_{j=1}^{n_{\mathrm{h}}} \frac{w_j\gamma_j}{\sigma_j}\left\{\sum_{k=1}^{n_y} v_{jk}y_{i-k} + \sum_{m=0}^{n_x-1} u_{jm}x_{i-m}\right\}. \tag{8.258}$$

Substituting the first-order probing expressions (8.240) and (8.241) yields

$$H_1(\Omega) = \frac{\sum_{j=1}^{n_h} \gamma_j w_j \frac{1}{\sigma_j^2} \sum_{m=0}^{n_x-1} u_{jm} \mathrm{e}^{-\mathrm{i}\Omega m \delta t} + \sum_{j=0}^{n_x-1} b_j \mathrm{e}^{-\mathrm{i}\Omega j \delta t}}{1 - \sum_{j=1}^{n_y} a_j \mathrm{e}^{-\mathrm{i}\Omega j \delta t} - \sum_{j=1}^{n_h} \gamma_j w_j \frac{1}{\sigma_j^2} \sum_{k=1}^{n_y} v_{jk} \mathrm{e}^{-\mathrm{i}\Omega k \delta t}}. \quad (8.259)$$

The second-order FRF H_2 is obtained as before:

$$\begin{aligned} H_2(\Omega_1, \Omega_2) = {} & \frac{1}{2!D} \sum_{j=1}^{n_h} w_j \gamma_j \\ & \times \Bigg\{ - \frac{1}{2\sigma_j^2} \sum_{k=0}^{n_x-1} \sum_{l=0}^{n_x-1} (\mathrm{e}^{-\mathrm{i}k\Omega_1\delta t} \mathrm{e}^{-\mathrm{i}l\Omega_2\delta t} + \mathrm{e}^{-\mathrm{i}k\Omega_2\delta t} \mathrm{e}^{-\mathrm{i}l\Omega_1\delta t}) \\ & + \frac{1}{\sigma_j^4} \sum_{k=0}^{n_x-1} \sum_{l=0}^{n_x-1} u_{jk} u_{jl} (\mathrm{e}^{-\mathrm{i}k\Omega_1\delta t} \mathrm{e}^{-\mathrm{i}l\Omega_2\delta t} + \mathrm{e}^{-\mathrm{i}k\Omega_2\delta t} \mathrm{e}^{-\mathrm{i}l\Omega_1\delta t}) \\ & - \frac{1}{2\sigma_j^2} \sum_{k=1}^{n_y} \sum_{l=1}^{n_y} H_1(\Omega_1) H_1(\Omega_2) \\ & \times (\mathrm{e}^{-\mathrm{i}k\Omega_1\delta t} \mathrm{e}^{-\mathrm{i}l\Omega_2\delta t} + \mathrm{e}^{-\mathrm{i}k\Omega_2\delta t} \mathrm{e}^{-\mathrm{i}l\Omega_1\delta t}) \\ & + \frac{1}{\sigma_j^4} \sum_{k=1}^{n_y} \sum_{l=1}^{n_y} v_{jk} v_{jl} H_1(\Omega_1) H_1(\Omega_2) \\ & \times (\mathrm{e}^{-\mathrm{i}k\Omega_1\delta t} \mathrm{e}^{-\mathrm{i}l\Omega_2\delta t} + \mathrm{e}^{-\mathrm{i}k\Omega_2\delta t} \mathrm{e}^{-\mathrm{i}l\Omega_1\delta t}) \\ & + \frac{1}{\sigma_j^4} \sum_{k=1}^{n_y} \sum_{l=0}^{n_x-1} v_{jk} u_{jl} (H_1(\Omega_1) \mathrm{e}^{-\mathrm{i}k\Omega_1\delta t} \mathrm{e}^{-\mathrm{i}l\Omega_2\delta t} \\ & + H_1(\Omega_2) \mathrm{e}^{-\mathrm{i}k\Omega_2\delta t} \mathrm{e}^{-\mathrm{i}l\Omega_1\delta t}) \Bigg\} \end{aligned} \quad (8.260)$$

where

$$D = 1 - \sum_{j=1}^{n_y} a_j \mathrm{e}^{-\mathrm{i}(\Omega_1+\Omega_2)j\delta t} - \sum_{j=1}^{n_h} \gamma_j w_j \frac{1}{\sigma_j^2} \sum_{k=1}^{n_y} \mathrm{e}^{-\mathrm{i}(\Omega_1+\Omega_2)k\delta t}. \quad (8.261)$$

8.11.4 Scaling the HFRFs

The network input vectors must be scaled prior to presentation to avoid saturation of the processing units. In the MLP program used with tanh activation functions, the y and x values are scaled between the interval $[0.8, -0.8]$. If a linear network is used then both are scaled between $[1, -1]$. The resulting HFRFs must therefore be scaled accordingly.

The scaling process gives the following, where the superscript 's' denotes scaled values and α and a are the x and y scaling factors respectively, and β and

b are the origin shifts.

$$x^{\mathrm{s}} = \alpha x + \beta \quad \text{and} \quad y^{\mathrm{s}} = ay + b. \tag{8.262}$$

The neural network program used here, MLP, calculates the scaled x values as

$$\begin{aligned} x^{\mathrm{s}} &= 1.6\left[\frac{x - x_{\min}}{x_{\max} - x_{\min}} - \frac{1}{2}\right] \\ &= \frac{1.6x}{x_{\max} - x_{\min}} - \frac{0.8(x_{\min} + x_{\max})}{x_{\max} - x_{\min}} \end{aligned} \tag{8.263}$$

which gives

$$\alpha = \frac{1.6}{x_{\max} - x_{\min}} \tag{8.264}$$

and similarly

$$a = \frac{1.6}{y_{\max} - y_{\min}}. \tag{8.265}$$

The dimensions of the nth kernel transform follow from

$$H_n(\Omega) = \frac{s\mathcal{F}[y(t)|_{n\Omega}]}{\mathcal{F}[x(t)|_{\Omega}]^n} = \frac{s\mathcal{F}[ay + b]}{\mathcal{F}[\alpha x + \beta]^n} \tag{8.266}$$

which holds from a dimensional point of view with s a constant.

As the constant offsets only affect the d.c. lines, this reduces to

$$= \frac{sa\mathcal{F}[y(t)]}{\alpha^n \mathcal{F}[x(t)]^n} \tag{8.267}$$

so the scaling relation is

$$H_n^{\mathrm{s}}(\Omega_1, \ldots, \Omega_n) = \frac{a}{\alpha^n} H_n(\Omega_1, \ldots, \Omega_n). \tag{8.268}$$

Therefore the true HFRF is given by

$$H_n(\Omega_1, \ldots, \Omega_n) = 1.6^{n-1} \frac{(y_{\max} - y_{\min})}{(x_{\max} - x_{\min})^n} H_n^{\mathrm{s}}(\Omega_1, \ldots, \Omega_n). \tag{8.269}$$

The network is therefore scale specific and may not perform well with data originating from very different excitation levels. The solution to this problem is for the training set to contain excitation levels spanning the range of interest.

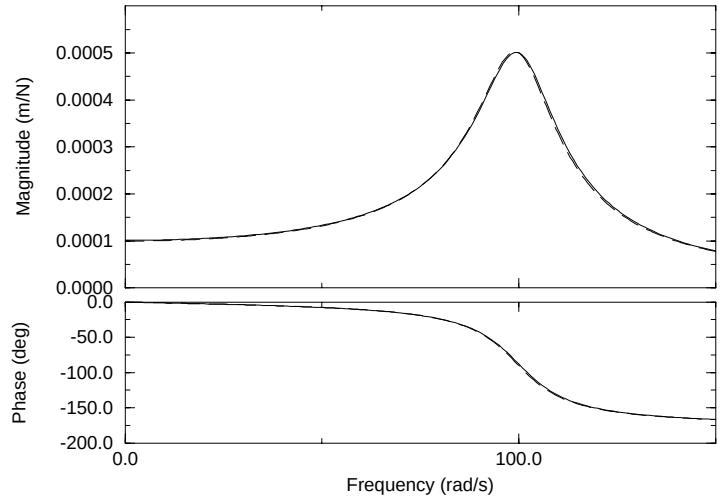


Figure 8.47. Exact (dashed) and estimated (full) H_1 from the Duffing oscillator, estimate from probing of 10:2:1 NARX MLP network.

8.11.5 Illustration of the theory

The asymmetric Duffing oscillator specified in (8.49) was chosen for the simulations with $m = 1$ kg, $c = 20$ N s m^{-1}, $k = 10^4$ N m^{-1}, $k_2 = 10^7$ N m^{-2} and $k_3 = 5 \times 10^9$ N m^{-3}. The differential equation of motion stepped forward using fourth-order Runge–Kutta. The excitation $x(t)$ was a white Gaussian sequence with zero mean and rms of 0.5, band-limited within the range 0–100 Hz. A time-step of 0.001 s was adopted and the data were subsampled by a factor of 5. This gave a final Δt of 0.005 s—corresponding to a sampling frequency of 200 Hz. For network training 1000 points of sampled input and displacement data were taken.

8.11.5.1 Results from MLP network

Using the time data described earlier, various networks were trained and tested. Once the weights were obtained, the previous formulae were used to compute the first- and second-order FRFs of the network.

Figure 8.47 shows the $H_1(\Omega)$ (full in the figure) that best approximated the theoretical result (shown dotted in the figure). The almost perfect overlay was obtained from a network with ten input units and two hidden units (i.e. 10:2:1). The network had converged to a model with an MPO error (see chapter 6) of 0.35. The comparison between measured data and MPO is given in figure 8.48.

The second-order FRF proved a little more difficult to estimate accurately. The best $H_2(\Omega_1, \Omega_2)$ estimation is compared to the theoretical kernel transform

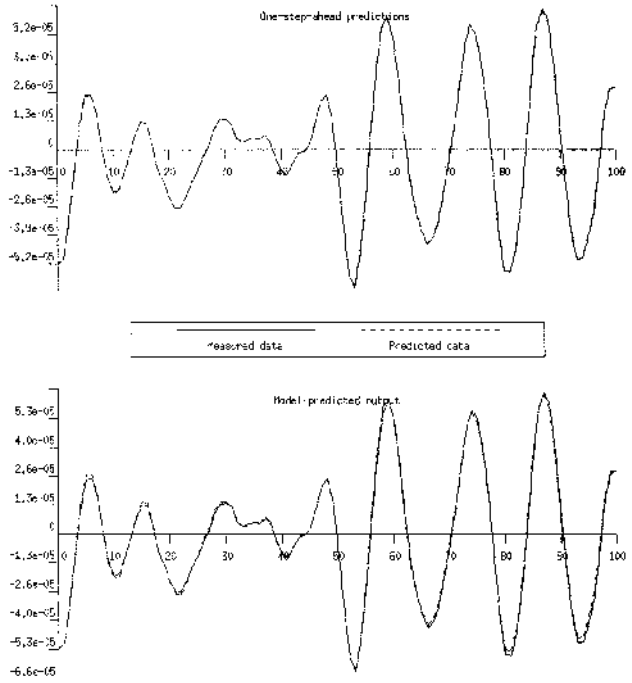


Figure 8.48. Comparison between the measured data and the NARX MLP network estimate for figure 8.47.

in figure 8.49 along its leading diagonal ($\Omega_1 = \Omega_2$). This was calculated from a 10:4:1 network trained to an MPO error of 0.27. The corresponding $H_1(\Omega_1)$ from the same network, shown in figure 8.50, shows a little discrepancy from theory. This is unfortunate from a system identification point of view as one would like a correct representation of the system at all orders of nonlinearity. That this happens appears to be due to the fact that the network can reproduce the signals with some confusion between the different ordered components y_n. This is discussed a little later.

8.11.5.2 Results from radial basis function networks

The RBF networks were trained and tested using the force and displacement data. A spread of results was observed, the best $H_1(\Omega)$ was given by a 6:2:1 network, trained to a 0.48 MPO error as shown in figure 8.51. A 4:4:1 network produced the best $H_2(\Omega_1, \Omega_2)$ after training to a 0.72 MPO error; this is shown in figure 8.52.

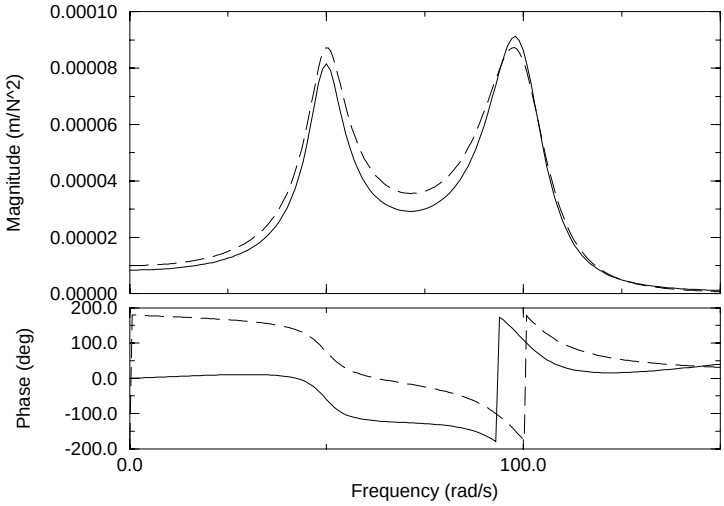


Figure 8.49. Exact (dashed) and estimated (full) principal diagonal for H_2 from the Duffing oscillator; an estimate from probing of a 10:4:1 NARX MLP network.

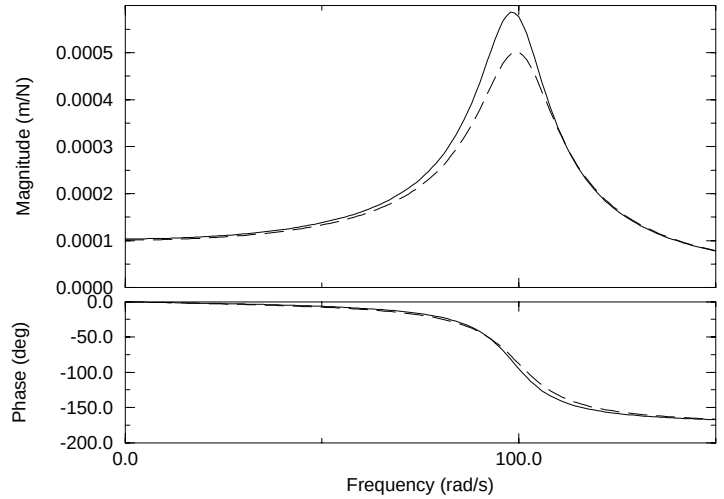


Figure 8.50. Exact (dashed) and estimated (full) H_1 from the Duffing oscillator; the estimate from probing of a 10:4:1 NARX MLP network.

8.11.5.3 Discussion

The results indicate possible *over-parametrization* by the network in modelling the time data rather than accurately modelling the system in question. In system

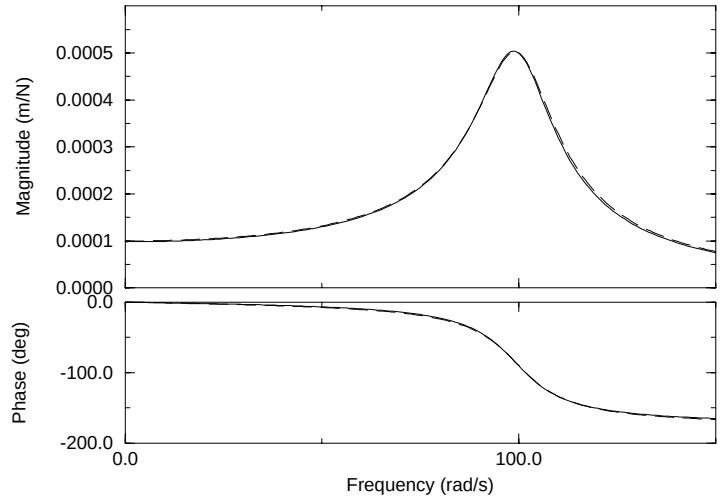


Figure 8.51. Exact (dashed) and estimated (full) H_1 from the Duffing oscillator; the estimate from probing of a 6:2:1 NARX RBF network.

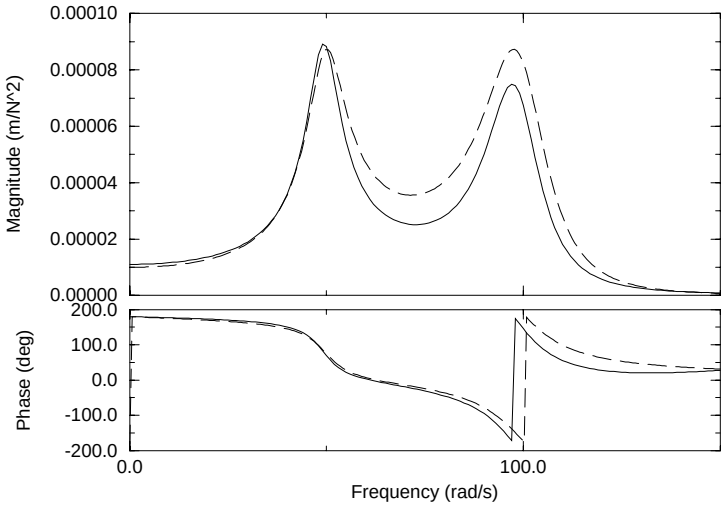


Figure 8.52. Exact (dashed) and estimated (full) principal diagonal for H_2 from the Duffing oscillator; the estimate from probing of a 4:4:1 NARX RBF network.

identification, over-parametrization often causes misleading results [231]. Over-parametrization is caused by the model having many more degrees of freedom than the system it is modelling. As the complexity of the network increases, its data-modelling abilities generally improve, up to a point. Beyond that point

the error on the training data continues to decrease but the error over a testing set begins to increase. Neural network users have long known this and *always* use a training and testing set for validation, and often use a training, testing *and* validation set. This may be necessary in this approach to system identification. Often the MPO error is taken as conclusive.

A principled approach to model testing may be able to inform the modeller that the model is indeed over-parametrized. The problem remaining is to remove the over-parametrization. It may be possible to accomplish this in two ways. First, there is a possibility of *regularization* [40], which seeks to ensure that the network weights do not evolve in an unduly correlated manner. The ideal situation would be if an analogue of the orthogonal LS estimator discussed in appendix E could be derived for neural networks. This appears to be a difficult problem and there is no current solution. The second possibility is *pruning*, where the network nodes and weights are tested for their significance and removed if they are found to contribute nothing to the model. This reduces the numbers of weights to be estimated for a given network structure.

8.12 The multi-input Volterra series

The single-input–single-output (SISO) Volterra series is by now established as a powerful tool in nonlinear system theory as previously discussed. In contrast, the multi-input version of the series appears to have received little attention since its inception (in the guise of the closely related Wiener series) in the work of Marmarelis and Naka [171]. This section aims to show that the theory is no more complicated and this is achieved by use of a case study.

When a nonlinear system is excited with more than one input, intermodulation terms arise in the response. In fact, the Volterra series still applies, however, a generalization is necessary. A short discussion is given here, details can be found in [279]. First of all, an extension of the notation is needed. A superscript is added to each Volterra kernel denoting the response point and the number of occurrences of each particular input relevant to the construction of that kernel is indicated; e.g. $h_5^{(j:aabbb)}(\tau_1,\ldots,\tau_5)$ represents a fifth-order kernel measured at response point j and having two inputs at point a and three at point b. For example, consider a nonlinear system excited at locations a and b with inputs $x^{(a)}(t)$ and $x^{(b)}(t)$. The expression for the response at a point j is the same as equation (8.3) in the single-input case. However, in the single-input case each nonlinear component $y_n^{(j)}(t)$ in (8.3) is expressed in terms of a single Volterra kernel; in the multi-input case, several kernels are needed. For the two-input case the first two components are given by:

$$y_1^{(j)}(t) = \int_{-\infty}^{+\infty} \mathrm{d}\tau\, h_1^{(j:a)}(\tau)x^{(a)}(t-\tau) + \int_{-\infty}^{+\infty} \mathrm{d}\tau\, h_1^{(j:b)}(\tau)x^{(b)}(t-\tau) \quad (8.270)$$

and

$$y_2^{(j)}(t) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \mathrm{d}\tau_1\,\mathrm{d}\tau_2\, h_2^{(j:aa)}(\tau_1,\tau_2)x^{(a)}(t-\tau_1)x^{(a)}(t-\tau_2)$$
$$+ \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \mathrm{d}\tau_1\,\mathrm{d}\tau_2\, \{h_2^{(j:ab)}(\tau_1,\tau_2) + h_2^{(j:ba)}(\tau_2,\tau_1)\}$$
$$\times\, x^{(a)}(t-\tau_1)x^{(b)}(t-\tau_2)$$
$$+ \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \mathrm{d}\tau_1\,\mathrm{d}\tau_2\, h_2^{(j:bb)}(\tau_1,\tau_2)x^{(b)}(t-\tau_1)x^{(b)}(t-\tau_2). \tag{8.271}$$

For convenience, the sum of cross-kernels is absorbed in a redefinition: $h_2^{(j:ab)} + h_2^{(j:ba)} \longrightarrow 2h_2^{(j:ab)}$. The corresponding frequency-domain quantities are defined as before. It can be shown that there is no longer total symmetry under permutations of the n symbols of h_n and H_n; a more restricted symmetry group applies. For example, taking the series to second order as before, $h_2^{(j:ba)}(\tau_1,\tau_2) \neq h_2^{(j:ba)}(\tau_2,\tau_1)$.

8.12.1 HFRFs for a continuous-time MIMO system

For the purposes of illustrating the general multi-input–multi-output (MIMO) theory, it is sufficient to consider a nonlinear 2DOF system with two inputs. The system here has equations of motion

$$m_1\ddot{y}^{(1)} + c_1\dot{y}^{(1)} + k_{11}y^{(1)} + k_{12}y^{(2)} + k'y^{(1)2} + k''y^{(1)3} = x^{(1)}(t) \tag{8.272}$$
$$m_2\ddot{y}^{(2)} + c_2\dot{y}^{(2)} + k_{21}y^{(1)} + k_{22}y^{(2)} = x^{(2)}(t). \tag{8.273}$$

The harmonic-probing algorithm extends straightforwardly to this case. The first-order FRFs are extracted as follows: substitute the probing expressions

$$x_{\mathrm{p}}^{(1)}(t) = \mathrm{e}^{\mathrm{i}\omega t}, \quad x_{\mathrm{p}}^{(2)}(t) = 0, \quad y_{\mathrm{p}}^{(1)}(t) = H_1^{(1:1)}\mathrm{e}^{\mathrm{i}\omega t}, \quad y_{\mathrm{p}}^{(2)}(t) = H_1^{(2:1)}\mathrm{e}^{\mathrm{i}\omega t} \tag{8.274}$$

into equations (8.272) and (8.273) and, in each case, equate the coefficients of $\mathrm{e}^{\mathrm{i}\omega t}$ on the left- and right-hand sides. Next, substitute the probing expressions

$$x_{\mathrm{p}}^{(1)}(t) = 0, \quad x_{\mathrm{p}}^{(2)}(t) = \mathrm{e}^{\mathrm{i}\omega t}$$
$$y_{\mathrm{p}}^{(1)}(t) = H_1^{(1:2)}(\omega)\mathrm{e}^{\mathrm{i}\omega t}, \quad y_{\mathrm{p}}^{(2)}(t) = H_1^{(2:2)}(\omega)\mathrm{e}^{\mathrm{i}\omega t} \tag{8.275}$$

and equate coefficients as before. This procedure results in four simultaneous equations for the unknown quantities $H_1^{(1:1)}$, $H_1^{(1:2)}$, $H_1^{(2:1)}$ and $H_1^{(2:2)}$. The H_1

matrix is obtained from

$$\begin{pmatrix} H_1^{(1:1)}(\omega) & H_1^{(1:2)}(\omega) \\ H_1^{(2:1)}(\omega) & H_1^{(2:2)}(\omega) \end{pmatrix} = \begin{pmatrix} -\omega^2 m_1 + \mathrm{i}\omega c_1 + k_{11} + k_{12} & -k_{12} \\ -k_{21} & -\omega^2 m_2 + \mathrm{i}\omega c_2 + k_{21} + k_{22} \end{pmatrix}^{-1}. \tag{8.276}$$

In order to display the FRFs, the linear parameters were chosen as $m_1 = m_2 = 1$, $c_1 = c_2 = 20$, $k_{11} = k_{22} = k_{12} = k_{21} = 10\,000$. For later, the nonlinear parameters were chosen as $k' = 1 \times 10^7$ and $k'' = 5 \times 10^9$. The H_1 magnitudes are displayed in figure 8.53. Only two of the functions are shown as the system is symmetrical and so $H_1^{(1:1)} = H_1^{(2:2)}$ and $H_1^{(1:2)} = H_1^{(2:1)}$ because of reciprocity. The continuous time H_1's are given by the full line.

The extraction of the second-order FRFs is a little more involved. First, the probing expressions

$$\begin{aligned} x_{\mathrm{p}}^{(1)}(t) &= \mathrm{e}^{\mathrm{i}\omega_1 t} + \mathrm{e}^{\mathrm{i}\omega_2 t} \\ x_{\mathrm{p}}^{(2)}(t) &= 0 \\ y_{\mathrm{p}}^{(1)}(t) &= H_1^{(1:1)}(\omega_1)\mathrm{e}^{\mathrm{i}\omega_1 t} + H_1^{(1:1)}(\omega_2)\mathrm{e}^{\mathrm{i}\omega_2 t} + 2H_2^{(1:11)}(\omega_1,\omega_2)\mathrm{e}^{\mathrm{i}(\omega_1+\omega_2)t} \\ y_{\mathrm{p}}^{(2)}(t) &= H_1^{(2:1)}(\omega_1)\mathrm{e}^{\mathrm{i}\omega_1 t} + H_1^{(2:1)}(\omega_2)\mathrm{e}^{\mathrm{i}\omega_2 t} + 2H_2^{(2:11)}(\omega_1,\omega_2)\mathrm{e}^{\mathrm{i}(\omega_1+\omega_2)t} \end{aligned} \tag{8.277}$$

are substituted into the equations (8.272) and (8.273) and the coefficients of the sum harmonic $\mathrm{e}^{\mathrm{i}(\omega_1+\omega_2)t}$ are extracted. Next, the exercise is repeated using the probing expressions

$$\begin{aligned} x_{\mathrm{p}}^{(1)}(t) &= 0 \\ x_{\mathrm{p}}^{(2)}(t) &= \mathrm{e}^{\mathrm{i}\omega_1 t} + \mathrm{e}^{\mathrm{i}\omega_2 t} \\ y_{\mathrm{p}}^{(1)}(t) &= H_1^{(1:2)}(\omega_1)\mathrm{e}^{\mathrm{i}\omega_1 t} + H_1^{(1:2)}(\omega_2)\mathrm{e}^{\mathrm{i}\omega_2 t} + 2H_2^{(1:22)}(\omega_1,\omega_2)\mathrm{e}^{\mathrm{i}(\omega_1+\omega_2)t} \\ y_{\mathrm{p}}^{(2)}(t) &= H_1^{(2:2)}(\omega_1)\mathrm{e}^{\mathrm{i}\omega_1 t} + H_1^{(2:2)}(\omega_2)\mathrm{e}^{\mathrm{i}\omega_2 t} + 2H_2^{(2:22)}(\omega_1,\omega_2)\mathrm{e}^{\mathrm{i}(\omega_1+\omega_2)t} \end{aligned} \tag{8.278}$$

and

$$\begin{aligned} x_{\mathrm{p}}^{(1)}(t) &= \mathrm{e}^{\mathrm{i}\omega_1 t} \\ x_{\mathrm{p}}^{(2)}(t) &= \mathrm{e}^{\mathrm{i}\omega_2 t} \\ y_{\mathrm{p}}^{(1)}(t) &= H_1^{(1:1)}(\omega_1)\mathrm{e}^{\mathrm{i}\omega_1 t} + H_1^{(1:2)}(\omega_2)\mathrm{e}^{\mathrm{i}\omega_2 t} + 2H_2^{(1:12)}(\omega_1,\omega_2)\mathrm{e}^{\mathrm{i}(\omega_1+\omega_2)t} \\ y_{\mathrm{p}}^{(2)}(t) &= H_1^{(2:1)}(\omega_1)\mathrm{e}^{\mathrm{i}\omega_1 t} + H_1^{(2:2)}(\omega_2)\mathrm{e}^{\mathrm{i}\omega_2 t} + 2H_2^{(2:12)}(\omega_1,\omega_2)\mathrm{e}^{\mathrm{i}(\omega_1+\omega_2)t}. \end{aligned} \tag{8.279}$$

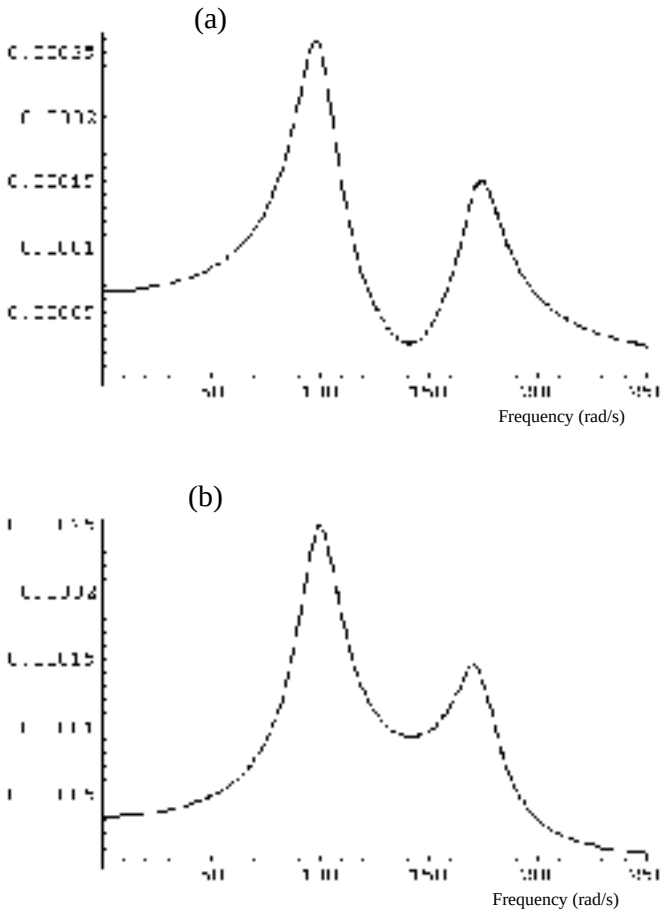


Figure 8.53. Comparison between HFRFs from a continuous-time system (full) and a discrete-time model (dashed): (*a*) $H_1^{(1:1)}(\omega)$; (*b*) $H_1^{(1:2)}(\omega)$.

This results in six equations for the variables $H_2^{(j:a\,b)}(\omega_1,\omega_2)$. The solution is

$$\begin{pmatrix} H_2^{(1:11)}(\omega_1,\omega_2) & H_2^{(1:12)}(\omega_1,\omega_2) & H_2^{(1:22)}(\omega_1,\omega_2) \\ H_2^{(2:11)}(\omega_1,\omega_2) & H_2^{(2:12)}(\omega_1,\omega_2) & H_2^{(2:22)}(\omega_1,\omega_2) \end{pmatrix} = \begin{pmatrix} H_1^{(1:1)}(\omega_1+\omega_2) & H_1^{(1:2)}(\omega_1+\omega_2) \\ H_1^{(2:1)}(\omega_1+\omega_2) & H_1^{(2:2)}(\omega_1+\omega_2) \end{pmatrix} \times \begin{pmatrix} S_2^{(1:11)}(\omega_1,\omega_2) & S_2^{(1:12)}(\omega_1,\omega_2) & S_2^{(1:22)}(\omega_1,\omega_2) \\ 0 & 0 & 0 \end{pmatrix} \quad (8.280)$$

where

$$S_2^{(1:ij)}(\omega_1,\omega_2) = -k' H_1^{(1:i)}(\omega_1) H_1^{(1:j)}(\omega_2). \quad (8.281)$$

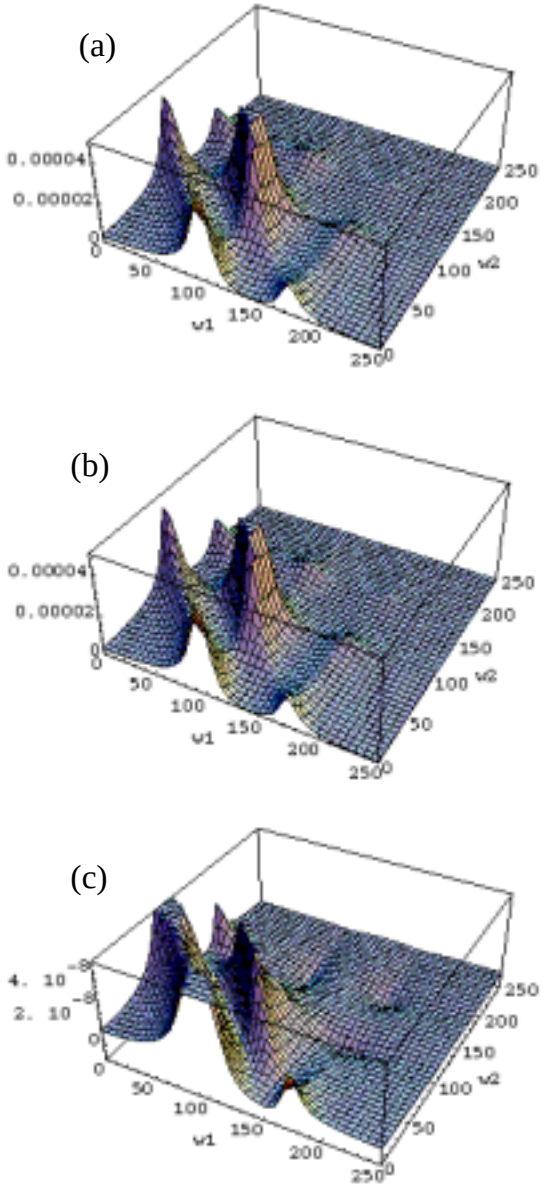


Figure 8.54. $H_2^{(1:11)}$ magnitude surface: (*a*) from the continuous-time equation of motion; (*b*) from a NARX model; (*c*) difference surface.

The magnitude surface for $H_2^{(1:11)}(\omega_1, \omega_2)$ is given in figure 8.54(*a*). To display the symmetry properties of the surface better, a contour map is given as

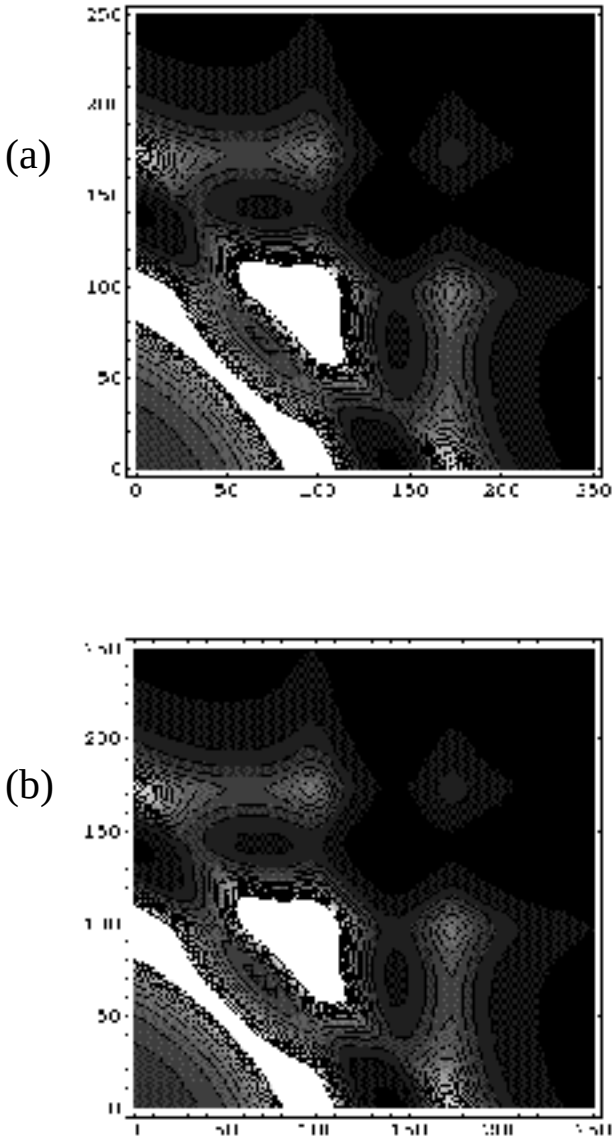


Figure 8.55. Contour map of $H_2^{(1:11)}$ magnitude surface: (*a*) from the continuous-time equation of motion; (*b*) from a NARX model.

figure 8.55(*a*). The magnitude of the cross-kernel HFRF $H_2^{(1:12)}(\omega_1, \omega_2)$ is given in figure 8.56(*a*). Note from the contour map in figure 8.57(*a*) that the symmetry about $\omega_1 = \omega_2$ is now absent.

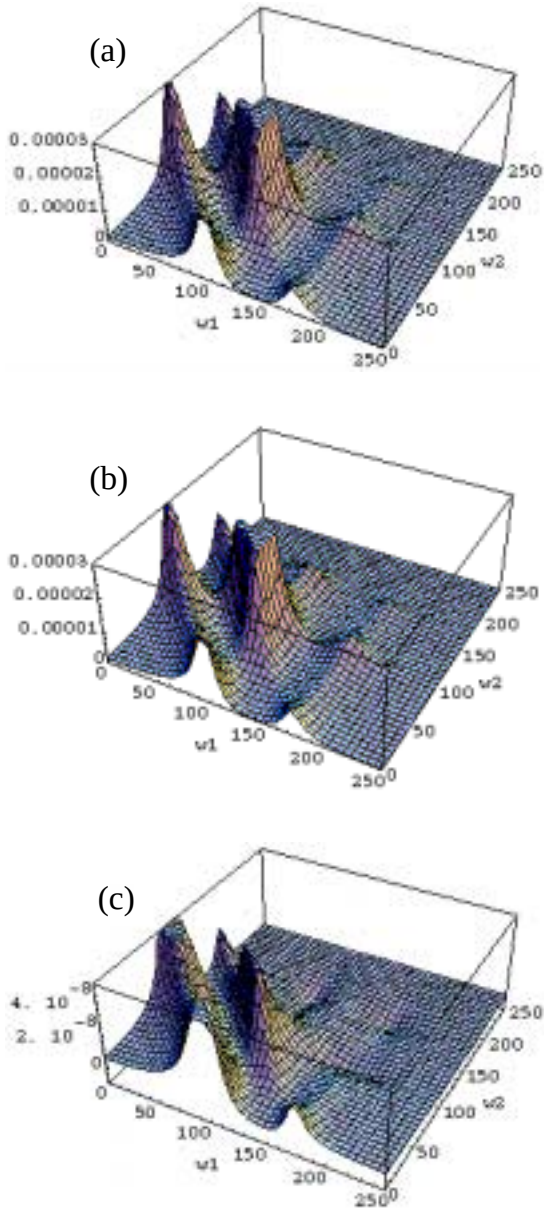
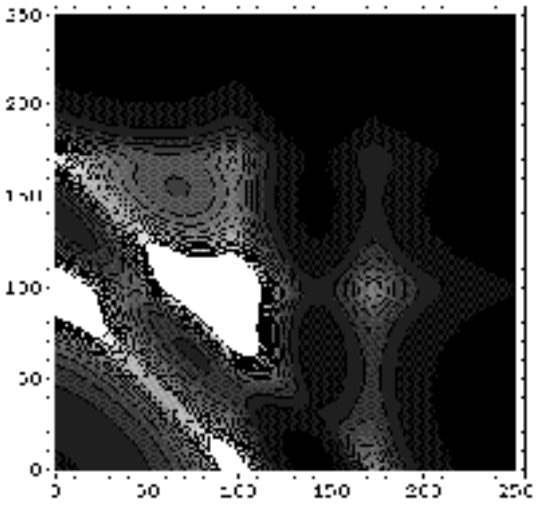


Figure 8.56. $H_2^{(1:12)}$ magnitude surface: (*a*) from the continuous-time equation of motion; (*b*) from a NARX model; (*c*) difference surface.

(a)



(b)

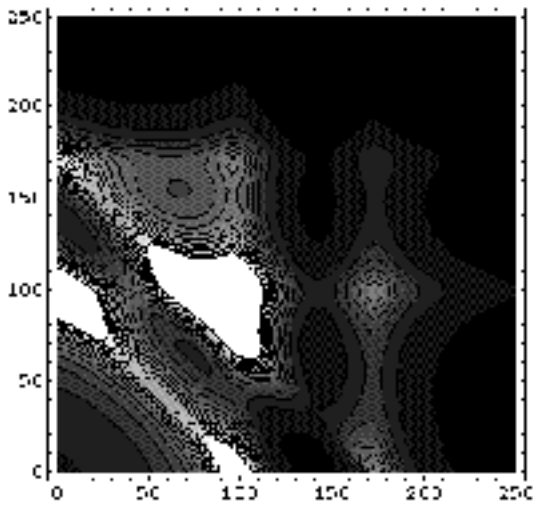


Figure 8.57. Contour map of $H_2^{(1:12)}$ magnitude surface: (*a*) from the continuous-time equation of motion; (*b*) from the NARX model.

8.12.2 HFRFs for a discrete-time MIMO system

In order to obtain a discrete-time system which compared with the continuous-time (8.272) and (8.273), these equations were integrated in time using a fourth-

order Runge–Kutta scheme. The two excitations $x^{(1)}(t)$ and $x^{(2)}(t)$ were two independent Gaussian noise sequences with rms 5.0, band-limited on the interval $(0, 200)$ Hz using a Butterworth filter. The time step was 0.001 corresponding to a sampling frequency of 1000 Hz. Using the resulting discrete force and displacement data, a NARX model was fitted. The model structure was easily guessed as

$$\begin{aligned} y_i^{(1)} &= \beta_1 y_{i-1}^{(1)} + \beta_2 y_{i-2}^{(1)} + \beta_3 y_{i-3}^{(1)} + \beta_4 y_{i-1}^{(2)} + \beta_5 y_{i-2}^{(2)} \\ &\quad + \beta_6 y_{i-1}^{(1)2} + \beta_7 y_{i-1}^{(1)3} + \beta_8 x_i^{(1)} + \beta_9 x_{i-1}^{(1)} + \beta_{10} x_{i-2}^{(1)} \qquad (8.282) \\ y_i^{(2)} &= \beta_{11} y_{i-1}^{(2)} + \beta_{12} y_{i-2}^{(2)} + \beta_{13} y_{i-3}^{(2)} + \beta_{14} y_{i-1}^{(1)} + \beta_{15} y_{i-2}^{(1)} \\ &\quad + \beta_{16} x_i^{(2)} + \beta_{17} x_{i-1}^{(2)} + \beta_{18} x_{i-2}^{(2)} \qquad (8.283) \end{aligned}$$

and the parameters were fitted using a linear LS algorithm. The actual values of the β_i coefficients will not be given here as they have little meaning. The model was tested by stepping it forward in time using the same excitations $x^{(1)}$ and $x^{(2)}$ and comparing the results with those from the original system. Figure 8.58 shows the comparisons for each DOF; it is impossible to distinguish the two curves (the NARX prediction is the broken line, the original data are the full line).

The HFRF extraction algorithm is almost identical to that for the continuous-time system, the same probing expressions and coefficient extraction procedures apply; the only difference is that where the derivative operator produces a prefactor $\mathrm{i}\omega$ when it meets the harmonic $\mathrm{e}^{-\mathrm{i}\omega t}$, the lag operator extracts a prefactor $\mathrm{e}^{\mathrm{i}\omega\Delta t}$, where Δt is the sampling interval.

The H_1 matrix for the discrete-time model is found to be

$$\begin{aligned} &\begin{pmatrix} H_1^{(1:1)}(\omega) & H_1^{(1:2)}(\omega) \\ H_1^{(2:1)}(\omega) & H_1^{(2:2)}(\omega) \end{pmatrix} \\ &\quad = \begin{pmatrix} 1 - \beta_1 \mathrm{e}^{-\mathrm{i}\omega\Delta t} - \beta_2 \mathrm{e}^{-\mathrm{i}2\omega\Delta t} - \beta_3 \mathrm{e}^{-\mathrm{i}3\omega\Delta t} & -\beta_4 \mathrm{e}^{-\mathrm{i}\omega\Delta t} - \beta_5 \mathrm{e}^{-\mathrm{i}2\omega\Delta t} \\ -\beta_{14} \mathrm{e}^{-\mathrm{i}\omega\Delta t} - \beta_{15} \mathrm{e}^{-\mathrm{i}2\omega\Delta t} & 1 - \beta_{11} \mathrm{e}^{-\mathrm{i}\omega\Delta t} - \beta_{12} \mathrm{e}^{-\mathrm{i}2\omega\Delta t} - \beta_{13} \mathrm{e}^{-\mathrm{i}3\omega\Delta t} \end{pmatrix}^{-1} \\ &\qquad \times \begin{pmatrix} \beta_8 + \beta_9 \mathrm{e}^{-\mathrm{i}\omega\Delta t} + \beta_{10} \mathrm{e}^{-\mathrm{i}2\omega\Delta t} & 0 \\ 0 & \beta_{16} + \beta_{17} \mathrm{e}^{-\mathrm{i}\omega\Delta t} + \beta_{18} \mathrm{e}^{-\mathrm{i}2\omega\Delta t} \end{pmatrix}. \end{aligned} \qquad (8.284)$$

The H_1s are shown in figure 8.53 as the broken lines, in comparison with the H_1s from the original continuous-time model. As would be expected from the accuracy of the NARX models, the two sets of H_1s agree almost perfectly.

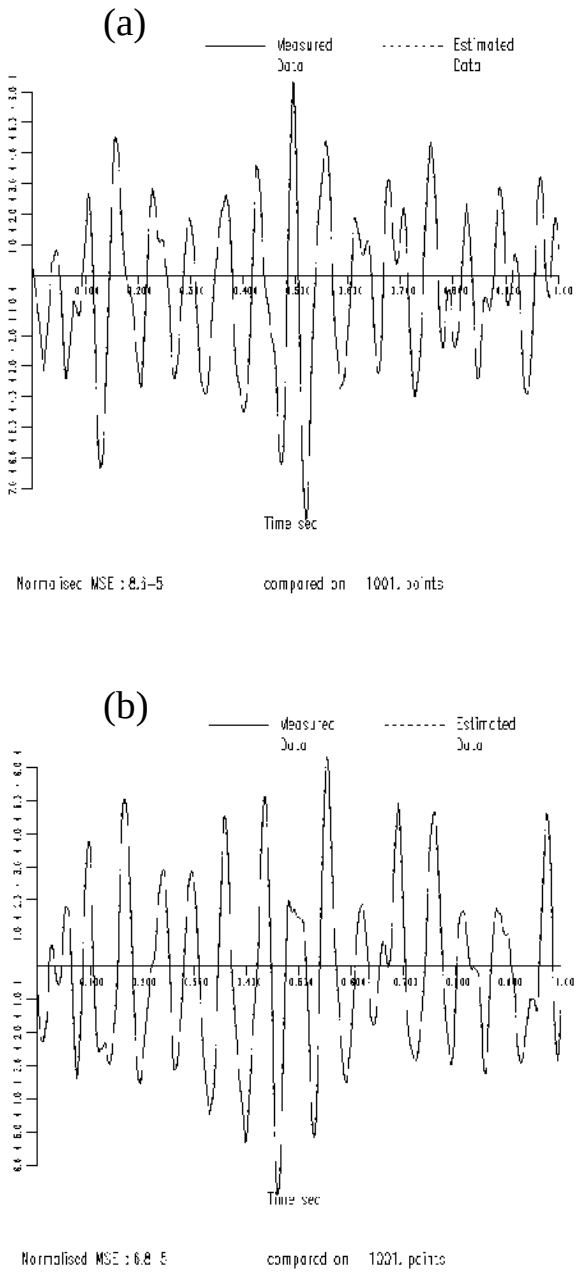


Figure 8.58. Comparison between the measured data and the NARX model-predicted output.

The H_2 matrix is found to be

$$\begin{aligned}
&\begin{pmatrix} H_2^{(1:11)}(\omega_1,\omega_2) & H_2^{(1:12)}(\omega_1,\omega_2) & H_2^{(1:22)}(\omega_1,\omega_2) \\ H_2^{(2:11)}(\omega_1,\omega_2) & H_2^{(2:12)}(\omega_1,\omega_2) & H_2^{(2:22)}(\omega_1,\omega_2) \end{pmatrix} \\
&\quad = \begin{pmatrix} H_1^{(1:1)}(\omega_1+\omega_2) & H_1^{(1:2)}(\omega_1+\omega_2) \\ H_1^{(2:1)}(\omega_1+\omega_2) & H_1^{(2:2)}(\omega_1+\omega_2) \end{pmatrix}^{-1} \\
&\qquad \times \begin{pmatrix} \beta_8 + \beta_9 \mathrm{e}^{-\mathrm{i}(\omega_1+\omega_2)\Delta t} + \beta_{10}\mathrm{e}^{-\mathrm{i}2(\omega_1+\omega_2)\Delta t} & 0 \\ 0 & \beta_{16} + \beta_{17}\mathrm{e}^{-\mathrm{i}(\omega_1+\omega_2)\Delta t} + \beta_{18}\mathrm{e}^{-\mathrm{i}2(\omega_1+\omega_2)\Delta t} \end{pmatrix} \\
&\qquad \times \begin{pmatrix} T_2^{(1:11)}(\omega_1,\omega_2) & T_2^{(1:12)}(\omega_1,\omega_2) & T_2^{(1:22)}(\omega_1,\omega_2) \\ 0 & 0 & 0 \end{pmatrix}
\end{aligned} \tag{8.285}$$

where

$$T_2^{(1:ij)}(\omega_1,\omega_2) = \beta_6 H_1^{(1:i)}(\omega_1) H_1^{(1:j)}(\omega_2)\mathrm{e}^{-\mathrm{i}(\omega_1+\omega_2)t}. \tag{8.286}$$

Examples of some of the H_2 surfaces for the discrete-time system can be found in figures 8.54–8.55 where they are compared with the corresponding continuous-time objects. As with the H_1 functions, the agreement is excellent.

The MIMO version of the Volterra series is no more difficult to apply than the standard SISO expansion. If multiple inputs are applied to a MDOF system, the Volterra cross-kernels or cross-HFRFs *must* be taken into account in any analysis, as they encode important information about intermodulations between the input signals.

Chapter 9

Experimental case studies

The previous chapters in this book have mainly concentrated on the theory of the methods, with occasional diversions which discuss experimental results. This might be regarded as unsatisfactory coverage of what is essentially a subject motivated by experiment and full-scale test. The aim of this final chapter is to provide a fuller context for the methods by describing a number of experimental case studies which apply the theory 'in anger' as it were.

The sections which follow can also be regarded as suggestions for nonlinear system identification *demonstrators*. The systems are recommended for their good behaviour and repeatability and provide a framework for the reader to explore possibilities for nonlinear identification.

9.1 An encastré beam rig

Throughout this book, the Duffing oscillator has been treated as a paradigm for nonlinear systems as innumerable researchers have done in the past. Its attraction is in its apparent simplicity. However, despite appearances, the system is capable of a wide range of behaviour which can be exploited for benchmarking system identification procedures. The problem with the Duffing system is that it is not trivial to construct a mechanical system for experimental studies which has the required simple equations of motion. The system discussed here has proved to be the most satisfactory in the authors' previous work. (Note that it is possible to simulate arbitrary nonlinear systems using analogue circuitry, and this has proved attractive occasionally [272]; however, laboratory systems are preferable as they provide a more realistic environment.)

The structure suggested here could hardly be simpler; it is a flexible beam with encastré (built-in or clamped) end conditions. As the following theory will show, the displacement response of this beam approximates that of a Duffing oscillator when high enough amplitude vibrations induce nonlinear strains. Further, by preloading the beam, the cubic response is made asymmetric and the equation of motion incorporates an additional quadratic stiffness term.

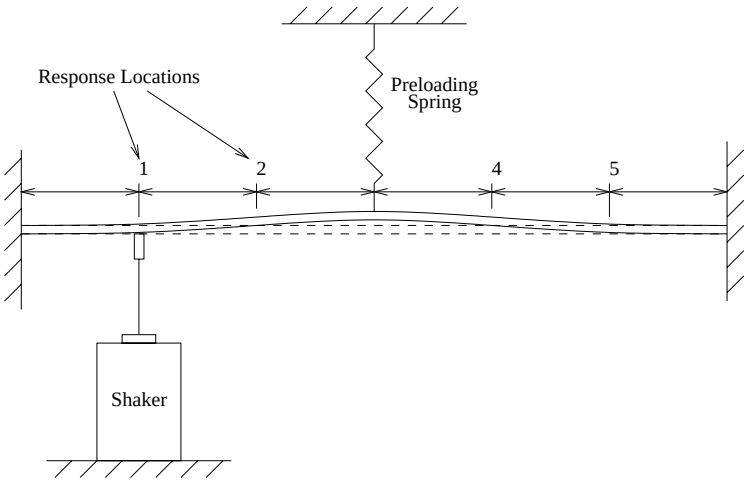


Figure 9.1. Diagram of the preloaded encastré beam experiment.

9.1.1 Theoretical analysis

Figure 9.1 shows a schematic diagram of the system where the beam is preloaded by the action of a spring. If the response is only measured at the centre of the beam, the system can be approximated by a SDOF model. The following analysis is essentially lifted from [110].

For a clamped–clamped beam, it is shown in [42] that the modeshape in the first mode of vibration is

$$v = \frac{y}{1.5881}\left[\cosh\left(\frac{\lambda x}{L}\right) - \cos\left(\frac{\lambda x}{L}\right) - \sigma\left(\sinh\left(\frac{\lambda x}{L}\right) - \sin\left(\frac{\lambda x}{L}\right)\right)\right] \tag{9.1}$$

where $\lambda = 4.7300$ and $\sigma = 0.9825$. x is the distance from one end of the beam and y is the displacement of the central point. Gifford makes—and justifies—the assumption that the profile

$$v = \frac{y}{2}\left[1 - \cos\left(\frac{2\pi x}{L}\right)\right] \tag{9.2}$$

is a good first approximation [110]. (It satisfies the boundary conditions of $v = \mathrm{d}v/\mathrm{d}x = 0$ at $x = 0$ and $x = L$ and also $v = \mathrm{d}v/\mathrm{d}x = 0$ at $x = L/2$.) If the preload force X induces a constant offset Δ to the motion, the resulting profile is

$$v = \frac{y + \Delta}{2}\left[1 - \cos\left(\frac{2\pi x}{L}\right)\right]. \tag{9.3}$$

The equation of motion can be derived using Lagrange's approach and requires the evaluation of the kinetic and potential energies for the system. It is assumed in the following that the damping can be ignored.

First, estimate the kinetic energy T

$$T = \int \mathrm{d}T = \int_0^L \mathrm{d}x \, \frac{1}{2} m' \left(\frac{\mathrm{d}u}{\mathrm{d}t} \right)^2 \tag{9.4}$$

where m' is the mass per unit length. Now

$$\frac{\mathrm{d}v}{\mathrm{d}t} = \frac{\partial v}{\partial y} \frac{\mathrm{d}y}{\mathrm{d}t} = \frac{\partial v}{\partial y} \dot{y} = \frac{1}{2} \left[1 - \cos \left(\frac{2\pi x}{L} \right) \right] \dot{y} \tag{9.5}$$

so the total kinetic energy is

$$T = \frac{m'}{8} \dot{y}^2 \int_0^L \mathrm{d}x \, \left[1 - \cos \left(\frac{2\pi x}{L} \right) \right]^2 = \frac{3m'L}{16} \dot{y}^2. \tag{9.6}$$

The next task is to compute the potential energy. For simplicity, it is assumed that the excitation $x(t)$ acts at the centre of the beam. The total strain energy is composed from four sources:

$$\begin{aligned} \text{total strain energy} = {} & \text{strain energy due to bending} \\ & + \text{strain energy due to tension} \\ & + \text{strain energy due to springs} \\ & - \text{work done on beam by external force } x(t) \end{aligned} \tag{9.7}$$

or

$$\begin{aligned} V(y) = {} & \frac{EI}{2} \left\{ \int_0^L \mathrm{d}x \left(\frac{\mathrm{d}^2 v}{\mathrm{d}x^2} \right)^2 - \int_0^L \mathrm{d}x \left(\frac{\mathrm{d}^2 v}{\mathrm{d}x^2} \right)^2 \Big|_{\Delta=0} \right\} \\ & + \frac{1}{2} EAL(\epsilon_{\mathrm{T}}^2 - \epsilon_0^2) + \frac{1}{2} k[(X+y)^2 - X^2] - \int \mathrm{d}y \, x \end{aligned} \tag{9.8}$$

where E is the Young's modulus, I is the second moment of area of the beam cross-section about the neutral axis, A is the cross-sectional area, ϵ_{T} is the longitudinal strain due to stretching of the beam (assumed uniform throughout the length), ϵ_0 is the tension strain in the equilibrium position, k is the stiffness of the springs and X is the initial preloaded displacement of the spring. (Note that the product of bending strain and tension strain integrates to zero over the length of the beam.)

The strain energy due to axial tension is computed as follows: assuming that the strain is uniform along the beam, $\epsilon_{\mathrm{T}} = \Delta L / L$, where ΔL is the extension. It follows that

$$\begin{aligned} \epsilon_{\mathrm{T}} = {} & \frac{1}{L} \int_0^L \mathrm{d}x \, \left(\sqrt{1 + \left(\frac{\mathrm{d}v}{\mathrm{d}x} \right)^2} - 1 \right) \approx \int_0^L \mathrm{d}x \, \left(1 + \frac{1}{2} \left(\frac{\mathrm{d}v}{\mathrm{d}x} \right)^2 - 1 \right) \\ = {} & \int_0^L \mathrm{d}x \, \frac{1}{2} \left(\frac{\mathrm{d}v}{\mathrm{d}x} \right)^2 \end{aligned} \tag{9.9}$$

and using (9.3) for the beam profile v gives

$$\epsilon_{\mathrm{T}} = \frac{\pi^2}{2L^3}(\Delta + y)^2 \int_0^L \mathrm{d}x \, \sin^2\left(\frac{2\pi x}{L}\right) = \frac{(\Delta + y)^2 \pi^2}{4L^2} \tag{9.10}$$

and the tension strain energy $V_{\mathrm{T}}(y)$ is

$$V_{\mathrm{T}}(y) = \frac{1}{2} EAL(\epsilon_{\mathrm{T}}^2 - \epsilon_0^2) = \frac{EA\pi^4}{32L^3}[(\Delta + y)^4 - \Delta^4]. \tag{9.11}$$

The next term needed is the strain energy due to bending $V_{\mathrm{B}}(y)$, from (9.8) this is

$$\begin{aligned} V_{\mathrm{B}}(y) &= \frac{EI}{2}\left\{ \int_0^L \mathrm{d}x \left(\frac{\mathrm{d}^2 v}{\mathrm{d}x^2}\right)^2 - \int_0^L \mathrm{d}x \left(\frac{\mathrm{d}^2 v}{\mathrm{d}x^2}\right)^2 \Big|_{\Delta=0} \right\} \\ &= EI[(\Delta + y)^2 - \Delta^2] \frac{\pi^4}{2L^4} \int_0^L \mathrm{d}x \, \cos^2\left(\frac{2\pi x}{L}\right) \\ &= \frac{EI\pi^4}{L^3}[(\Delta + y)^2 - \Delta^2]. \end{aligned} \tag{9.12}$$

So the total potential energy is, from (9.8)

$$\begin{aligned} V(y) = \frac{EA\pi^4}{32L^3}[(\Delta + y)^4 - \Delta^4] + \frac{EI\pi^4}{L^3}[(\Delta + y)^2 - \Delta^2] \\ + \tfrac{1}{2}k[(X + y)^2 - X^2] - xy. \end{aligned} \tag{9.13}$$

Now forming the Lagrangian $L = T - V$, the equation of motion follows from Lagrange's equation

$$\frac{\partial L}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{y}} = 0. \tag{9.14}$$

After a little work, one finds

$$\begin{aligned} \frac{3m'L}{8}\ddot{y} + \left[\frac{EA\pi^4}{8L^3} + \frac{2EI\pi^4}{L^3} + 2k\right] y + \frac{3\Delta EA\pi^4}{8L^2} y^2 + \frac{EA\pi^4}{8L^3} y^3 \\ + \frac{2EI\Delta\pi^4}{L^3} + \frac{EA\pi^4\Delta^3}{8L^3} + kX = x(t). \end{aligned} \tag{9.15}$$

Note that the force X produces the offset Δ at equilibrium and the response y is measured with respect to this equilibrium position. This means that $y = 0$ if $x = 0$. This means all the constants must sum to zero so

$$\frac{3m'L}{8}\ddot{y} + \left[\frac{3\Delta^2 EA\pi^4}{8L^3} + \frac{2EI\pi^4}{L^3} + k\right] y + \frac{3\Delta EA\pi^4}{8L^2} y^2 + \frac{EA\pi^4}{8L^3} y^3 = x(t) \tag{9.16}$$

or

$$m\ddot{y} + k_1 y + k_2 y^2 + k_3 y^3 = x(t) \tag{9.17}$$

is the final equation of motion, which is that of an asymmetric undamped Duffing oscillator as required. Several observations can be made.

(1) For small oscillations with $y^2 \ll y$, the linear natural frequency of the system is

$$\omega_n = \frac{1}{2\pi}\sqrt{\frac{3\Delta^2 EA\pi^4 + 16EI\pi^4 + 16kL^3}{12m'L^4\pi^2}}. \tag{9.18}$$

(2) The coefficient of y^3 is positive, so the system is a hardening cubic for large excitations. The quadratic term, despite its positive coefficient will produce a softening effect as discussed in chapter 3.

(3) If $\Delta = 0$, equation (9.16) becomes

$$\frac{3m'L}{8}\ddot{y} + \left[\frac{2EI\pi^4}{L^3} + k\right] y + \frac{EA\pi^4}{8L^3}y^3 = x(t) \tag{9.19}$$

or

$$m\ddot{y} + k_1 y + k_3 y^3 = x(t) \tag{9.20}$$

and the system is a classical undamped Duffing oscillator.

(4) It follows from the analysis that the nonlinearity is purely a result of the tension strain energy, i.e. large forces cause the beam to change its length because it is axially restricted due to the encastré boundary conditions.

9.1.2 Experimental analysis

In order to illustrate the behaviour of the beam, the Volterra series and the corresponding higher-order FRFs (HFRFs) will be applied. The following analysis is due to Storer and can be found in more detail in [237].

Before results are presented, one or two points are worth mentioning. First, a simple bright mild steel beam presents an experimental problem in the sense that the damping is extremely low—approximately 0.1% of critical. This leads to problems with frequency resolution (which can be alleviated somewhat by using zoomed measurements), which can limit the accuracy of FRF-based methods. A simple solution to this problem, and the one adopted here, is to add damping to the beam in the form of a constrained viscoelastic layer. In this case, thick foam tape of thickness $\frac{1}{8}$ in was added to one side of the beam and constrained using $\frac{1}{32}$ in thick steel. The damping layer was added to the central 20 in region of the beam. This had the effect of raising the damping in the first few modes to 2% of critical and this was considered acceptable.

The dimensions of the beam and the associated geometrical and material constants are given in table 9.1.

If the preload is set to zero, the constants of the Duffing oscillator model of the beam from (9.19) are given in table 9.2.

Table 9.1. Material and geometrical constants for steel beam.

Quantity	Symbol	Magnitude	Units
Length	L	1.2	m
Width	w	25.4	mm
Thickness	t	3.17	mm
Area	$A = wt$	8.052×10^{-5}	m^2
Second moment of area	$I = wt^3/12$	6.743×10^{-11}	m^4
Density	ρ	7800	kg m^{-3}
Mass per unit length	ρwt	0.628	kg
Young's modulus	E	2.011×10^{11}	N m^{-2}

Table 9.2. Duffing oscillator model constants for steel beam.

Quantity	Magnitude	Units
m	0.283	kg
k_1	1432.9	N m^{-1}
k_3	1.066×10^8	N m^{-3}
$f_n = \omega_n/2\pi = \sqrt{k_1/m}$	11.2	Hz

9.1.2.1 Linear analysis

The first test on the structure was to find the natural frequencies with and without preload. For the dynamical tests, accelerometers were placed at constant intervals of 0.2 m as shown on figure 9.1. Note that the accelerometers add mass of 0.03 kg each which is not negligible compared to the 0.78 kg mass of the beam, the natural frequencies are modified accordingly. The preload was induced in this test by a single tension spring.

In order to identify the natural frequencies the system was subjected to broadband random vibration of rms 1 V and the resulting measured accelerations were passed to a spectrum analyser, in order to compute the averaged FRF. The basic components of the experiment are:

(1) Broadband random noise generator—to produce the excitation signal. In this case a DIFA/Scadas data acquisition system was used.
(2) Power amplifier—to drive the shaker with the noise signal. In this case a Gearing and Watson amplifier.
(3) Electrodynamic shaker—to transmit the random force signal to the structure. Here, a Ling V201 shaker was used.
(4) Force gauge—to measure the force transmitted to the beam. A Kistler force gauge, type 911 was used here.

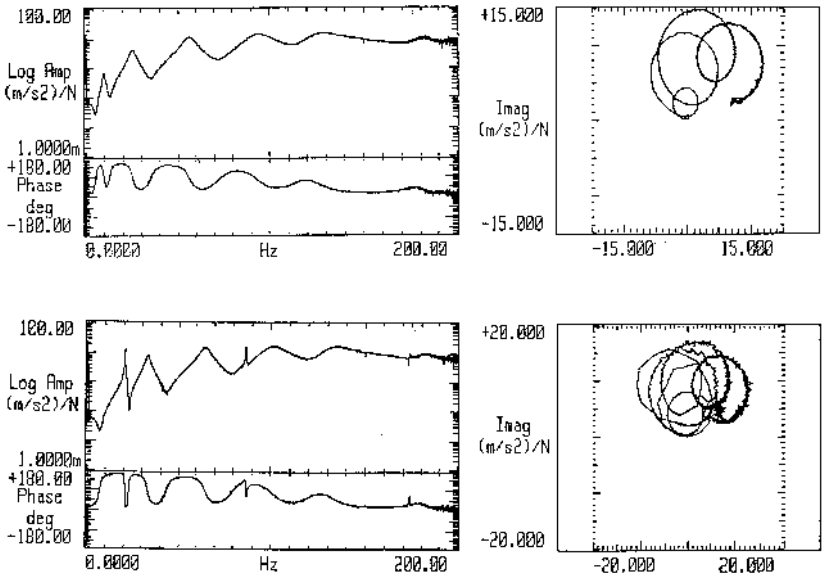


Figure 9.2. Accelerance FRF measured using random excitation with 10 averages from the direct location: (*a*) without preload; (*b*) with preload.

(5) At least one accelerometer to measure the acceleration response. Endevco accelerometers were used here.
(6) At least one charge amplifier—to amplify the signals from the accelerometer. In this case, Bruel and Kjaer.
(7) A two-channel (at least) spectrum analyser or frequency response function analyser to compute the FRFs. The DIFA/Scadas system used here is driven by LMS software which has a full FRF analysis capability.

The accelerance FRFs for the beam with and without preload are shown in figure 9.2. The acceleration is measured from the direct response point (point 1 in figure 9.1) and the FRFs are the result of 10 averages. There are predominantly five natural frequencies up to 200 Hz. They are tabulated in table 9.3; the preload was adjusted to give a lateral deflection $\Delta = 5$ mm.

As expected, the natural frequencies rise substantially when the preload is applied. In addition, there is an extra very sharp resonance at 86.7 Hz when the spring is tightened; this is a wave resonance of the tension spring (it is unclear why this resonance only appeared when the spring was tensioned).

9.1.2.2 FRF distortion and Hilbert transform

Random excitation is optimal for producing FRFs for linear systems. However, as discussed in chapter 2, stepped-sine excitation is superior for the analysis of

Table 9.3. Experimental natural frequencies of beam—unloaded and loaded.

Natural frequency (Hz)	Without preload	With preload
f_1	9.7	20.4
f_2	25.4	32.9
f_3	56.7	64.9
f_4	95.3	101.6
f_5	127.2	134.4

nonlinear systems. The second set of results given here are for a stepped-sine test of the preloaded beam. The same experimental components as described earlier are needed for the test. In addition, the FRF analyser should have stepped-sine or swept-sine capability—fortunately this is common.

Figure 9.3 shows the accelerance FRFs from the direct point on the beam when the sinusoidal forcing amplitudes are 0.1 V and 0.4 V. (Note that the amplitudes are voltages as specified at the signal generator, the magnitude of the force in Newtons at the beam is a matter for measurement.) In this case, the interpretation is simply that there is a 'low' level of excitation and a 'high'. There is noticeable distortion in the 0.4 V FRF (dotted). In fact there is a bifurcation or 'jump' at the first resonance as discussed in chapters 2 and 3. The second and third resonances shift upwards in frequency when the excitation is increased, and this is consistent with the hardening nature of the nonlinearity. In contrast, the first resonance shifts down and this is explained by the presence of the softening quadratic stiffness term induced by the preload. Note that the Duffing oscillator analogy is not precise because the system under investigation is a MDOF. However, a full MDOF analysis of the system would still be expected to reveal quadratic and cubic nonlinearities.

Note that the equipment, which is routinely used for linear structural analysis, is already giving information about the presence and type of the nonlinearity via the FRF distortion.

Given an FRF from stepped-sine excitation another useful and relatively simple test for linearity is given by the Hilbert transform (chapter 4). Three algorithms for computing the transform are given in section 4.4 of chapter 4. The Hilbert transform of the 0.1 V FRF is given in figure 9.4. There is marked distortion indicating nonlinearity even in this FRF at the lower level of forcing. In particular, the first mode shifts downwards consistent with softening. Note that the FRF distortion measure is simpler to see if two FRFs at two levels are available for comparison and this is essentially a test of homogeneity. The Hilbert transform only needs one FRF to be measured.

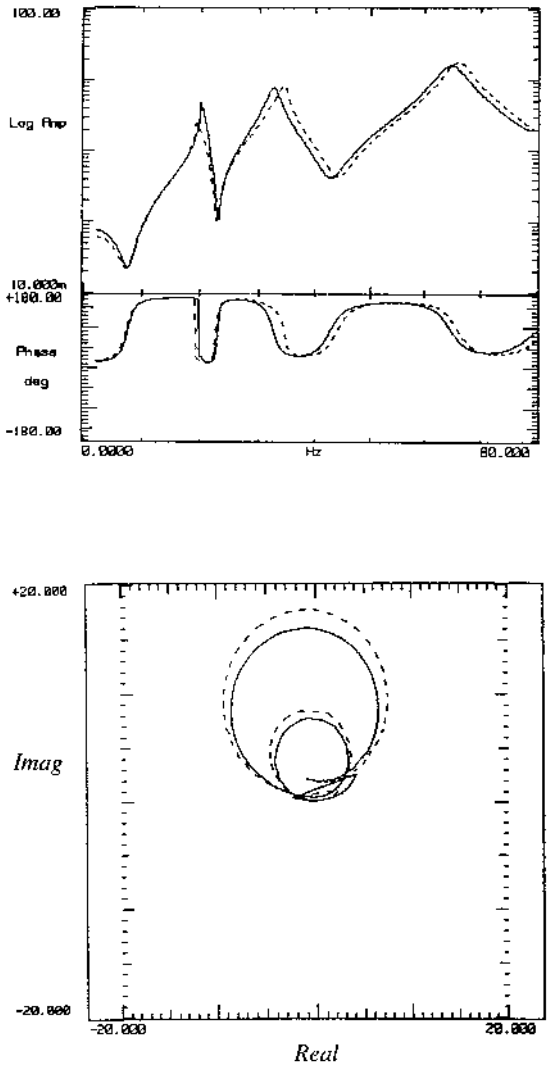


Figure 9.3. Accelerance FRF measured from the preloaded beam at the direct location using sine excitation with amplitudes 0.1 V (full) and 0.4 V (broken).

9.1.2.3 Static test

The engineer should never scorn simplicity. For systems with polynomial stiffness nonlinearities, it may be possible to extract the stiffness characteristic from a static test. In this case, it was carried out using the simple arrangement shown in figure 9.5. The load deflection curves in the unloaded and preloaded state are

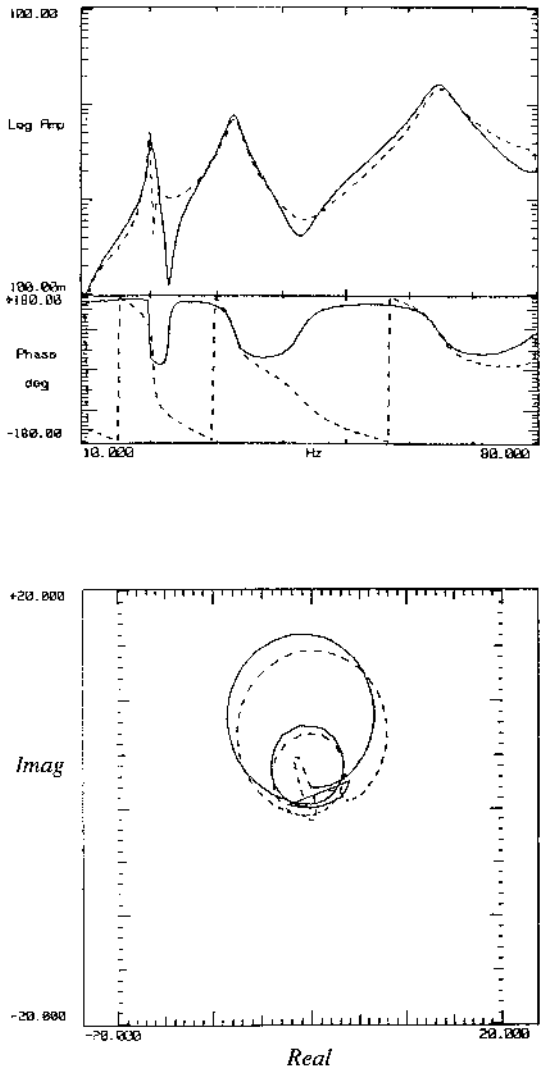


Figure 9.4. Accelerance FRF (full) and Hilbert transform (broken) measured at location 1 using sine excitation with an amplitude of 0.1 V (lower level).

given in figures 9.6 and 9.7. Note the symmetry of the curve for the un-preloaded beam.

Mathematical expressions for the stiffness curve can be obtained by fitting polynomial curves using the least-squares method. This is discussed in more detail in the context of stiffness sections in section 7.3. When a cubic of the form

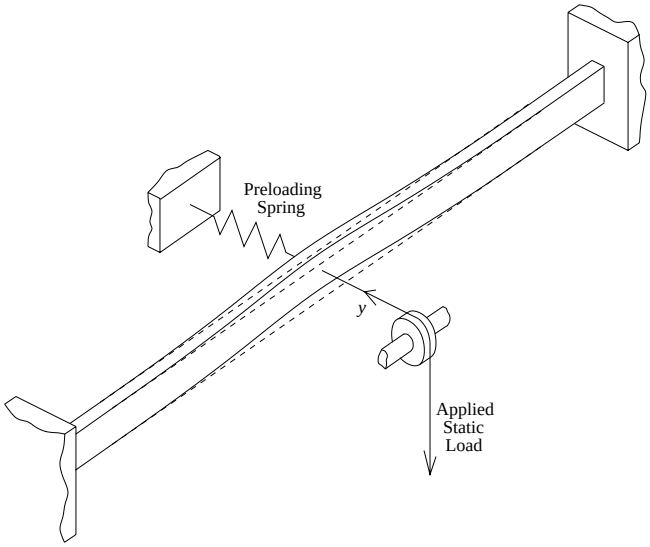


Figure 9.5. Arrangement of the static test on the preloaded beam.

$f(y) = \alpha_1 y + \alpha_3 y^3$ was fitted, the resulting coefficients were $\alpha_1 = 1710.0$ N m^{-1} and $\alpha_3 = 0.634 \times 10^8$ N m^{-3}. These can be compared with the estimates $k_1 = 1432.9$ N m^{-1} and $k_3 = 1.066 \times 10^8$ N m^{-3}, from the simple theory of the previous section. The results are quite close given the naivety of the theory (i.e. the SDOF assumption and the approximate deflection shape (9.2)).

For the preloaded beam an asymmetric curve is fitted: $f(y) = \beta_1 y + \beta_2 y^2 + \beta_3 y^3$. The results are $\beta_1 = 8650.0$ N m^{-1}, $\beta_2 = 1.2 \times 10^6$ N m^{-2} and $\beta_3 = 0.602 \times 10^8$ N m^{-3}. Note that due to the preload, $\alpha_1 \neq \beta_1$ (see equations (9.16) and (9.19)).

9.1.2.4 Higher-order FRFs

This section describes how HFRFs can be used to fit parametric models. The basis of the method is given in section 8.10; there it is shown that in equations (8.219) to (8.221), under sinusoidal excitation $x(t) = X \sin(\Omega t)$,

$$H_1(\Omega) = \frac{Y(\Omega)}{X} + \mathrm{O}(X^2) \tag{9.21}$$

$$H_2(\Omega, \Omega) = \frac{2Y(2\Omega)}{X^2} + \mathrm{O}(X^2) \tag{9.22}$$

$$H_3(\Omega, \Omega, \Omega) = \frac{4Y(3\Omega)}{X^3} + \mathrm{O}(X^2). \tag{9.23}$$

So, a stepped-sine test can yield information about the diagonals of the HFRFs. Now, as the expressions for the diagonals of H_2 and H_3

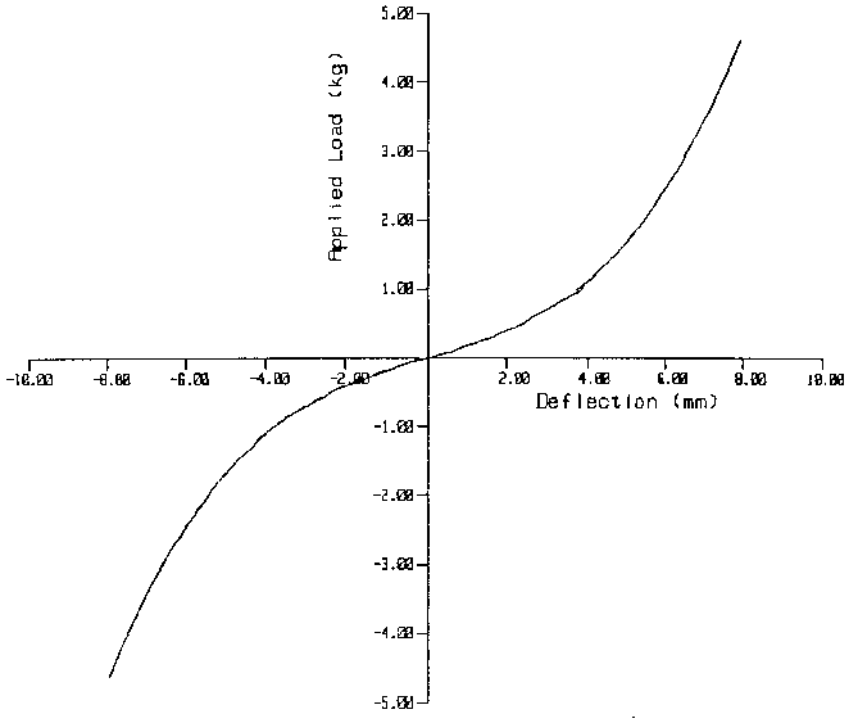


Figure 9.6. Load deflection curve (full) measured from beam without preload: a polynomial curve-fit of the form $f(y) = \alpha_1 y + \alpha_3 y^3$ is superimposed (broken).

(equations (8.206) and (8.207) respectively) contain all the structural parameters of interest up to third order in the system equations of motion, curve-fitting to the diagonal HFRFs will yield estimates of the parameters and thus provide a model.

For reference, the diagonal variants of (8.206) and (8.207) are included here:

$$\begin{aligned} H_2^{rs}(\omega,\omega) = & \sum_{m=1}^{N} (\omega^2 c_{2mm} - k_{2mm}) H_1^{sm}(2\omega) H_1^{rm}(\omega)^2 \\ & - \sum_{m=1}^{N-1} \sum_{n=m+1}^{N} (\omega^2 c_{2mn} - k_{2mn}) [H_1^{sm}(2\omega) - H_1^{sn}(2\omega)] \\ & \times [H_1^{rm}(\omega) - H_1^{rn}(\omega)]^2 \end{aligned} \tag{9.24}$$

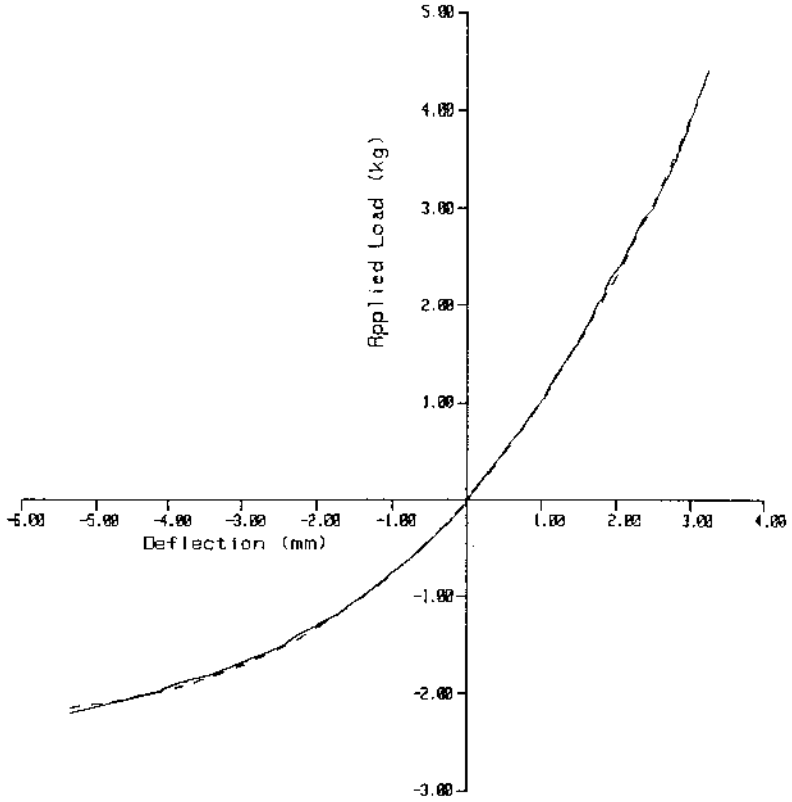


Figure 9.7. Load deflection curve (full) measured from beam with preload: a polynomial curve-fit of the form $f(y) = \beta_1 y + \beta_2 y^2 + \beta_3 y^3$ is superimposed (broken).

and

$$
\begin{aligned}
H_3^{rs}(\omega,\omega,\omega) = {} & 2\sum_{m=1}^{N} H_1^{sm}(3\omega)[2\omega^2 c_{2mm} - k_{2mm}]H_1^{rm}(\omega)H_2^{rm}(\omega,\omega) \\
& + 2\sum_{m=1}^{N-1}\sum_{n=m+1}^{N} [H_1^{sm}(3\omega) - H_1^{sn}(3\omega)][2\omega^2 c_{2mn} - k_{2mn}] \\
& \times [H_1^{rm}(\omega) - H_1^{rn}(\omega)][H_2^{rm}(\omega,\omega) - H_2^{rn}(\omega,\omega) \\
& + \sum_{m=1}^{N} (i\omega^3 c_{3mm} - k_{3mm})H_1^{sm}(3\omega)H_1^{rm}(\omega)^3 \\
& - \sum_{m=1}^{N-1}\sum_{n=m+1}^{N} (\mathrm{i}\omega^3 c_{3mn} - k_{3mn})[H_1^{sm}(3\omega) - H_1^{sn}(3\omega)]
\end{aligned}
$$

$$\times [H_1^{rm}(\omega) - H_1^{rn}(\omega)]^3. \tag{9.25}$$

The equipment needed for such a test is essentially as before. However, matters are considerably simplified if the signal generator and acquisition device are matched as follows. Suppose a sinusoid at frequency Ω is input, the sampling frequency for acquisition must be set at a multiple of Ω in order to avoid leakage in the Fourier transforms which are needed to extract the harmonics. Any leakage will result in magnitude and phase errors in both the fundamental and the harmonics. The suggested strategy is this. Generate a forcing signal $X\sin(\Omega t)$. Set the sampling frequency to $\Omega/2^N$ with N in the range 4–7, this gives between 16 (well above the Nyquist limit even for the third harmonic) and 128 points per cycle. Accumulate 2^N points of the response at a time and Fourier transform. The resulting spectrum will have the d.c. component as the first spectral line, the response at the fundamental $Y(\Omega)$ as the second, the response at second harmonic $Y(2\Omega)$ as the third, etc. Note that the spectral line will specify both amplitude and phase of the response as required. This is the strategy followed in [237] and illustrated here.

At each new frequency, the forcing signal will elicit a transient response. This procedure should be repeated until, at each frequency, the magnitude and phase of the fundamental and the harmonics has stabilized.

If anti-aliasing filters or any form of filters are in use, the low-pass cut-off must be set high enough to allow the highest harmonic required to pass unimpeded.

It is critical that the forcing signal is a sinusoid, any higher harmonic content on the input will invalidate the equations (9.21)–(9.23). Such harmonic content can be introduced by nonlinearities in the shaker for example. In some cases, it may be necessary to use closed-loop feedback control in order to ensure a constant amplitude sinusoidal force.

When this strategy was applied to the encastré beam, the results were as shown in figure 9.8; the excitation amplitude is 0.1 V. Note the presence in H_2 of the peaks at half the first and second resonant frequencies, this type of phenomenon is shown in figure 8.15 and discussed in section 8.4 for a SDOF system. Figure 9.9 shows a comparison between diagonal HFRFs measured at 0.1 V and 0.4 V. They are not coincident, and this shows that the measured HFRFs are not amplitude invariants, as they should be. The reason is the presence of the $\mathrm{O}(X^2)$ term in equations (9.21) and (9.23). The test requires a balance in amplitude between exciting the relevant nonlinearities and introducing amplitude dependence.

Once the HFRFs are obtained, a parametric model can be extracted by curve-fitting. However, there is a more direct approach than fitting to the diagonals, as described in the next section. The diagonals are then used in the process of model validation.

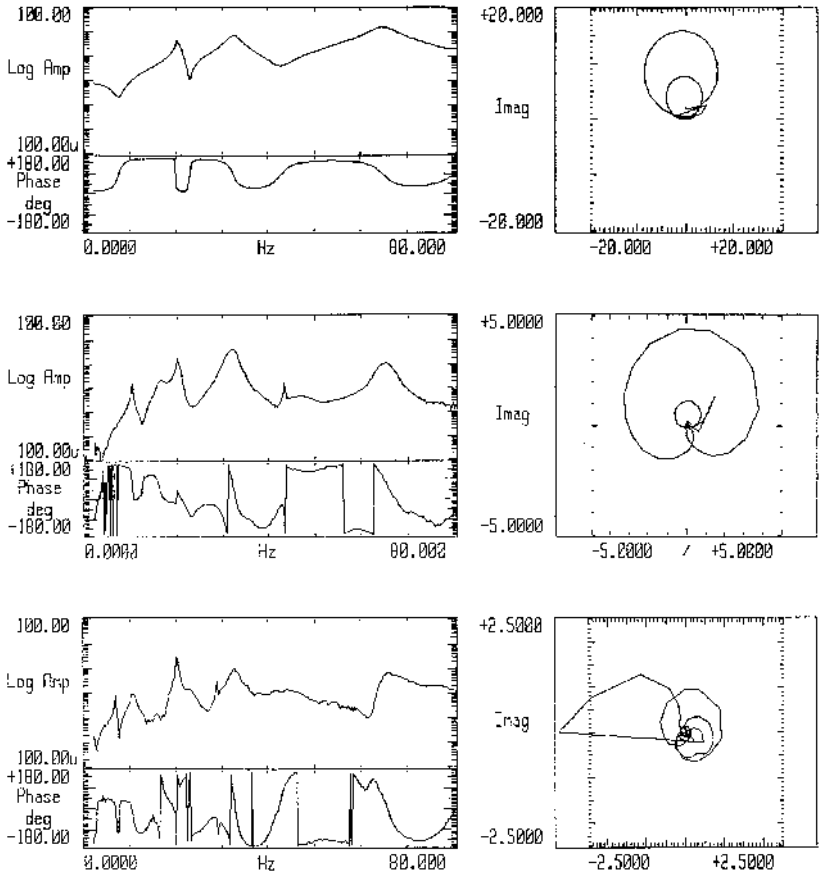


Figure 9.8. First-, second- and third-order diagonal accelerance HFRFs measured using 0.1 V sine excitation from the preloaded beam at location 1.

9.1.2.5 Direct parameter estimation (DPE)

Storer [237] adopted a DPE approach to the modelling, the relevant discussion is in section 7.5[1]. Because five response points are used, the model is assumed to be 5DOF. A broadband random excitation was used at a high level of 4 V in order to excite the nonlinearities. The excitation was band-limited between 6 and

[1] Storer actually also used a variant of DPE based in the frequency domain to identify the system at low excitation. The fundamental objects are the measured FRFs. These are used to generate, via inverse Fourier transform, time-data for the DPE curve-fit. The advantages of the method are that the integrations needed to estimate velocity and displacement can be carried out in the frequency domain and also that noise effects are reduced by using averaged quantities. The method can only be used, of course, for linear structures or at low excitations when the effect of nonlinearities can be neglected.

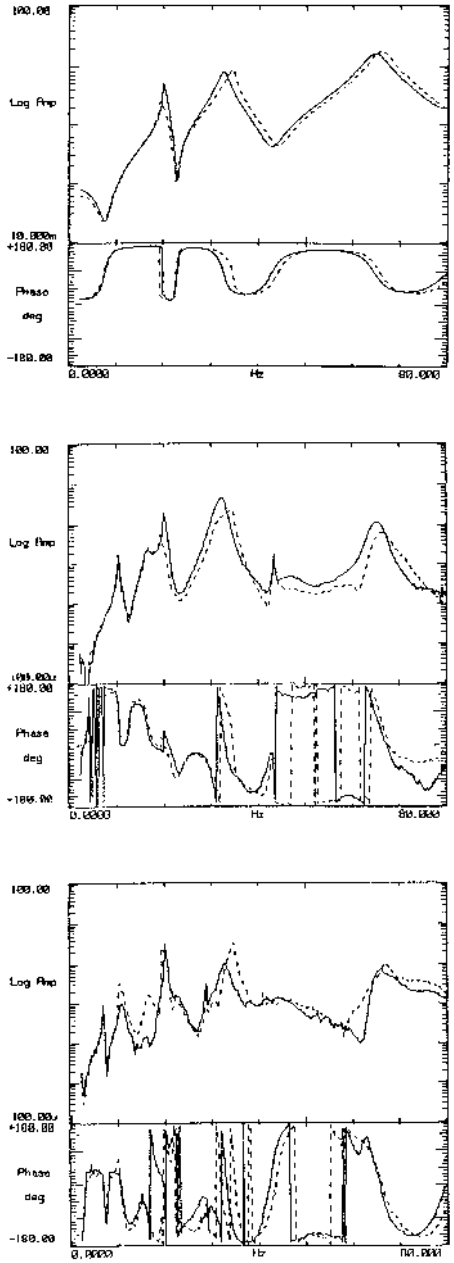


Figure 9.9. First-, second- and third-order diagonal accelerance HFRFs measured using 0.1 V (full) and 0.4 V (broken) sine excitation from the preloaded beam at location 1.

80 Hz. The low frequency cut-off was set to alleviate problems with integration of the data as described in appendix I. After the integration process (trapezium rule), the data were filtered in the interval 6–300 Hz in order to capture the appropriate harmonics. The linear mass, damping and stiffness parameters obtained from the parameter estimation were

$$[m] = \begin{pmatrix} 0.234 & 0 & 0 & 0 & 0 \\ 0 & 0.233 & 0 & 0 & 0 \\ 0 & 0 & 0.114 & 0 & 0 \\ 0 & 0 & 0 & 0.211 & 0 \\ 0 & 0 & 0 & 0 & 0.264 \end{pmatrix} \text{[kg]}$$

$$[c] = \begin{pmatrix} 19.29 & 23.56 & 28.64 & -22.53 & -26.20 \\ 5.26 & 17.65 & 9.11 & 9.94 & 15.60 \\ -2.48 & -2.26 & 17.81 & 2.81 & 9.18 \\ -22.80 & -52.30 & -72.00 & 37.70 & -65.50 \\ 61.47 & 14.32 & 257.2 & 94.39 & 215.7 \end{pmatrix} \text{[N s m}^{-1}\text{]}$$

$$[k] = \begin{pmatrix} 63.64 & 11.51 & 7.52 & -32.16 & -7.38 \\ 11.51 & 42.72 & 18.55 & -26.58 & -31.73 \\ 7.52 & 18.55 & 0.64 & -12.95 & -15.60 \\ -32.16 & -26.58 & 12.95 & 3.60 & 17.01 \\ -7.38 & -31.73 & -15.60 & 17.01 & 34.56 \end{pmatrix} \times 10^3 \text{ [N m}^{-1}\text{]}. \tag{9.26}$$

Note the enforced symmetry of the stiffness matrix, this is required in order to estimate all parameters using only one excitation point as described in section 7.5. The nonlinear parameters are given in table 9.4 (there is no significance cut-off).

The direct method of validating these parameters is to assess the normalized mean-square errors (MSE) between the model data and the measured data. For DOF 1–5 the MSE values on the accelerations were found to be 0.4207, 0.1591, 9.0797, 0.9197 and 0.8638. All indicate excellent agreement except for the 3DOF which is nonetheless acceptable. Because the HFRF diagonals have been measured, there is a alternative means of validation. Figure 9.10 shows the measured diagonal HFRFs compared with the results predicted using the parametric model. The results are very reasonable.

In conclusion, the preloaded encastré beam is a versatile rig which can be tuned simply to give symmetric or asymmetric characteristics. The methods illustrated here span a broad range of nonlinear system identification techniques and they by no means exhaust the possibilities of the rig.

9.2 An automotive shock absorber

This section contrasts the preceding one by concentrating on one method of analysis, but exploring it in more depth. The system of interest is an automotive

Table 9.4. Nonlinear parameters from DPE method on the preloaded beam.

Links		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	0.703	1.251
4	5	0.532	15.05
5	ground	0.552	−6.139

damper or shock absorber which can be purchased at any motor spares depot. As will be shown, the instrumentation is a little more involved, but can be assembled using components available in most well-equipped dynamics laboratories.

The background for the analysis is given in section 7.4.5 which briefly shows the identification of a monofrequency restoring force surface for a shock absorber. The discussion there related more to the motivation of the analysis than the actual experimental procedure, the discussion here will reverse the priorities. The overall object of the analysis is to improve on the standard characterizations of absorbers, namely the work (figure 9.11) and characteristic (figure 9.12) diagrams.

9.2.1 Experimental set-up

The procedures of chapter 7 are applied here to a number of sets of test data measured and provided by Centro Ricerche FIAT (CRF) of Torino, Italy. The data were obtained using the experimental facilities of the vehicle test group at CRF. The apparatus is shown in figure 7.51. The components are as follows:

(1) A fixed frame to which the shock absorber can be attached. This should be pretested to ensure that no natural frequencies of the frame intrude into the

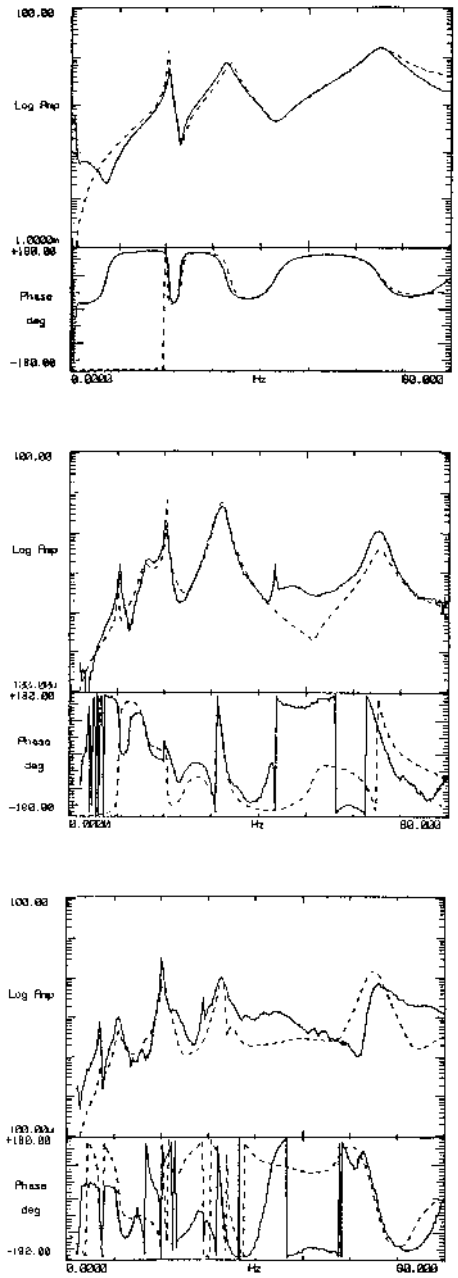


Figure 9.10. Comparison between the higher-order diagonal FRFs measured at the direct location (full) and the FRF diagonals regenerated using the identified parameters (broken).

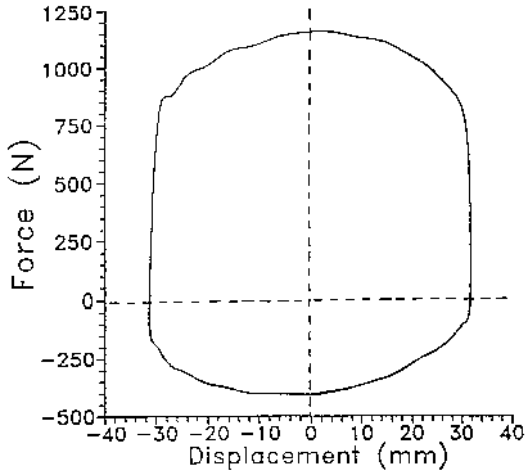


Figure 9.11. Typical shock absorber work diagram.

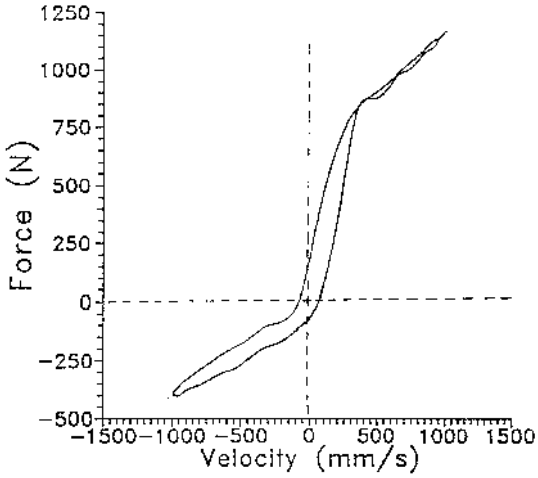


Figure 9.12. Typical shock absorber characteristic diagram.

frequency interval of the test [50].

(2) A load cell or force transducer terminated at the top of the frame to measure the restoring force of the absorber.
(3) A displacement transducer to measure the displacement of the base end of the absorber. (Alternatively, one can use an accelerometer, but this was sub-optimal here—see later.)
(4) An actuator. This can be electrodynamic or hydraulic. In this case because

of the high forces and strokes required, a hydraulic actuator was used.

(5) A controlled signal source which can deliver a sinusoidal displacement of a prescribed frequency and amplitude to the base of the absorber. In this case closed-loop control was used.

(6) A time-data acquisition system.

In the tests, the shock absorber was allowed to move only in the vertical direction, this constraint was necessary to ensure that the absorber behaved as if it were a SDOF system. The top of the absorber was grounded against a load cell which allowed the measurement of the force transmitted by the absorber. The system was excited by applying a hydraulic actuator to the free base of the shock absorber. Fixed to the base were a displacement transducer and an accelerometer, the velocity of the base was the only unmeasured state variable.

In order to minimize temperature effects, each test was carried out as quickly as possible to avoid internal heating in the absorber.

If one assumes that the inertia of the shock absorber is concentrated at the grounded end, a useful simplification is obtained as discussed in section 7.4.3, the relevant equation for the system is

$$f(y, \dot{y}) = x(t) \tag{9.27}$$

i.e. it is the restoring force itself which is measured by the load cell.

Shock absorbers are known to be frequency dependent., i.e.

$$f = f(y, \dot{y}, \omega) \tag{9.28}$$

where ω is the frequency of excitation. In order to allow the possibility of investigating frequency effects, a sinusoidal response was required. A closed-loop control strategy was adopted whereby the excitation signal was modified adaptively until a harmonic displacement output was obtained[2].

If the displacement response of a system is a single harmonic, force data are only available above the corresponding phase trajectory which is simply an ellipse. For this reason periodic signals are not the best choice for generating restoring force surfaces; ideally, force data are required which are evenly distributed in the phase plane. In order to meet this requirement, several tests were carried out at each frequency and each subtest was for a different response amplitude. For each frequency, this procedure gave force data over a set of

[2] In reality the shock absorber will be subjected to random vibrations, so there are essentially two modelling strategies: first, one can identify the system using random vibration and fit a non-physical model like the hyperbolic tangent model or polynomial model described in chapter 7. The second approach is to identify a model for one fixed frequency and assume the frequency dependence of the system is small. (A possible third approach is to identify monofrequency models for several frequencies and pass between them somehow in a manner dependent on the instantaneous frequency of the excitation. The problem here is to sensibly specify an instantaneous frequency for a random signal.) The second approach was taken by FIAT; however, tests were carried out at several frequencies in order to decide which gave the most widely applicable model and also to allow the possibility of examining the variations between models.

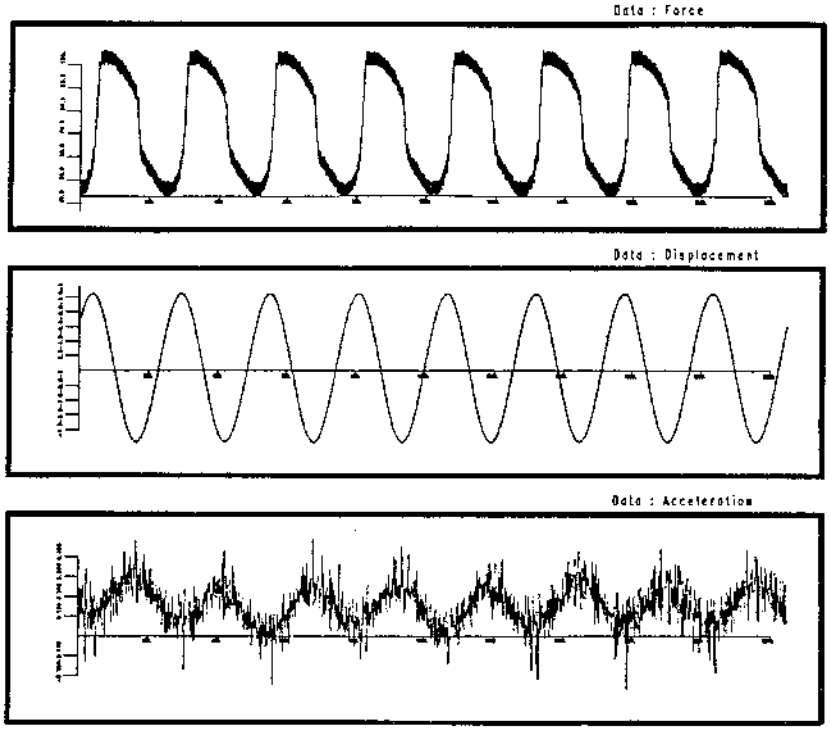


Figure 9.13. Measured data from the **F** shock absorber: 1 Hz, 1 mm subtest.

concentric curves in the phase plane. This allowed the construction of a force surface for each test frequency. Comparison of such surfaces could indicate whether the absorber under test was frequency dependent. At a given frequency the range of displacements which can be obtained is limited by the maximum velocity which the actuator can produce. At high frequencies lower amplitudes can be reached as the maximum test velocity is equal to the product of the maximum test displacement and the frequency. The frequencies at which tests were carried out were 1, 5, 10, 15, 20 and 30 Hz although only the 1 and 10 Hz tests are discussed here. Much more detail can be found in [273].

The data were supplied in a series of subtest files, each of which contained eight cycles of the sampled force, displacement and acceleration signals. Each file channel contained 2048 points, i.e. 256 points per cycle.

An example of the subtest data, that for the 1 Hz, 1 mm amplitude test is shown in figure 9.13. Both the force signal and the acceleration signal were polluted by high-frequency noise. The corruption of the acceleration signal is very marked; this was due to hydraulic noise from the actuator. A simple digital

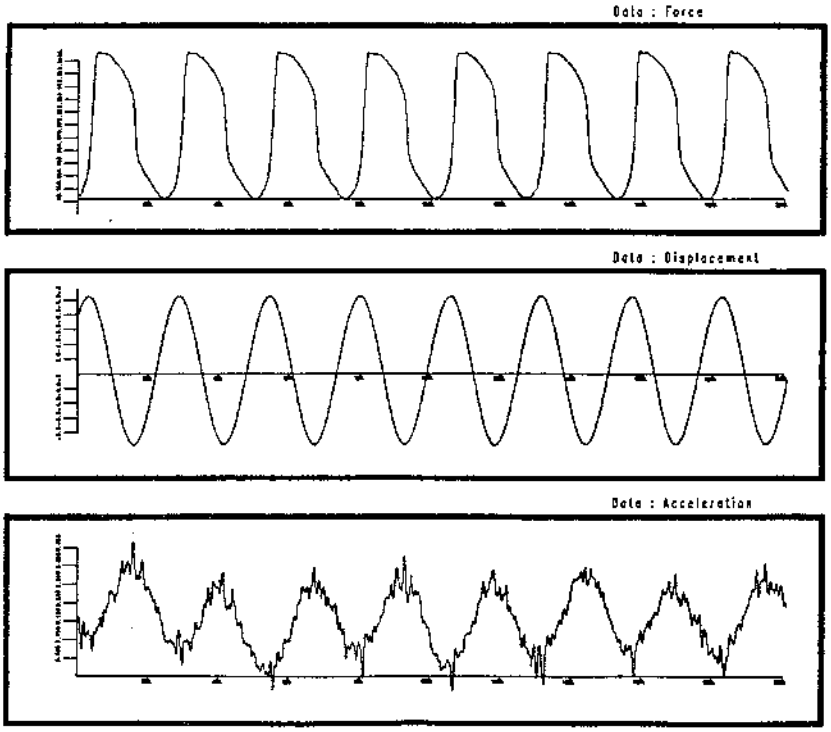


Figure 9.14. Measured data from the **F** shock absorber after smoothing: 1 Hz, 1 mm subtest.

filter ('smoothing by 5s and 3s' [129]), which is represented numerically by

$$y_i = \tfrac{1}{15}(x_{i+3} + 2x_{i+2} + 3x_{i+1} + 3x_i + 3x_{i-1} + 2x_{i-2} + x_{i-3}) \qquad (9.29)$$

was used to remove the noise component from the force and acceleration signals. The gain of this filter falls off quite rapidly with frequency; however, because the experimental sampling rate was so high, there was no appreciable modification of the higher harmonic content of the force signal, as one can immediately see from figure 9.14 which shows the data from figure 9.13 after smoothing (the filter in (9.29) was applied several times). Another important point regarding the use of this filter is that it is *not* zero-phase, and actually causes a small backward time-shift in the force and acceleration signals which is evident on careful comparison of figures 9.13 and 9.14. In order to maintain the synchronization of the signals, which is crucial to the restoring force method, the measured displacement data are also passed through the filter. This simultaneous filtering of the input and output data is not strictly valid for nonlinear systems as the resulting input–output process may be different to the original. It is justified here by the fact that the only

possible modification to the harmonic content of the signals would be to the force data and, as previously observed, there is no visible modification. As a further check, surfaces were obtained for both smoothed and unsmoothed data and no appreciable difference was found beyond the presence of noise on the surface for the unsmoothed data.

A further problem was the estimation of the unmeasured velocity data. These could have been obtained in two ways as described in appendix I:

(1) by numerically integrating the measured acceleration data; and
(2) by numerically differentiating the measured displacement data.

Of the two possibilities it is usually preferable to integrate for reasons discussed in appendix I. However, for a number of reasons differentiation appeared to be the appropriate choice here, those reasons being:

(1) Although numerical differentiation amplifies high-frequency noise, in this case the displacement signal was controlled and reasonably noise-free. Further, the displacement signal was passed through the smoothing filter. Consequently, there was little noise to amplify.
(2) Differentiation may not be zero-phase and can introduce delays. However, delays, if induced, are generally small and are usually only a problem if multiple differentiations are carried out. Because the method is so sensitive to the presence of delays, a number of tests were carried out in which data from differentiation and data from integration were compared by looking at the positions of the zero-crossings; in each case there was no appreciable phase difference.
(3) The acceleration signal was corrupted by both low- and high-frequency noise. It is clear from figure 9.14 that the high-frequency component was not completely removed by smoothing. However, integration itself is a smoothing operation so the high-frequency component is usually removed in passing to velocity. The main problem comes from the low-frequency component which can introduce significant trends into the estimated velocity (figure 9.15). These trends need to be removed, making integration more troublesome to implement than differentiation.

In summary, all force surfaces obtained in this section were obtained using velocities from differentiation of measured displacements. The main reservation is that, in actual road tests using instrumented vehicles, it is usually only accelerometers which are installed. In this case, the experimental strategy should be arranged in order to ensure acceleration data of an appropriate quality.

After calculation of the velocities, each set of processed subtest files were assembled into test files, each containing data with a common test frequency but a range of amplitudes. At this point the data files contained up to 25 000 points with many cycles repeated; to speed the analysis they were decimated by factors of 2 or 3 in order to obtain test files with 6000–8000 points. The results of the analysis are summarized in the following section.

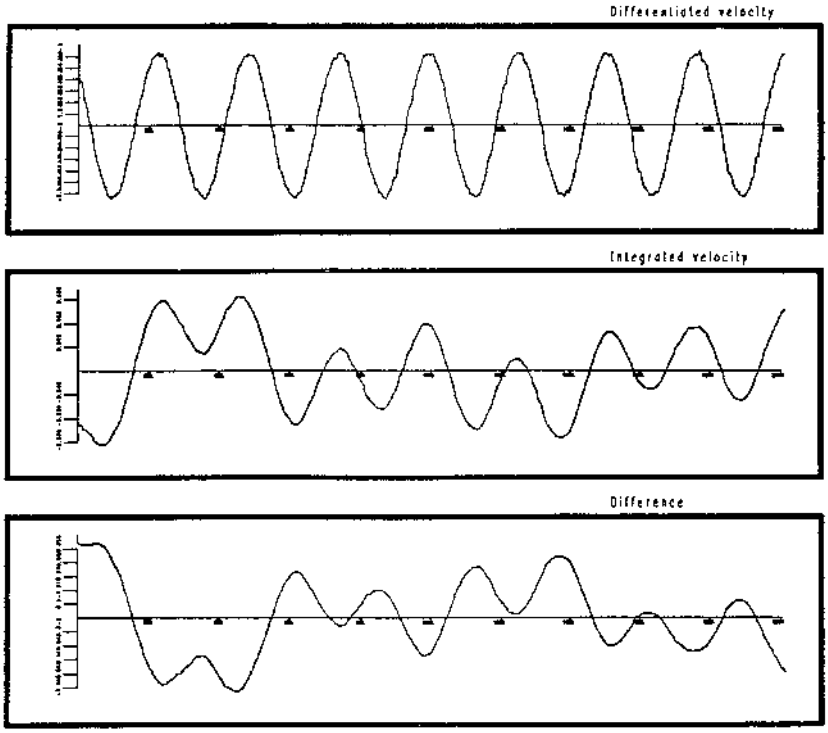


Figure 9.15. Velocity estimates from differentiation and integration for the **F** shock absorber after data smoothing: 1 Hz, 1 mm subtest.

9.2.2 Results

Two types of shock absorbers were tested. The first type (labelled **F**) is used in the front suspension of a FIAT vehicle, the second type (labelled **R**) is used for the rear suspension in the same vehicle. Each absorber was tested at six frequencies, 1, 5, 10, 15, 20, and 30 Hz, a total of 24 tests. For the sake of brevity, the following discussion is limited to the 10 Hz tests for each absorber, the full set of results can be found in [273]. In fact, the qualitative characteristics of each absorber changed very little over the range of frequencies.

*9.2.2.1 Front absorber (**F**)*

This absorber can be used to illustrate one point common to both the absorbers regarding the control procedure used to generate a sinusoidal displacement signal. If one considers the phase trajectories at low frequency (figure 9.16 for 1 Hz), one observes that they are a set of concentric ellipses (they appear to be circles because of the scaling of the plotting routines). This shows clearly that the control was

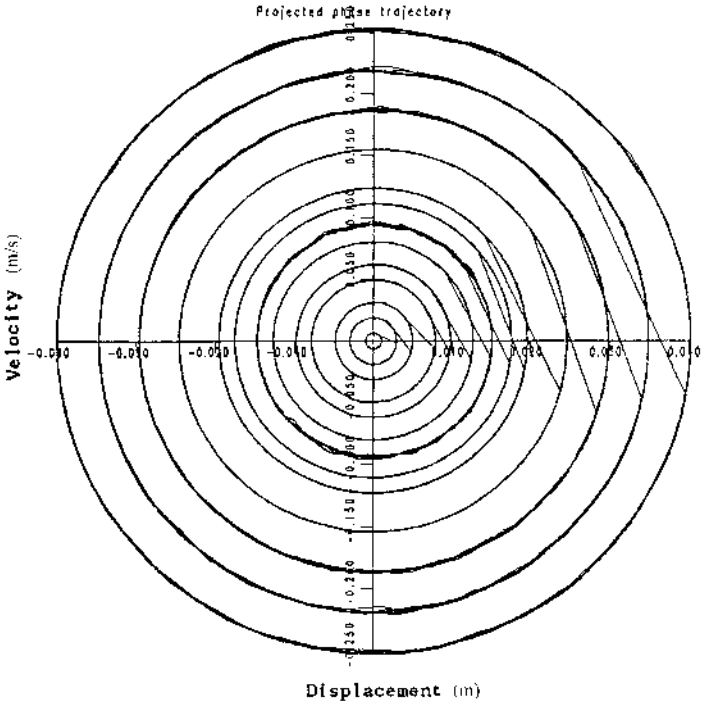


Figure 9.16. Phase trajectories from the **F** shock absorber after smoothing and velocity estimation: 1 Hz test.

successful in producing a harmonic displacement. However, at higher frequencies (figure 9.17 for 20 Hz), where control was more difficult, the trajectories were far from circular. This does not cause any problems for applying the restoring force method which is insensitive to the type of excitation used. However, it will cause a little uncertainty regarding the interpretation. If the absorbers are frequency dependent, one may not be observing the pure characteristics of the system at that frequency if the response contains an appreciable component formed from higher harmonics.

The phase trajectories for the 10 Hz test are shown in figure 9.18, all but the highest amplitudes give a circular curve, indicating that the control has been successful.

The characteristic diagram is shown in figure 9.19, and gives a very clear indication of the type of damping nonlinearity. The characteristic is almost one of bilinear damping, i.e. constant high damping (i.e. high $c_{\mathrm{eq}} = \frac{\partial f}{\partial \dot{y}}$ in rebound, low c_{eq} in compression). This is modified slightly at high forcing levels when the damping increases a little. The presence of a finite area enclosed by the characteristic diagram shows that the system has some position dependence

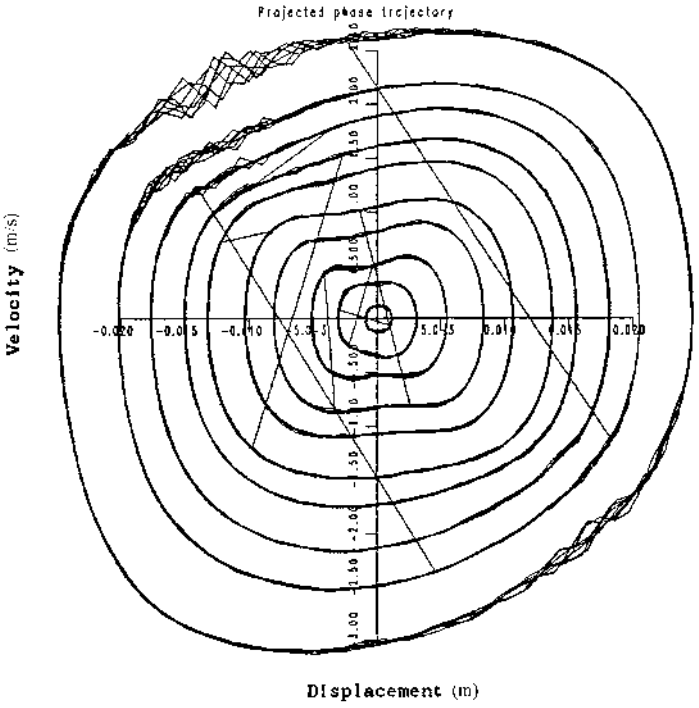


Figure 9.17. Phase trajectories from the **F** shock absorber after smoothing and velocity estimation: 20 Hz test.

or stiffness. As the frequency increases, the enclosed area increases slightly, indicating perhaps a small degree of nonlinearity in the position dependence. The work diagram is shown in figure 9.20. The slight positive correlation between force and displacement asymmetry also indicates the presence of a small stiffness characteristic.

The restoring force surface and the associated contour map are given in figures 9.21 and 9.22. Some comments on the range of validity of the surface are required here. The interpolation procedure which was used to construct the data is only exact for a linear system as discussed in section 7.2.2, consequently in regions away from the main body of the data, the fastest that the interpolated value or interpolant can grow is linearly. For these tests, the regions in the corners of the plotting grids are furthest away from the data, which are essentially confined within a circular region in the centre of the plotting square. As a result, the interpolated surface can only grow linearly in the corners. However, the corners correspond to regions of simultaneous large positions and velocities and are therefore areas where the force will be likely to show nonlinear behaviour. In summary, the corner regions of the force surfaces are not to be trusted. A rough

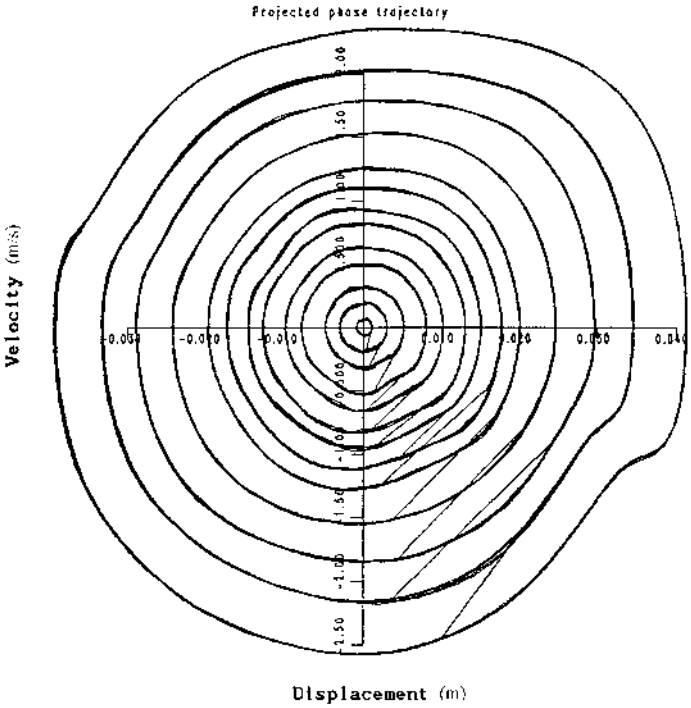


Figure 9.18. Phase trajectories from the **F** shock absorber after smoothing and velocity estimation: 10 Hz test.

indicator of the range of validity is given by the broken circle in figure 9.22; one should think of this circle as superimposed on all subsequent contour maps. The force surface shown in figure 7.52 which comes from the analysis of this test is zoomed and therefore does not show the edge effects.

The restoring force surface shows a very clear bilinear characteristic as one would expect from the previous discussion. Even clearer is the contour map; the contours, which are concentrated in the positive velocity half-plane, are almost parallel and inclined at a small angle to the $\dot{y} = 0$ axis showing that the position dependence of the absorber is small and nominally linear.

The tests at all other frequencies provided confirmation for the previous conclusions. It is difficult to make a direct comparison between surfaces at different frequencies because different extents of the phase plane are explored in each test. However, if the topography of the surfaces does not change, e.g. no new features appear suddenly at an identifiable common point in two surfaces, then there is evidence for independence of frequency. This condition is necessary but not sufficient for frequency independence; one should check that the actual force values above a given point are not changing in some systematic fashion.

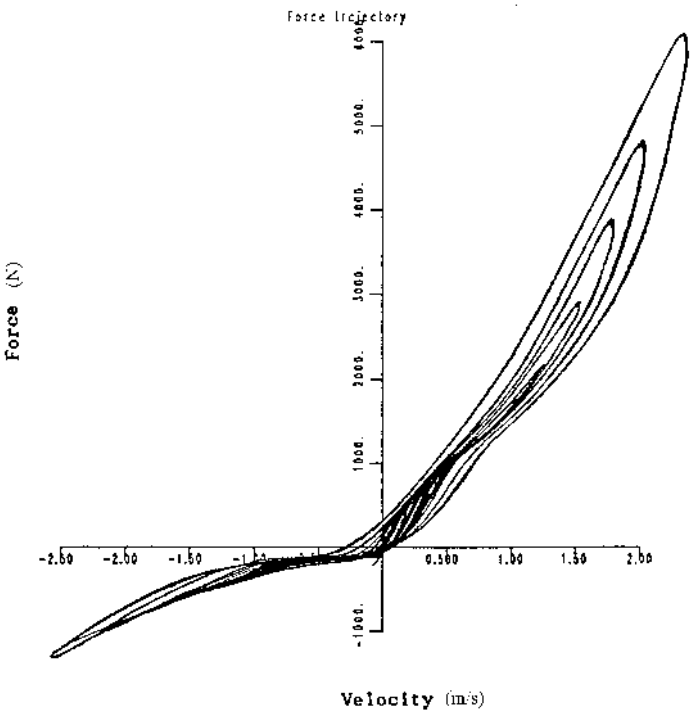


Figure 9.19. Characteristic diagram for the **F** shock absorber: 10 Hz test.

This type of analysis was carried out for the data in [274].

In summary, absorber **F** has a small linear stiffness and a strong bilinear damping characteristic. Further, there is no evidence of displacement–velocity coupling effects.

*9.2.2.2 Rear absorber (**R**)*

The phase trajectories for the 10 Hz test are given in figure 9.23. As before, the control is good up to high amplitudes.

The characteristic diagram is shown in figure 9.24; it displays a sort of soft Coulomb characteristic[3]. Note the shape of the 'hysteresis' loop, it is widest in the region where the velocity reverses. This may indicate that for **R** the stiffness is more significant at low velocities than at high. The work diagram (figure 9.25) provides essentially the same information as before, i.e. a small

[3] By a 'soft' friction characteristic it is simply meant that the transition between negative and positive forces does not occur instantaneously as the phase trajectory passes through $\dot{y} = 0$. Instead, the transition is more gradual and the characteristic resembles the hyperbolic tangent as discussed in section 7.6.2.

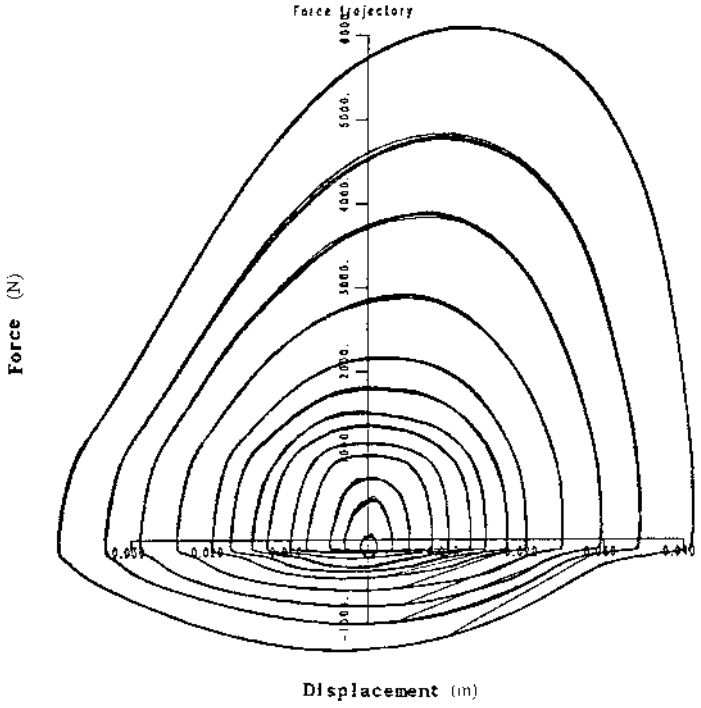


Figure 9.20. Work diagram for the **F** shock absorber: 10 Hz test.

stiffness component is indicated. A more interesting feature of the diagram is the presence of small discontinuities in the force at points near where the velocities change sign (i.e. near extrema in the displacement), this type of behaviour is very much reminiscent of the backlash characteristic described in [262]. It is associated with a pause in the motion of some types of absorber at the end of a stroke, where a small time is needed for pressures in the internal chambers to equalize before valves can open, flow can take place and the absorber can move again.

The restoring force surface and its associated contour map are shown in figures 9.26 and 9.27; in both cases the soft friction characteristic is clearly visible. The contours in figure 9.27 are parallel to the displacement axis at high velocities and inclined at an angle for low velocities, this confirms the observation made above that the stiffness forces are more significant at low velocities.

In summary, absorber **R** has a soft Coulomb damping characteristic. The stiffness characteristics change as the velocity increases, indicating a small displacement–velocity coupling.

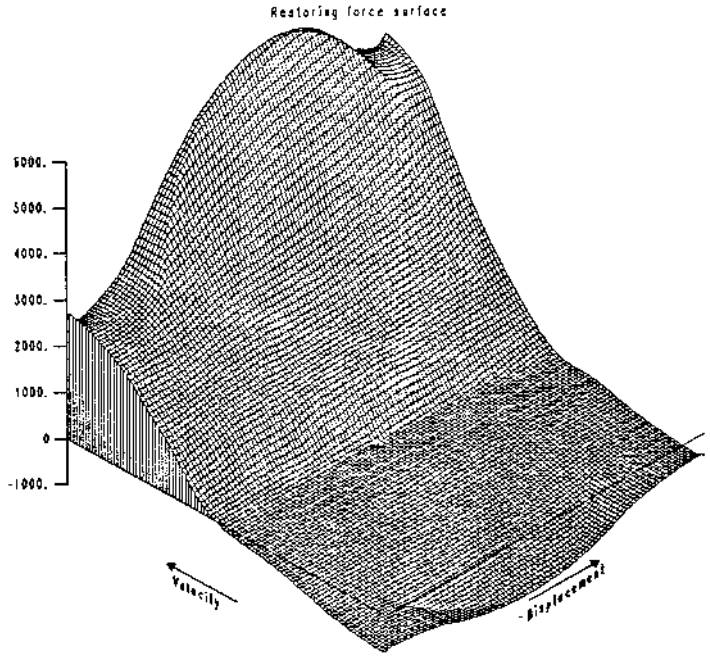


Figure 9.21. Estimated restoring force surface for the **F** shock absorber: 10 Hz test.

9.2.3 Polynomial modelling

As discussed in section 7.4.5, there are essentially two ways by which one might model the behaviour of a shock absorber. First, one might try to determine the equations of motion of the system using the laws of fluid dynamics, thermodynamics etc. Once the equations are known, one can identify the absorber by using measured data to estimate the coefficients of the terms in the equations. The second approach is to simply assume that the terms in the equations can be approximated by a heuristic model like the hyperbolic tangent or by a sum of polynomial terms say, and using measured data, estimate best-fit coefficients for the model. Following the latter approach means that one will learn nothing about the physics of the system but will obtain a model which will behave like the true system. This is clearly adequate for simulation purposes.

The complexity of the shock absorbers means that the first of these approaches is very difficult; Lang's physical model [157] mentioned in section 7.4.5 of this book required the specification of 87 independent parameters. Because of this, the parametric approach was adopted here. This is perfectly reasonable as the models are only required for simulation purposes. In fact the models were used as components in multi-body simulations of the response of a full automobile.

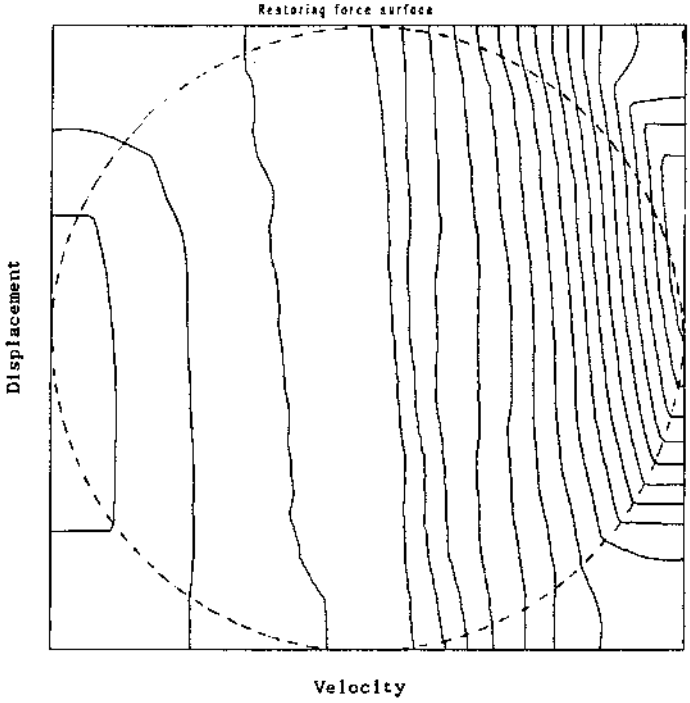


Figure 9.22. Contour map of the estimated restoring force surface for the **F** shock absorber: 10 Hz test.

The polynomial models fitted to the damping sections in section 7.6.2 are inadequate for the systems discussed here as there is non-trivial coupling between displacement and velocity. In order to overcome this a general multinomial model was used of the form

$$f_{\mathrm{m}}(y,\dot{y}) = \sum_{i=0}^{n}\sum_{j=0}^{m} a_{ij} y^i \dot{y}^j. \tag{9.30}$$

This shall be referred to as a (m,n)-model. The terms themselves have no physical meaning (except that a_{01} and a_{10} approximate the linear damping and stiffness at low levels of excitation), but their sum will reproduce the topography of the force surface.

The parameters here were obtained using a standard LS estimation scheme (chapter 6) based on the singular-value decomposition [209]. In each case, 2000 points each of displacement, velocity and force data were used. As each test file contained approximately 7000 points, about 5000 points were discarded. As the data were arranged in cycles of increasing amplitude, taking the first 2000 points would have restricted attention to low amplitude data, so the points in the files were 'randomized' before selecting points for the estimation procedure. It was

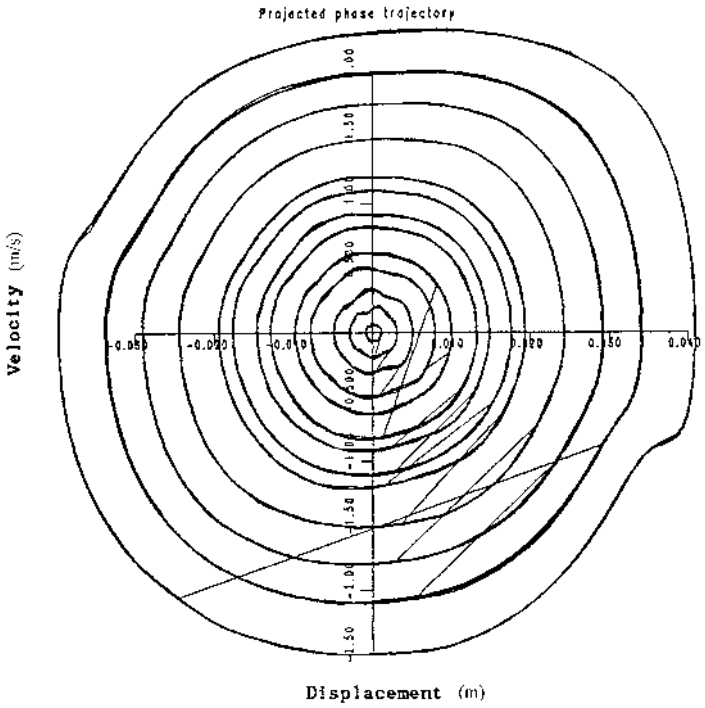


Figure 9.23. Phase trajectories from the **R** shock absorber after smoothing and velocity estimation: 10 Hz test.

decided to simply 'shuffle' the data in some systematic fashion. The method used was to apply 20 consecutive Farot shuffles [107] to the data. An advantage of shuffling rather than truly randomizing is that one need not store the original data; it can be reformed by performing a sequence of inverse shuffles. After shuffling, the first 2000 points in the file were selected.

Having obtained a model it is necessary to validate it. The most direct means of doing this is to use the measured displacement and velocity data to predict the force time-history which can then be compared with the true force values. This comparison was produced in the form of a plot of the shuffled data overlaid by the corresponding model data. An objective measure of the goodness of fit was provided by the normalized mean square error on the force or MSE($\hat{f}$) as defined in chapter 6. Significance testing of the coefficients was also carried out to simplify the models; the significance threshold for the models was set at 0.01%.

As a simple visual test of model validity, one can generate the restoring force surface corresponding to the model data. The simplest way is to use the model to evaluate the polynomial model function over the the original plotting grid and display that. Unfortunately, the model is *not* valid in the corners of the grid and

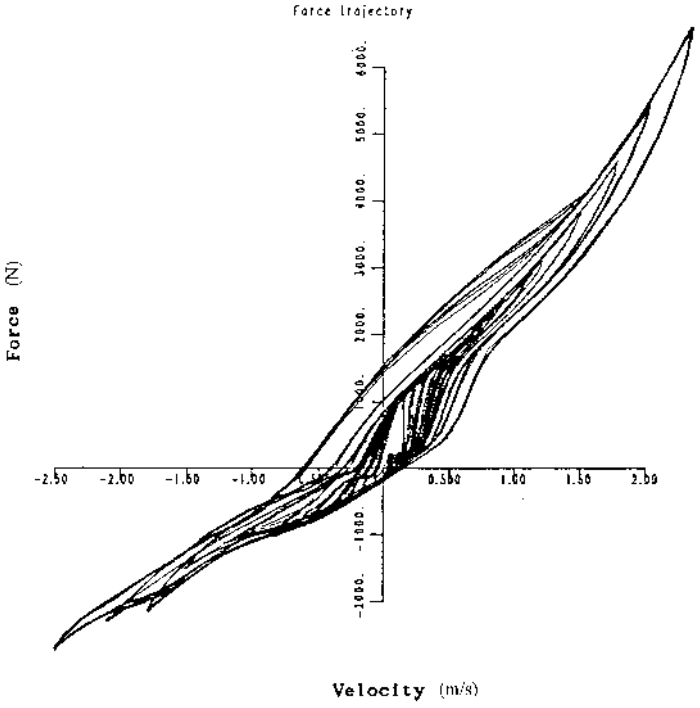


Figure 9.24. Characteristic diagram for the **R** shock absorber: 10 Hz test.

can give very misleading results in these regions. This phenomenon is illustrated via a couple of examples. It is deceptive to compare the surfaces over a zoomed region as one should always be aware of the input-sensitivity of the model.

The first model considered is that for the 1 Hz test of shock absorber **F**. The model fit to the force time-data is shown in figure 9.28. The model MSE of 2.06 indicates good agreement. The restoring force surface constructed from the data is compared with the surface defined by the model polynomial over the same region, in figure 9.29. The agreement appears to be reasonable although the surface constructed from the model polynomial suffers from distortion in the corner regions. The closeness of the two surfaces over the relevant portion of the phase plane is shown much more clearly in a comparison of the contour maps (figure 9.30).

9.2.4 Conclusions

The results obtained here allow a number of conclusions: first the restoring force surface approach provides a valuable means of characterizing the nonlinear behaviour of a shock absorber. The surface provides information in an easily

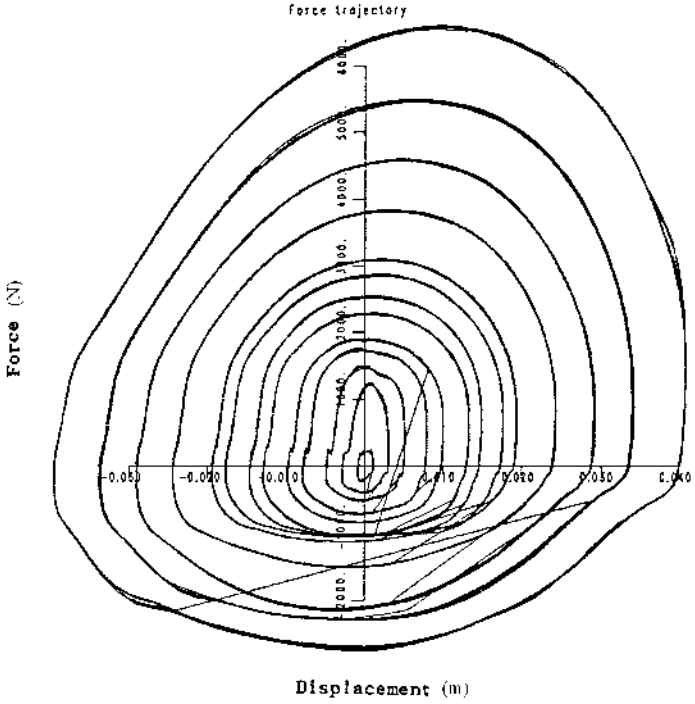


Figure 9.25. Work diagram for the **R** shock absorber: 10 Hz test.

interpretable form. Further as it does not depend on any *a priori* model structure it is truly non-parametric. The method is ideally suited to the systems as the experimental procedure allows major simplifications, i.e. SDOF behaviour, neglect of the inertia etc.

It is possible to fit parametric models to the data which, certainly for the shock absorbers considered here, capture most of the essential aspects of their behaviour. The models obtained are acceptable for use in simulation studies provided they are not used outside their range of validity.

A strategy based on differentiating to estimate the missing velocity data is shown to be effective. The main reason for this is that the acceleration data were not needed due to neglect of the inertia effects. If acceleration were needed, an integration-based procedure might possibly be preferred which avoids differentiating twice.

9.3 A bilinear beam rig

The motivation for this, the third of the demonstrators, was originally to construct a system with localized damage in order to benchmark fault detection algorithms.

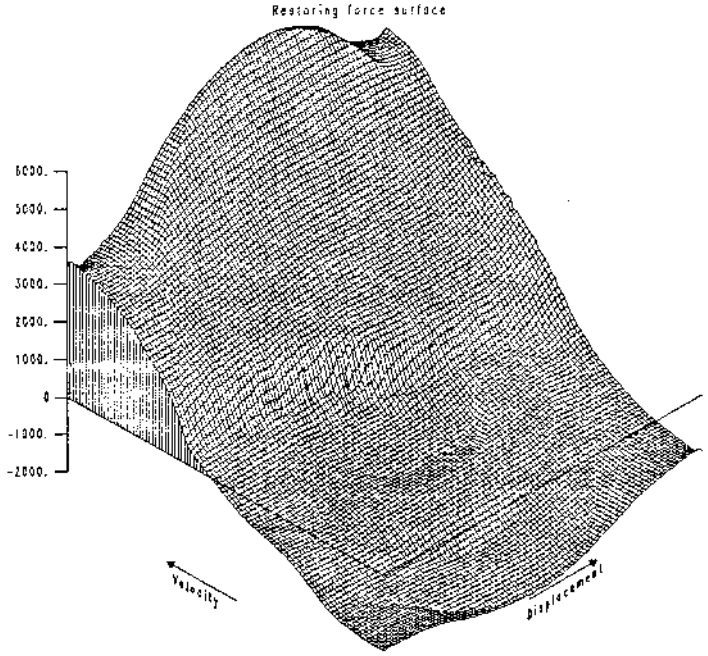


Figure 9.26. Estimated restoring force surface for the **R** shock absorber: 10 Hz test.

In practice, the system proved to show very interesting nonlinear behaviour and it is therefore appropriate to discuss it here. The system is fairly simple to construct and requires only basic equipment which is nonetheless sufficient to carry out various forms of analysis.

9.3.1 Design of the bilinear beam

In designing an experimental system for fault detection trials, it was considered undesirable to grow actual fatigue cracks in the structure, for reasons which will become clear later. The main assumption was that a sufficiently narrow gap should provide an adequate representation of a crack for the purposes of dynamics. A simple beam with rectangular cross-section was adopted as the underlying structure. For simplicity, free–free conditions were chosen for the dynamic tests described here. The beam was suspended from elastic supports attached at the nodal points for the first (flexural) mode of vibration (figure 9.31). Essentially the same instrumentation was needed as that in section 9.1.

The beam was fabricated from aluminium with the dimensions 1 m long by 50 mm wide by 12 mm deep. As it was intended that techniques would ultimately be developed for the *location* of faults within structures, it was decided that the gap would be introduced as part of a removable element which could be fixed to

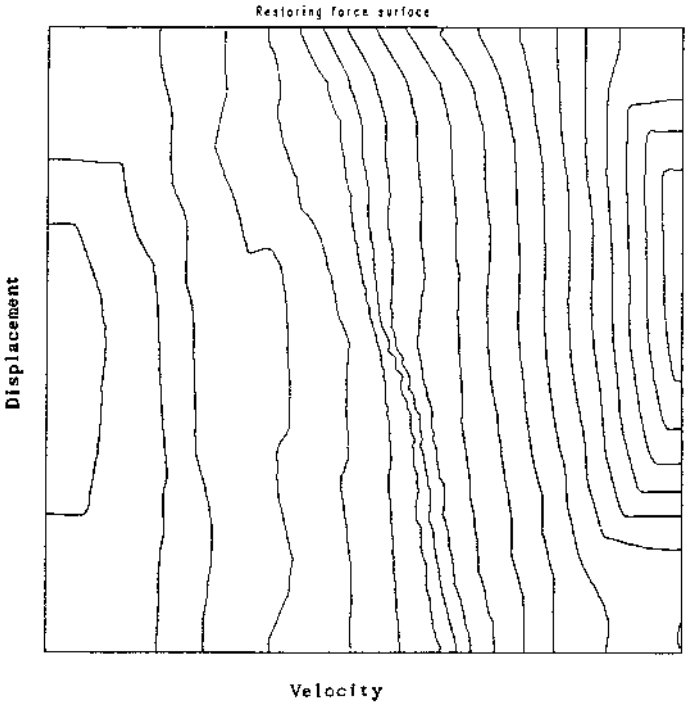


Figure 9.27. Contour map of the estimated restoring force surface for the **R** shock absorber: 10 Hz test.

the beam. For the purposes of nonlinear analysis, this is also desirable as the gap size can also be changed without difficulty.

Designing the gap element was not a trivial exercise and various structures were tried. The first gap element considered was that shown in figure 9.32. Two steel inserts were fixed by screws into a rectangular channel; the depth of the channel was initially half the depth of the beam.

With a gap of 0.05 mm, a static deflection test was carried out with the beam simply supported at the nodal lines and loaded centrally; the results are shown in figure 9.33. As required, the system showed a bilinear characteristic; however, the ratio of the stiffnesses was not very high at 1.25. An investigation was carried out into methods of increasing this ratio. In the following discussion, the stiffness when the gap is closed is denoted by k_2 and denoted by k_1 if the gap is open. If k_2 is assumed to be fixed by the material and geometry, the only way to increase the stiffness ratio $k' = k_2/k_1$ is to decrease k_1. An elementary calculation based

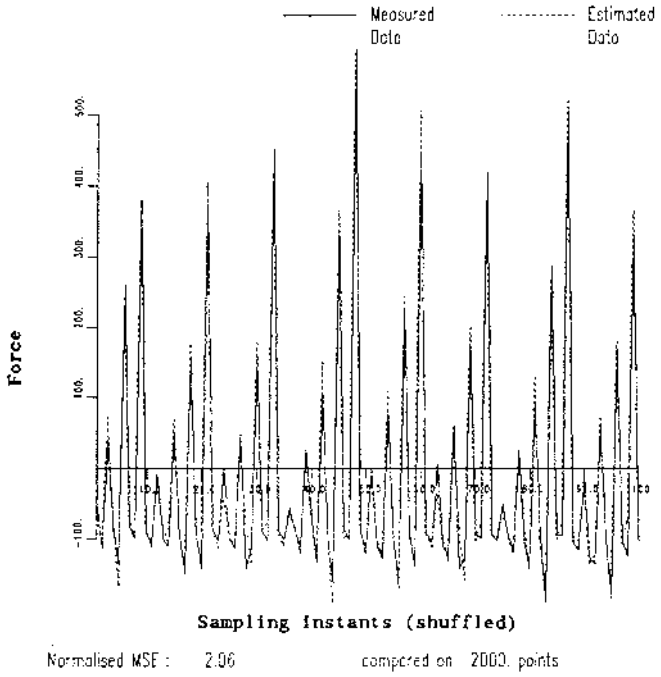


Figure 9.28. Comparison between measured force data and model data for the **F** absorber: 1 Hz test.

on the theory of elasticity yields the following expression for k_1:

$$\frac{1}{k_1} = \frac{1}{8EI}\left[\frac{L^3}{3} + \left(\frac{d}{d_1}\right)^3 L^2\delta\right] \tag{9.31}$$

where L is the length of the beam. Note that this is independent of t.

In order to reduce the stiffness ratio, d_1 must be decreased or δ must be increased. It is not practical to reduce d_1 as this would cause the beam to fail very quickly due to fatigue. Unfortunately, if δ is increased, a *trilinear* stiffness characteristic is obtained because the inserts bend when the gap closes but not when it opens (figure 9.34). A static load/deflection test for the beam with $\delta = 46$ mm gave the results show in figure 9.35. The assumption that k_2 is constant also fails since the inserts are not bonded to the beam.

In order to obtain a sufficiently high value for δ, the configuration shown in figure 9.36 was adopted. Elementary beam theory calculations gave the equations

$$\frac{1}{k_1} = \frac{L^3}{Ebd^3}\left(\frac{1}{2} + 12\frac{\delta}{L}\right) + \mathrm{O}\left[\left(\frac{\delta}{L}\right)^2\right] \tag{9.32}$$

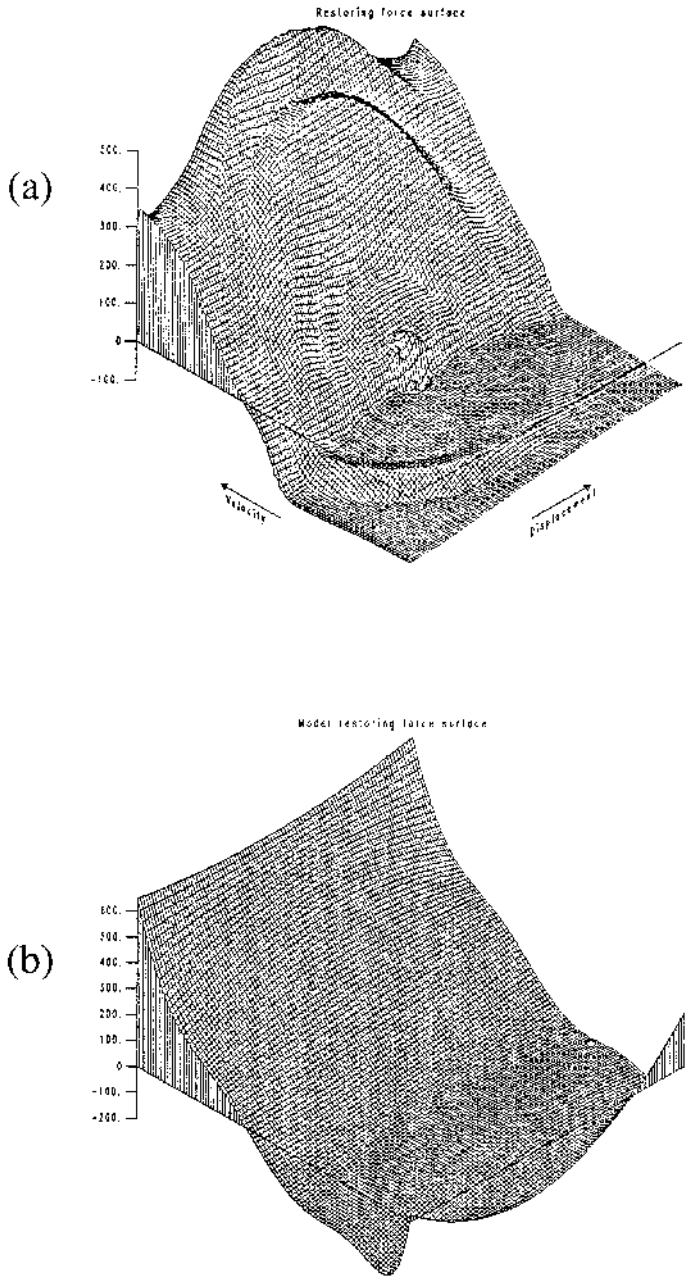


Figure 9.29. Comparison between estimated (*a*) force surface and (*b*) model surface, for the **F** absorber: 1 Hz test.

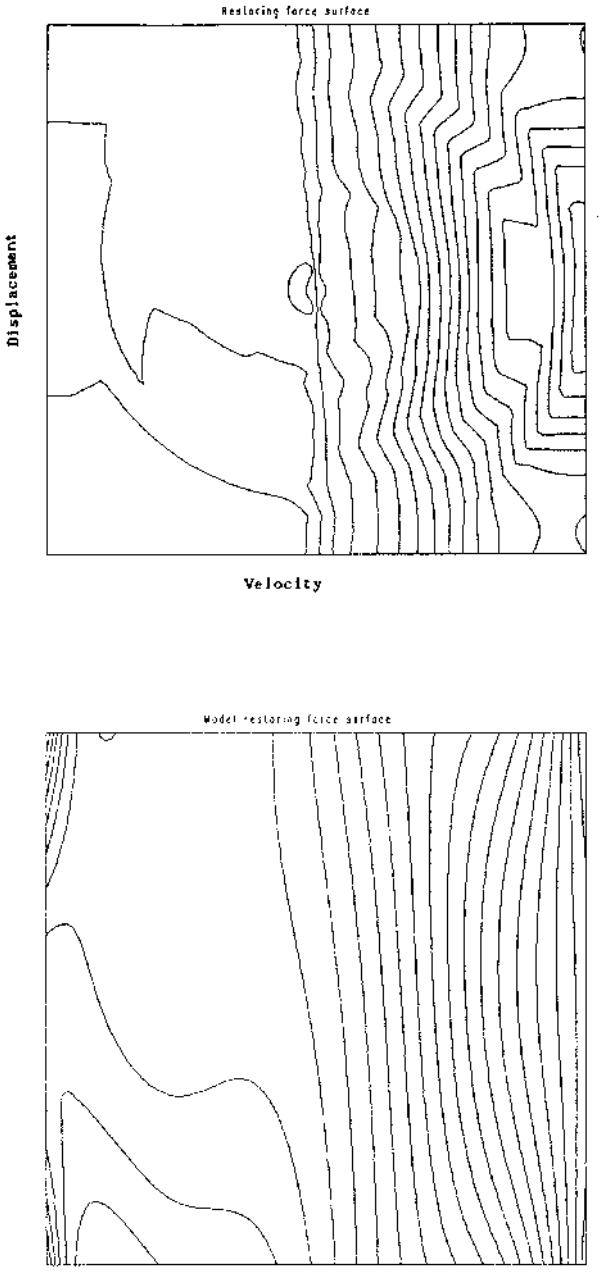


Figure 9.30. Comparison between contour maps of the estimated (*a*) force surface and (*b*) model surface, for the **F** absorber: 1 Hz test.

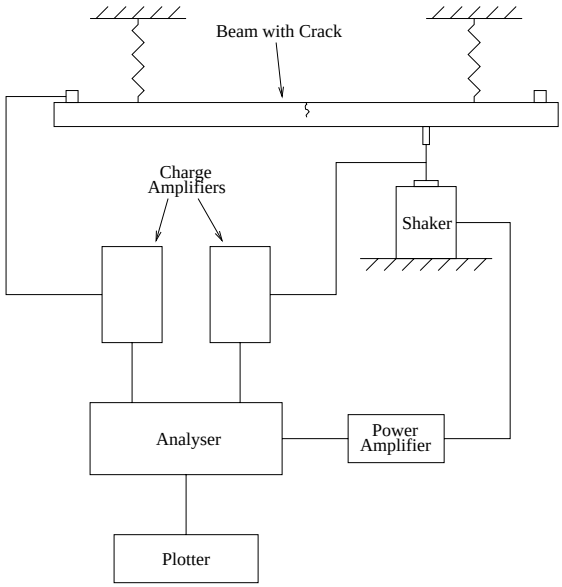


Figure 9.31. Bilinear beam experiment using a gap element to induce the nonlinearity.

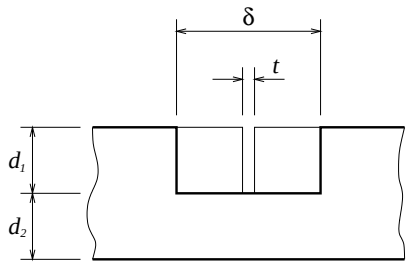


Figure 9.32. Arrangement of first gap element in beam.

$$\frac{1}{k_2} = \frac{L^3}{Ebd^3}\left(\frac{1}{2} + \frac{4}{3}\frac{\delta}{L}\right) + \mathrm{O}\left[\left(\frac{\delta}{L}\right)^2\right]. \tag{9.33}$$

To first order, the stiffness ratio is, therefore,

$$k' = \frac{\left(\frac{1}{2} + 12\frac{\delta}{L}\right)}{\left(\frac{1}{2} + \frac{4}{3}\frac{\delta}{L}\right)}. \tag{9.34}$$

Taking $\delta = 0.046$ m and $L = 0.885$ m gives a stiffness ratio of 1.96. This was considered high enough, so a beam and inserts were fabricated with these dimensions. A static load-deflection test on the resulting structure gave the results in figure 9.37; the bilinear characteristic is very clear. The actual value of k' for

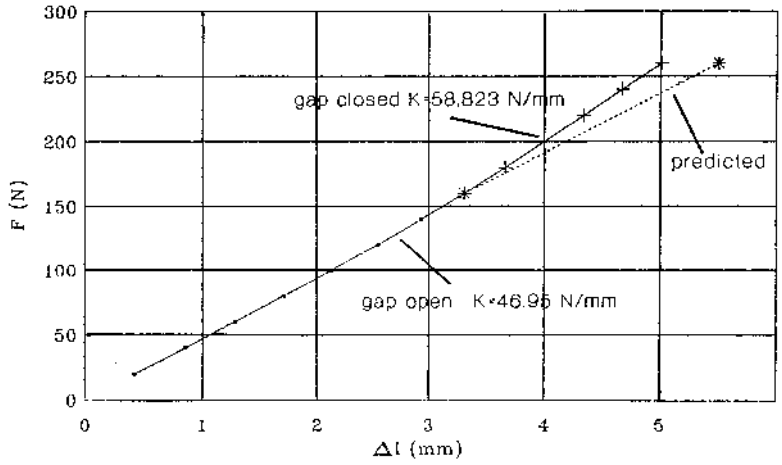


Figure 9.33. Static load-deflection curve for beam with first gap element: channel width 12 mm and gap 0.05 mm.

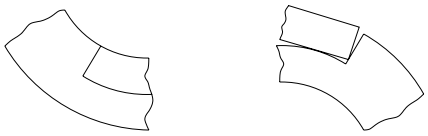


Figure 9.34. Bending of the inserts in the first gap element for positive and negative beam curvature.

the structure was 2.85; the second-order effects in equations (9.32) and (9.33) are clearly substantial.

The insert geometry of figure 9.36 is thus validated as a simple but effective means of introducing a bilinear stiffness characteristic in a beam. It remains to be established that such a gap geometry produces *dynamical* behaviour characteristic of a bilinear system or for that matter a genuine fatigue crack.

9.3.2 Frequency-domain characteristics of the bilinear beam

The experimental configuration for the dynamic analysis of the nonlinear beam is given in figure 9.31. Two accelerometers were placed at equal distances from the ends of the beam in order to maintain symmetry, only one was used for measurements.

In order to extract the frequency characteristics of the system, a stepped-sine excitation was used as discussed in section 9.1.2. Attention was restricted to the neighbourhood of the first (flexural) mode at 39 Hz. At each frequency in the range 20–60 Hz, the steady-state response amplitude was obtained. The

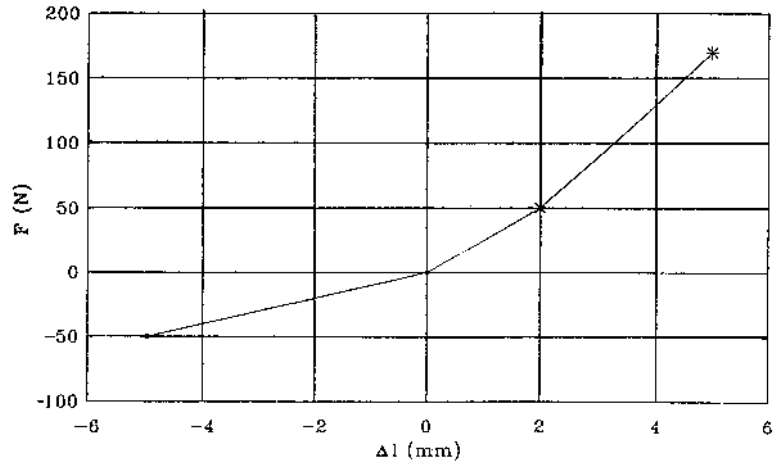


Figure 9.35. Static load-deflection curve for beam with first gap element: extended channel width.

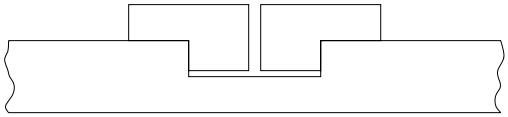


Figure 9.36. Modified arrangement of inserts.

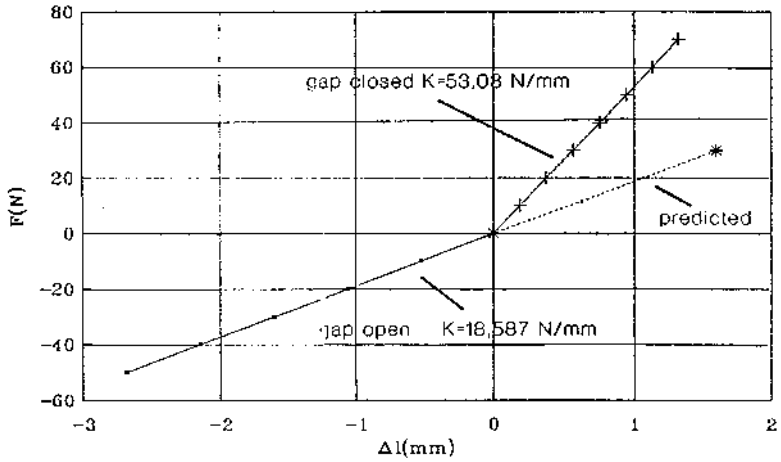


Figure 9.37. Static load-deflection curve of beam with modified inserts.

results are given in figures 9.38(*a*)–(*c*) for low, medium and high amplitudes of excitation. Note the discontinuities A and B in the FRFs in figures 9.38(*b*) and (*c*). At higher levels of excitation, as the frequency increases, the FRF initially follows the low-excitation (linear system) response curve. The discontinuity A occurs at the frequency where the response first causes the gap to close at that level of input force. The frequency of the feature A decreases with increasing input amplitude as the deflection required to close the gap is constant and can therefore be reached further away from the linear resonance. The second discontinuity is reminiscent of the 'jump phenomenon' encountered in the study of Duffing's equation (chapter 3); in fact a downward sweep through the frequencies for the same input amplitude shows the jump occurring at a lower frequency than might be expected. The sudden drop in the response causes the gap to cease closing and the response again follows the curve for the system at low excitation.

In order to compare the response of the bilinear beam to that of a truly cracked beam, a fatigue crack was generated in a beam of the same dimensions. A sharp notch was machined at one face of the beam to initiate crack growth and a hole was drilled through the centre of the beam to serve as a crack stop. Cyclic loading was applied until the required crack appeared.

The FRF for the cracked beam is given in figure 9.39. The results are very similar to those from the beam with the gap element. However, the change in gradient which gives rise to feature A is not discontinuous in this case. Where the gap element could be assumed to be fully closed or fully open at a given time, plastic deformation of the actual crack surfaces meant that it closed gradually, giving a continuous variation in the stiffness. The results are still very encouraging as they indicate that the gap element provides a good representation of a true fatigue crack even though the stiffness ratio is low in the latter case (≈ 1.25 as estimated earlier). The advantages of using the gap element are manifold. Most importantly, the element can easily be moved about the structure; also, any experimental results are much more repeatable—the act of growing the crack having introduced a uncontrolled plastic region into the beam. Finally, about 20 hr of cyclic loading were required before the crack appeared and the beam failed very shortly afterwards.

If excitation of higher modes than the first can be minimized in some way, it might be expected that a simple SDOF bilinear system could represent the behaviour of the bilinear beam. In this case the analysis of section 3.8.2 (see figure 3.9) is applicable and harmonic balance can be used to give the approximate form of the FRF:

$$\Lambda(\omega) = \frac{1}{k + \frac{(k'-k)}{\pi}\left\{\frac{\pi}{2} - \sin^{-1}\left(\frac{y_c}{Y}\right) - \frac{y_c}{Y}\sqrt{Y^2 - y_c^2}\right\} - m\omega^2 + \mathrm{i}c\omega} \tag{9.35}$$

If $Y > y_c$, with a simple linear FRF for $Y < y_c$.

The FRF is obtained by specifying an amplitude X and computing the corresponding Y for each ω over the range of interest. Figure 9.40 shows the

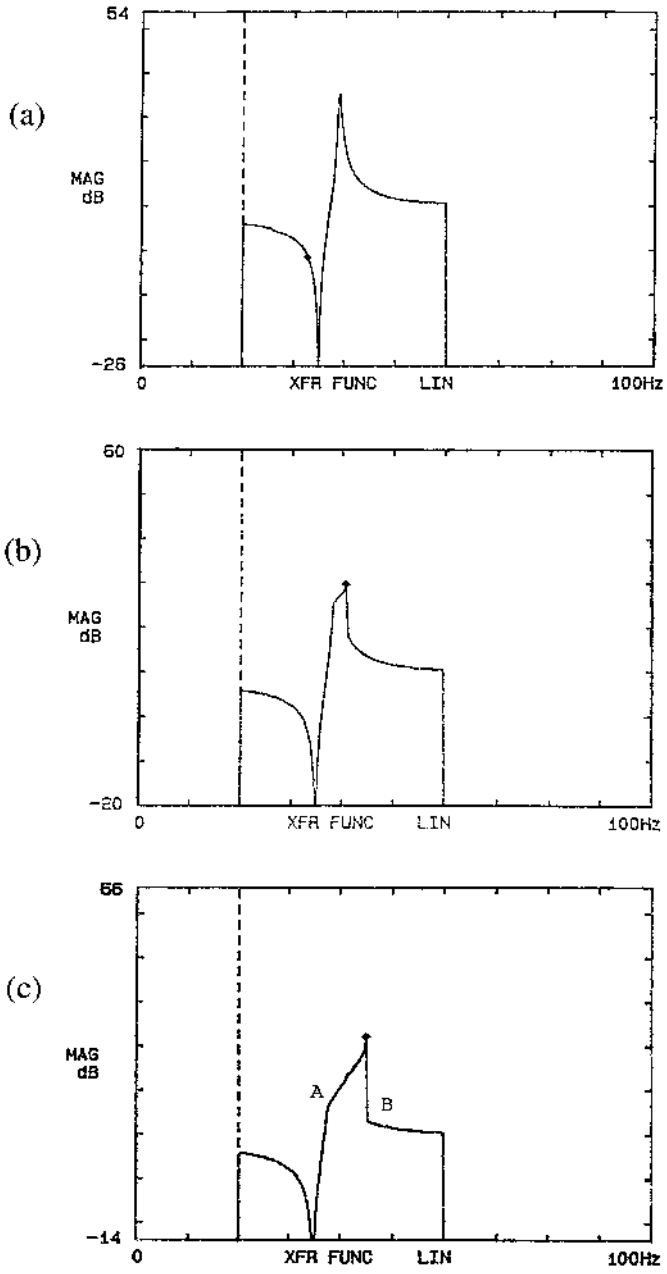


Figure 9.38. Frequency response of beam with gap element: (*a*) low amplitude; (*b*) moderate amplitude; (*c*) high amplitude.

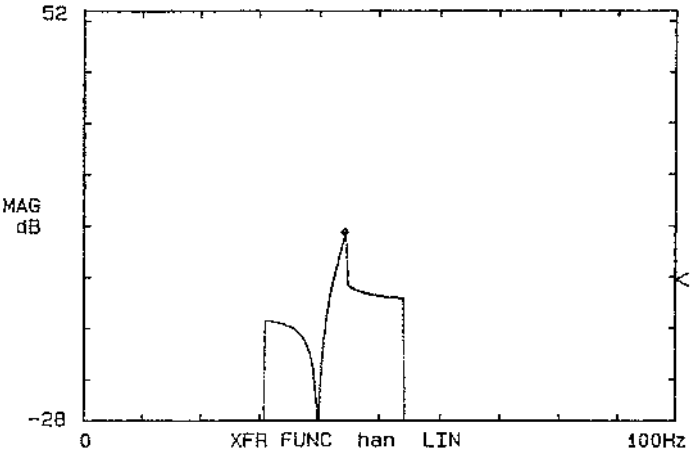


Figure 9.39. Frequency response of cracked beam at high amplitude.

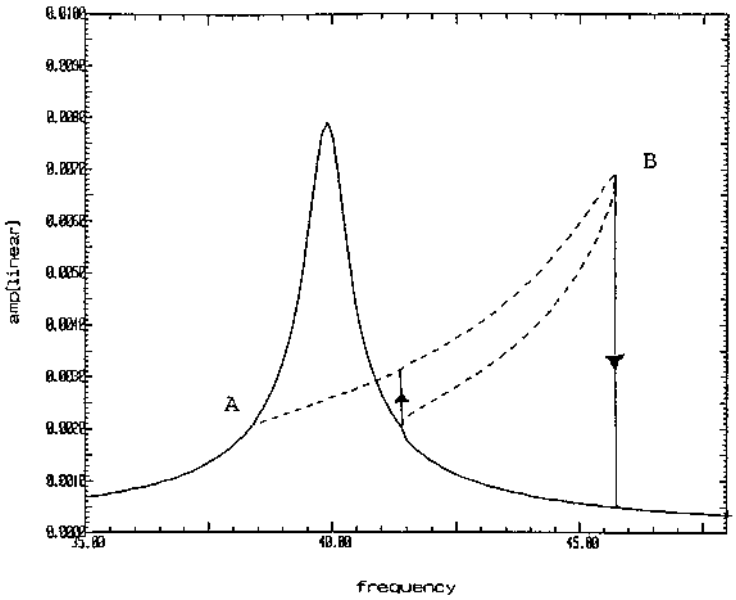


Figure 9.40. Frequency response at low and high amplitudes of an SDOF analytical model of bilinear stiffness with $k' = 2$.

computed FRF for a bilinear system with a stiffness ratio of 2 and a linear natural frequency of 40 Hz. For low values of X, the response is computed from the linear FRF alone, the result is shown by the solid line in figure 9.40. If a high value of X

is used such that the condition $Y > y_c$ is met, the dotted curve is obtained. Note that three solutions are possible over a certain range of frequencies; however, only the upper branch is stable. If the frequency sweeps up, the response follows the upper branch of the dotted curve until at point B this solution ceases to exist, the response drops down to the linear response curve. If the frequency sweeps down, the response follows the linear curve until the condition $Y = y_c$ is met (at the same height as feature A); after this the response follows the 'nonlinear' curve until the point A is reached.

It will be noted that the analytical FRF curve bears a remarkable resemblance to that from the beam with gap element (figure 9.38(*c*)). The results of this section establish the close correspondence in the frequency domain between a beam with a gap, a beam with a fatigue crack and an SDOF bilinear oscillator (if the first mode alone is excited for the beams). This justifies the experimental study of beams with gap elements for damage purposes also. Before proceeding to model the beam, the following section briefly considers the correspondence between these three systems in the time domain.

9.3.3 Time-domain characteristics of the bilinear beam

This section shows that the correspondence between the beams and SDOF bilinear system is also demonstrable in the time domain.

When excited with a harmonic excitation at low amplitude, all three systems responded with a sinusoid at the forcing frequency as expected. The behaviour at higher levels of excitation is more interesting.

First, the beam with the gap was harmonically excited at a frequency below the first (non-rigid) resonance. The resulting response signal is given in figure 9.41(*a*), note the substantial high-frequency component. A numerical simulation was carried out using an SDOF bilinear system with the same resonant frequency and excitation frequency; a fourth-order Runge–Kutta integration routine was used and the results are given in figure 9.41(*b*). The characteristics of the two traces are very similar (allowing for the fact that the two plots have opposite orientations and are scaled differently), the main difference is the high-frequency content in figure 9.41(*a*). It will be shown a little later that this component of the response is due to the nonlinear excitation of higher modes of vibration in the beam. Because the simulated system is SDOF, it can only generate a high-frequency component through harmonics and these are not sufficiently strong here.

For the second set of experiments the beams with a gap element and with a fatigue crack were harmonically excited at frequencies close to the first (non-rigid) resonance, the resulting responses are given in figures 9.42(*a*) and (*b*). Allowing for the orientation of the plots, the responses are very similar in form. In order to facilitate comparison, a low-pass filter has been applied in order to remove the high-frequency component which was visible in figure 9.41(*a*). When the simulated SDOF bilinear system was excited at a frequency close to its

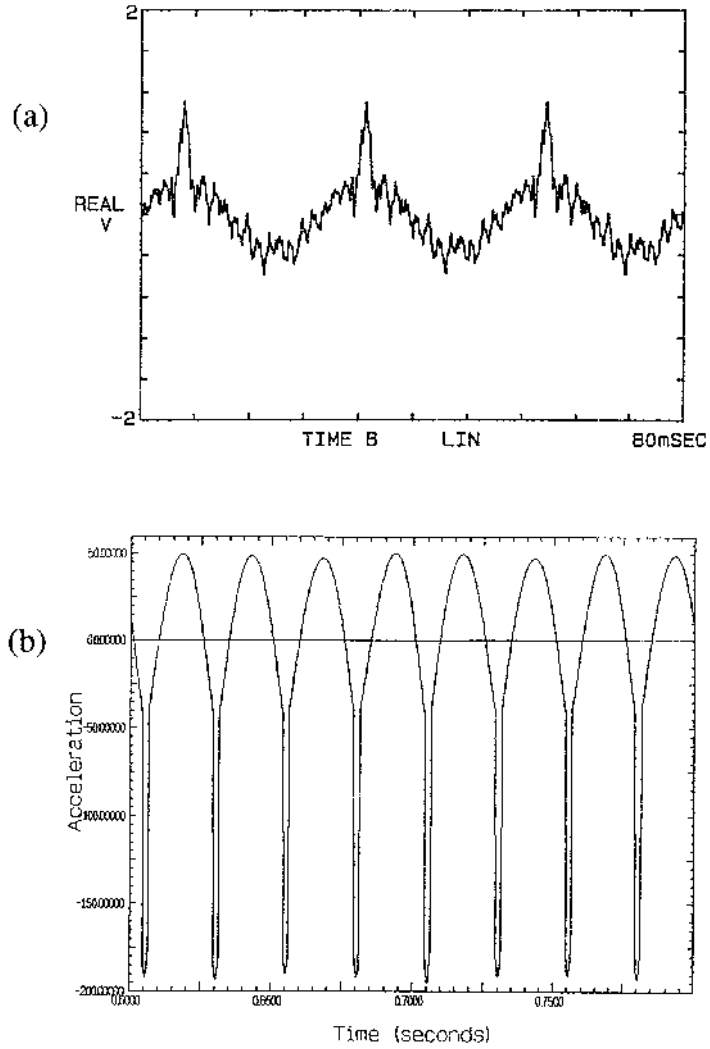


Figure 9.41. Time response of systems under harmonic excitation below resonance: (*a*) beam with gap element; (*b*) SDOF simulation.

resonance, the results shown in figure 9.42(*c*) were obtained. Again, disregarding the scaling and orientation of the plot, the results are very similar to those from the two experimental systems.

This study reinforces the conclusions drawn at the end of the previous section—there is a close correspondence between the responses of the three systems under examination. Another possible means of modelling the beams is provided

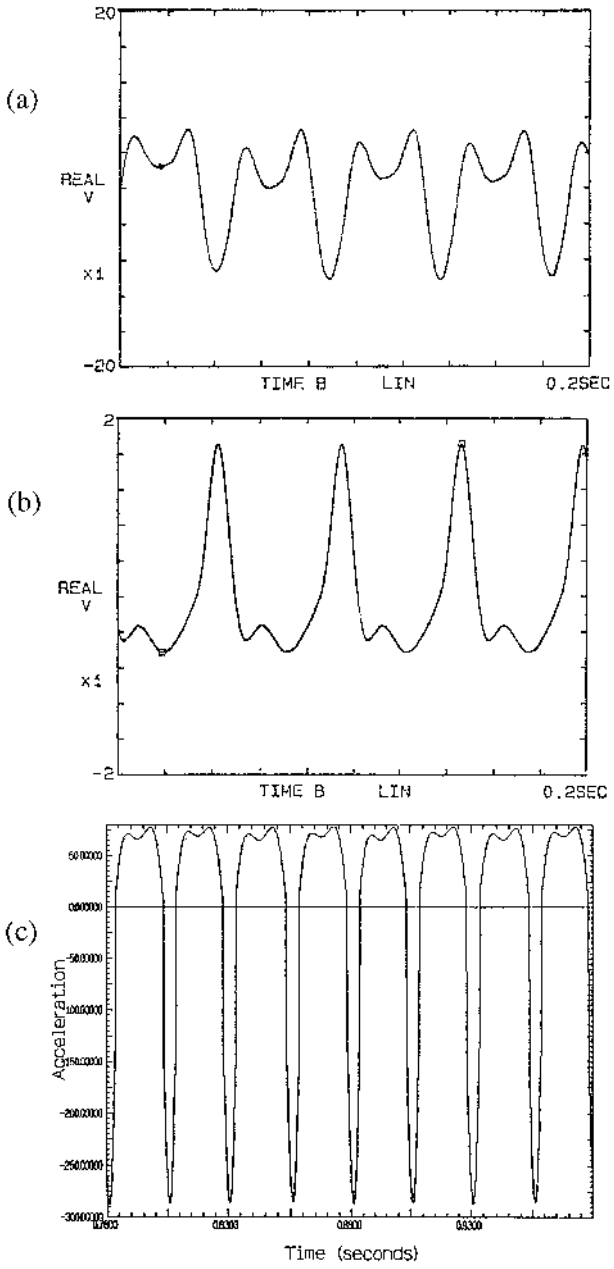


Figure 9.42. Time response of systems under harmonic excitation around resonance: (*a*) beam with gap element; (*b*) cracked beam; (*c*) SDOF simulation.

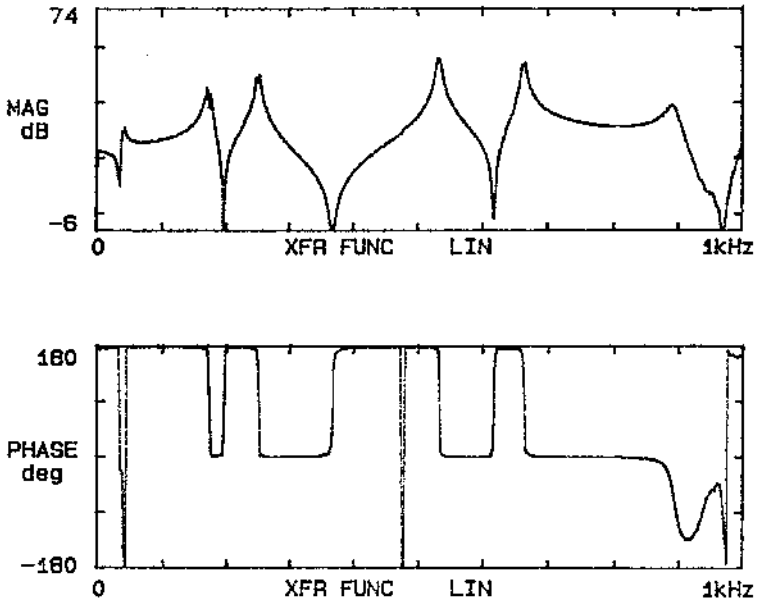


Figure 9.43. FRF of beam with gap element under low-level random excitation.

by finite element analysis and a number of preliminary results are discussed in [220].

9.3.4 Internal resonance

An attempt was made to generate SDOF behaviour by exciting the bilinear beam (with gap) with a band-limiting random force centred on a single mode. First, the beam was excited by a broadband random signal at a low enough level to avoid exciting the nonlinearity. The resulting FRF is given in figure 9.43; the first three natural frequencies were 42.5, 175 and 253 Hz. As the system has free–free boundary conditions it has rigid-body modes which should properly be called the first modes; for convenience it is adopted as a convention that numbering will begin with the first non-rigid-body mode. The rigid-body modes are not visible in the accelerance FRF in figure 9.43 as they are strongly weighted out of the acceleration response. However, their presence is signalled by the anti-resonance before the 'first' mode at 42.5 Hz.

When the system is excited at its first natural frequency by a sinusoid at low amplitude, the acceleration response is a perfect sinusoid. The corresponding response spectrum is a single line at the response frequency, confirming that the system is behaving linearly. When the excitation level is increased to the point when the gap closes during a forcing cycle, the response is far from sinusoidal,

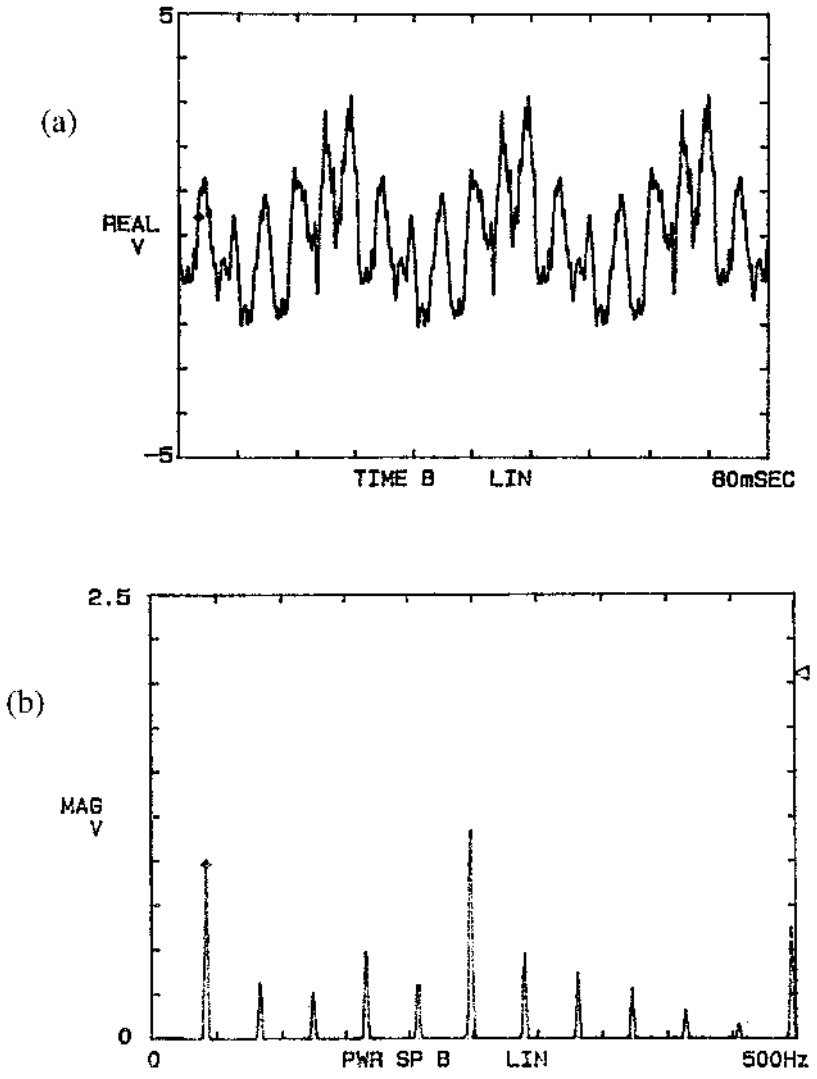


Figure 9.44. Acceleration response of beam with gap element under high level harmonic excitation: (*a*) time response; (*b*) response spectrum.

as shown in figure 9.44(*a*). The higher harmonic content of the response is considerable and this is clearly visible in the spectrum given in figure 9.44(*b*). At first it appears a little unusual that the component at the sixth harmonic is *stronger* than the fundamental component; however, this is explicable in terms of the MDOF nature of the beam.

Note that the second natural frequency is close to four times the first ($175 \approx 4 \times 42.5 = 170$), while the third is nearly six times the first ($253 \approx 6 \times 42.5 = 255$); this has rather interesting consequences. Exciting the system with a band-limited input centred about the first natural frequency was supposed to elicit an effectively SDOF response in order to compare with SDOF simulation and allow a simple model. This argument depended on the harmonics in the response of the nonlinear system not coinciding with the resonances of the underlying linear system. In such a situation, 'internal resonances' can occur where energy is transferred between resonant frequencies. The standard analysis of such resonances has been discussed in many textbooks and monographs, e.g. [222] and will not be repeated here.

The bilinear system discussed here is capable of behaviour characteristic of weak or strong behaviour depending on the excitation. In fact, internal resonances can occur even under conditions of weak nonlinearity; a simple argument based on the Volterra series can provide some insight.

As described in chapter 8, the magnitude of the fundamental response is largely governed by the size of $H_1(\Omega)$. $H_1(\Omega)$ is simply the FRF of the underlying linear system and is well known to have an expansion of the form:

$$H_1(\Omega) = \frac{\prod_{j=1}^{n_z}(\Omega - \omega_{zj})}{\prod_{j=1}^{n_p}(\Omega - \omega_{pj})} \tag{9.36}$$

where n_z is the number of *zeroes* ω_{zj} and n_p is the number of *poles* ω_{pj}. It is, of course, the poles which generate the maxima or resonances in the FRF; if the forcing frequency is near ω_{pi} say, $H_1(\Omega)$ is large and the response component at Ω is correspondingly large. Similarly, if $H_n(\Omega, \ldots, \Omega)$ is large, there will be a large output component at the nth harmonic $n\Omega$. It can be shown for a range of structural systems that

$$H_n(\Omega, \ldots, \Omega) = f[H_1(\omega), H_1(2\Omega), \ldots, H_1((n-1)\Omega)]H_1(n\Omega) \tag{9.37}$$

where the function f depends on the particular nonlinear system (see equation (8.216) for an example for H_2). This means that if $n\Omega$ is close to *any* of the poles of H_1, H_n will be large and there will be a correspondingly large output at the harmonic $n\Omega$. In general *all* harmonics will be present in the response of nonlinear systems, notable exceptions to this rule are systems with symmetric nonlinearities for which all *even*-order FRFs vanish.

This is how a spectrum like that in figure 9.44(*b*) might occur. Consider the component at the sixth harmonic; it has already been remarked that six times the first natural frequency of the system is close to the third natural frequency. If the excitation is at ω_{p1}, i.e. at the first resonance, then $H_1(6\omega_{p1}) \approx H_1(\omega_{p3})$ will be large and so therefore will $H_6(\omega_{p1}, \ldots, \omega_{p1})$; a correspondingly large component will be observed in the output at sixth harmonic. This can be regarded as a purely *nonlinear* excitation of the third natural frequency. A similar argument applies to the fourth harmonic in figure 9.44(*b*); this is elevated because it coincides with

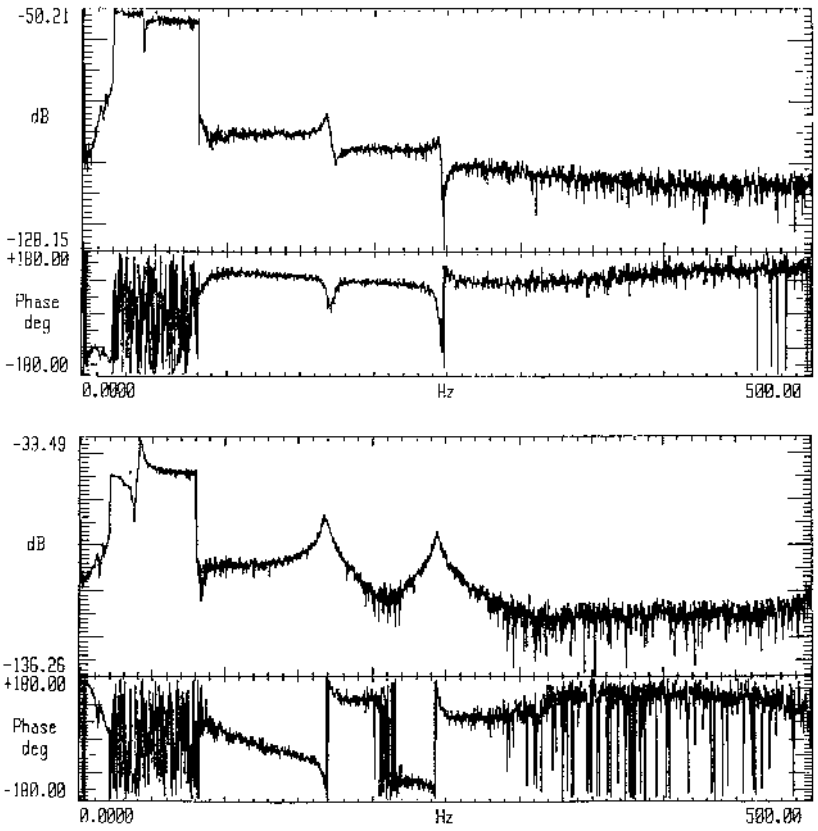


Figure 9.45. Spectra from beam under low level random excitation: (*a*) force; (*b*) acceleration.

the second natural frequency of the system, i.e. $H_1(4\omega_{\mathrm{p}1}) \approx H_1(\omega_{\mathrm{p}2})$. Because of these nonlinear effects the beam system cannot be regarded as an SDOF with a resonance at the first natural frequency even if a harmonic excitation is used because energy is always transferred to the higher modes of the system. It is impossible to circumvent this by exciting in an interval around the second natural frequency. The reason is that the centre of the beam where the gap element is located is situated at a node of the second mode, and it is impossible to cause the gap to close. An excitation band-limited around the first natural frequency was therefore selected in the knowledge that this might cause difficulties for subsequent SDOF modelling.

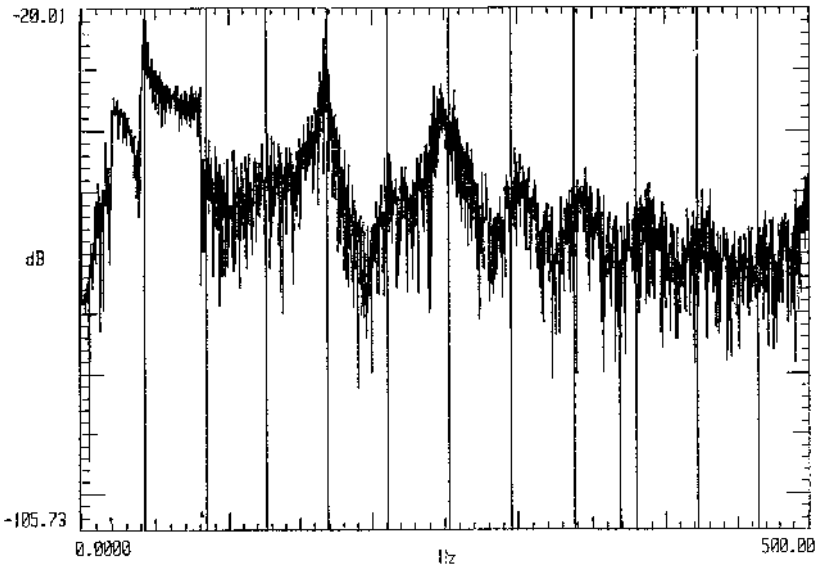


Figure 9.46. Acceleration spectrum from beam under high level random excitation showing energy transfer to the second and third modes.

9.3.5 A neural network NARX model

The system was first excited with a band-limited random signal in the range 20–80 Hz at a low level which did not cause the gap to close. Sampling was carried out at 2000 Hz and the force and acceleration data were saved for identification. The input and output spectra for the system are given in figure 9.45. Excitation of the second and third modes is clearly minimal. The level of excitation was then increased up to the point where the gap closed frequently and the high-frequency content increased visibly. Figure 9.46 shows the acceleration spectrum marked with multiples of the first resonance and it shows clearly the nonlinear energy transfer to the second and third modes at fourth and sixth harmonic of the first mode.

There is a problem with the usual identification strategy in that it uses force and displacement data. If the acceleration data are twice integrated, it is necessary to use a high-pass filter to remove integration noise as discussed in chapter 7. However, this clearly removes any d.c. component which should be present in the displacement if the restoring force has even components. This is the case with the bilinear system. Because of this, a model was fitted to the force–acceleration process. The model selected was a neural network NARX model as described in chapter 6.

As the data were oversampled, it was subsampled by a factor of 12, yielding

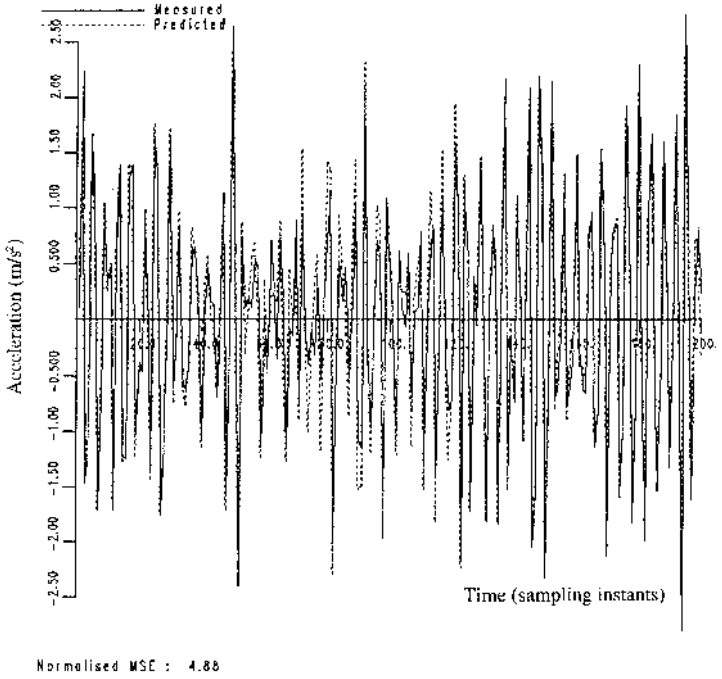


Figure 9.47. Comparison between measured acceleration data and that predicted by the neural network NARX model for high level random excitation.

a sampling frequency of 167 Hz; 598 input–output pairs were obtained. The network was trained to output the current acceleration when presented with the last four sampled forces and accelerations. The network converged on a model and a comparison between the system output and that predicted by the network is given in figure 9.47, here the MSE of 4.88 is respectable. Some minor improvement was observed on using modelling networks trained with a wider range of lagged forces and accelerations (i.e. six of each and eight of each). The improvements did not justify the added complexity of the networks.

9.4 Conclusions

The three systems described in this chapter can be used to illustrate a broad range of nonlinear behaviours. In particular, the beam rigs are extremely simple to construct and require only instrumentation which should be found in any dynamics laboratory. In the course of discussing these systems, essentially all of the techniques described in earlier chapters have been illustrated; namely

- harmonic distortion (chapters 2 and 3),

- FRF distortion (chapters 2 and 3),
- Hilbert transforms (chapters 4 and 5),
- NARX models and neural networks (chapter 6),
- restoring force surfaces (chapter 7) and
- Volterra series and HFRFs (chapter 8).

It will hopefully be clear to the reader that the experimental and analytical study of nonlinear systems is not an arcane discipline, but an essential extension of standard linear vibration analysis well within the reach of all with access to basic dynamics equipment and instrumentation. The conclusions of chapter 1 introduced the idea of a 'toolbox' for the analysis of nonlinear structural systems. Hopefully this book will have convinced the reader that the toolbox is far from empty. Some of the techniques discussed here will stand the test of time, while others will be superseded by more powerful methods—the subject of nonlinear dynamics is continually evolving. In the introduction it was suggested that structural dynamicists working largely with linear techniques should at least be informed about the presence and possible consequences of nonlinearity. This book will hopefully have placed appropriate methods within their reach.

Chapter 6

System identification—discrete time

6.1 Introduction

One can regard dynamics in abstract terms as the study of certain sets. For example: for single-input–single-output (SISO) systems, the set is composed of three objects; $D = \{x(t), y(t), S[\,]\}$ where $x(t)$ is regarded as a stimulus or input function of time, $y(t)$ is a response or output function and $S[\,]$ is a functional which maps $x(t)$ to $y(t)$ (figure 6.1 shows the standard diagrammatic form). In fact, there is redundancy in this object; given any two members of the set, it is possible, in principle, to determine the third member. This simple fact serves to generate almost all problems of interest in structural dynamics, they fall into three classes:

Simulation. Given $x(t)$ and an appropriate description of $S[\,]$ (i.e. a differential equation if x is given as a function; a difference equation if x is given as a vector of sampled points), construct $y(t)$. The solution of this problem is not trivial. However, in analytical terms, the solution of differential equations, for example, is the subject of innumerable texts, and will not be discussed in detail here, [227] is a good introduction. If the problem must be solved numerically, [209] is an excellent reference.

Deconvolution. Given $y(t)$ and an appropriate description of $S[\,]$, construct $x(t)$. This is a so-called *inverse problem of the first kind* [195] and is subject to numerous technical difficulties even for linear systems. Most importantly, the solution will not generally be unique and the problem will often be ill-posed in other senses. The problem is not discussed any further here, the reader can refer to a number of works, [18, 242, 246] for further information.

System Identification. Given $x(t)$ and $y(t)$, construct an appropriate representation of $S[\,]$. This is the *inverse problem of the second kind* and forms the subject of this chapter and the one that follows. Enough basic theory will be presented to allow the reader to implement a number of basic strategies.

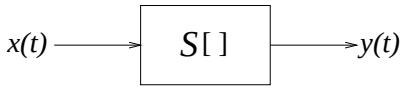


Figure 6.1. Standard block diagram representation of single-input single-output (SISO) system.

There are a number of texts on system identification which can be consulted for supporting detail: [167, 231, 168] are excellent examples.

To expand a little on the definition of system identification, consider a given physical system which responds in some measurable way $y_{\mathrm{s}}(t)$ when an external stimulus or excitation $x(t)$ is applied, a mathematical model of the system is required which responds with an identical output $y_{\mathrm{m}}(t)$ when presented with the same stimulus. The model will generally be some functional which maps the input $x(t)$ to the output $y_{\mathrm{m}}(t)$.

$$y_{\mathrm{m}}(t) = S[x](t). \tag{6.1}$$

If the model changes when the frequency or amplitude characteristics of the excitation change, it is said to be *input-dependent*. Such models are unsatisfactory in that they may have very limited predictive capabilities.

The problem of system identification is therefore to obtain an appropriate functional $S[\,]$ for a given system. If *a priori* information about the system is available, the complexity of the problem can be reduced considerably. For example, suppose that the system is known to be a continuous-time linear single degree-of-freedom dynamical system; in this case the form of the equation relating the input $x(t)$ and the response $y(t)$ is known to be (the subscripts on y will be omitted where the meaning is clear from the context)

$$m\ddot{y} + c\dot{y} + ky = x(t). \tag{6.2}$$

In this case the implicit structure of the functional $S[\,]$ is known and the only unknowns are the coefficients or parameters m, c and k; the problem has been reduced to one of *parameter estimation*. Alternatively, rewriting equation (6.2) as

$$(Ly)(t) = x(t) \tag{6.3}$$

where L is a second-order linear differential operator, the solution can be written as

$$y(t) = (L^{-1}x)(t) = \int \mathrm{d}\tau\, h(\tau - t)x(\tau) \tag{6.4}$$

which explicitly displays $y(t)$ as a *linear* functional of $x(t)$. Within this framework, the system is identified by obtaining a representation of the function $h(t)$ which has been introduced in earlier chapters as the impulse response or

Green's function for the system. It has also been established that in structural dynamics, $h(t)$ is usually obtained via its Fourier transform $H(\omega)$ which is the system transfer function

$$H(\omega) = \frac{Y(\omega)}{X(\omega)} \tag{6.5}$$

where $X(\omega)$ and $Y(\omega)$ are the Fourier transforms of $x(t)$ and $y(t)$ respectively and $H(\omega)$ is the standard

$$H(\omega) = \frac{1}{-m\omega^2 + \mathrm{i}c\omega + k} \tag{6.6}$$

and $H(\omega)$ is completely determined by the three parameters m, c and k as expected. This striking duality between the time- and frequency-domain representations for a linear system means that there are a number of approaches to linear system identification based in the different domains. In fact, the duality extends naturally to nonlinear systems where the analogues of both the impulse response and transfer functions can be defined. This representation of nonlinear systems, and its implications for nonlinear system identification, will be discussed in considerable detail in chapter 8.

6.2 Linear discrete-time models

It is assumed throughout the following discussions that the structure detection problem has been reduced to the selection of a number of terms linear in the unknown parameters[1]. This reduces the problem to one of parameter estimation and in this particular case allows a solution by well-known *least-squares* methods. A discussion of the mathematical details of the parameter estimation algorithm is deferred until a little later; the main requirement is that measured time data should be available for each term in the model equation which has been assigned a parameter. In the case of equation (6.2), records are needed of displacement $y(t)$, velocity $\dot{y}(t)$, acceleration $\ddot{y}(t)$ and force $x(t)$ in order to estimate the parameters. From the point of view of an experimenter who would require considerable instrumentation to acquire the data, a simpler approach is to adopt the discrete-time representation of equation (6.2) as discussed in chapter 1. If the input force and output displacement signals are sampled at regular intervals of time Δt, records of data $x_i = x(i\Delta t)$ and $y_i = y(i\Delta t)$ are obtained for $i = 1, \ldots, N$ and are related by equation (1.67):

$$y_i = a_1 y_{i-1} + a_2 y_{i-2} + b_1 x_{i-1}. \tag{6.7}$$

This linear difference equation is only one of the possible discrete-time representations of the system in equation (6.2). The fact that it is not unique

[1] Note the important fact that the model being linear in the parameters in no way restricts the approach to *linear systems*. The majority of all the nonlinear systems discussed so far are linear in the parameters.

is a consequence of the fact that there are many different discrete representations of the derivatives. The discrete form (6.7) provides a representation which is as accurate as the approximations (1.64) and (1.65) used in its derivation. In the time series literature this type of model is termed 'Auto-Regressive with eXogenous inputs' (ARX). To recap, the term 'auto-regressive' refers to the fact that the present output value is partly determined by or regressed on previous output values. The regression on past input values is indicated by the words 'exogenous inputs' (the term exogenous arose originally in the literature of econometrics, as did much of the taxonomy of time-series models)[2].

Through the discretization process, the input–output functional of equation (6.1) has become a linear input–output function with the form

$$y_i = F(y_{i-1}, y_{i-2}; x_{i-1}). \tag{6.8}$$

The advantage of adopting this form is that only the two states x and y need be measured in order to estimate all the model parameters a_1, a_2 and b_1 in (6.7) and thus identify the system. Assuming that the derivatives are all approximated by discrete forms similar to equations (1.64) and (1.65), it is straightforward to show that a general linear system has a discrete-time representation

$$y_i = \sum_{j=1}^{n_y} a_j y_{i-j} + \sum_{j=1}^{n_x} b_j x_{i-j} \tag{6.9}$$

or

$$y_i = F(y_{i-1}, \ldots, y_{i-n_y}; x_{i-1}, \ldots, x_{i-n_x}). \tag{6.10}$$

As before, all the model parameters $a_1, \ldots, a_{n_y}, b_1, \ldots, b_{n_x}$ can be estimated using measurements of the x and y data only. The estimation problem is discussed in the following section.

6.3 Simple least-squares methods

6.3.1 Parameter estimation

Having described the basic structure of the ARX model, the object of the present section is to give a brief description of the least-squares methods which can be used to estimate the model parameters. Suppose a model of the form (6.7) is required for a set of measured input and output data $\{x_i, y_i; i = 1, \ldots, N\}$. Taking measurement noise into account one has

$$y_i = a_1 y_{i-1} + a_2 y_{i-2} + b_1 x_{i-1} + \zeta_i \tag{6.11}$$

[2] Note that there is a small contradiction with the discussion of chapter 1. There the term 'moving-average' was used to refer to the regression on past inputs. In fact, the term is more properly used when a variable is regressed on past samples of a noise signal. This convention is adopted in the following. The AR part of the model is the regression on past outputs y, the X part is the regression on the measured *eXogenous* inputs x and the MA part is the regression on the unmeasurable noise states ζ. Models containing only the deterministic x and y terms are therefore referred to as ARX.

where the residual signal ζ_i is assumed to contain the output noise and an error component due to the fact that the parameter estimates may be incorrect. (The structure of the signal ζ is critical to the analysis; however, the discussion is postponed until later in the chapter.) The *least-squares* estimator finds the set of parameter estimates which minimizes the error function

$$J = \sum_{i=1}^{N} \zeta_i^2. \tag{6.12}$$

The parameter estimates obtained will hopefully reduce the residual sequence to measurement noise only.

The problem is best expressed in terms of matrices. Assembling each equation of the form (6.7) for $i = 3, \ldots, N$ into a matrix equation gives

$$\begin{pmatrix} y_3 \\ y_4 \\ \vdots \\ y_{\mathrm{N}} \end{pmatrix} = \begin{pmatrix} y_2 & y_1 & x_2 \\ y_3 & y_2 & x_3 \\ \vdots & \vdots & \vdots \\ y_{\mathrm{N}-1} & y_{\mathrm{N}-2} & x_{\mathrm{N}-1} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ b_1 \end{pmatrix} + \begin{pmatrix} \zeta_3 \\ \zeta_4 \\ \vdots \\ \zeta_{\mathrm{N}} \end{pmatrix} \tag{6.13}$$

or

$$\{Y\} = [A]\{\beta\} + \{\zeta\} \tag{6.14}$$

in matrix notation. As usual, matrices shall be denoted by square brackets, column vectors by curly brackets. $[A]$ is called the design matrix, $\{\beta\}$ is the vector of parameters and $\{\zeta\}$ is the residual vector. In this notation the sum of squared errors is

$$J(\{\beta\}) = \{\zeta\}^{\mathrm{T}}\{\zeta\} = (\{Y\}^{\mathrm{T}} - \{\beta\}^{\mathrm{T}}[A]^{\mathrm{T}})(\{Y\} - [A]\{\beta\}). \tag{6.15}$$

Minimizing this expression with respect to variation of the parameters proceeds as follows. The derivatives of J w.r.t. the parameters are evaluated and set equal to zero, the resulting linear system of equations yields the parameter estimates. Expanding (6.15) gives

$$J(\{\beta\}) = \{Y\}^{\mathrm{T}}\{Y\} - \{Y\}^{\mathrm{T}}[A]\{\beta\} - \{\beta\}^{\mathrm{T}}[A]^{\mathrm{T}}\{Y\} + \{\beta\}^{\mathrm{T}}[A]^{\mathrm{T}}[A]\{\beta\} \tag{6.16}$$

and differentiating with respect to $\{\beta\}^{\mathrm{T}}$, yields[3]

$$\frac{\partial J(\{\beta\})}{\partial \{\beta\}^{\mathrm{T}}} = -[A]^{\mathrm{T}}\{Y\} + [A]^{\mathrm{T}}[A]\{\beta\} \tag{6.17}$$

[3] Note that for the purposes of matrix calculus, $\{\beta\}$ and $\{\beta\}^{\mathrm{T}}$ are treated as independent. This is no cause for alarm; it is no different from treating z and z^* as independent in complex analysis. If the reader is worried, the more laborious calculation in terms of matrix elements is readily seen to yield the same result.

and setting the derivative to zero gives the well-known *normal equations* for the best parameter estimates $\{\hat{\beta}\}$:

$$[A]^{\mathrm{T}}[A]\{\hat{\beta}\} = [A]^{\mathrm{T}}\{Y\} \tag{6.18}$$

which are trivially solved by

$$\{\hat{\beta}\} = ([A]^{\mathrm{T}}[A])^{-1}[A]^{\mathrm{T}}\{Y\} \tag{6.19}$$

provided that $[A]^{\mathrm{T}}[A]$ is invertible. In practice, it is not necessary to invert this matrix in order to obtain the parameter estimates. In fact, solutions which avoid this are preferable in terms of speed [102, 209]. However, as shown later, the matrix $([A]^{\mathrm{T}}[A])^{-1}$ contains valuable information. A stable method of solution like LU decomposition [209] should always be used.

In practice, direct solution of the normal equations via (6.19) is not recommended as problems can arise if the matrix $[A]^{\mathrm{T}}[A]$ is close to singularity. Suppose that the right-hand side of equation (6.19) has a small error $\{\delta Y\}$ due to round-off say, the resulting error in the estimated parameters is given by

$$\{\delta\beta\} = ([A]^{\mathrm{T}}[A])^{-1}[A]^{\mathrm{T}}\{\delta Y\}. \tag{6.20}$$

As the elements in the inverted matrix are inversely proportional to the determinant of $[A]^{\mathrm{T}}[A]$, they can be arbitrarily large if $[A]^{\mathrm{T}}[A]$ is close to singularity. As a consequence, parameters with arbitrarily large errors can be obtained. This problem can be avoided by use of more sophisticated techniques. The near-singularity of the matrix $[A]^{\mathrm{T}}[A]$ will generally be due to correlations between its columns (recall that a matrix is singular if two columns are equal), i.e. correlations between model terms. It is possible to transform the set of equations (6.19) into a new form in which the columns of the design matrix are uncorrelated, thus avoiding the problem. Techniques for accomplishing this will be discussed in Appendix E.

6.3.2 Parameter uncertainty

Because of random errors in the measurements, different samples of data will contain different noise components and consequently they will lead to slightly different parameter estimates. The parameter estimates therefore constitute a random sample from a population of possible estimates; this population being characterized by a probability distribution. Clearly, it is desirable that the expected value of this distribution should coincide with the true parameters. If such a condition holds, the parameter estimator is said to be *unbiased* and the necessary conditions for this situation will be discussed in the next section. Now, given that the unbiased estimates are distributed about the true parameters, knowledge of the variance of the parameter distribution would provide valuable

information about the possible scatter in the estimates. This information turns out to be readily available; the *covariance matrix* $[\Sigma]$ for the parameters is defined by

$$[\Sigma](\{\hat{\beta}\}) = E[(\{\hat{\beta}\} - E[\{\hat{\beta}\}]) \cdot (\{\hat{\beta}\} - E[\{\hat{\beta}\}])^{\mathrm{T}}] \tag{6.21}$$

where the quantities with carets are the estimates and the expectation E is taken over all possible estimates. The diagonal elements of this matrix, σ_{ii}^2, are the variances of the parameter estimates $\hat{\beta}_i$.

Under the assumption that the estimates are unbiased and therefore $E[\{\hat{\beta}\}] = \{\beta\}$ where $\{\beta\}$ are now the true parameters, then

$$[\Sigma](\{\hat{\beta}\}) = E[(\{\hat{\beta}\} - \{\beta\}) \cdot (\{\hat{\beta}\} - \{\beta\})^{\mathrm{T}}]. \tag{6.22}$$

Now, substituting equation (6.14) containing the *true* parameters into equation (6.19) for the estimates, yields

$$\{\hat{\beta}\} = \{\beta\} + ([A]^{\mathrm{T}}[A])^{-1}[A]^{\mathrm{T}}\{\zeta\} \tag{6.23}$$

or, trivially

$$\{\hat{\beta}\} - \{\beta\} = ([A]^{\mathrm{T}}[A])^{-1}[A]^{\mathrm{T}}\{\zeta\} \tag{6.24}$$

which can be immediately substituted into (6.22) to give

$$[\Sigma] = E[([A]^{\mathrm{T}}[A])^{-1}[A]^{\mathrm{T}}\{\zeta\}\{\zeta\}^{\mathrm{T}}[A]([A]^{\mathrm{T}}[A])^{-1}]. \tag{6.25}$$

Now, it has been assumed that the only variable which changes from measurement to measurement if the excitation is repeated exactly is $\{\zeta\}$. Further, if $\{\zeta\}$ is independent of $[A]$, i.e. independent of x_i and y_i etc, then in this particular case

$$[\Sigma] = ([A]^{\mathrm{T}}[A])^{-1}[A]^{\mathrm{T}}E[\{\zeta\}\{\zeta\}^{\mathrm{T}}][A]([A]^{\mathrm{T}}[A])^{-1}]. \tag{6.26}$$

In order to proceed further, more assumptions must be made. First assume that the noise process $\{\zeta\}$ is zero-mean, i.e. $E[\{\zeta\}] = 0$. In this case the expectation in equation (6.26) is the covariance matrix of the noise process, i.e.

$$E[\{\zeta\}\{\zeta\}^{\mathrm{T}}] = [E[\zeta_i\zeta_j]] \tag{6.27}$$

and further assume that

$$E[\zeta_i\zeta_j] = \sigma_\zeta^2\delta_{ij} \tag{6.28}$$

where σ_ζ^2 is the variance of the residual sequence ζ_i and δ_{ij} is the Kronecker delta. Under this condition, the expression (6.26) collapses to

$$[\Sigma] = \sigma_\zeta^2([A]^{\mathrm{T}}[A])^{-1}. \tag{6.29}$$

The standard deviation for each estimated parameter is, therefore,

$$\sigma_i = \sigma_\zeta\sqrt{([A]^{\mathrm{T}}[A])_{ii}^{-1}}. \tag{6.30}$$

Now, if the parameter distributions are Gaussian, standard theory [17] yields a 95% confidence interval of $\{\hat{\beta}\} \pm 1.96\{\sigma\}$, i.e. there is a 95% probability that the true parameters fall within this interval.

6.3.3 Structure detection

In practice, it is unusual to know which terms should be in the model. This is not too much of a problem if the system under study is known to be linear; the number of possible terms is a linear function of the numbers of the lags n_y, n_x and n_e. However, it will be shown later that if the system is nonlinear, the number of possible terms increases *combinatorially* with increasing numbers of time lags. In order to reduce the computational load on the parameter estimation procedure it is clearly desirable to determine which terms should be included. With this in mind, a naive solution to the problem of *structure detection* can be found for simple least-squares parameter estimation. As the initial specification of an ARX model (6.9) includes all lags up to orders n_x and n_y, the model-fitting procedure needs to include some means of determining which of the possible terms are significant so that the remainder can safely be discarded. In order to determine whether a term is an important part of the model, a significance factor can be defined as follows. Each model term $\theta(t)$, e.g. $\theta(t) = y_{i-2}$ or $\theta(t) = x_{i-5}$, can be used on its own to generate a time series which will have variance σ_θ^2. The significance factor s_θ is then defined by

$$s_\theta = 100\frac{\sigma_\theta^2}{\sigma_y^2} \tag{6.31}$$

where σ_y^2 is the variance of the estimated output, i.e. the sum of all the model terms. Roughly speaking, s_θ is the percentage contributed to the model variance by the term θ. Having estimated the parameters the significance factors can be determined for each term; all terms which contribute less than some threshold value $s_{\min}$ to the variance can then be discarded. This procedure is only guaranteed to be effective if one works with an uncorrelated set of model terms. If the procedure were used on terms with intercorrelations one might observe two or more terms which appear to have a significant variance which actually cancelled to a great extent when added together. The more advanced least-squares methods described in appendix E allow the definition of an effective term selection criterion—namely the *error reduction ratio* or *ERR*.

6.4 The effect of noise

In order to derive the parameter uncertainties in equation (6.30), it was necessary to accumulate a number of assumptions about the noise process ζ. It will be shown in this section, that these assumptions have much more important consequences. Before proceeding, a summary will be made:

(1) It is assumed that ζ is zero-mean:

$$E[\{\zeta\}] = E[\zeta_i] = 0. \tag{6.32}$$

(2) It is assumed that ζ is uncorrelated with the process variables:

$$E[[A]^{\mathrm{T}}\{\zeta\}] = 0. \tag{6.33}$$

(3) The covariance matrix of the noise is assumed to be proportional to the unit matrix:

$$E[\zeta_i \zeta_j] = \sigma_\zeta^2 \delta_{ij}. \tag{6.34}$$

Now, the last assumption merits further discussion. It can be broken down into two main assertions:

(3a)

$$E[\zeta_i \zeta_j] = 0, \quad \forall i \neq j. \tag{6.35}$$

That is, the value of ζ at the time indexed by i is uncorrelated with the values at all other times. This means that there is no repeating structure in the data and it is therefore impossible to predict future values of ζ on the basis of past measurements. Such a sequence is referred to as *uncorrelated*.

The quantity $E[\zeta_i \zeta_j]$ is essentially the autocorrelation function of the signal ζ. Suppose i and j are separated by k lags, i.e. $j = i - k$, then

$$E[\zeta_i \zeta_j] = E[\zeta_i \zeta_{i-k}] = \phi_{\zeta\zeta}(k) \tag{6.36}$$

and the assumption of no correlation, can be written as

$$\phi_{\zeta\zeta}(k) = \sigma_\zeta^2 \delta_{k0} \tag{6.37}$$

where δ_{k0} is the Kronecker delta which is zero unless $k = 0$ when it is unity.

Now, it is a well-known fact, that the Fourier transform of the autocorrelation is the power spectrum; in this case the relationship is simpler to express in continuous time, where

$$\phi_{\zeta\zeta}(\tau) = E[\zeta(t)\zeta(t-\tau)] = \frac{P}{2\pi}\delta(\tau) \tag{6.38}$$

and P is the power spectral density of the signal. The normalization is chosen to give a simple result in the frequency domain. $\delta(\tau)$ is the Dirac δ-function.

One makes use of the relation

$$\begin{aligned} \mathcal{F}[\phi_{\zeta\zeta}(\tau)] &= \int_{-\infty}^{\infty} \mathrm{d}\tau \, \mathrm{e}^{-\mathrm{i}\omega\tau} E[\zeta(t)\zeta(t+\tau)] \\ &= E\left[\int_{-\infty}^{\infty} \mathrm{d}\tau \, \mathrm{e}^{-\mathrm{i}\omega\tau} \zeta(t)\zeta(t+\tau)\right] \\ &= E[Z(\omega)Z(\omega)] = S_{\zeta\zeta}(\omega) \end{aligned} \tag{6.39}$$

where $Z(\omega)$ is the spectrum of the noise process. The manifest fact that $\phi_{\zeta\zeta}(\tau) = \phi_{\zeta\zeta}(-\tau)$ has also been used earlier.

For the assumed form of the noise (6.38), it now follows that

$$S_{\zeta\zeta}(\omega) = P. \tag{6.40}$$

So the signal ζ contains equal proportions of all frequencies. For this reason, such signals are termed *white noise*. Note that a mathematical white noise process cannot be realized physically as it would have infinite power and therefore infinite variance[4].

(3b) It is assumed that $E[\zeta_i^2]$ takes the same value for all i. That is, the variance σ_ζ^2 is constant over time. This, together with the zero-mean condition amounts to an assumption that ζ is *weakly stationary*. Weak stationarity of a signal simply means that the first two statistical moments are time-invariant. *True* or *strong* stationarity would require all moments to be constant.

So to recap, in order to estimate the parameter uncertainty, it is assumed that the noise process ζ is white uncorrelated weakly stationary noise and uncorrelated with the process variables x_i and y_i. The question is: Is this assumption justified?

Consider the continuous-time form (6.2) and assume that the output measurement only is the sum of a *clean* part $y_c(t)$ which satisfies the equation of motion and a noise component $e(t)$ which satisfies all the previously described assumptions. (In the remainder of this book, the symbol e will be reserved for such noise processes, ζ will be used to denote the generic noise process.)

$$y(t) = y_c(t) + e(t). \tag{6.41}$$

The equation of motion for the measured quantity is

$$m\ddot{y} + c\dot{y} + ky = x(t) - m\ddot{e} - c\dot{e} - ke \tag{6.42}$$

or, in discrete time,

$$y_i = a_1 y_{i-1} + a_2 y_{i-2} + b_1 x_{i-1} - e_i + a_1 e_{i-1} + a_2 e_{i-2}. \tag{6.43}$$

So the noise process ζ_i of (6.14) is actually formed from

$$\zeta_i = -e_i + a_1 e_{i-1} + a_2 e_{i-2} \tag{6.44}$$

and the covariance matrix for this process takes the form (in matrix terms)

$$[E[\zeta_i \zeta_j]] = \sigma_e^2 \begin{pmatrix} 1+a_1^2+a_2^2 & a_1(a_2-1) & -a_2 & 0 & \dots \\ a_1(a_2-1) & 1+a_1^2+a_2^2 & a_1(a_2-1) & -a_2 & \dots \\ -a_2 & a_1(a_2-1) & 1+a_1^2+a_2^2 & a_1(a_2-1) & \dots \\ 0 & -a_2 & a_1(a_2-1) & 1+a_1^2+a_2^2 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \tag{6.45}$$

[4] This is why the relation (6.40) does not contain the variance. If one remains in discrete-time with (6.37), the power spectrum is obtained from the discrete Fourier transform

$$S_{\zeta\zeta}(j) = \sum_{j=1}^{N} \phi_{\zeta\zeta}(k) \mathrm{e}^{\mathrm{i}k\Delta t j \Delta\omega} \Delta t = \sum_{j=1}^{N} \sigma_\zeta^2 \delta_{k0} \mathrm{e}^{\mathrm{i}k\Delta t j \Delta\omega} \Delta t = \sigma_\zeta^2 \Delta t = \frac{\sigma_\zeta^2}{N\Delta\omega} = \frac{\sigma_\zeta^2}{2\omega_\mathrm{N}}$$

which is the power spectral density (ω_N is the Nyquist frequency). Note that a signal which satisfies (6.37) has finite power. Where there is likely to be confusion, signals of this form will be referred to as *discrete white*.

Such a process will not have a constant power spectrum. The signal contains different proportions at each frequency. As a result it is termed *coloured* or *correlated* noise. If the noise is coloured, the simple relations for the parameter uncertainties are lost. Unfortunately there are also more serious consequences which will now be discussed. In order to simplify the discussion, a simpler model will be taken. a_2 shall be assumed zero (this makes the normal equations a 2×2 system which can be solved by hand), and the noise process will take the simplest coloured form possible. So

$$y_i = ay_{i-1} + bx_{i-1} - e_i + ce_{i-1} \tag{6.46}$$

and e_i satisfies all the appropriate assumptions and its variance is σ_e^2. The processes x_i and y_i are assumed stationary with respective variances σ_x^2 and σ_y^2 and x_i is further assumed to be an uncorrelated noise process. Now suppose the model takes no account of correlated measurement noise, i.e. a form

$$y_i = \hat{a}y_{i-1} + \hat{b}x_{i-1} + e_i' \tag{6.47}$$

is assumed. The normal equations (6.18) for the estimates $\hat{a}$ and $\hat{b}$ can be shown to be

$$\begin{pmatrix} \sum_{i=1}^N y_{i-1}^2 & \sum_{i=1}^N y_{i-1}x_{i-1} \\ \sum_{i=1}^N y_{i-1}x_{i-1} & \sum_{i=1}^N x_{i-1}^2 \end{pmatrix} \begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^N y_i y_{i-1} \\ \sum_{i=1}^N y_i x_{i-1} \end{pmatrix}. \tag{6.48}$$

Dividing both sides of the equations by $N-1$ yields

$$\begin{pmatrix} E[y_{i-1}^2] & E[y_{i-1}x_{i-1}] \\ E[y_{i-1}x_{i-1}] & E[x_{i-1}^2] \end{pmatrix} \begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix} = \begin{pmatrix} E[y_i y_{i-1}] \\ E[y_i x_{i-1}] \end{pmatrix}. \tag{6.49}$$

In order to evaluate the estimates, it is necessary to compute a number of expectations, although the calculation is a little long-winded, it is instructive and so is given in detail.

(1) First $E[y_{i-1}^2]$ is needed. This is straightforward as $E[y_{i-1}^2] = E[y_i^2] = \sigma_y^2$ due to stationarity. Similarly $E[x_{i-1}^2] = \sigma_x^2$.

(2)

$$\begin{aligned} E[y_{i-1}x_{i-1}] &= E[(ay_{i-2} + bx_{i-2} - e_{i-1} + ce_{i-2})x_{i-1}] \\ &= aE[y_{i-2}x_{i-1}] + bE[x_{i-2}x_{i-1}] \\ &\quad - E[e_{i-1}x_{i-1}] + cE[e_{i-2}x_{i-1}]. \end{aligned}$$

Now, the first expectation vanishes because x_{i-1} is uncorrelated noise and it is impossible to predict it from the past output y_{i-2}. The second expectation vanishes because x_i is uncorrelated and the third and fourth expectations vanish because e_i is uncorrelated with x. In summary, $E[y_{i-1}x_{i-1}] = 0$.

(3)

$$\begin{aligned} E[y_i y_{i-1}] &= E[(ay_{i-1} + bx_{i-1} - e_i + ce_{i-1})y_{i-1}] \\ &= aE[y_{i-1}y_{i-1}] + bE[x_{i-1}y_{i-1}] - E[e_i y_{i-1}] + cE[e_{i-1}y_{i-1}]. \end{aligned}$$

The first expectation is already known to be σ_y^2. The second is zero because the current input is unpredictable given only the current output. The fourth expectation is zero because the current noise e_i is unpredictable from the past output. This leaves $E[e_{i-1}y_{i-1}]$ which is

$$\begin{aligned} E[e_{i-1}y_{i-1}] &= aE[e_{i-1}y_{i-2}] + bE[e_{i-1}x_{i-2}] \\ &\quad - E[e_{i-1}e_{i-1}] + cE[e_{i-1}e_{i-2}] \\ &= -\sigma_e^2 \end{aligned}$$

So finally, $E[y_i y_{i-1}] = a\sigma_y^2 - c\sigma_e^2$.

(4)

$$\begin{aligned} E[y_i x_{i-1}] &= aE[y_{i-1}x_{i-1}] + bE[x_{i-1}x_{i-1}] - E[e_i x_{i-1}] + cE[e_{i-1}x_{i-1}] \\ &= b\sigma_x^2. \end{aligned}$$

Substituting all of these results into the normal equations (6.46) yields

$$\begin{pmatrix} \sigma_y^2 & 0 \\ 0 & \sigma_x^2 \end{pmatrix} \begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix} = \begin{pmatrix} a\sigma_y^2 - c\sigma_e^2 \\ b\sigma_x^2 \end{pmatrix} \tag{6.50}$$

and these are trivially solved to give the estimates:

$$\hat{a} = a - c\frac{\sigma_e^2}{\sigma_y^2}, \quad \hat{b} = b. \tag{6.51}$$

So, although the estimate for b is correct, the estimate for a is in error. Because this argument is in terms of expectations, it means that this error will occur *no matter how much data are measured*. In the terminology introduced earlier, the estimate is *biased*. The bias only disappears under two conditions.

(1) First, in the limit as the noise-to-signal ratio goes to zero. This is expected.
(2) Second, if $c = 0$, and this is the condition for ζ to be uncorrelated white noise.

The conclusion is that coloured measurement noise implies biased parameter estimates. The reason is that the model (6.47) assumes that the only non-trivial relationships are between the input and output processes. In fact there is structure within the noise process which is not accounted for. In order to eliminate the bias, it is necessary to take this structure into account and estimate a model for the noise process—a *noise model*. In the previous example, the measurement noise ζ_i is regressed on past values of a white noise process, i.e. it is a *moving average* or

MA model in the terminology introduced in chapter 1. The general noise model of this type takes the form

$$\zeta_i = \sum_{j=0}^{n_e} c_j e_{i-j}. \tag{6.52}$$

A more compact model can sometimes be obtained by assuming the more general ARMA form

$$\zeta_i = \sum_{j=1}^{n_\zeta} d_j \zeta_{i-j} + \sum_{j=0}^{n_e} c_j e_{i-j}. \tag{6.53}$$

So, some remarks are required on the subject of parameter estimation if a noise model is necessary. First of all a structure for the model must be specified, then the situation is complicated by the fact that the noise signal is unmeasurable. In this case, an initial fit is made to the data without a noise model, the model predicted output is then subtracted from the measured output to give an estimate of the noise signal. This allows the re-estimation of parameters, including now the noise model parameters. The procedure—fit model–predict output–estimate noise signal—is repeated until the parameters converge.

6.5 Recursive least squares

The least-squares algorithm described in the last section assumes that all the data are available for processing at one time. It is termed the *batch* or *off-line* estimator. In many cases it will be interesting to monitor the progress of a process in order to see if the parameters of the model change with time. Such a situation is not uncommon—a rocket burning fuel or a structure undergoing failure will both display time-varying parameters. In the latter case, monitoring the parameters could form the basis of a non-destructive damage evaluation system. It is clear that some means of tracking time variation could prove valuable. A naive approach consists of treating the data as a new batch every time a new measurement becomes available and applying the off-line algorithm. This is computationally expensive as a matrix inverse is involved and, in some cases, might not be fast enough to track changes. Fortunately, it is possible to derive an *on-line* or *recursive* version of the least-squares algorithm which does not require a matrix inverse at each step. The derivation of this algorithm is the subject of this section [5].

First, assume the general ARX form for the model as given in equation (6.9). If n measurements have already been accumulated; the form of the least-squares problem is

$$\{Y\}_n = [A]_n\{\beta\} + \{\zeta\}_n \tag{6.54}$$

[5] The derivation can be expressed in terms of the so-called *matrix inversion lemma* as discussed in [168]. However, the derivation presented here is considered more instructive, it follows an argument presented in [30].

with solution

$$\{\hat{\beta}\}_n = ([A]_n^{\mathrm{T}}[A]_n)^{-1}[A]_n^{\mathrm{T}}\{Y\}_n. \tag{6.55}$$

Now, if new measurements for x and y, become available, the problem becomes

$$\begin{pmatrix} \{Y\}_n \\ y_{n+1} \end{pmatrix} = \begin{pmatrix} [A]_n \\ \{\phi\}_{n+1}^{\mathrm{T}} \end{pmatrix} \{\beta\} + \begin{pmatrix} \{\zeta\}_n \\ \zeta_{n+1} \end{pmatrix} \tag{6.56}$$

with

$$\{\phi\}_{n+1}^{\mathrm{T}} = (y_n, \ldots, y_{n-n_y}, x_{n-1}, \ldots, x_{n-n_x+1}) \tag{6.57}$$

and this has the updated solution

$$\{\hat{\beta}\}_{n+1} = \left[\begin{pmatrix} [A]_n & \{\phi\}_{n+1} \end{pmatrix} \begin{pmatrix} [A]_n \\ \{\phi\}_{n+1}^{\mathrm{T}} \end{pmatrix} \right]^{-1} \begin{pmatrix} [A]_n & \{\phi\}_{n+1} \end{pmatrix} \begin{pmatrix} \{Y\}_n \\ y_{n+1} \end{pmatrix} \tag{6.58}$$

or, on expanding,

$$\{\hat{\beta}\}_{n+1} = ([A]_n^{\mathrm{T}}[A]_n + \{\phi\}_{n+1}\{\phi\}_{n+1}^{\mathrm{T}})^{-1}([A]_n\{Y\}_n + \{\phi\}_{n+1}y_{n+1}). \tag{6.59}$$

Now define $[P]_n$:

$$[P]_n = ([A]_n^{\mathrm{T}}[A]_n)^{-1} \tag{6.60}$$

and note that this is nearly the covariance matrix for the parameters, in fact

$$[\Sigma] = \sigma_\zeta^2[P]. \tag{6.61}$$

(The matrix $[P]$ is often referred to as the covariance matrix and this convention will be adopted here. If confusion is likely to arise in an expression, the distinction will be drawn.) With the new notation, the update rule (6.59) becomes trivially

$$\{\hat{\beta}\}_{n+1} = ([P]_n + \{\phi\}_{n+1}\{\phi\}_{n+1}^{\mathrm{T}})^{-1}([A]_n\{Y\}_n + \{\phi\}_{n+1}y_{n+1}) \tag{6.62}$$

and taking out the factor $[P]_n$ gives

$$\{\hat{\beta}\}_{n+1} = [P]_n(I + \{\phi\}_{n+1}\{\phi\}_{n+1}^{\mathrm{T}}[P]_n)^{-1}([A]_n\{Y\}_n + \{\phi\}_{n+1}y_{n+1}). \tag{6.63}$$

Note that the first bracket is simply $[P]_{n+1}$, expanding this with the binomial theorem yields

$$\begin{aligned} [P]_{n+1} &= [P]_n(I - \{\phi\}_{n+1}\{\phi\}_{n+1}^{\mathrm{T}}[P]_n + (\{\phi\}_{n+1}\{\phi\}_{n+1}^{\mathrm{T}}[P]_n)^2 - \cdots) \\ &= [P]_n(I - \{\phi\}_{n+1}[1 - \{\phi\}_{n+1}^{\mathrm{T}}[P]_n\{\phi\}_{n+1} \\ &\quad + (\{\phi\}_{n+1}^{\mathrm{T}}[P]_n\{\phi\}_{n+1})^2 - \cdots]\{\phi\}_{n+1}^{\mathrm{T}}[P]_n) \\ &= [P]_n \left(I - \frac{\{\phi\}_{n+1}\{\phi\}_{n+1}^{\mathrm{T}}[P]_n}{1 + \{\phi\}_{n+1}^{\mathrm{T}}[P]_n\{\phi\}_{n+1}} \right). \end{aligned} \tag{6.64}$$

So

$$\{\hat{\beta}\}_{n+1} = [P]_n \left(I - \frac{\{\phi\}_{n+1}\{\phi\}_{n+1}^{\mathrm{T}}[P]_n}{1 + \{\phi\}_{n+1}^{\mathrm{T}}[P]_n\{\phi\}_{n+1}} \right) ([A]_n\{Y\}_n + \{\phi\}_{n+1}y_{n+1}) \tag{6.65}$$

which expands to

$$\begin{aligned}\{\hat{\beta}\}_{n+1} = {} & [P]_n[A]_n^{\mathrm{T}}\{Y\}_n - \frac{[P]_n\{\phi\}_{n+1}\{\phi\}_{n+1}^{\mathrm{T}}[P]_n}{1 + \{\phi\}_{n+1}^{\mathrm{T}}[P]_n\{\phi\}_{n+1}}[A]_n^{\mathrm{T}}\{Y\}_n \\ & + [P]_n\{\phi\}_{n+1}^{\mathrm{T}}y_{n+1} - \frac{[P]_n\{\phi\}_{n+1}\{\phi\}_{n+1}^{\mathrm{T}}[P]_n}{1 + \{\phi\}_{n+1}^{\mathrm{T}}[P]_n\{\phi\}_{n+1}}\{\phi\}_{n+1}y_{n+1}.\end{aligned} \tag{6.66}$$

Now, noting that (6.55) can be written in the form

$$\{\hat{\beta}\}_n = [P]_n[A]_n^{\mathrm{T}}\{Y\}_n \tag{6.67}$$

equation (6.66) can be manipulated into the form

$$\{\hat{\beta}\}_{n+1} = \{\hat{\beta}\}_n + \{K\}_{n+1}(y_{n+1} - \{\phi\}_{n+1}^{\mathrm{T}}\{\hat{\beta}\}_n) \tag{6.68}$$

where the *Kalman gain* $\{K\}$ is defined by

$$\{K\}_{n+1} = \frac{[P]_n\{\phi\}_{n+1}}{1 + \{\phi\}_{n+1}^{\mathrm{T}}[P]_n\{\phi\}_{n+1}} \tag{6.69}$$

and the calculation is complete; equations (6.68) and (6.69), augmented by (6.64), constitute the update rules for the off-line or recursive least-squares (RLS) algorithm[6].

The iteration is started with the estimate $\{\beta\}_0 = \{0\}$. $[P]$ is initialized diagonal with large entries; the reason for this is that the diagonal elements of $[P]$ are proportional to the standard deviations in the parameter estimates, so starting with large entries encodes the fact that there is little confidence in the initial estimate.

The object of this exercise was to produce an iterative algorithm which could track variations in parameters. Unfortunately this is not possible with

[6] Note that equation (6.68) takes the form

$$\text{new estimate} = \text{old estimate} + \text{gain} \times \text{prediction error}.$$

Anticipating the sections and appendices on neural networks, it can be stated that this is simply the backpropagation algorithm for the linear-in-the-parameters ARX model considered as an almost trivial neural network (figure 6.2). The gain vector $\{K\}$ can therefore be loosely identified with the gradient vector

$$\frac{\partial J(\{\beta\})}{\partial \{\beta\}^{\mathrm{T}}}.$$

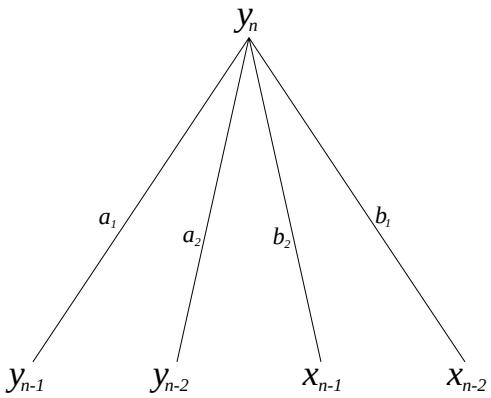


Figure 6.2. An ARX system considered as a linear neural network.

this algorithm as it stands. The iterative procedure is actually obtained directly from (6.19), and after N iterations the resulting parameters are identical to those which would be obtained from the off-line estimate using the N sets of measurements. The reason for this is that the recursive procedure remembers *all* past measurements and weights them equally. Fortunately, a simple modification exists which allows past data to be weighted with a factor which decays exponentially with time, i.e. the objective function for minimization is

$$J_{n+1} = \lambda J_n + (y_{n+1} - \{\phi\}_{n+1}^{\mathrm{T}}\{\hat{\beta}\})^2 \tag{6.70}$$

where λ is a *forgetting factor*, i.e. if $\lambda < 1$, past data are weighted out.

The required update formulae are [167]

$$\{K\}_{i+1} = \frac{[P]_i\{\phi\}_{i+1}}{\lambda + \{\phi_{i+1}\}^{\mathrm{T}}[P]_i\{\phi_{i+1}\}} \tag{6.71}$$

$$[P]_{i+1} = \frac{1}{\lambda}(1 - \{K\}_{i+1}\{\phi\}_{i+1}^{\mathrm{T}})[P]_i \tag{6.72}$$

with (6.68) unchanged. In this formulation the parameter estimates can keep track of variation in the true system parameters. The smaller λ is, the faster the procedure can respond to changes. However, if λ is too small the estimates become very susceptible to spurious variations due to measurement noise. A value for λ in the range 0.9–0.999 is usually adopted.

When the measurements are noisy, the RLS method is well known to give biased estimates and more sophisticated approaches are needed. The double least-squares (DLS) method [67] averages the estimates of two approaches, one that tends to give a positive damping bias and a second that usually gives a negative damping bias. The DLS technique has been shown to work well on simulated structural models based on the ARX [67]. The on-line formulation is very similar

to RLS, the update rules are

$$\{K\}_{i+1} = \frac{[P]_i\{\psi\}_{i+1}}{\lambda + \{\phi_{i+1}\}^{\mathrm{T}}[P]_i\{\psi_{i+1}\}} \tag{6.73}$$

with (6.72) and (6.64) unchanged. The vector $\{\phi\}_{i+1}$ is defined as before, but a new *instrument* vector is needed:

$$\{\psi\}_{n+1}^{\mathrm{T}} = (y_{n+1} + y_n, \ldots, y_{n+1-n_y} + y_{n-n_y}, x_{n-1}, \ldots, x_{n-n_x+1}). \tag{6.74}$$

Another approach, the instrumental variables (IV) method, uses the same update rule, but sets the instruments as time-delayed samples of output. Such a delay theoretically removes any correlations of the noise which lead to bias. In the IV formulation

$$\{\psi\}_{n+1}^{\mathrm{T}} = (y_{n-p}, \ldots, y_{n-p-n_y}, x_{n-1}, \ldots, x_{n-n_x+1}) \tag{6.75}$$

where p is the delay.

6.6 Analysis of a time-varying linear system

The methods described in the previous section are illustrated here with a simple case study. The time-varying system studied is a vertical plastic beam with a built-in end—a cantilever. At the free end is a pot of water. During an experiment, the mass of the system could be changed by releasing the water into a receptacle below. Figure 6.3 shows the experimental arrangement. The instrumentation needed to carry out such an experiment is minimal. Essentially all that is required is two sensors and some sort of acquisition system. The input sensor should be a force gauge. The output sensor could be a displacement, velocity or acceleration sensor—the relative merits and demerits of each are discussed in the following section. There are presently many inexpensive computer-based data capture systems, many based on PCs, which are perfectly adequate for recording a small number of channels. The advantage of using a computer-based system is that the signal processing can be carried out in software. If Fourier transforms are possible, the acquisition system is fairly straightforwardly converted to an FRF analyser.

In order to make the system behave as far as possible like a SDOF system, it was excited with a band-limited random force covering only the first natural frequency. The acceleration was measured with an accelerometer at the free end. In order to obtain the displacement signal needed for modelling, the acceleration was integrated twice using the trapezium rule. Note that the integration of time data is not a trivial matter and it will be discussed in some detail in appendix I. During the acquisition period the water was released. Unfortunately it was impossible to locate this event in time with real precision. However, it was nominally in the centre of the acquisition period so that the parameter estimator

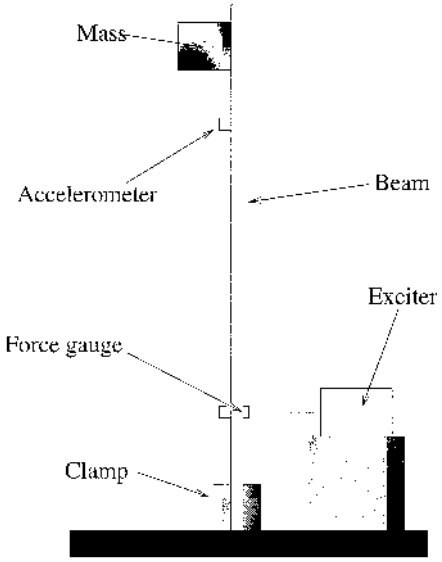


Figure 6.3. Experimental arrangement for a time-varying cantilever experiment.

was allowed to 'warm-up'. (Note also that the integration routine removes a little data from the beginning and end of the record.) Another slight problem was caused by the fact that it was impossible to release the water without communicating some impulse to the system.

The model structure (6.9) was used as it is appropriate to a SDOF system. In general the *minimal* model needed for a N degree-of-freedom system is

$$y_i = \sum_{j=1}^{2N} a_j y_{i-j} + \sum_{j=1}^{2N-1} b_j x_{i-j} \tag{6.76}$$

and this is minimal because it assumes the simplest discretization rule for the derivatives.

A minor problem with discrete-time system identification for the structural dynamicist is that the model coefficients have no physical interpretation. However, although it is difficult to convert the parameters to masses, dampings and stiffnesses, it is relatively straightforward to obtain frequencies and damping ratios [152]. One proceeds via the *characteristic polynomial*

$$\phi(p) = 1 - \sum_{j=1}^{2N} a_j p^{2N-j} \tag{6.77}$$

whose roots (the poles of the model) are given by

$$p_j = \exp^{\Delta t(-\zeta_j \omega_{nj} \pm \mathrm{i}\omega_{\mathrm{d}j})}. \tag{6.78}$$

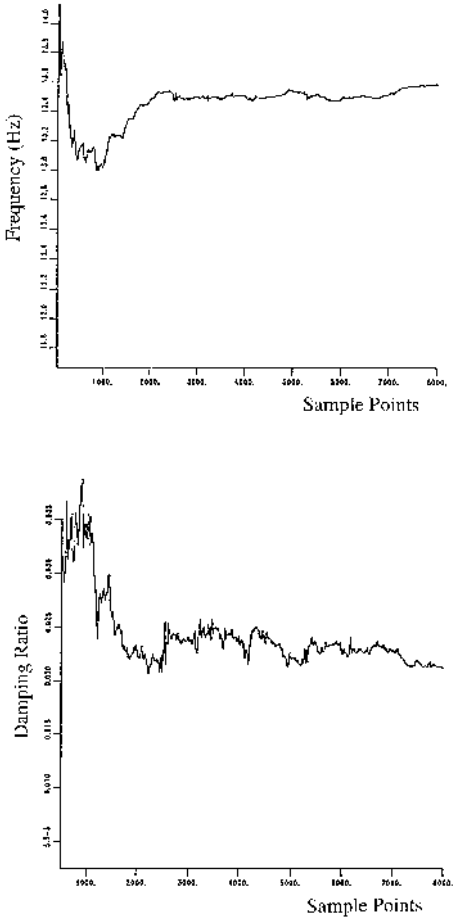


Figure 6.4. Identified parameters from the experimental cantilever beam with water, $\lambda = 1$: (*a*) frequency; (*b*) damping ratio.

The frequency and damping for the system with water are shown in figure 6.4. In this case, the system was assumed to be time-invariant and a forgetting factor $\lambda = 1$ was used. After an initial disturbance, the estimator settles down to the required constant value. The situation is similar when the system is tested without the water (figure 6.5). In the final test (figure 6.6), the water was released about 3000 samples into the record. A forgetting factor of 0.999 was used, note that this value need not be very far from unity. As expected, the natural frequency jumps between two values. The damping ratio is disturbed during the transition region but returns to the correct value afterwards.

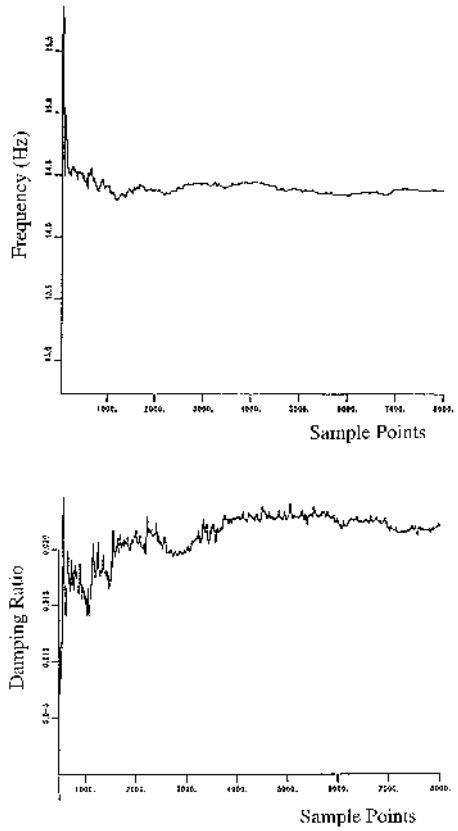


Figure 6.5. Identified parameters from the experimental cantilever beam without water, $\lambda = 1$: (*a*) frequency; (*b*) damping ratio.

In the next chapter, methods for directly extracting physical parameters are presented.

6.7 Practical matters

The last section raised certain questions about the practice of experimentation for system identification. This section makes a number of related observations.

6.7.1 Choice of input signal

In the system identification literature, it is usually said that an input signal must be *persistently exciting* if it is to be of use for system identification. There are numerous technical definitions of this term of varying usefulness [231]. Roughly

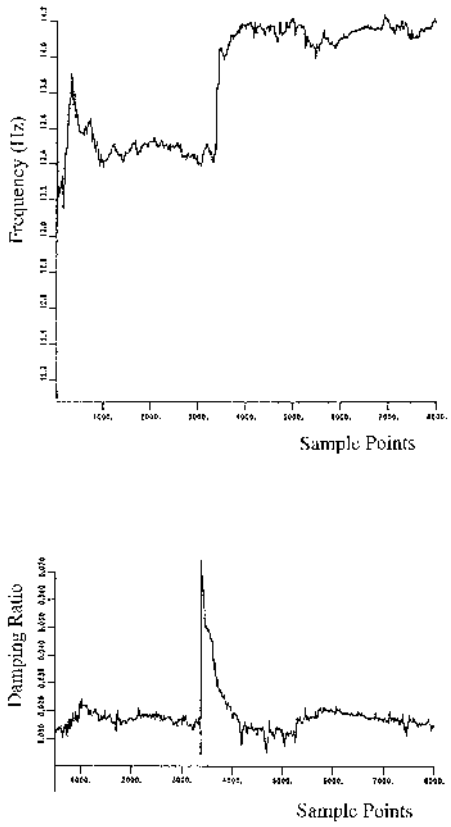


Figure 6.6. Identified parameters from the experimental time-varying cantilever beam, $\lambda = 0.999$: (*a*) frequency; (*b*) damping ratio.

speaking, the term means that the signal should have enough frequency coverage to excite all the modes of interest. This is the only consideration in linear system identification. The situation in nonlinear system identification is slightly different; there, one must also excite the nonlinearity. In the case of polynomial nonlinearity, the level of excitation should be high enough that all terms in the polynomial contribute to the restoring force. In the case of Coulomb friction, the excitation should be *low* enough that the nonlinearity is exercised. For piecewise linear stiffness or damping, all regimes should be covered.

The more narrow-band a signal is, the less suitable it is for identification. Consider the limit—a single harmonic $X\sin(\omega t - \phi)$. The standard SDOF oscillator equation (6.2) becomes

$$-m\omega^2 Y\sin(\omega t) + c\omega Y\cos(\omega t) + kY\sin(\omega t) = X\sin(\omega t - \phi). \qquad (6.79)$$

Now, it is a trivial fact that

$$-(m\omega^2+\alpha)Y\sin(\omega t)+c\omega Y\cos(\omega t)+(k-\alpha)Y\sin(\omega t)=X\sin(\omega t-\phi) \tag{6.80}$$

is identically satisfied with α arbitrary. Therefore, the system

$$\left(m-\frac{\alpha}{\omega^2}\right)\ddot{y}+c\dot{y}+(k-\alpha)y=X\sin(\omega t-\phi) \tag{6.81}$$

explains the input–output process just as well as the true (6.2). This is simply a manifestation of linear dependence, i.e. there is the relation

$$\ddot{y}=-\omega^2 y \tag{6.82}$$

and this will translate into discrete time as

$$y_i+(\omega^2-2)y_{i-1}+y_{i-2}=0. \tag{6.83}$$

So the sine-wave is unsuitable for linear system identification. If one consults [231], one finds that the sine-wave is only persistently exciting of the very lowest order. Matters are improved by taking a sum of N_h sinusoids

$$x(t)=\sum_{i=1}^{N_h}C_i\sin(\omega_i t) \tag{6.84}$$

and it is a simple matter to show that the presence of even two sinusoids is sufficient to break the linear dependence (6.82) (although the two frequencies should be reasonably separated).

In the case of a nonlinear system, the presence of harmonics is sufficient to break linear dependence even if a single sinusoid is used, i.e.

$$y(t)=A_1\sin(\omega t)+A_3\sin(3\omega t)+\cdots \tag{6.85}$$

$$\ddot{y}(t)=-\omega^2 A_1\sin(\omega t)+-9\omega^2 A_3\sin(3\omega t)+\cdots. \tag{6.86}$$

However, the input is still sub-optimal [271B].

6.7.2 Choice of output signal

This constitutes a real choice for structural dynamicists as the availability of the appropriate sensors means that it is possible to obtain displacement, velocity or acceleration data.

For a linear system, the choice is almost arbitrary, differentiation of (6.2) yields the equations of motion for the linear SDOF system if velocity or acceleration is observed.

$$m\ddot{v}+c\dot{v}+kv=\dot{x}(t) \tag{6.87}$$

and

$$m\ddot{a} + c\dot{a} + ka = \ddot{x}(t) \tag{6.88}$$

which result in discrete-time forms

$$v_i = a_1 v_{i-1} + a_2 v_{i-2} = b_1 x_{i-1} + b_2 x_{i-2} \tag{6.89}$$

and

$$a_i = a_1 a_{i-1} + a_2 a_{i-2} = b_0 x_i + b_1 x_{i-1} + b_2 x_{i-2} \tag{6.90}$$

which are a little more complicated than (6.7). The only slight difference is a few more lagged x terms and the present of the current input x_i in the acceleration form. Note also that the coefficients of the AR part are unchanged. This might be expected as they specify the characteristic polynomial from which the frequencies and dampings are obtained.

If the system is nonlinear, i.e. Duffing's system (anticipating (6.94)), the situation is different. On the one hand, the harmonics of the signal are weighted higher in the velocity and even more so in the acceleration, and this might suggest that these forms are better for fitting nonlinear terms. On the other hand the equations of motion become considerably more complex. For the velocity state, the Duffing system has the equation

$$m\dot{v} + cv + +k_1\left(\int^t \mathrm{d}\tau\, v(\tau)\right) + k_2\left(\int^t \mathrm{d}\tau\, v(\tau)\right)^2 + k_3\left(\int^t \mathrm{d}\tau\, v(\tau)\right)^3 = x(t) \tag{6.91}$$

or

$$m\ddot{v} + c\dot{v} + k_1 v + v\left(\int^t \mathrm{d}\tau\, v(\tau)\right)\left\{2k_2 + 3k_3\left(\int^t \mathrm{d}\tau\, v(\tau)\right)\right\} = \dot{x}(t) \tag{6.92}$$

either form being considerably more complicated than (6.94). The equation of motion for the acceleration data is more complicated still. It is known that it is difficult to fit time-series models with polynomial terms to force–velocity or force–acceleration data from a Duffing oscillator system [58].

In the case of the Duffing system, the simplest structure is obtained if all three states are measured and used in the modelling. This is the situation with the direct parameter estimation approach discussed in the next chapter.

6.7.3 Comments on sampling

The choice of sampling frequency is inseparable from the choice of input bandwidth. Shannon's criterion [129] demands that the sampling frequency should be higher than twice the frequency of interest to avoid aliasing. In the case of a linear system, this means twice the highest frequency in the input. In the case of a nonlinear system, the frequency should also capture properly the appropriate number of harmonics. Having said this, the effect of aliasing on system identification for discrete-time systems is not clear.

Surprisingly, it is also possible to oversample for the purposes of system identification. Ljung [167] summarizes his discussion on over-sampling as follows.

- 'Very fast sampling leads to numerical problems, model fits in high-frequency bands, and poor returns for hard work.'
- 'As the sampling interval increases over the natural time constants of the system, the variance (of parameter estimates) increases drastically.' (In fact, he shows analytically for a simple example that the parameter variance tends to infinity as the sampling interval Δt tends to zero [167] p 378.)
- 'Optimal choices of Δt for a fixed number of samples will lie in the range of the time constants of the system. These are, however, not known, and overestimating them may lead to very bad results.'

Comprehensive treatments of the problem can also be found in [119] and [288]. A useful recent reference is [146].

It is shown in [277] that there is a very simple explanation for oversampling. As the sampling frequency increases, there comes a point where the estimator can do better by establishing a simple linear interpolation than it can by finding the true model. An approximate upper bound for the *over-sampling frequency* is given by

$$f_{\mathrm{s}} = \pi 32^{\frac{1}{4}} \gamma^{\frac{1}{2}} f_{\max} \tag{6.93}$$

for high signal-to-noise ratios γ.

(This result can only be regarded as an existence result due to the fact that the signal-to-noise ratio would not be known in practice.)

6.7.4 The importance of scaling

In the previous discussion of normal equations, it was mentioned that the conditioning and invertibility of the information matrix $[A]^{\mathrm{T}}[A]$ is critical. The object of this short section is to show how scaling of the data is essential to optimize the condition of this matrix. The discussion will be by example, data are simulated from a linear SDOF system (6.2) and a discrete-time Duffing oscillator (6.95).

It is assumed that the model structure (6.7) is appropriate to linear SDOF data, so the design matrix would take the form given in (6.13). A system with a linear stiffness of $k = 10^4$ was taken for the example, and this meant that an input force $x(t)$ with rms 0.622, generated a displacement response with rms 5.87×10^{-5}. There is consequently a large mismatch between the scale of the first two columns of $[A]$ and the third. This mismatch is amplified when $[A]$ is effectively squared to form the information matrix

$$\begin{pmatrix} 0.910 \times 10^{-4} & 0.344 \times 10^{-5} & 0.188 \times 10^{-2} \\ 0.940 \times 10^{-4} & 0.346 \times 10^{-5} & 0.144 \times 10^{-2} \\ 0.114 & 0.144 \times 10^{-2} & 0.389 \times 10^{3} \end{pmatrix}.$$

The condition of this matrix can be assessed by evaluating the singular values and in this case they are found to be 388.788, 1.302×10^{-4} and 5.722×10^{-8}. The *condition number* is defined as the ratio of the maximum-to-minimum singular value and in this case is 6.80×10^{9}. Note that if one rejects singular values on the basis of proportion, a high condition number indicates a high probability of rejection and hence deficient effective rank. The other indicator of condition is the determinant; this can be found from the product of singular values and in this case is 2.90×10^{-9}, quite low.

A solution to this problem is fairly straightforward. If there were no scale mismatch between the columns in $[A]$, the information matrix would be better conditioned. Therefore, *one should always divide each column by its standard deviation*, the result in this case is a scaled information matrix

$$\begin{pmatrix} 0.264 \times 10^{5} & 0.100 \times 10^{4} & 0.515 \times 10^{2} \\ 0.273 \times 10^{5} & 0.100 \times 10^{4} & 0.396 \times 10^{2} \\ 0.312 \times 10^{4} & 0.396 \times 10^{2} & 0.100 \times 10^{4} \end{pmatrix}$$

and this has singular values 19.1147, 997.314 and 997.314. The condition number is 1996.7 and the determinant is 7.27×10^{8}. There is clearly no problem with condition.

To drive home the point, consider a Duffing system: one of the columns in the design matrix contains y^3, which will certainly exaggerate the scale mismatch. Simulating 1000 points of input–output data for such a system gives an information matrix,

$$\begin{pmatrix} 0.344 \times 10^{-5} & 0.343 \times 10^{-5} & 0.289 \times 10^{-13} & 0.186 \times 10^{-2} \\ 0.343 \times 10^{-5} & 0.345 \times 10^{-5} & 0.289 \times 10^{-13} & 0.142 \times 10^{-2} \\ 0.289 \times 10^{-13} & 0.289 \times 10^{-13} & 0.323 \times 10^{-21} & 0.132 \times 10^{-10} \\ 0.186 \times 10^{-2} & 0.142 \times 10^{-2} & 0.132 \times 10^{-10} & 0.389 \times 10^{3} \end{pmatrix}$$

with singular values 389.183, $4.843\,77 \times 10^{-6}$, $1.502\,26 \times 10^{-8}$ and $1.058\,79 \times 10^{-22}$. The condition number of this matrix is 3.676×10^{24} and the determinant is 3.0×10^{-33}. In order to see what the effect of this sort of condition is, the inverse of the matrix was computed using the numerically stable LU decomposition in single precision in FORTRAN. When the product of the matrix and inverse was computed, the result was

$$\begin{pmatrix} 1.000 & 0.000 & -55.34 & 0.000 \\ 0.000 & 1.000 & -15.75 & 0.000 \\ 0.000 & 0.000 & 1.000 & 0.000 \\ 0.000 & -0.020 & 8192.0 & 1.000 \end{pmatrix}$$

so the inverse is seriously in error. If the information matrix is scaled, the singular values become 2826.55, 1001.094, 177.984 and 177.984, giving a condition number of 608.0 and a determinant of 2.34×10^{9}. The inverse was computed

and the check matrix was

$$\begin{pmatrix} 1.000 & 0.000 & 0.000 & 0.000 \\ 0.000 & 1.000 & 0.000 & 0.000 \\ 0.000 & 0.000 & 1.000 & 0.000 \\ 0.000 & 0.000 & 0.000 & 1.000 \end{pmatrix}$$

as required. This example shows that without appropriate scaling, the normal equations approach can fail due to condition problems. Scaling also produces marked improvements if the other least-squares techniques are used.

6.8 NARMAX modelling

All the discussion so far has concerned linear systems. This does not constitute a restriction. The models described are all *linear in the parameters* so linear least-squares methods suffice. The models can be extended to nonlinear systems without changing the algorithm as will be seen. Arguably the most versatile approach to nonlinear discrete-time systems is the NARMAX (nonlinear auto-regressive moving average with eXogenous inputs) methodology which has been developed over a considerable period of time by S A Billings and numerous co-workers. An enormous body of work has been produced; only the most basic overview can be given here. The reader is referred to the original references for more detailed discussions, notably [59, 60, 149, 161, 162].

The extension of the previous discussions to nonlinear systems is straightforward. Consider the Duffing oscillator represented by

$$m\ddot{y} + c\dot{y} + ky + k_3 y^3 = x(t) \tag{6.94}$$

i.e. the linear system of (6.2) augmented by a cubic term. Assuming the simplest prescriptions for approximating the derivatives as before, one obtains, in discrete time,

$$y_i = a_1 y_{i-1} + a_2 y_{i-2} + b_1 x_{i-1} + c y_{i-1}^3 \tag{6.95}$$

where a_1,a_2 and b_1 are unchanged from (6.7) and

$$c = \frac{\Delta t^2 k_3}{m}. \tag{6.96}$$

The model (6.95) is now termed a NARX (nonlinear ARX) model. The regression function $y_i = F(y_{i-1}, y_{i-2}; x_{i-1})$ is now nonlinear; it contains a cubic term. However, the model is still linear in the parameters which have to be estimated, so *all* of the methods previously discussed still apply.

If *all* terms of order three or less were included in the model structure, i.e. $(y_{i-2})^2 x_{i-1}$ etc a much more general model would be obtained (these more complicated terms often arise, particularly if nonlinear damping is present):

$$y_i = F^{(3)}(y_{i-1}, y_{i-2}; x_{i-1}) \tag{6.97}$$

(the superscript denotes the highest-order product terms) which would be sufficiently general to represent the behaviour of any dynamical systems with nonlinearities up to third order, i.e. containing terms of the form $\dot{y}^3$, $\dot{y}^2 y$ etc.

The most general polynomial NARX model (including products of order $\leq n_p$) is denoted by

$$y_i = F^{(n_p)}(y_{i-1}, \ldots, y_{i-n_y}; x_{i-1}, \ldots, x_{i-n_x}). \tag{6.98}$$

It has been proved in the original papers by Leontaritis and Billings [161, 162], that under very mild assumptions, any input–output process has a representation by a model of the form (6.98). If the system nonlinearities are polynomial in nature, this model will represent the system well for all levels of excitation. If the system nonlinearities are not polynomial, they can be approximated arbitrarily accurately by polynomials over a given range of their arguments (Weierstrass approximation theorem [228]). This means that the system can be accurately modelled by taking the order n_p high enough. However, the model would be input-sensitive as the polynomial approximation required would depend on the data. This problem can be removed by including non-polynomial terms in the NARX model as described in [33].

For example, consider the equation of motion of the forced simple pendulum

$$\ddot{y} + \sin y = x(t) \tag{6.99}$$

or, in discrete time,

$$y_i = a_1 y_{i-1} + a_2 y_{i-2} + b_1 x_{i-1} + c \sin(y_{i-1}). \tag{6.100}$$

The most compact model of this system will be obtained by including a basis term $\sin(y_{i-1})$ rather than approximating by a polynomial in y_{i-1}.

The preceding analysis unrealistically assumes that the measured data are free of noise—this condition is relaxed in the following discussion. However, as before, it is assumed that the noise signal $\zeta(t)$ is additive on the output signal $y(t)$. This constituted no restriction when the system was assumed to be linear but is generally invalid for a nonlinear system. As shown later, if the system is nonlinear the noise process can be very complex; multiplicative noise terms with the input and output are not uncommon, but can be easily accommodated by the algorithms described earlier and in much more detail in [161, 162, 149, 60].

Under the previous assumption, the measured output has the form

$$y(t) = y_c(t) + \zeta(t). \tag{6.101}$$

where $y_c(t)$ is again the 'clean' output from the system. If the underlying system is the Duffing oscillator of equation (6.94), the equation satisfied by the measured data is now

$$m\ddot{y} + c\dot{y} + ky + k_3 y^3 - m\ddot{\zeta} - c\dot{\zeta} - k\zeta - k_3(\zeta^3 + 3y^2\zeta + 3y\zeta^2 = x(t) \tag{6.102}$$

and the corresponding discrete-time equation will contain terms of the form ζ_{i-1}, ζ_{i-2}, $\zeta_{i-1}y_{i-1}^2$ etc. Note that even simple additive noise on the output introduces cross-product terms if the system is nonlinear. Although these terms all correspond to unmeasurable states they must be included in the model. If they are ignored the parameter estimates will generally be biased. The system model (6.98) is therefore extended again by the addition of the noise model and takes the form

$$y_i = F^{(3)}(y_{i-1}, y_{i-2}; x_{i-2}; \zeta_{i-1}, \zeta_{i-2}) + \zeta_i. \tag{6.103}$$

The term 'moving-average' referring to the noise model should now be understood as a possibly nonlinear regression on past values of the noise. If a general regression on a fictitious uncorrelated noise process $e(t)$ is incorporated, one obtains the final general form

$$y_i = F^{(n_p)}(y_{i-1}, \ldots, y_{i-n_y}; x_{i-1}, \ldots, x_{i-n_x}; e_{i-1}, \ldots, e_{i-n_e}) + e_i. \tag{6.104}$$

This type of model is the generic NARMAX model.

A completely parallel theory has been developed for the more difficult case of time-series analysis where only measured outputs are available for the formulation of a model; this is documented in [244].

The structure detection can be carried out using the significance statistic of the NARMAX model—the ERR statistic (E.32)—in essentially two ways:

Forward selection. The model begins with no terms. All one-term models are fitted and the term which gives the greatest ERR, i.e. the term which accounts for the most signal variance is retained. The process is iterated, at each step including the term with greatest ERR and is continued until an acceptable model error is obtained.

Backward selection. The model begins with all *terms* and at the first step, the term with smallest ERR is deleted. Again the process is iterated until the accepted error is obtained.

Forward selection is usually implemented as it requires fitting smaller models. To see how advantageous this might be, note that the number of terms in a generic NARMAX model is roughly

$$\sum_{i=0}^{n_p} \frac{(n_y + n_x + n_e)^{n_p}}{n_p!} \tag{6.105}$$

with the various lags etc as previously defined.

6.9 Model validity

Having obtained a NARMAX model for a system, the next stage in the identification procedure is to determine if the structure is correct and the

parameter estimates are unbiased. It is important to know if the model has successfully captured the system dynamics so that it will provide good predictions of the system output for different input excitations, or if it has simply fitted the model to the data; in which case it will be of little use since it will only be applicable to one data set. Three basic tests of the validity of a model have been established [29], they are now described in increasing order of stringency. In the following, y_i denotes a measured output while $\hat{y}_i$ denotes an output value predicted by the model.

6.9.1 One-step-ahead predictions

Given the NARMAX representation of a system

$$y_i = F^{(n_p)}(y_{i-1}, \ldots, y_{i-n_y}; x_{i-1}, \ldots, x_{i-n_x}; e_{i-1}, \ldots, e_{i-n_e}) + e_i \quad (6.106)$$

the one-step-ahead (OSA) prediction of y_i is made using measured values for all past inputs *and* outputs. Estimates of the residuals are obtained from the expression $\hat{e}_i = y_i - \hat{y}_i$, i.e.

$$\hat{y}_i = F^{(n_p)}(y_{i-1}, \ldots, y_{i-n_y}; x_{i-1}, \ldots, x_{i-n_x}; \hat{e}_{i-1}, \ldots, \hat{e}_{i-n_e}). \quad (6.107)$$

The OSA series can then be compared to the measured outputs. Good agreement is clearly a necessary condition for model validity.

In order to have an objective measure of the goodness of fit, the normalized mean-square error (MSE) is introduced; the definition is

$$\text{MSE}(\hat{y}) = \frac{100}{N\sigma_y^2}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \quad (6.108)$$

where the caret denotes an estimated quantity. This MSE has the following useful property; if the mean of the output signal $\overline{y}$ is used as the model, i.e. $\hat{y}_i = \overline{y}$ for all i, the MSE is 100.0, i.e.

$$\text{MSE}(\hat{y}) = \frac{100}{N\sigma_y^2}\sum_{i=1}^{N}(y_i - \overline{y})^2 = \frac{100}{\sigma_y^2}\cdot\sigma_y^2 = 100. \quad (6.109)$$

Experience shows that an MSE of less than 5.0 indicates good agreement while one of less than 1.0 reflects an excellent fit.

6.9.2 Model predicted output

In this case, the inputs are the only measured quantities used to generate the model output, i.e.

$$\hat{y}_i = F^{(n_p)}(\hat{y}_{i-1}, \ldots, \hat{y}_{i-n_y}; x_{i-1}, \ldots, x_{i-n_x}; 0, \ldots, 0). \quad (6.110)$$

The zeroes are present because the prediction errors will not generally be available when one is using the model to predict output. In order to avoid a misleading transient at the start of the record for $\hat{y}$, the first n_y values of the measured output are used to start the recursion. As before, the estimated outputs must be compared with the measured outputs, with good agreement a necessary condition for accepting the model. It is clear that this test is stronger than the previous one; in fact the OSA predictions can be excellent in some cases when the model-predicted output (MPO) shows complete disagreement with the measured data.

6.9.3 Correlation tests

These represent the most stringent of the validity checks. The appropriate reference is [34]. The correlation function $\phi_{uv}(k)$ for two sequences of data u_i and v_i is defined as usual by

$$\phi_{uv} = E(u_i v_{i+k}) \approx \frac{1}{N-k} \sum_{i=1}^{N-k} u_i v_{i+k}. \tag{6.111}$$

In practice, normalized estimates of all the previous correlation functions are obtained using

$$\hat{\phi}_{uv}(k) = \frac{\frac{1}{N-k}\sum_{i=1}^{N-k} u_i v_{i+k}}{\{E(u_i^2)E(v_i^2)\}^{\frac{1}{2}}}, \quad k \geq 0 \tag{6.112}$$

with a similar expression for $k < 0$. The normalized expression is used because it allows a simple expression for the 95% confidence interval for a zero result, namely $\pm 1.96/\sqrt{(N)}$. The confidence limits are required because the estimate of ϕ_{uv} is made only on a finite set of data; as a consequence it will never be truly zero. The model is therefore considered adequate if the correlation functions described earlier fall within the 95% confidence limits. These limits are indicated by a broken line when the correlation functions are shown later.

For a linear system it is shown in [34], that necessary conditions for model validity are

$$\phi_{ee}(k) = \delta_{0k} \tag{6.113}$$

$$\phi_{xe}(k) = 0, \quad \forall k. \tag{6.114}$$

The first of these conditions is true only if the residual sequence e_i is a white-noise sequence. It is essentially a test of the adequacy of the noise model whose job it is to reduce the residuals to white noise. If the noise model is correct, the system parameters should be free from bias. The second of these conditions states that the residual signal is uncorrelated with the input sequence x_i, i.e. the model has completely captured the component of the measured output which is correlated with the input. Another way of stating this requirement is that the residuals should be unpredictable from the input.

In the case of a nonlinear system it is sometimes possible to satisfy these requirements even if the model is invalid. It is shown in [34] that an exhaustive test of the fitness of a nonlinear model requires the evaluation of three additional correlation functions. The extra conditions are

$$\phi_{e(ex)}(k) = 0, \quad \forall k \geq 0 \tag{6.115}$$

$$\phi_{x^{2\prime}e}(k) = 0, \quad \forall k \tag{6.116}$$

$$\phi_{x^{2\prime}e^2}(k) = 0, \quad \forall k. \tag{6.117}$$

The prime which accompanies the x^2 indicates that the mean has been removed.

6.9.4 Chi-squared test

One final utility can be mentioned. If the model fails the validity tests one can compute a statistic η as in [60] for a given term *not* included in the model to see if it should be present. The test is specifically developed for nonlinear systems and is based on chi-squared statistics. A number of values of the statistic for a specified term are plotted together with the 95% confidence limits. If values of the statistic fall *outside* the limits, the term should be included in the model and it is necessary to re-estimate parameters accordingly. Examples of all the test procedures described here will be given in the following section.

6.9.5 General remarks

Strict model validation requires that the user have a separate set of testing data from that used to form the model. This is to make sure that the identification scheme has learnt the underlying model and not simply captured the features of the data set. The most rigorous approach demands that the testing data have a substantially different form from the estimation data. Clearly different amplitudes can be used. Also, different excitations can be used. For example if the model is identified from data from Gaussian white-noise excitation, the testing data could come from PRBS (pseudo-random binary sequence) or chirp.

6.10 Correlation-based indicator functions

Having established the normalized correlation functions in the last section, it is an opportune moment to mention two simple correlation tests which can signal nonlinearity by manipulating measured time data. If records of both input x and output y are available, it can be shown that the correlation function

$$\phi_{x^2y'}(k) = E[x_i y'_{i+k}] \tag{6.118}$$

vanishes for all τ if and only if the system is linear [35]. The prime signifies that the mean has been removed from the signal.

If only sampled outputs are available, it can be shown that under certain conditions [31], the correlation function

$$\phi_{y'y'^2}(k) = E[y'_{i+k}(y'_i)^2] \tag{6.119}$$

is zero for all k if and only if the system is linear. In practice, these functions will never be identically zero; however, confidence intervals for a zero result can be calculated straightforwardly. As an example the correlation functions for acceleration data from an offset bilinear system at both low and high excitation are shown in figure 6.7; the broken lines are the 95% confidence limits for a zero result. The function in figure 6.7(*b*) indicates that the data from the high excitation test arise from a nonlinear system. The low excitation test did not excite the nonlinearity and the corresponding function (figure 6.7(*a*)) gives a null result as required.

There are a number of caveats associated with the latter function. It is a necessary condition that the third-order moments of the input vanish and all even-order moments exist. This is not too restrictive in practice; the conditions hold for a sine wave or a Gaussian noise sequence for example. More importantly, the function (6.119) as it stands only detects even nonlinearity, e.g. quadratic stiffness. In practice, to identify odd nonlinearity, the input signal should contain a d.c. offset, i.e. a non-zero mean value. This offsets the output signal and adds an even component to the nonlinear terms, i.e.

$$y^3 \longrightarrow (y + \overline{y})^3 = y^3 + 3y^2\overline{y} + 3y\overline{y}^2 + \overline{y}^3. \tag{6.120}$$

A further restriction on (6.119) is that it cannot detect odd damping nonlinearity[7], as it is not possible to generate a d.c. offset in the velocity to add an odd component to the nonlinearity. Figure 6.8 shows the correlation function for a linear system and a system with Coulomb friction, the function fails to signal nonlinearity. (Note that the coherence function in the latter case showed a marked decrease which indicated strong nonlinearity.)

6.11 Analysis of a simulated fluid loading system

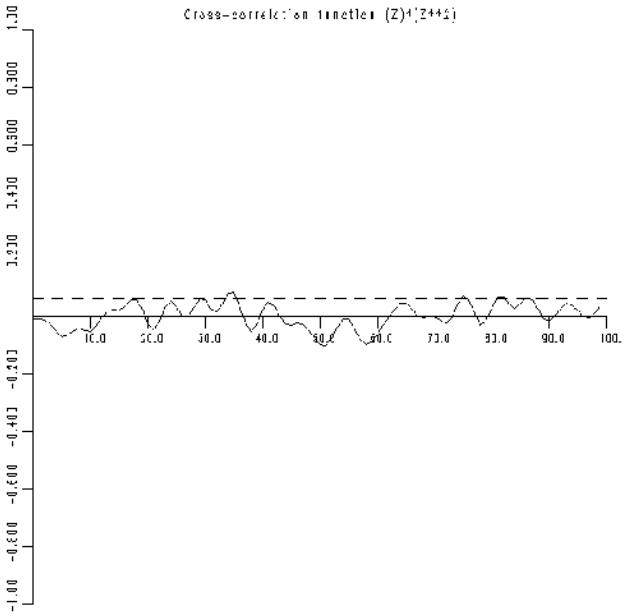
In order to demonstrate the concepts described in previous sections, the techniques are now applied to simulated data from the Morison equation, which is used to predict forces on offshore structures [192],

$$F(t) = \tfrac{1}{2}\rho D C_{\mathrm{d}} u|u| + \tfrac{1}{4}\pi\rho D^2 C_{\mathrm{m}}\dot{u} \tag{6.121}$$

where $F(t)$ is the force per unit axial length, $u(t)$ is the instantaneous flow velocity, ρ is water density and D is diameter; C_{d} and C_{m} are the dimensionless drag and inertia coefficients. The first problem is to determine an appropriate

[7] The authors would like to thank Dr Steve Gifford for communicating these results to them [112] and giving permission for their inclusion.

(a)



(b)

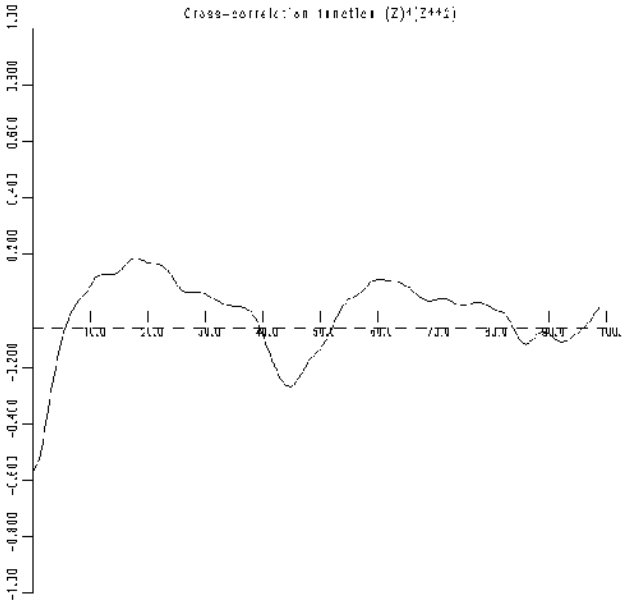
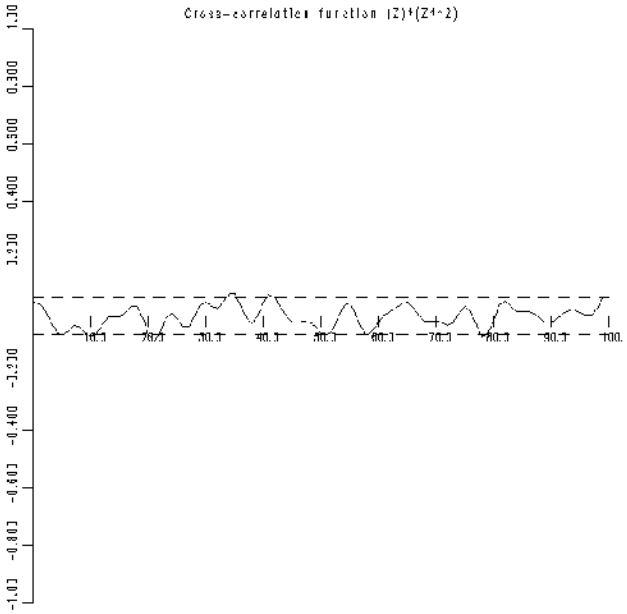


Figure 6.7. Correlation function for a bilinear system with the discontinuity offset in displacement: (*a*) low excitation; (*b*) high excitation.

(a)



(b)

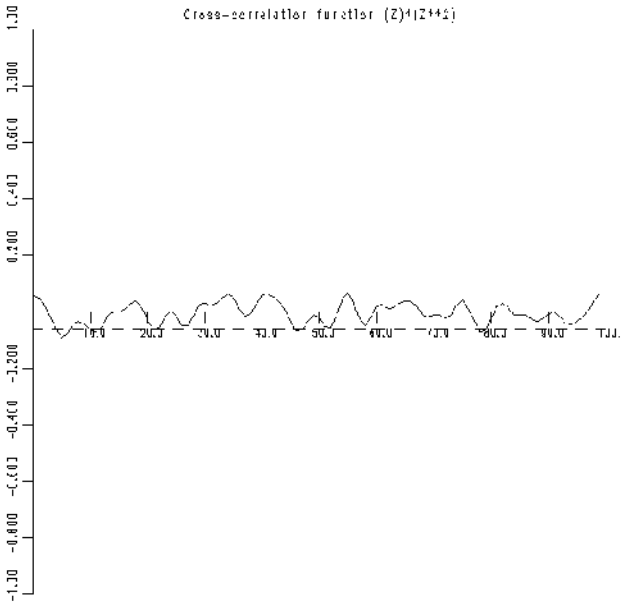


Figure 6.8. Correlation functions for: (*a*) linear system; (*b*) Coulomb friction system.

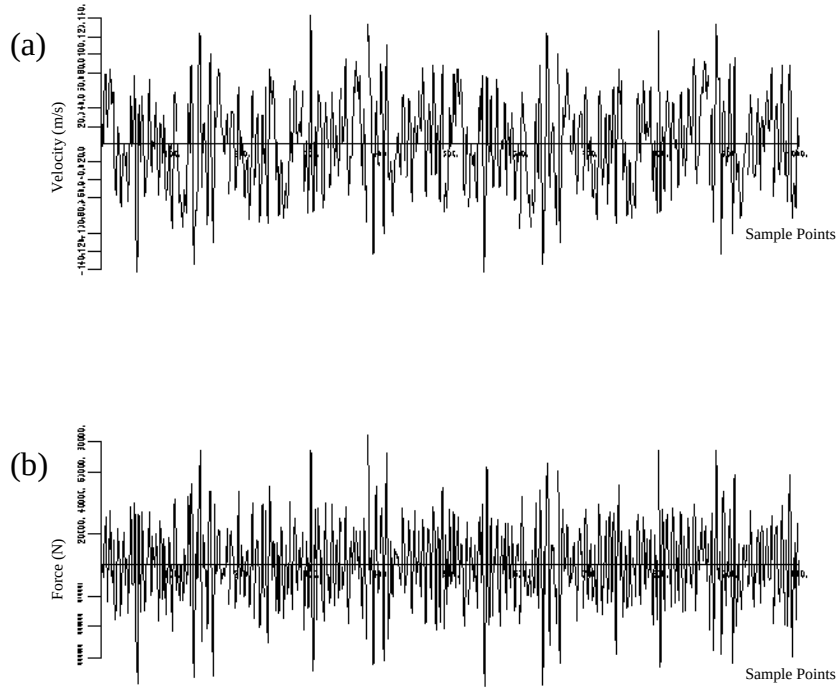


Figure 6.9. Simulated velocity and force signals for fluid loading study.

discrete-time form. The conditions $\rho = 1$, $D = 2$, $C_{\mathrm{d}} = \frac{3}{2}$ and $C_{\mathrm{m}} = 2$ are imposed giving the equation

$$F(t) = 2\pi\dot{u} + \tfrac{3}{2}u(t)|u(t)| \tag{6.122}$$

where $F(t)$ is the system output and $u(t)$ will be the input. Using the forward difference approximation to the derivative, the discrete form

$$F_i = \frac{2\pi}{\Delta t}(u_i - u_{i-1}) + \frac{3}{2}u_i|u_i| \tag{6.123}$$

is obtained. The basic form of the NARMAX procedures used here utilizes polynomial model terms. For the sake of simplicity, the $u|u|$ term in the simulation model is replaced by a cubic approximation

$$u_i|u_i| = \alpha u_i + \beta u_i^3 + \mathrm{O}(u_i^5). \tag{6.124}$$

The coefficients α and β are obtained by a simple least-squares argument. Substituting (6.124) into (6.123) yields the final NARMAX form of Morison's equation

$$F_i = \left(\frac{3\alpha}{2} + \frac{2\pi}{\Delta t}\right)u_i - \frac{2\pi}{\Delta t}u_{i-1} + \frac{3\beta}{2}u_i^3 \tag{6.125}$$

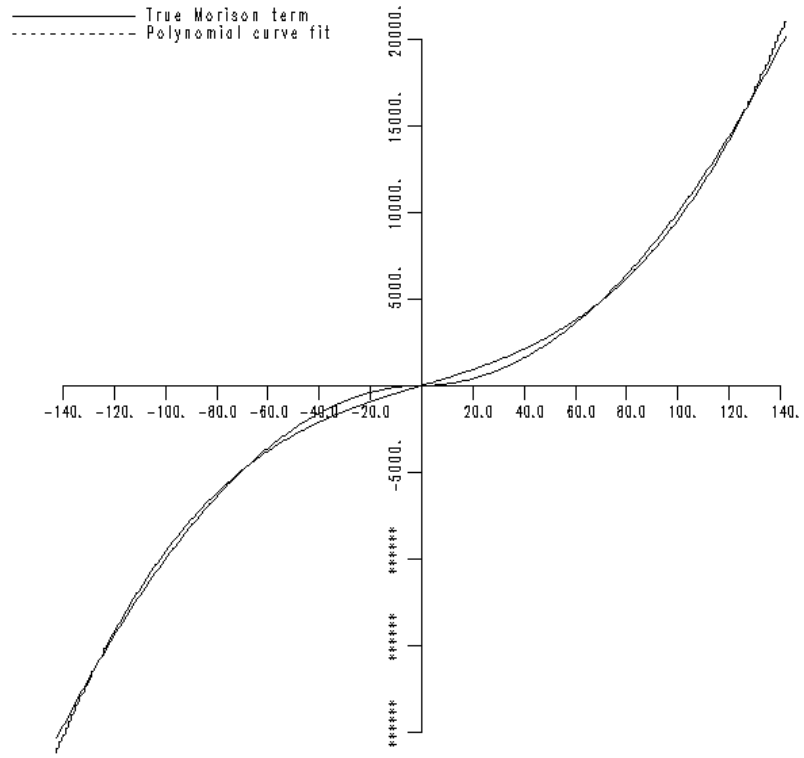


Figure 6.10. Comparison between $u|u|$ and cubic approximation.

or

$$F_i = a_1 u_i + a_2 u_{i-1} + a_3 u_i^3. \tag{6.126}$$

This is the model which was used for the simulation of force data. A velocity signal was used which had a uniform spectrum in the range 0–20 Hz. This was obtained by generating 50 sinusoids each with an amplitude of 10.0 units spaced uniformly in frequency over the specified range; the phases of the sinusoids were taken to be random numbers uniformly distributed on the interval $[0, 2\pi]$. The sampling frequency was chosen to be 100 Hz, giving five points per cycle of the highest frequency present. The amplitude for the sinusoids was chosen so that the nonlinear term in (6.126) would contribute approximately 13% to the total variance of F. The simulated velocity and force data are displayed in figure 6.9. In order to show the accuracy of the cubic approximation (6.124) over the range of velocities generated, the function $u|u|$ is plotted in figure 6.10 together with the cubic curve fit; the agreement is very good so a fifth-order term in the NARMAX

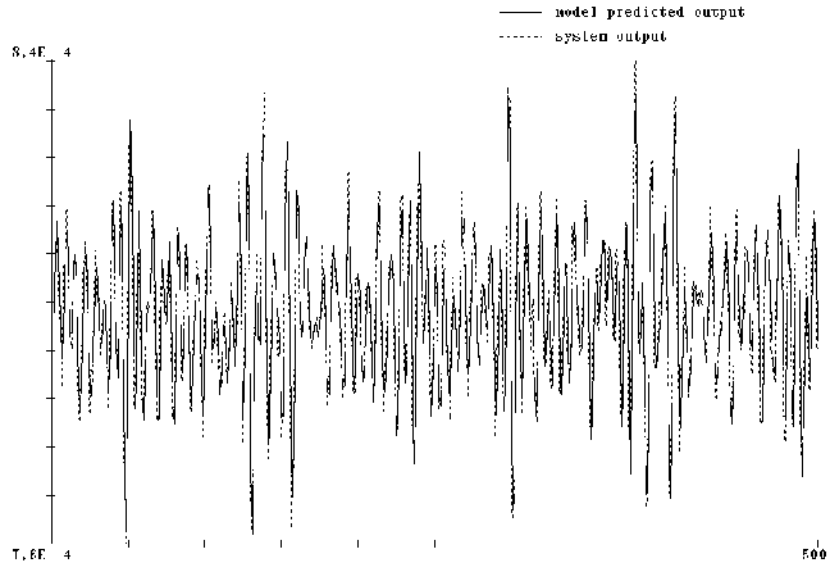


Figure 6.11. Fluid-loading study: model predicted output for linear process model—no noise model.

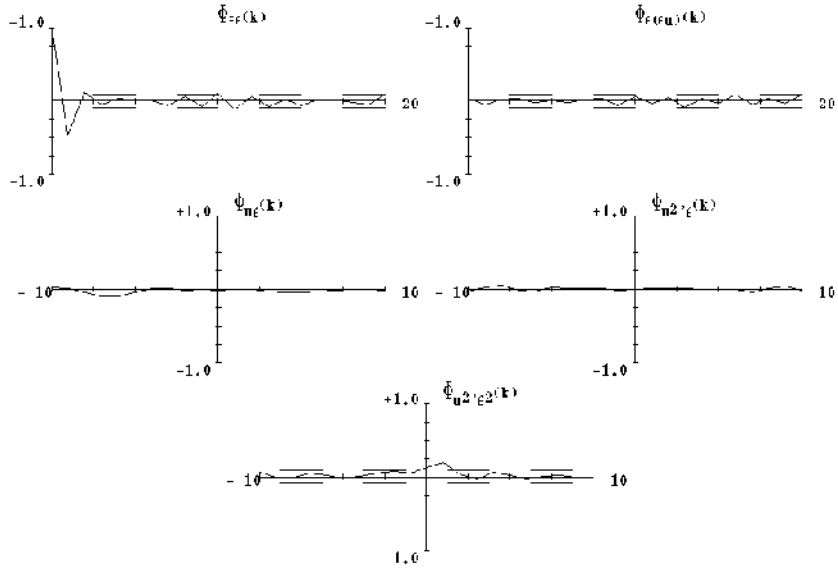


Figure 6.12. Fluid-loading study: correlation tests for linear process model—no noise model.

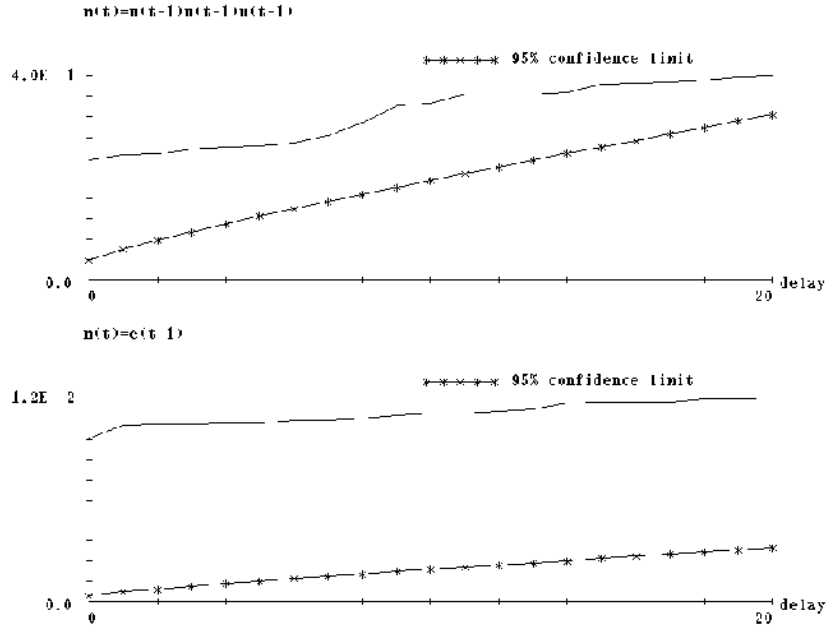


Figure 6.13. Fluid-loading study: chi-squared tests for linear process model—no noise model.

model is probably not needed. The values of the exact NARMAX coefficients for the data were $a_1 = 697.149$, $a_2 = -628.32$ and $a_3 = 0.007\,67$.

In order to demonstrate fully the capabilities of the procedures, a *coloured* noise signal ζ was added to the force data. The noise model chosen was

$$\zeta_i = 0.222\,111 e_{i-1} - e_{i-2} + e_{i-3} \tag{6.127}$$

where e_i was a Gaussian white-noise sequence. The variance of $e(t)$ was chosen in such a way that the overall signal-to-noise ratio $\sigma_{\mathrm{F}}/\sigma_{\zeta}$ would be equal to 5.0. This corresponds to the total signal containing approximately 17% noise. This is comparatively low, a benchtest study described in [270] showed that the NARMAX procedures could adequately identify Morison-type systems with the signal-to-noise ratio as high as unity.

The first attempt to model the data assumed the linear structure

$$F_i = a_1 u_i + a_2 u_{i-1}. \tag{6.128}$$

The resulting parameter estimates were $a_1 = 745.6$ and $a_2 = -631.18$ with standard deviations $\sigma_{\mathrm{a}_1} = 7.2$ and $\sigma_{\mathrm{a}_2} = 7.2$. The estimated value of a_1 is 7.0 standard deviations away from the true parameter; this indicates bias. The reason for the overestimate is that the u_i^3 term which should have been included in the

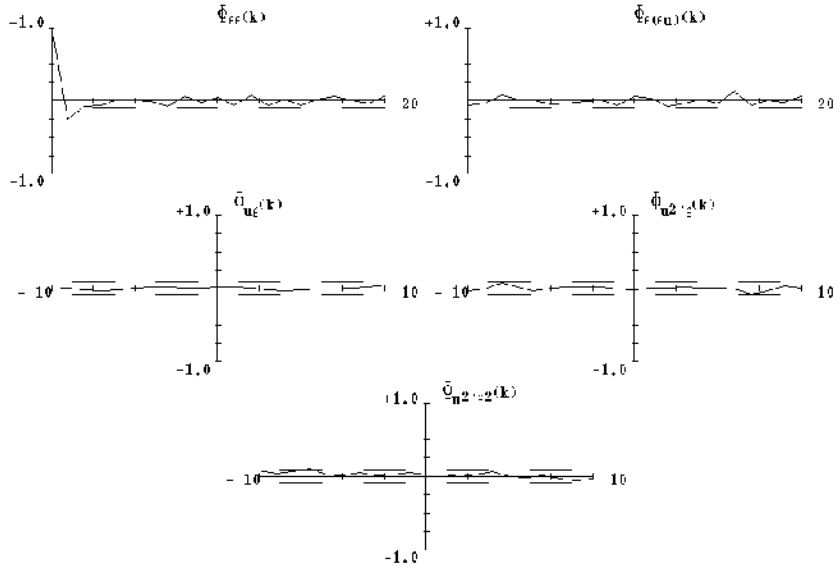


Figure 6.14. Fluid-loading study: correlation tests for nonlinear process model with linear noise model.

Table 6.1. Parameter table for Morison model of Christchurch Bay data.

Model term	Parameter	ERR	Standard deviation
u_i	$0.880\,80\mathrm{e}+03$	$0.187\,64\mathrm{e}-01$	$0.203\,44\mathrm{e}+02$
u_{i-1}	$-0.845\,93\mathrm{e}+03$	$0.385\,39\mathrm{e}+00$	$0.200\,08\mathrm{e}+02$
u_i^3	$0.339\,83\mathrm{e}+02$	$0.381\,32\mathrm{e}+00$	$0.219\,13\mathrm{e}+01$

model is strongly correlated with the u_i term; as a consequence the NARMAX model can represent some of the nonlinear behaviour by adding an additional u_i component. It is because of effects like this that data from nonlinear systems can sometimes be adequately represented by linear models. However, such models will be input-dependent as changing the level of input would change the amount contributed by the nonlinear term and hence the estimate of a_1.

The OSA predictions for the model were observed to be excellent. The MPO, shown in figure 6.11, also agreed well with the simulation data. However, if the correlation tests are consulted (figure 6.12), both ϕ_{ee} and $\phi_{u^2 e}$ show excursions outside the 95% confidence interval. The first of these correlations indicates that the system noise is inadequately modelled, the second shows that the model does not take nonlinear effects correctly into account. This example shows clearly the

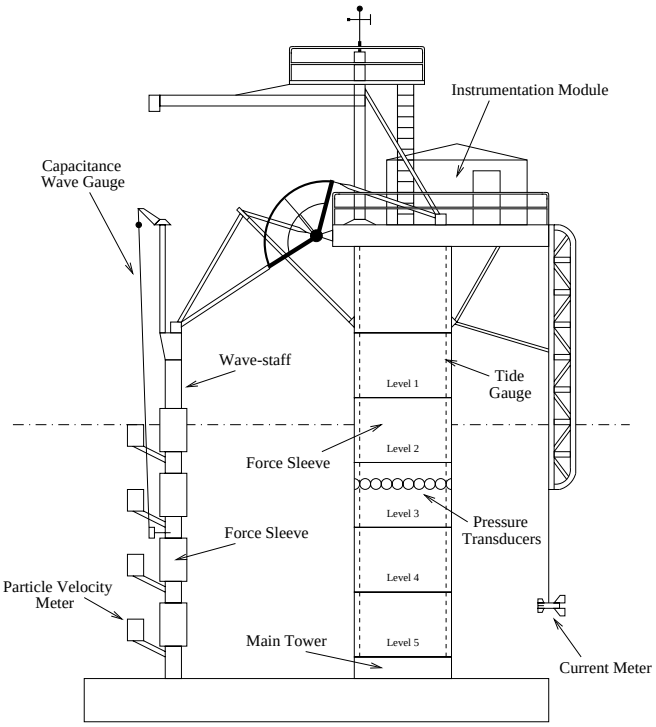


Figure 6.15. Schematic diagram of the Christchurch Bay tower.

utility of the correlation tests. Figure 6.13 shows the results of chi-squared tests on the terms u_i^3 and e_{i-1}; in both cases the plots are completely outside the 95% confidence interval; this shows that these terms should have been included in the model. A further test showed that the e_{i-2} term should also have been included.

In the second attempt to identify the system, the correct process model was assumed:

$$F_i = a_1 u_i + a_2 u_{i-1} + a_3 u_i^3 \tag{6.129}$$

but no noise model was included. The resulting parameter estimates were $a_1 = 693.246$, $a_2 = -628.57$ and $a_3 = 0.079$ with standard deviations $\sigma_{a_1} = 9.1$, $\sigma_{a_2} = 6.7$ and $\sigma_{a_3} = 0.0009$. The inclusion of the nonlinear term in the model has removed the principal source of the bias on the estimate of a_1 and all estimates are now within one standard deviation of the true results. The one-step-ahead predictions and model predicted outputs for this model showed no visible improvements over the linear model. However, the correlation test showed $\phi_{u^2 e}$ to be within the confidence interval, indicating that the nonlinear behaviour is now correctly captured by the model. As expected $\phi_{ee}(k)$ is still non-zero for $k > 0$ indicating that a noise model is required. This conclusion was reinforced

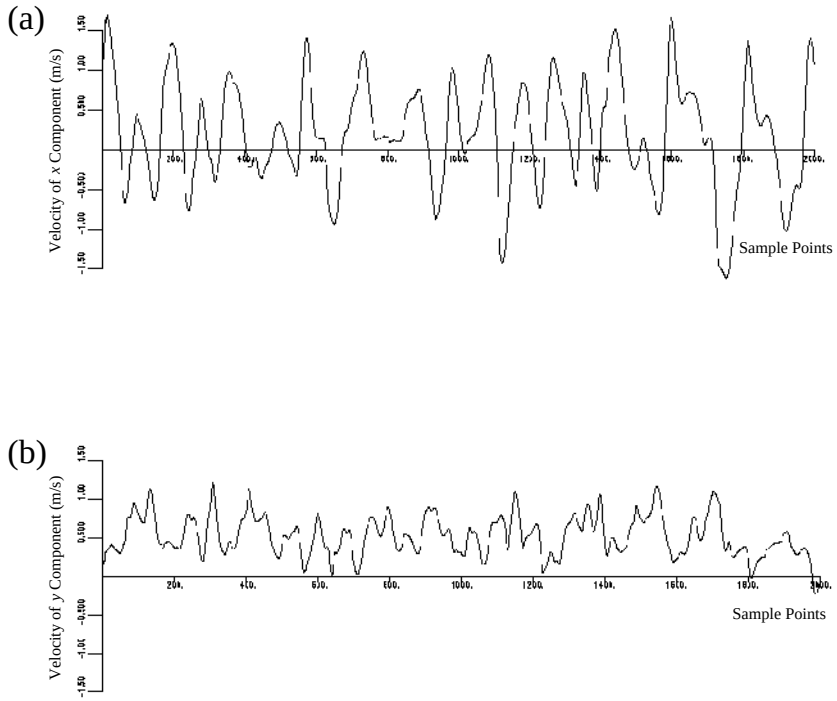


Figure 6.16. X- and Y-components of the velocity signal for a sample of Christchurch Bay data.

by the chi-squared tests for e_{i-1} and e_{i-2} which showed that these terms should be included.

The final attempt to model the system used the correct nonlinear structure and included a noise model with linear terms e_{i-1} and e_{i-2}. The correlation tests (figure 6.14) improved but still showed a slight excursion outside the confidence limits for $\phi_{ee}(k)$ at $k = 1$. Generally, if $\phi_{ee}(k)$ leaves the confidence interval at lag k, a term e_{i-k} should be included in the model. In this case the tests show that the term in e_{i-1} could be improved.

This simulation illustrates nicely the suitability of NARMAX procedures for the study of time data. More importantly it shows the need for the correlation tests; it is not sufficient to look at agreement between model predicted data and measured data. The estimation procedures can still allow a good representation of a given data set even if the model structure is wrong, simply by biasing the parameter estimates for the terms present. However, in this case the model is simply a curve fit to a specific data set and will be totally inadequate for prediction on different inputs.

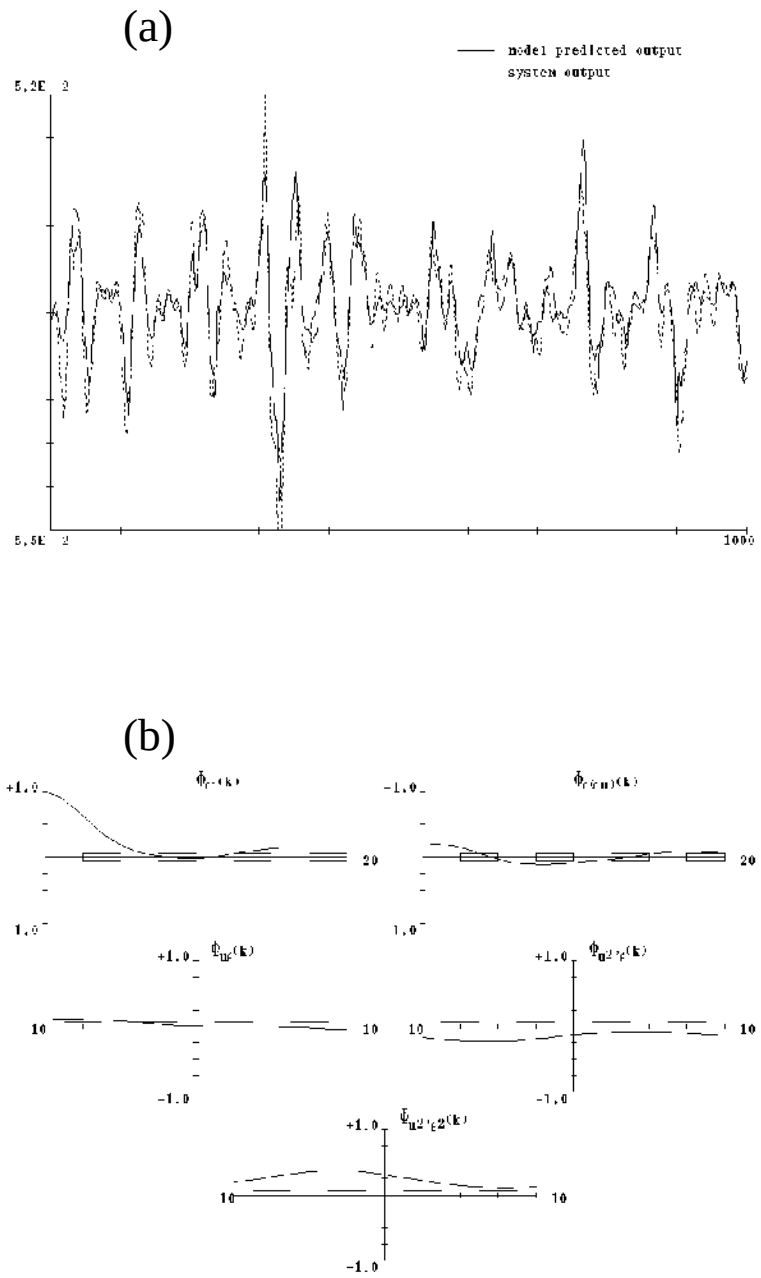


Figure 6.17. Discrete Morison equation model fit to the Christchurch Bay data: (*a*) model-predicted output; (*b*) correlation tests.

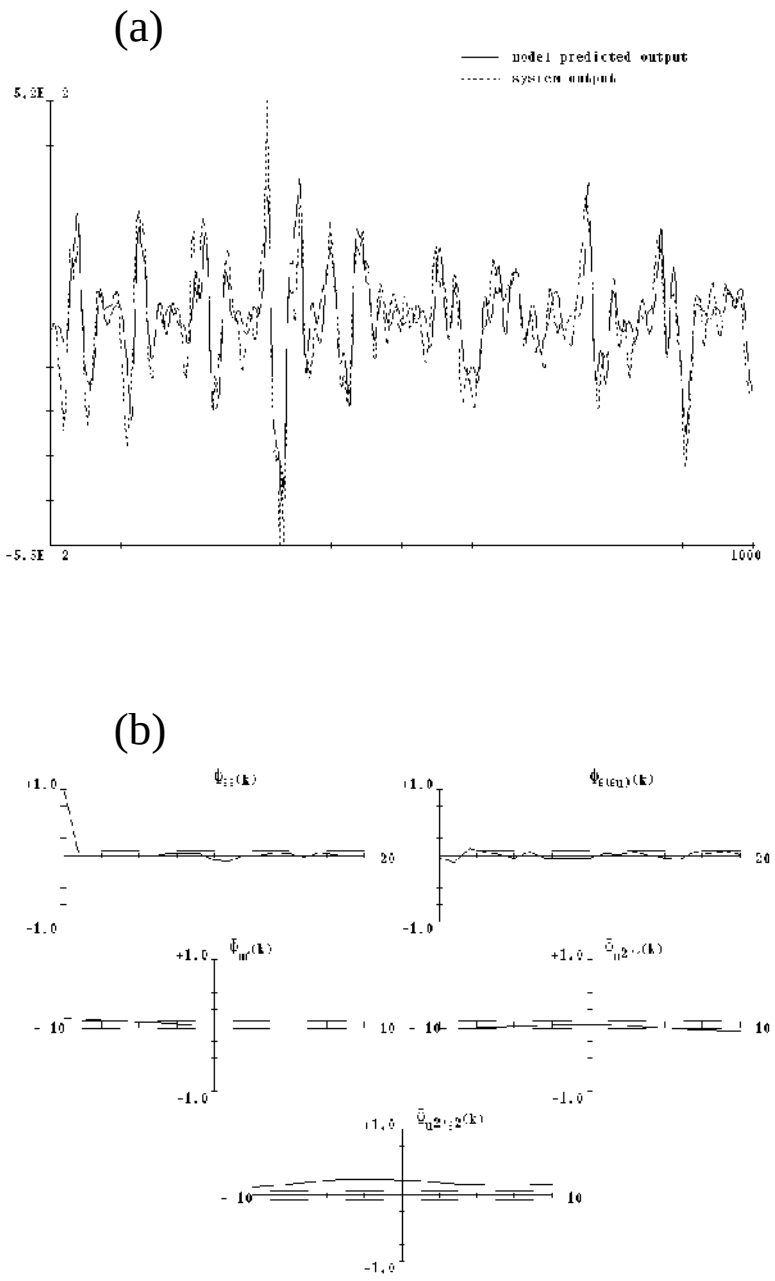


Figure 6.18. NARMAX model fit to the Christchurch Bay data: (*a*) model-predicted output; (*b*) correlation tests.

Table 6.2. Parameter table for NARMAX model of Christchurch Bay data.

Model term	Parameter
F_{i-1}	$0.198\mathrm{e}+01$
F_{i-2}	$-0.126\mathrm{e}+01$
F_{i-3}	$0.790\mathrm{e}-01$
F_{i-4}	$0.395\mathrm{e}+00$
F_{i-5}	$-0.328\mathrm{e}+00$
F_{i-6}	$0.111\mathrm{e}+00$
u_i	$0.119\mathrm{e}+03$
u_{i-1}	$-0.300\mathrm{e}+03$
u_{i-2}	$0.323\mathrm{e}+03$
u_{i-3}	$-0.155\mathrm{e}+03$
u_{i-4}	$-0.946\mathrm{e}+01$
u_{i-5}	$0.273\mathrm{e}+02$
F_{i-3}^2	$-0.193\mathrm{e}-03$
$F_{i-2}F_{i-5}$	$0.137\mathrm{e}-03$
F_{i-1}^3	$-0.232\mathrm{e}-05$
$F_{i-1}^2F_{i-4}$	$-0.193\mathrm{e}-05$
$F_{i-1}u_{i-4}^2$	$-0.221\mathrm{e}+00$
$F_{i-4}u_i^2$	$0.188\mathrm{e}+00$
$F_{i-3}u_iu_{i-4}$	$-0.457\mathrm{e}+00$
$F_{i-2}u_{i-3}^2$	$0.466\mathrm{e}+00$
$F_{i-1}F_{i-2}u_i$	$-0.731\mathrm{e}-03$
$F_{i-1}^2u_{i-4}$	$0.482\mathrm{e}-03$
$u_{i-3}u_{i-4}^2$	$0.437\mathrm{e}+02$
$u_iu_{i-4}^2$	$0.158\mathrm{e}+03$
$u_{i-1}u_{i-4}^2$	$-0.196\mathrm{e}+03$
F_{i-2}^3	$0.101\mathrm{e}-04$
$F_{i-1}F_{i-2}^2$	$-0.222\mathrm{e}-04$
$F_{i-1}^2F_{i-3}$	$0.483\mathrm{e}-05$
$F_{i-1}^2F_{i-2}$	$0.120\mathrm{e}-04$

6.12 Analysis of a real fluid loading system

In this section the NARMAX model structure is fitted to forces and velocities measured on the Christchurch Bay Tower which was constructed to test (amongst other things) fluid loading models in a real directional sea environment. The tower is shown in figure 6.15 and is described in considerable more detail in [39].

The tower was instrumented with pressure transducers and velocity meters. The data considered here were measured on the small diameter wave staff (Morison's equation is only really appropriate for slender members). Substantial wave heights were observed in the tests (up to 7 m) and the sea was directional with a prominent current. The velocities were measured with calibrated perforated

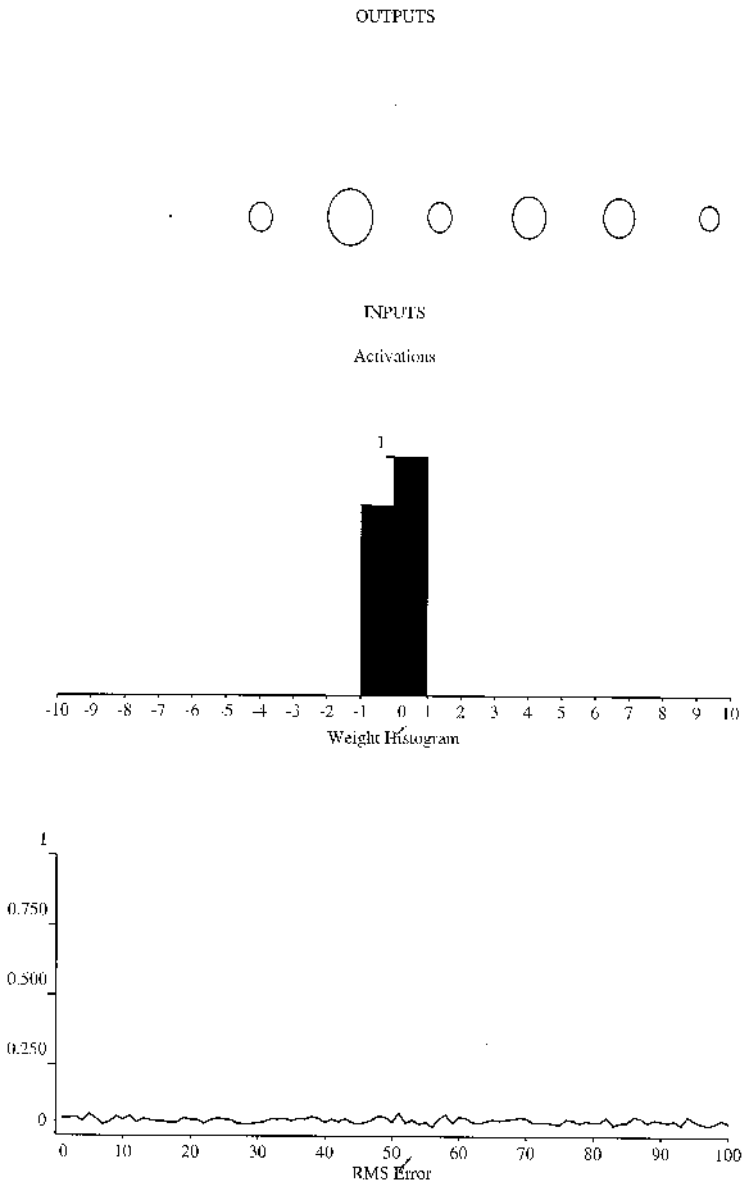


Figure 6.19. State of neural network after training on linear system identification problem: network outputs, weight histogram and rms error curve.

ball meters attached at a distance of 1.228 m from the cylinder axis. This will not give the exact velocity at the centre of the force sleeve unless waves are

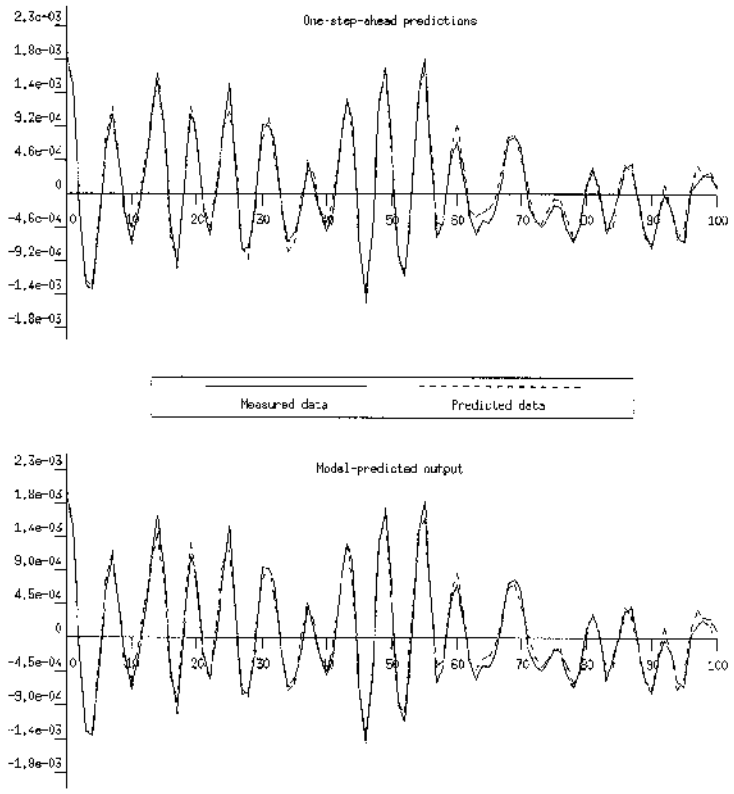


Figure 6.20. OSA and MPO predictions for linear system identification example using a neural network.

unidirectional with crests parallel to the line joining the velocity meter to the cylinder. This is called the Y-direction and the normal to this, the X-direction. The waves are, however, always varying in direction so data were chosen here from an interval when the oscillatory velocity in the X-direction was large and that in the Y-direction small. A sample of 1000 points fitting these criteria is shown in figure 6.16. It can be seen that the current is mainly in the Y-direction. In this case the velocity ball is upstream of the cylinder and interference by the wake on the ball will be as small as possible with this arrangement. Clearly the data are not of the same quality as those in the previous section and should provide a real test of the method.

As in the previous section, the discrete form of Morison's equation was fitted to the data to serve as a basis for comparison. The coefficients are presented in table 6.1. Note that the coefficients of u_i and u_{i-1} are almost equal and opposite indicating that they constitute the discretization of an inertia term $\dot{u}$. The MSE

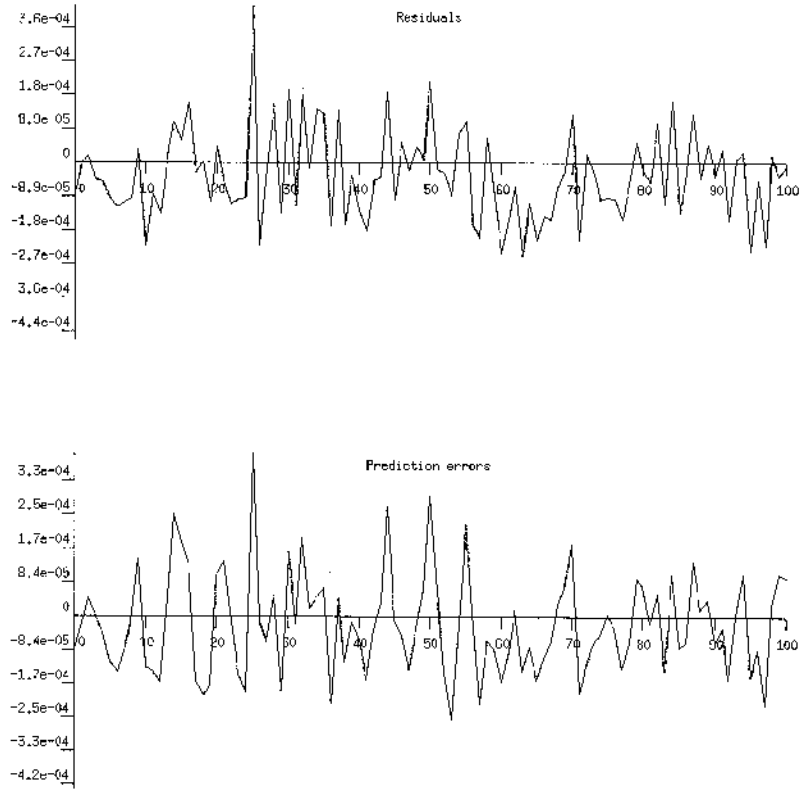


Figure 6.21. Residuals and prediction errors for linear system identification example using a neural network.

for the model is 21.43 which indicates significant disagreement with reality[8]. The MPO is shown in figure 6.17 together with the correlation tests. One concludes that the model is inadequate.

The data were then analysed using the structure detection algorithm to determine which terms should be included in the model. A linear noise model was included. The resulting model is given in table 6.2.

A complex model was obtained which includes terms with no clear physical interpretation. (This model is probably over-complex and could be improved by careful optimization. However, it suffices to illustrate the main points of the argument.) The fact that such a model is required can be offered in support of the conclusion that the inadequacy of Morison's equation is due to gross vortex shedding effects which can even be observed in simplified experimental

[8] In order to compare the effectiveness of the noise model, the MSE is computed here using the residuals instead of the prediction errors.

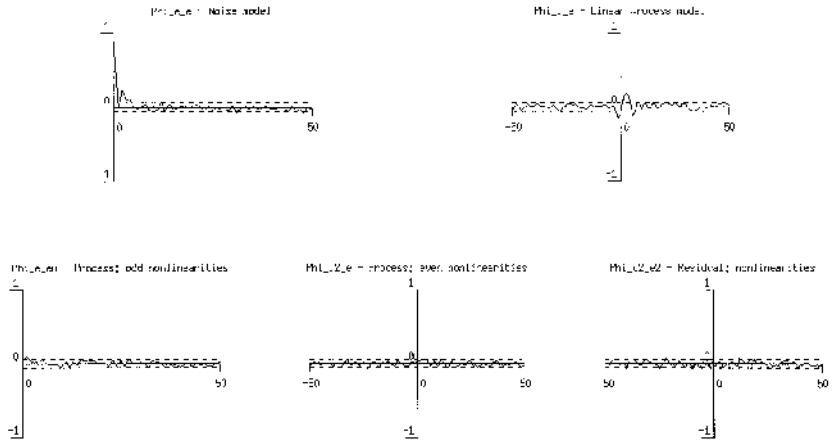


Figure 6.22. Correlation tests for linear system identification example using a neural network.

conditions [199]. The MPO and correlation tests are shown in figure 6.18. Although the validity tests show a great deal of improvement, the MPO appears to be worse. This is perfectly understandable; one of the effects of correlated noise (indicated by the function ϕ_{ee} in figure 6.17) is to bias the model coefficients so that the model fits the data rather than the underlying system. In this case the MPO is actually accounting for some of the system noise; this is clearly incorrect. When the noise model is added to reduce the noise to a white sequence, the unbiased model no longer predicts the noise component and the MPO appears to represent the data less well. This is one reason why the MSE adopted here makes use of the residual sequence e_i rather than the prediction errors ζ_i. In this case, the MSE is 0.75 which shows a marked improvement over the Morison equation. The fact that the final correlation function in figure 6.18 still indicates problems with the model can probably be attributed to the time-dependent phase relationship between input and output described earlier.

6.13 Identification using neural networks

6.13.1 Introduction

The problem of system identification in its most general form is the construction of the functional $S[\,]$ which maps the inputs of the system to the outputs. The problem has been simplified considerably in the discussion so far by assuming that a linear-in-the-parameters model with an appropriate structure can be used. Either an *a priori* structure is assumed or clever structure detection is needed. An alternative approach would be to construct a complete 'black-box' representation

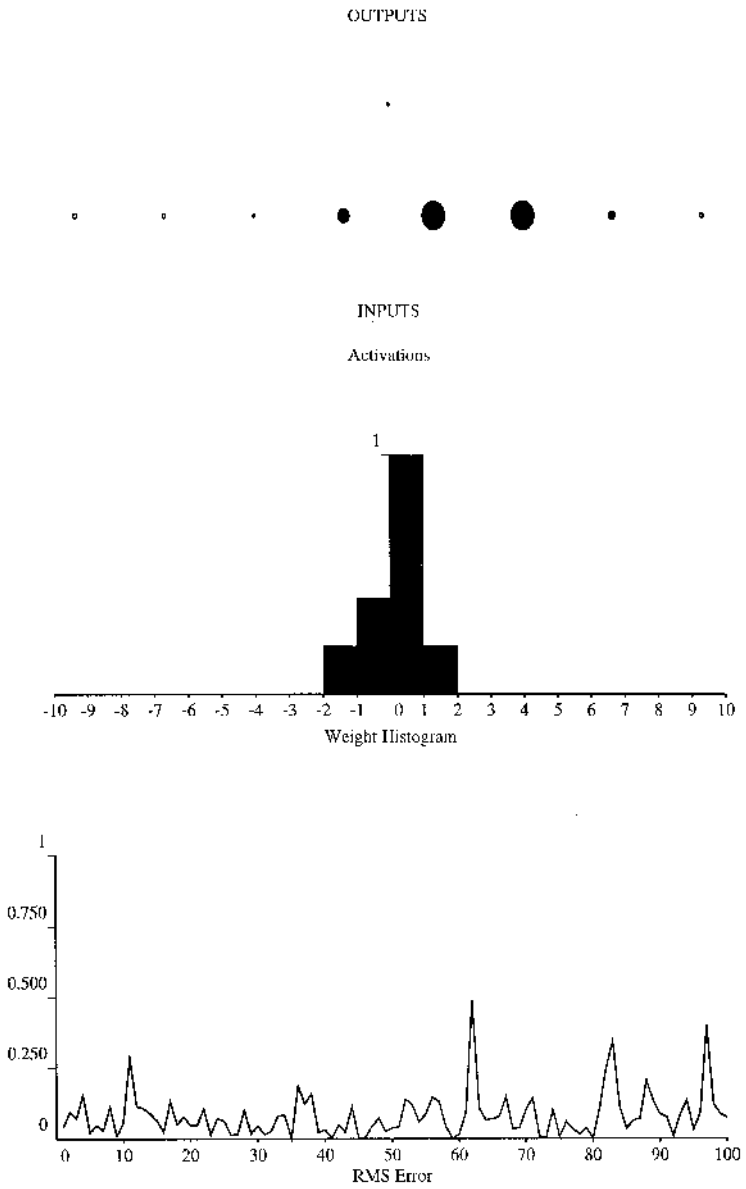


Figure 6.23. Final network state for the linear neural network model of the Duffing oscillator.

on the basis of the data alone. Artificial neural networks have come into recent prominence because of their ability to learn input–output relationships

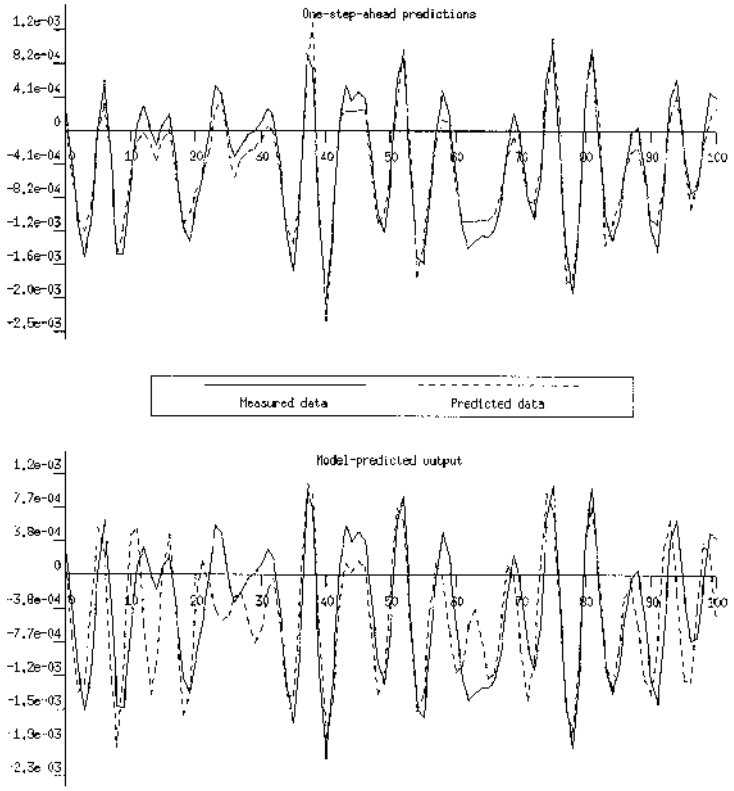


Figure 6.24. OSA and MPO predictions for the linear neural network model of the the Duffing oscillator.

by training on measured data and they appear to show some promise for the system identification problem. Appendix F gives a detailed discussion of the historical development of the subject, ending with descriptions of the most often used forms—the multi-layer perceptron (MLP) and radial basis function (RBF). In order to form a model with a neural network it is necessary to specify the form of the inputs and outputs; in the case of the MLP and RBF, the NARX functional form (6.98) is often used:

$$y_i = F(y_{i-1}, \ldots, y_{i-n_y}; x_{i-1}, \ldots, x_{i-n_x}) \tag{6.130}$$

except that the superscript n_p is omitted as the model is *not* polynomial. In the case of the MLP with a linear output neuron, the appropriate structure for a SDOF

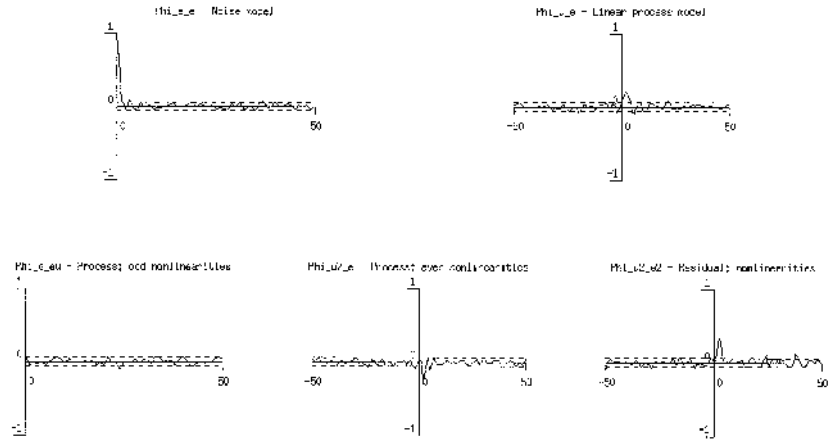


Figure 6.25. Correlation tests for the linear neural network model of the Duffing oscillator.

system is

$$y_i = s + \sum_{j=1}^{n_h} w_j \tanh\left(\sum_{k=1}^{n_y} v_{jk} y_{i-k} + \sum_{m=0}^{n_x-1} u_{jm} x_{i-m} + b_j\right) \tag{6.131}$$

or, if a nonlinear output neuron is used,

$$y_i = \tanh\left[s + \sum_{j=1}^{n_h} w_j \tanh\left(\sum_{k=1}^{n_y} v_{jk} y_{i-k} + \sum_{m=0}^{n_x-1} u_{jm} x_{i-m} + b_j\right)\right]. \tag{6.132}$$

For the RBF network

$$y_i = s + \sum_{j=1}^{n_h} w_j \exp\left\{-\frac{1}{2\sigma_j^2}\left[\sum_{k=1}^{n_y}(y_{i-k} - v_{jk})^2 + \sum_{m=0}^{n_x-1}(x_{i-m} - u_{jm})^2\right]\right\} + \underbrace{\sum_{j=1}^{ny} a_j y_{i-j} + \sum_{j=0}^{n_x-1} b_j x_{i-j}}_{\text{from linear connections}} \tag{6.133}$$

where the quantities v_{jk} and u_{jm} are the hidden node centres and the σ_i is the standard deviation or *radius* of the Gaussian at hidden node i. The first part of this expression is the standard RBF network.

Some of the earliest examples of the use of neural networks for system identification and modelling are the work of Chu *et al* [64] and Narendra and Parthasarathy [194]. Masri *et al* [179, 180] are amongst the first structural dynamicists to exploit the techniques. The latter work is interesting because it

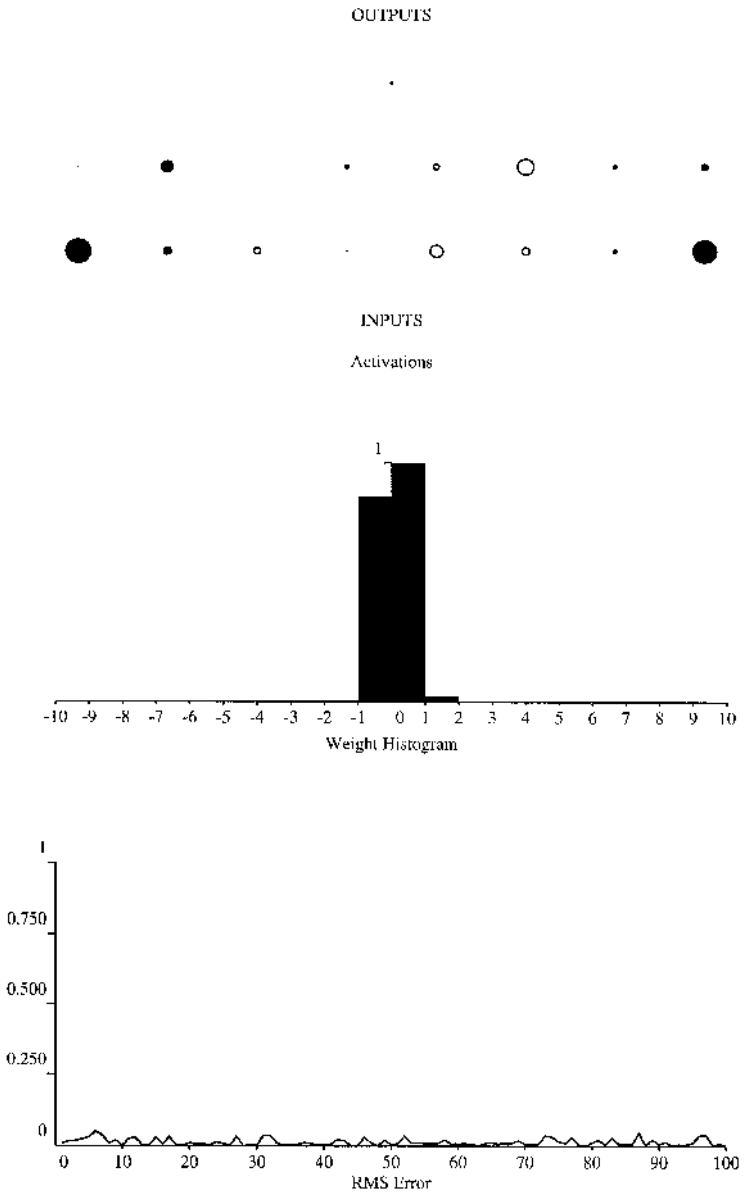


Figure 6.26. Final neural network state for the nonlinear model of the Duffing oscillator.

demonstrates 'dynamic neurons' which are said to increase the utility of the MLP structure for modelling dynamical systems. The most comprehensive programme of work to date is that of Billings and co-workers starting with [36] for the MLP

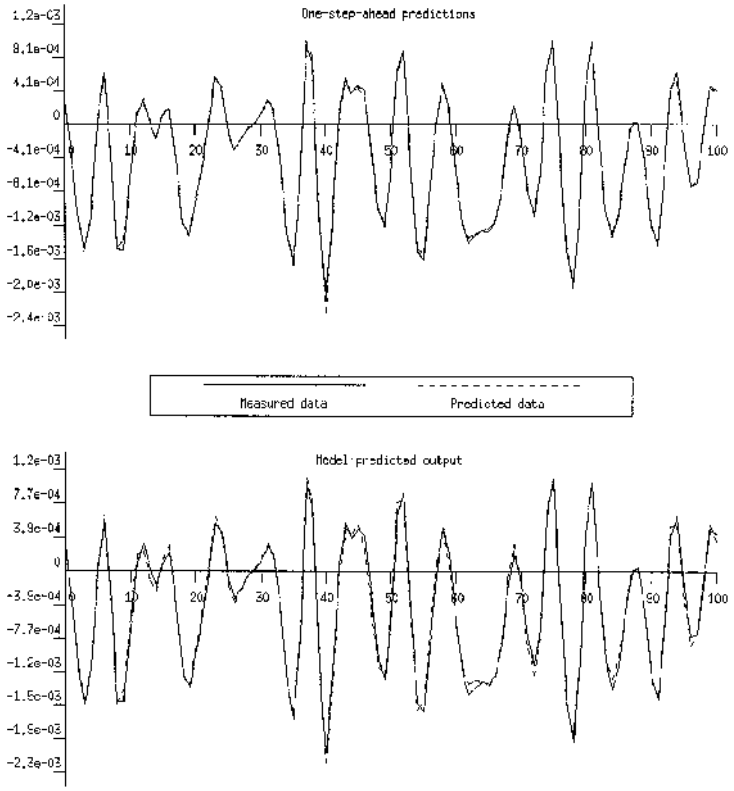


Figure 6.27. OSA and MPO predictions for the nonlinear neural network model of the Duffing oscillator.

structure and [62] for the RBF.

The use of the neural network will be illustrated with a couple of case studies, only the MLP results will be shown.

6.13.2 A linear system

The data consists of 999 pairs of input–output data for a linear dynamical system with equation of motion

$$\ddot{y} + 20\dot{y} + 10^4 y = x(t) \tag{6.134}$$

where $x(t)$ is a zero-mean Gaussian sequence of rms 10.0. (The data were obtained using a fourth-order Runge–Kutta routine to step the differential equation forward in time.) The output data are corrupted by zero-mean Gaussian white noise. A structure using four lags in both input and output were chosen.

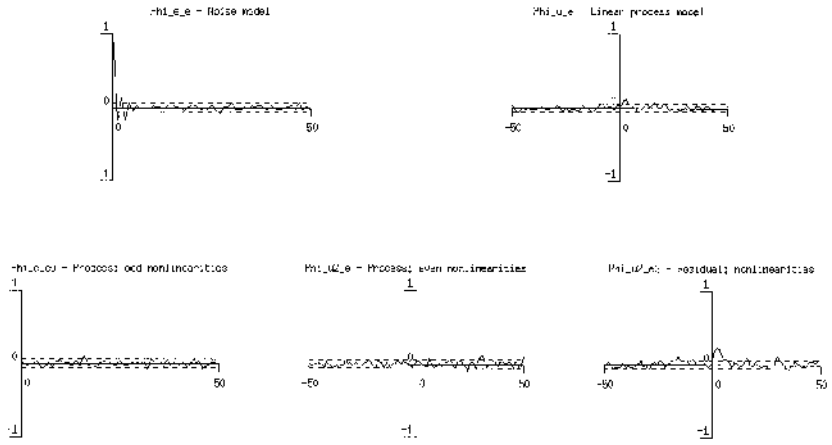


Figure 6.28. Correlation tests for the nonlinear neural network model of the Duffing oscillator.

The network activation function was taken as *linear*, forcing the algorithm to fit an ARX model. Because of this, the network did not need hidden units. The network was trained using 20 000 presentations of individual input–output pairs at random from the training set. The training constants are not important here. The state of the network at the end of training is shown in figure 6.19. The top graph shows the activations (neuronal outputs) over the network for the last data set presented. The centre plot shows the numerical distribution of the weights over the network. The final plot is most interesting and shows the evolution of the network error in the latest stages of training.

After training, the network was tested. Figure 6.20 shows some of the OSA and MPO predictions. Figure 6.21 shows the corresponding residuals and prediction errors. Finally, figure 6.22 shows the correlation test. The results are fairly acceptable. The MSEs are 3.09 for the OSA and 3.44 for the MPO.

6.13.3 A nonlinear system

The data for this exercise consisted of 999 pairs of input–output points (x–y) for the nonlinear Duffing oscillator system

$$\ddot{y} + 20\dot{y} + 10^4 y + 10^7 y^2 + 5 \times 10^9 y^3 = x(t). \tag{6.135}$$

As before, the data were generated using a Runge–Kutta procedure. In this case, the data are *not* corrupted by noise.

6.13.3.1 A linear model

It is usual in nonlinear system identification to fit a linear model first. This gives information about the degree of nonlinearity and also provides guidance on the appropriate values for the lags n_y and n_x. As this is a single-degree-of-freedom (SDOF) system like that in the first exercise, one can expect reasonable results using the same lag values. A linear network was tried first.

The final state of the network is saved after the 20 000 presentations; the result is given in figure 6.23. The MSEs reported by the procedure are 8.72 for the OSA and 41.04 for the MPO which are clearly unacceptable. Figures 6.24 and 6.25, respectively, show the predictions and correlation tests.

6.13.3.2 A nonlinear model

This time a nonlinear network but with a linear output neuron was used. Eight hidden units were used. The final network state is shown in figure 6.26. The rms error shows a vast improvement on the linear network result (figure 6.23). This is reflected in the network MSEs which were 0.34 (OSA) and 3.10 (MPO). The network predictions are given in figure 6.27 and the correlation tests in figure 6.28.

It is shown in [275] that the neural network structures discussed here can represent a broad range of SDOF nonlinear systems, with continuous or discontinuous nonlinearities. This is one of the advantages of the neural network approach to identification; a 'black box' is specified which can be surprisingly versatile. The main disadvantage is that the complex nature of the network generally forbids an analytical explanation of why training sometimes fails to converge to an appropriate global minimum. For modelling purposes, it is unfortunate that the structure detection algorithms which prove so powerful in the NARMAX approach cannot be implemented, although 'pruning' algorithms are being developed which allow some simplification of the network structures. The network structure and training schedule must be changed if a different set of lagged variables is to be used.

Appendix A

A rapid introduction to probability theory

Chapter 2 uses some ideas from probability theory relating, in particular, to probability density functions. A background in probability theory is required for a complete understanding of the chapter; for those without the necessary background, the required results are collected together in this appendix. The arguments are not intended to be rigorous. For a complete mathematical account of the theory the reader can consult one of the standard texts [13, 97, 124].

A.1 Basic definitions

The *probability* $P(E)$ of an event E occurring in a given situation or experimental trial is defined as

$$P(E) = N(S) \xrightarrow{\text{limit}} \infty \frac{N(E)}{N(S)} \tag{A.1}$$

where $N(S)$ is the number of times the situation occurs or the experiment is conducted, and $N(E)$ is the number of times the event E follows. Clearly $1 \geq P(E) \geq 0$ with $P(E) = 1$ asserting the certainty of event E and $P(E) = 0$ indicating its impossibility. In a large number of throws of a true die, the result 6 would be expected $\frac{1}{6}$ of the time, so $P(6) = \frac{1}{6}$.

If two events E_1 and E_2 are *mutually exclusive* then the occurrence of one precludes the occurrence of the other. In this case, it follows straightforwardly from (A.1) that

$$P(E_1 \cup E_2) = P(E_1) + P(E_2) \tag{A.2}$$

where the symbol $\cup$ represents the logical 'or' operation, so $P(E_1 \cup E_2)$ is the probability that event E_1 'or' event E_2 occurs. If E_1 and E_2 are not mutually exclusive, a simple argument leads to the relation

$$P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2) \tag{A.3}$$

where the symbol $\cap$ represents the logical 'and' operation.

If a set of mutually exclusive events $\{E_1, \ldots, E_{\rm N}\}$ is exhaustive in the sense that one of the E_i *must* occur, it follows from the previous definitions that

$$P(E_1 \cup E_2 \cup \cdots \cup E_{\rm N}) = P(E_1) + P(E_2) + \cdots + P(E_{\rm N}) = 1. \qquad \text{(A.4)}$$

So in throwing a die

$$P(1) + P(2) + P(3) + P(4) + P(5) + P(6) = 1 \qquad \text{(A.5)}$$

(in an obvious notation). Also, if the die is true, all the events are equally likely

$$P(1) = P(2) = P(3) = P(4) = P(5) = P(6) \qquad \text{(A.6)}$$

and these two equations show that $P(6) = \frac{1}{6}$ as asserted earlier.

Two events E_1 and E_2 are *statistically independent* or just *independent* if the occurrence of one *in no way* influences the probability of the other. In this case

$$P(E_1 \cap E_2) = P(E_1) \times P(E_2). \qquad \text{(A.7)}$$

A.2 Random variables and distributions

The outcome of an individual throw of a die is completely unpredictable. However, the value obtained has a definite probability which can be determined. Variables of this type are referred to as *random variables*. In the example cited, the random variable can only take one of six values; it is therefore referred to as *discrete*. In the following discussion, it will also be necessary to consider random variables which can take a continuous range of values.

Imagine a party where a group of guests have been driven by boredom to make a bet on the height of the next person to arrive. Assuming no cheating on any of their parts, this is a continuous random variable. Now, if they are to make the most of their guesses, they should be guided by probability. It can safely be assumed that $P(3\ \mathrm{m}) = 0$ and $P(0.1\ \mathrm{m}) = 0$. (Heights will always be specified in metres from now on and the units will be dropped.) However, if it is assumed that all intermediate values are possible, this gives an infinity of outcomes. A rough argument based on (A.4) gives [1]

$$P(h_1) + P(h_2) + \cdots + P(h_i) + \cdots = 1 \qquad \text{(A.8)}$$

and the individual probabilities must all be zero. This agrees with common sense; if one person guesses 1.8 m, there is no real chance of observing *exactly* this value, any sufficient precise measurement would show up a discrepancy. If individual probabilities are all zero, how can statistical methods be applied? In practice, to avoid arguments the party guests would probably specify a range of heights

[1] In fact there is an uncountable infinity of outcomes so they cannot actually be ordered in sequence as the equation suggests.

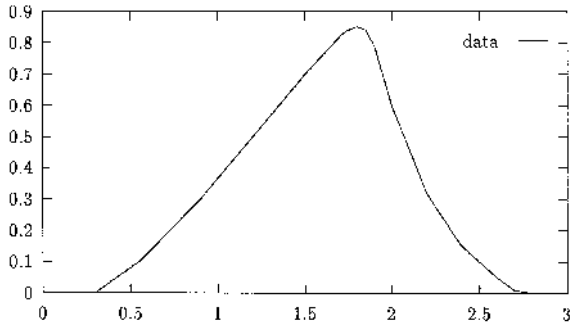


Figure A.1. Probability density function for the party guessing game.

centred on a particular value. This points to the required mathematical structure; for a random variable X, the *probability density function* (PDF) $p(x)$ is defined by

$$p(x) \text{ is the probability that } X \text{ takes a value between } x \text{ and } x + \mathrm{d}x.$$

So what will $p(x)$ look like? Well it has already been established that $P(0.1) = P(3) = 0$. It would be expected that the most probable height (a meaningful definition of 'most probable' will be given later) would be around 1.8 m so $P(1.8)$ would be a maximum. The distribution would be expected to rise smoothly up to this value and decrease steadily above it. Also, if children are allowed, values 60 cm smaller than the most probable height will be more likely than values 60 cm higher. Altogether this will give a PDF like that in figure A.1.

Now, suppose a party guest gives the answer 1.75 ± 0.01. What is the probability that the height falls within this *finite* range? Equation (A.2) implies the need for a summation of the probabilities of all possible values. For a continuum of values, the analogue of the summation (A.2) is an integral, so

$$P(X = x; 1.74 \leq x \leq 1.76) = \int_{1.74}^{1.76} p(x)\,\mathrm{d}x. \tag{A.9}$$

In general

$$P(X = x; a \leq x \leq b) = \int_{a}^{b} p(x)\,\mathrm{d}x. \tag{A.10}$$

Geometrically, this probability is represented by the area under the PDF curve between a and b (the shaded area in figure A.2).

The total area under the curve, i.e. the probability of X taking *any* value must be 1. In analytical terms

$$\int_{-\infty}^{\infty} p(x)\,\mathrm{d}x = 1. \tag{A.11}$$

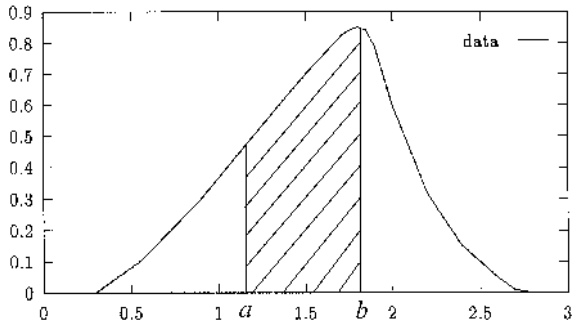


Figure A.2. Probability of a value in the interval a to b.

Note that this condition requires that $p(x) \longrightarrow 0$ as $x \longrightarrow \pm\infty$.

The party guests can therefore establish probabilities for their guesses. The question of how to optimize the guess using the information from the PDF is answered in the next section which shows how to compute the *expected value* of a random variable.

Note that the random variable need not be a scalar. Suppose the party guests had attempted to guess height and weight, two random variables. Their estimate would be an example of a two-component *random vector* $\boldsymbol{X} = (X_1, X_2)$. The probability density function $p(\boldsymbol{X})$ is defined exactly as before

$$p(\boldsymbol{x}) \text{ is the probability that } \boldsymbol{X} \text{ takes a value between } \boldsymbol{x} \text{ and } \boldsymbol{x} + \mathrm{d}\boldsymbol{x}.$$

The PDF is sometimes written in the form $p(X_1, X_2)$ and is referred to as the *joint probability density function* between X_1 and X_2.

The N-dimensional analogue of (A.10) is the multiple integral

$$\begin{aligned} P(\boldsymbol{X} = \boldsymbol{x}; a_1 \leq x_1 \leq b_1, \ldots, a_\mathrm{N} \leq x_\mathrm{N} \leq b_\mathrm{N}) \\ = \int_{a_1}^{b_1} \ldots \int_{a_\mathrm{N}}^{b_\mathrm{N}} p(\boldsymbol{x}) \, \mathrm{d}x_1 \ldots \mathrm{d}x_\mathrm{N}. \end{aligned} \tag{A.12}$$

Random vectors are very important in the theory of statistical pattern recognition; measurement/feature vectors are random vectors.

Suppose a two-component random vector is composed of two statistically independent variables X_1 and X_2 with individual PDFs p_1 and p_2, then by (A.7),

$$\begin{aligned} P(X_1 = x_1; x_1 \in [a_1, b_1]) \times P(X_2 = x_2; x_2 \in [a_2, b_2]) \\ = \left\{ \int_{a_1}^{b_1} p_1(x_1) \, \mathrm{d}x_1 \right\} \left\{ \int_{a_2}^{b_2} p_2(x_2) \, \mathrm{d}x_2 \right\} \\ = \int_{a_1}^{b_1} \int_{a_2}^{b_2} p_1(x_1) p_1(x_2) \, \mathrm{d}x_1 \, \mathrm{d}x_2 \end{aligned} \tag{A.13}$$

and, according to (A.12), this is equal to

$$\int_{a_1}^{b_1}\int_{a_2}^{b_2} p_j(x_1, x_2)\,\mathrm{d}x_1\,\mathrm{d}x_2 \tag{A.14}$$

where $p_j(\boldsymbol{x})$ is the joint PDF. As the last two expressions are equal *for all* values of a_1, a_2, b_1, b_2, it follows that

$$p_j(\boldsymbol{x}) = p_j(x_1, x_2) = p_1(x_1)p_2(x_2) \tag{A.15}$$

which is the analogue of (A.7) for continuous random variables. Note that this is *only* true if X_1 and X_2 are independent. In the general N-dimensional case, the joint PDF will factor as

$$p_j(\boldsymbol{x}) = p_1(x_1)p_2(x_2)\ldots p_{\mathrm{N}}(x_{\mathrm{N}}). \tag{A.16}$$

A.3 Expected values

Returning to the party guessing game of previous sections, suppose that the guests have equipped themselves with a PDF for the height (possibly computed from the heights of those already present). The question arises as to how they can use this information in order to compute a best guess or *expected value* for the random variable.

In order to simplify matters, consider first a discrete random variable, the outcome of a throw of a die. In this case, if the die is true, each outcome is equally likely and it is not clear what is meant by expected value. Consider a related question: if a die is cast N_{c} times, what is the expected value of the sum? This is clearly

$$N(1)\times 1 + N(2)\times 2 + N(3)\times 3 + N(4)\times 4 + N(5)\times 5 + N(6)\times 6 \tag{A.17}$$

where $N(i)$ is the expected number of occurrences of the value i as an outcome. If N_{c} is small, say 12, statistical fluctuations will have a large effect and two occurrences of each outcome could not be relied on. However, for a true die, there is no better guess as to the numbers of each outcomes. If N_{c} is large, then

$$N(i) \approx P(i)\times N_{\mathrm{c}} \tag{A.18}$$

and statistical fluctuations will have a much smaller effect. There will be a corresponding increase in confidence in the expected value of the sum, which is now

$$E(\text{sum of } N_{\mathrm{c}} \text{ die casts}) = \sum_{i=1}^{6} N_{\mathrm{c}} P(i) i \tag{A.19}$$

where E is used to denote the expected value, E will also sometimes be referred to as the expectation operator. This last expression contains a quantity independent

of N_c which can quite reasonably be defined as the expected value of a single cast. If

$$E(\text{sum of } N_c \text{ die casts}) = N_c \times E(\text{single cast}) \tag{A.20}$$

then

$$E(\text{single cast}) = \sum_{i=1}^{6} P(i)i \tag{A.21}$$

and this is simply a sum over the possible outcomes with each term weighted by its probability of occurrence. This formulation naturally deals with the case of a biased die ($P(i) \neq P(j), i \neq j$) in the same way as for a true die. In general then

$$E(X) = \sum_{x_i} P(X = x_i)x_i \tag{A.22}$$

where the random variable can take any of the discrete values x_i.

For the throw of a true die

$$E(\text{single cast}) = \tfrac{1}{6} \times 1 + \tfrac{1}{6} \times 2 + \tfrac{1}{6} \times 3 + \tfrac{1}{6} \times 4 + \tfrac{1}{6} \times 5 + \tfrac{1}{6} \times 6 = 3.5 \tag{A.23}$$

and this illustrates an important fact, that the expected value of a random variable need not be one of the allowed values for the variable. Also, writing the last expression as

$$E(\text{single cast}) = \frac{1 + 2 + 3 + 4 + 5 + 6}{6} \tag{A.24}$$

it is clear that the expected value is simply the *arithmetic mean* taken over the possible outcomes or values of the random variable. This formulation can *only* be used when all outcomes are equally likely. However, the expected value of a random variable X will often be referred to as the *mean* and will be denoted $\overline{x}$.

The generalization of (A.22) to the case of continuous random variables is straightforward and simply involves the replacement of the weighted sum by a weighted integral. So

$$\overline{X} = E(X) = \int_{-\infty}^{\infty} xp(x)\,\mathrm{d}x \tag{A.25}$$

where $p(x)$ is the PDF for the random variable X. Note that the integral need only be taken over the range of possible values for x. However, the limits are usually taken as $-\infty$ to ∞ as any values outside the valid range have $p(x) = 0$ and do not contribute to the integral in any case.

It is important to note that the expected value of a random variable is *not* the same as the peak value of its PDF. The distribution in figure A.3 provides a simple counterexample.

The mean is arguably the most important statistic of a PDF. The other contender is the standard deviation which also conveys much useful information.

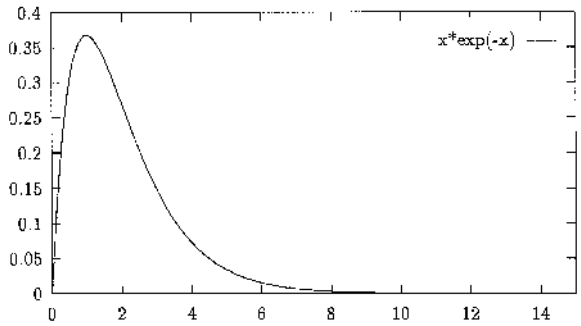


Figure A.3. Expected value is not the same as the PDF maximum.

Consider the party game again. For all intents and purposes, the problem has been solved and all of the guests should have made the same guess. Now, when the new guest arrives and is duly measured, it is certain that there will be some error in the estimate (the probability that the height coincides with the expected value is zero). The question arises as to how good a guess is the mean[2].

Consider the two probability distributions in figure A.4. For the distribution in figure A.4(*a*), the mean is always a good guess, for that in figure A.4(*b*), the mean would often prove a bad estimate. In statistical terms, what is required is the expected value $E(\epsilon)$ of the error $\epsilon = X - \overline{x}$. Pursuing this

$$E(\epsilon) = E(X - \overline{x}) = E(X) - E(\overline{x}) \tag{A.26}$$

because E is a linear operator (which is obvious from (A.25)). Further, on random variables the E operator extracts the mean, on ordinary numbers like $\overline{x}$ it has no effect (the expected value of a number must always be that number). So

$$E(X) - E(\overline{x}) = \overline{x} - \overline{x} = 0 \tag{A.27}$$

and the final result, $E(\epsilon) = 0$, is not very informative. This arises because positive and negative errors are equally likely so the expected value is zero. The usual means of avoiding this problem is to consider the expected value of the error-squared, i.e. $E(\epsilon^2)$. This defines the statistic σ^2 known as the variance:

$$\begin{aligned} \sigma^2 &= E(\epsilon^2) = E((X - \overline{x})^2) \\ &= E(X^2) - E(2\overline{x}X) + E(\overline{x}^2) \end{aligned}$$

[2] This is an important question. In the system identification theory discussed in chapter 6, systems are often modelled by assuming a functional form for the equations of motion and then finding the values of the equation's constants which best fit the measured data from tests. Because of random measurement noise, different sets of measured data will produce different parameter estimates. The estimates are actually samples from a population of possible estimates and it is assumed that the true values of the parameters correspond to the expected values of the distribution. It is clearly important to know, given a particular estimate, how far away from the expected value it is.

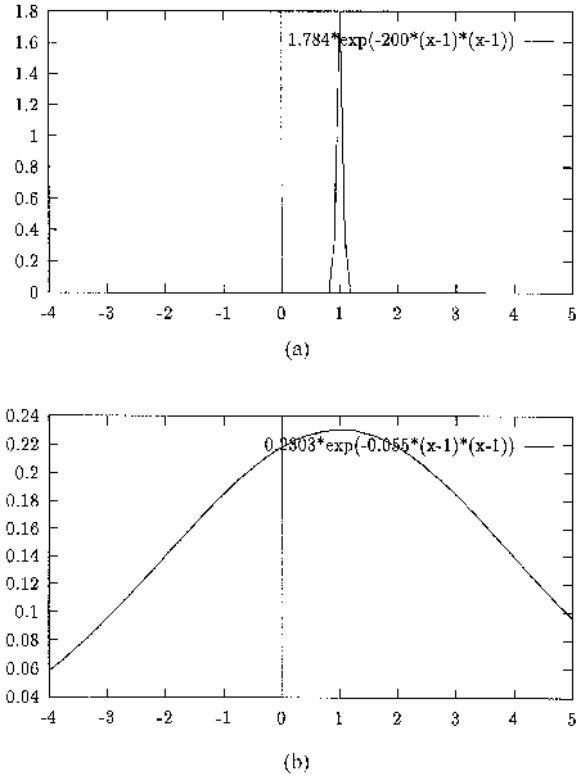


Figure A.4. Distributions with (*a*) small variance, (*b*) high variance.

$$= E(X^2) - 2\overline{x}E(X) + \overline{x}^2$$
$$= E(X^2) - 2\overline{x}^2 + \overline{x}^2 \tag{A.28}$$
$$\sigma^2 = E(X^2) - \overline{x}^2. \tag{A.29}$$

Equation (A.29) is often given as an alternative definition of the variance. In the case of equally probable values for X, (A.12) reduces, via (A.9) to the expression

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - \overline{x})^2 \tag{A.30}$$

where the $x_i, i = 1, \ldots, N$ are the possible values taken by X[3].

In the case of X a continuous random variable, (A.25) shows that the

[3] Actually this expression for the variance is known to be *biased*. However, changing the denominator from N to $N - 1$ remedies the situation. Clearly, the bias is a small effect as long as N is large.

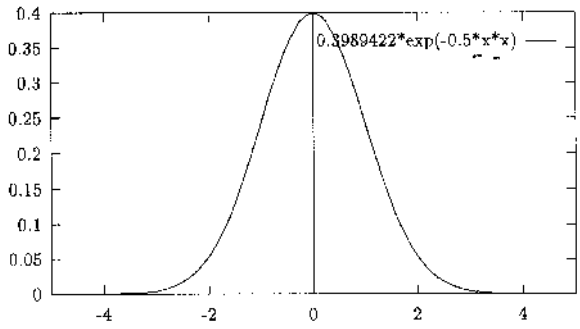


Figure A.5. The Gaussian distribution $N(0, 1)$.

appropriate form for the variance is

$$\sigma^2 = \int_{-\infty}^{\infty} (x - \overline{x})^2 p(x)\,\mathrm{d}x. \tag{A.31}$$

The *standard deviation* σ is simply the square root of the variance. It can therefore be interpreted as the expected root-mean-square (rms) error in using the mean as a guess for the value of a random variable. It clearly gives a measure of the width of a probability distribution in much the same way as the mean provides an estimate of where the centre is.

A.4 The Gaussian distribution

The Gaussian or normal distribution is arguably the most important of all. One of its many important properties is that its behaviour is fixed completely by a knowledge of its mean and variance. In fact the functional form is

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2}\left(\frac{x - \overline{x}}{\sigma}\right)^2\right\}. \tag{A.32}$$

This is sometimes denoted $N(\overline{x}, \sigma)$. (It is straightforward to show from (A.25) and (A.31) that the parameters $\overline{x}$ and σ in (A.32) truly are the distribution mean and standard deviation.) As an example, the Gaussian $N(0, 1)$ is shown in figure A.5.

One of the main reasons for the importance of the Gaussian distribution is provided by the *Central Limit Theorem* [97] which states (roughly): If $X_i, i = 1, \ldots, N$ are N independent random variables, possibly with completely different distributions, then the random variable X_Σ formed from the sum

$$X_\Sigma = X_1 + X_2 + \cdots + X_\mathrm{N} \tag{A.33}$$

has a Gaussian distribution. Much of system identification theory assumes that measurement noise has a Gaussian density function. If the noise arises from a number of independent mechanisms and sources, this is partly justified by the central limit theorem. The main justification is that it is usually the only way to obtain analytical results.

The Gaussian is no less important in higher dimensions. However, the generalization of (A.32) to random vectors $\overline{X} = (X_1, \ldots, X_n)$ requires the introduction of a new statistic, the *covariance* $\sigma_{X_i X_j}$ defined by

$$\sigma_{X_i X_j} = E((X_i - \overline{x}_i)(X_j - \overline{x}_j)) \tag{A.34}$$

which measures the degree of correlation between the random variables X_i and X_j. Consider two independent random variables X and Y.

$$\sigma_{XY} = E((X - \overline{x})(Y - \overline{y})) = \int\int (x - \overline{x})(y - \overline{y}) p_j(x, y)\,\mathrm{d}x\,\mathrm{d}y. \tag{A.35}$$

Using the result (A.15), the joint PDF p_j factors

$$\begin{aligned}\sigma_{XY} &= \int\int (x - \overline{x})(y - \overline{y}) p_x(x) p_y(y)\,\mathrm{d}x\,\mathrm{d}y \\ &= \left\{ \int (x - \overline{x}) p_x(x)\,\mathrm{d}x \right\} \times \left\{ \int (y - \overline{y}) p_y(y)\,\mathrm{d}y \right\} \\ &= 0 \times 0.\end{aligned} \tag{A.36}$$

So $\sigma_{XY} \neq 0$ indicates a degree of interdependence or correlation between X and Y. For a random vector, the information is encoded in a matrix—the *covariance matrix* $[\Sigma]$—where

$$\Sigma_{ij} = E((X_i - \overline{x}_i)(X_j - \overline{x}_j)). \tag{A.37}$$

Note that the diagonals are the usual variances

$$\Sigma_{ii} = \sigma^2_{X_i}. \tag{A.38}$$

As in the single-variable case, the vector of means $\{\overline{x}\}$ and the covariance matrix $[\Sigma]$ completely specify the Gaussian PDF. In fact

$$p(\{x\}) = \frac{1}{(2\pi)^{\frac{N}{2}} \sqrt{|\Sigma|}} \exp\left\{ -\frac{1}{2} (\{x\} - \{\overline{x}\})^{\mathrm{T}} [\Sigma]^{-1} (\{x\} - \{\overline{x}\}) \right\} \tag{A.39}$$

for an N-component random vector $\boldsymbol{X}$. $|\Sigma|$ denotes the determinant of the matrix Σ.

Appendix B

Discontinuities in the Duffing oscillator FRF

As discussed in chapter 3, discontinuities are common in the composite FRFs of nonlinear systems and fairly simple theory suffices to estimate the positions of the jump frequencies ω_{low} and ω_{high}, at least for the first-order harmonic balance approximation. The approach taken is to compute the discriminant of the cubic equation (3.10) which indicates the number of real solutions [38].

In a convenient notation, (3.10) is

$$a_3 Y^6 + a_2 Y^4 + a_1 Y^2 + a_0 = 0 \tag{B.1}$$

Now, dividing by a_3 and making the transformation $Y^2 = z - a_2/(3a_3)$ yields the normal form

$$z^3 + pz + q = 0 \tag{B.2}$$

and the discriminant D is then given by

$$D = -4p^3 - 27q^2. \tag{B.3}$$

Now, the original cubic (3.10) has three real solutions if $D \geq 0$ and only one if $D < 0$. The bifurcation points are therefore obtained by solving the equation $D = 0$. For Duffings equation (3.3) this is an exercise in computer algebra and the resulting discriminant is

$$\begin{aligned} D = \frac{256}{729k_3^6}(&-64c^2k^4\omega^2 - 128c^4k^2\omega^4 + 256c^2k^3m\omega^4 - 64c^6\omega^6 \\ &+ 256c^4km\omega^6 - 384c^2k^2m^2\omega^6 - 128c^4m^2\omega^8 + 256c^2km^3\omega^8 \\ &- 64c^2m^4\omega^{10} - 48k^3k_3X^2 - 432c^2kk_3X^2\omega^2 + 144k^2k_3mX^2\omega^2 \\ &+ 432c^2k_3mX^2\omega^4 - 144kk_3m^2X^2\omega^4 + 48k_3m^3X^2\omega^6 - 243k_3^2X^4). \end{aligned} \tag{B.4}$$

Table B.1. 'Exact' jump frequencies in rad s^{-1} for upward sweep.

	Damping coefficient c			
Forcing X	0.01	0.03	0.1	0.3
0.1	3.04	1.86	1.23	—
0.3	5.16	3.04	1.78	—
1.0	9.36	5.44	3.04	1.86
3.0	16.18	9.36	5.16	3.04
10.0	29.52	17.06	9.36	5.44
30.0	51.13	29.52	16.18	9.36
100.0	93.34	53.89	29.52	17.06

Table B.2. Estimated jump frequencies and percentage errors (bracketed) in rad s^{-1} for upward sweep.

	Damping coefficient c			
Forcing X	0.01	0.03	0.1	0.3
0.1	3.03	1.85	1.23	—
	(−0.32)	(−0.54)	(0.00)	—
0.3	5.15	3.03	1.77	—
	(−0.19)	(−0.32)	(−0.56)	—
1.0	9.33	5.42	3.03	1.86
	(−0.32)	(−0.36)	(−0.32)	(0.00)
3.0	16.13	9.33	5.15	3.03
	(−0.30)	(−0.32)	(−0.19)	(−0.32)
10.0	29.44	17.01	9.33	5.42
	(−0.29)	(−0.29)	(−0.32)	(−0.36)
30.0	50.98	29.44	16.13	9.33
	(−0.29)	(−0.27)	(−0.30)	(−0.32)
100.0	93.07	53.73	29.44	17.01
	(−0.30)	(−0.30)	(−0.27)	(−0.29)

As bad as this looks, it is just a quintic in ω^2 and can have at most five independent solutions for ω. In fact in all the cases examined here, it had two real roots and three complex. The lowest real root is the bifurcation point for a downward sweep and the highest is the bifurcation point for an upward sweep. The equation $D = 0$ is solved effortlessly using computer algebra. However, note that an analytical solution is possible using elliptic and hypergeometric functions [148].

The study by Friswell and Penny [104] computed the bifurcation points of the FRF, not in the first harmonic balance approximation but for a multi-harmonic series solution. They obtained excellent results using expansions up to third and fifth harmonic for the response. Newton's method was used to solve the equations obtained as even up to third harmonic the expressions are exceedingly complex.

The values $m = k = k_3 = 1$ were chosen here for the Duffing oscillator. This is because [104] presents bifurcation points for a ninth-harmonic solution to (3.3) with these parameters and these can therefore be taken as reference data. A range of c and X values were examined. The results for the upward sweep only are given here, for the downward sweep the reader can consult [278]. The 'exact' values from [104] are given in table B.1 for a range of damping coefficient values.

The estimated bifurcation points obtained from the discriminant for the upward sweep are given in table B.2. Over the examples given, the percentage errors range from -0.19 to -0.56. This compares very well with the results of [104] which ranged from -0.29 to -0.33.

Appendix C

Useful theorems for the Hilbert transform

C.1 Real part sufficiency

Given the FRF for a causal system, equations (4.17) and (4.18) show that the real part can be used to reconstruct the imaginary part and *vice versa*. It therefore follows that all the system characteristics are encoded in each part separately. Thus, it should be possible to arrive at the impulse response using the real part or imaginary part of the FRF alone.

From (4.4)

$$g(t) = g_{\text{even}}(t) + g_{\text{odd}}(t) \tag{C.1}$$

and from (4.7)

$$g(t) = g_{\text{even}}(t)(1 + \epsilon(t)) \tag{C.2}$$

or

$$g(t) = g_{\text{even}}(t) \times 2\Theta(t) \tag{C.3}$$

where $\Theta(t)$ is the Heaviside unit-step function. Finally

$$g(t) = 2\mathcal{F}^{-1}\{\text{Re}\, G(\omega)\}\Theta(t) \tag{C.4}$$

which shows that the real part alone of the FRF is sufficient to form the impulse response. A similar calculation gives

$$g(t) = 2\mathcal{F}^{-1}\{\text{Im}\, G(\omega)\}\Theta(t). \tag{C.5}$$

C.2 Energy conservation

The object of this exercise is to determine how the total energy of the system, as encoded in the FRF, is affected by Hilbert transformation. If the energy functional is defined as usual by

$$\int_{-\infty}^{\infty} \text{d} \circ |f(\circ)|^2. \tag{C.6}$$

Then, Parseval's theorem

$$\int_{-\infty}^{\infty} \mathrm{d}t\, |g(t)|^2 = \int_{-\infty}^{\infty} \mathrm{d}\omega\, |G(\omega)|^2 \tag{C.7}$$

where $G(\omega) = \mathcal{F}\{g(t)\}$, shows that energy is conserved under the Fourier transform.

Taking the Hilbert transform of $G(\omega)$ yields

$$\mathcal{H}\{G(\omega)\} = \tilde{G}(\omega) \tag{C.8}$$

and an application of Parseval's theorem gives

$$\int_{-\infty}^{\infty} \mathrm{d}\omega\, |\tilde{G}(\omega)|^2 = \int_{-\infty}^{\infty} \mathrm{d}t\, |\mathcal{F}^{-1}\{\tilde{G}(\omega)\}|^2. \tag{C.9}$$

By the definition of the Hilbert transform

$$\tilde{G}(\omega) = G(\omega) * -\frac{1}{\pi\omega} \tag{C.10}$$

and taking the inverse Fourier transform yields

$$\mathcal{F}^{-1}\{\tilde{G}(\omega)\} = g(t)\epsilon(t) \tag{C.11}$$

so

$$|\mathcal{F}^{-1}\{\tilde{G}(\omega)\}|^2 = |g(t)\epsilon(t)|^2 = |g(t)|^2 \tag{C.12}$$

as $|\epsilon(t)|^2 = 1$. Substituting this result into (4.26) and applying Parseval's theorem once more gives

$$\int_{-\infty}^{\infty} \mathrm{d}\omega\, |\tilde{G}(\omega)|^2 = \int_{-\infty}^{\infty} \mathrm{d}t\, |g(t)|^2 = \int_{-\infty}^{\infty} \mathrm{d}\omega\, |G(\omega)|^2. \tag{C.13}$$

Thus, energy is also conserved under Hilbert transformation.

C.3 Commutation with differentiation

Given a function $G(\omega)$, it can be shown that the Hilbert transform operator $\mathcal{H}$ commutes with the derivative operator $\mathrm{d}/\mathrm{d}\omega$ under fairly general conditions, i.e.

$$\mathcal{H}\left\{\frac{\mathrm{d}G}{\mathrm{d}\omega}\right\} = \frac{\mathrm{d}\tilde{G}}{\mathrm{d}\omega} \tag{C.14}$$

Consider

$$\frac{\mathrm{d}\tilde{G}}{\mathrm{d}\omega} = \frac{\mathrm{d}}{\mathrm{d}\omega}\left(-\frac{1}{\mathrm{i}\pi}\int_{-\infty}^{\infty} \mathrm{d}\Omega\, \frac{G(\Omega)}{\Omega-\omega}\right) = -\frac{1}{\mathrm{i}\pi}\int_{-\infty}^{\infty} \mathrm{d}\Omega\, \frac{\mathrm{d}}{\mathrm{d}\omega}\left(\frac{G(\Omega)}{\Omega-\omega}\right) \tag{C.15}$$

assuming differentiation and integration commute. Elementary differentiation yields

$$\frac{\mathrm{d}\tilde{G}}{\mathrm{d}\omega} = -\frac{1}{\mathrm{i}\pi}\int_{-\infty}^{\infty} \mathrm{d}\Omega\, \frac{G(\Omega)}{(\Omega-\omega)^2}. \tag{C.16}$$

Now

$$\mathcal{H}\left\{\frac{\mathrm{d}G}{\mathrm{d}\omega}\right\} = -\frac{1}{\mathrm{i}\pi}\int_{-\infty}^{\infty} \mathrm{d}\Omega\, \frac{\frac{\mathrm{d}G}{\mathrm{d}\Omega}}{\Omega-\omega} \tag{C.17}$$

and integrating by parts yields

$$-\frac{1}{\mathrm{i}\pi}\int_{-\infty}^{\infty} \mathrm{d}\Omega\, \frac{\frac{\mathrm{d}G}{\mathrm{d}\Omega}}{\Omega-\omega} = \left[-\frac{1}{\mathrm{i}\pi}\frac{G(\Omega)}{\Omega-\omega}\right]_{\Omega=-\infty}^{\Omega=\infty} + \frac{1}{\mathrm{i}\pi}\int_{-\infty}^{\infty} \mathrm{d}\Omega\, G(\Omega)\frac{\mathrm{d}}{\mathrm{d}\Omega}\left(\frac{1}{\Omega-\omega}\right). \tag{C.18}$$

Now, assuming that the first term (the boundary term) vanishes, i.e. $G(\omega)$ has fast enough fall-off with ω (it transpires that this is a vital assumption in deriving the Hilbert transform relations anyway—see chapter 5), simple differentiation shows

$$\mathcal{H}\left\{\frac{\mathrm{d}G}{\mathrm{d}\omega}\right\} = -\frac{1}{\mathrm{i}\pi}\int_{-\infty}^{\infty} \mathrm{d}\Omega\, \frac{G(\Omega)}{(\Omega-\omega)^2} \tag{C.19}$$

and together with (C.16), this establishes the desired result (C.14).

An identical argument in the time domain suffices to prove

$$\mathcal{H}\left\{\frac{\mathrm{d}g(t)}{\mathrm{d}t}\right\} = \frac{\mathrm{d}\tilde{g}}{\mathrm{d}t} \tag{C.20}$$

with an appropriate time-domain definition of $\mathcal{H}$.

Having established the Fourier decompositions (4.79) and (4.84), it is possible to establish three more basic theorems of the Hilbert transform.

C.4 Orthogonality

Considered as objects in a vector space, it can be shown that the scalar product of an FRF or spectrum $G(\omega)$ with its associated Hilbert transform $\tilde{G}(\omega)$ vanishes, i.e.

$$\langle G, \tilde{G}\rangle = \int_{-\infty}^{\infty} \mathrm{d}\omega\, G(\omega)\tilde{G}(\omega) = 0. \tag{C.21}$$

Consider the integral, using the Fourier representation (4.79), one has $\mathcal{H}\circ\mathcal{F} = \mathcal{F}\circ\in$, so

$$\int_{-\infty}^{\infty} \mathrm{d}\omega\, G(\omega)\tilde{G}(\omega) = \int_{-\infty}^{\infty} \mathrm{d}\omega \int_{-\infty}^{\infty} \mathrm{d}t\, \mathrm{e}^{-\mathrm{i}\omega t} g(t) \int_{-\infty}^{\infty} \mathrm{d}\tau\, \mathrm{e}^{-\mathrm{i}\omega\tau}\epsilon(\tau)g(\tau). \tag{C.22}$$

A little rearrangement yields

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{d}t\mathrm{d}\tau \left\{ \int_{-\infty}^{\infty} \mathrm{d}\omega \, \mathrm{e}^{-\mathrm{i}\omega(t+\tau)} \right\} \epsilon(\tau) g(t) g(\tau) \tag{C.23}$$

the bracketed expression is a δ-function $2\pi\delta(t+\tau)$ (appendix D). Using the projection property, the expression becomes

$$2\pi \int_{-\infty}^{\infty} \mathrm{d}t \, \epsilon(-t) g(t) g(-t) = -2\pi \int_{-\infty}^{\infty} \mathrm{d}t \, \epsilon(t) g(t) g(-t). \tag{C.24}$$

The integrand is clearly odd, so the integral vanishes. This establishes the desired result (C.21). An almost identical proof suffices to establish the time-domain orthogonality:

$$\langle g, \tilde{g} \rangle = \int_{-\infty}^{\infty} \mathrm{d}t \, g(t) \tilde{g}(t) = 0. \tag{C.25}$$

C.5 Action as a filter

The action of the Hilbert transform on time functions factors as

$$\mathcal{H} = -\mathcal{F}^{-1} \circ \in \circ \mathcal{F} \tag{C.26}$$

as derived in chapter 4. This means that, following the arguments of [289], it can be interpreted as a filter with FRF

$$H(\omega) = -\epsilon(\omega) \tag{C.27}$$

i.e. all negative frequency components remain unchanged, but the positive frequency components suffer a sign change. (It is immediately obvious now why the Hilbert transform exchanges sines and cosines. Energy conservation also follows trivially.) Each harmonic component of the original signal is shifted in phase by $\pi/2$ radians and multiplied by i[1].

Now consider the action on a sine wave

$$\begin{aligned} \mathcal{H}\{\sin(\alpha t)\} &= -\mathcal{F}^{-1} \circ \in \circ \mathcal{F}\{\sin(\alpha t)\} \\ &= -\mathcal{F}^{-1} \circ \in \left\{ \frac{1}{2\mathrm{i}} (\delta(\omega-\alpha) - \delta(\omega+\alpha)) \right\} \\ &= -\mathcal{F}^{-1} \left\{ \frac{1}{2\mathrm{i}} (\delta(\omega-\alpha) + \delta(\omega+\alpha)) \right\} \\ &= \mathrm{i}\cos(\alpha t) \end{aligned} \tag{C.28}$$

[1] With the traditional time-domain definition of the Hilbert transform, the filter action is to phase shift all frequency components by $-\pi/2$.

a result wich could have been obtained by phase-shifting by $\pi/2$ and multiplying by i. The same operation on the cosine wave yields,

$$\mathcal{H}\{\cos(\alpha t)\} = -\mathrm{i}\sin(\alpha t). \tag{C.29}$$

Now suppose the functions are premultiplied by an exponential decay with a time constant long compared to their period $2\pi/\alpha$. Relations similar to (C.28) and (C.29) will hold:

$$\mathcal{H}\{\mathrm{e}^{-\lambda t}\sin(\alpha t)\} \approx \mathrm{i}\mathrm{e}^{-\lambda t}\cos(\alpha t) \tag{C.30}$$

$$\mathcal{H}\{\mathrm{e}^{-\lambda t}\cos(\alpha t)\} \approx -\mathrm{i}\mathrm{e}^{-\lambda t}\sin(\alpha t) \tag{C.31}$$

for sufficiently small λ. This establishes the result used in section 4.7.

The results (C.28) and (C.29) hold only for $\alpha > 0$. If $\alpha < 0$, derivations of the type given for (C.28) show that the signs are inverted. It therefore follows that

$$\mathcal{H}\{\sin(\alpha t)\} = i\epsilon(\alpha)\cos(\alpha t) \tag{C.32}$$

$$\mathcal{H}\{\cos(\alpha t)\} = -i\epsilon(\alpha)\sin(\alpha t). \tag{C.33}$$

These results are trivially combined to yield

$$\mathcal{H}\{\mathrm{e}^{\pm\mathrm{i}\alpha t}\} = -\epsilon(\alpha)\mathrm{e}^{\pm\mathrm{i}\alpha t}. \tag{C.34}$$

C.6 Low-pass transparency

Consider the Hilbert transform of a time-domain product $m(t)n(t)$ where the spectra $M(\omega)$ and $N(\omega)$ do not overlap and M is low-pass and N is high-pass, then

$$\mathcal{H}\{m(t)n(t)\} = m(t)\,\mathcal{H}\{n(t)\} \tag{C.35}$$

i.e. the Hilbert transform passes through the low-pass function. The proof given here follows that of [289].

Using the spectral representations of the function, one has

$$m(t)n(t) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\mathrm{d}\omega\,\mathrm{d}\Omega\,M(\omega)N(\Omega)\mathrm{e}^{-\mathrm{i}(\omega+\Omega)t}. \tag{C.36}$$

Applying the Hilbert transform yields

$$\mathcal{H}\{m(t)n(t)\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\mathrm{d}\omega\,\mathrm{d}\Omega\,M(\omega)N(\Omega)\mathcal{H}\{\mathrm{e}^{-\mathrm{i}(\omega+\Omega)t}\} \tag{C.37}$$

and by (C.34)

$$\mathcal{H}\{m(t)n(t)\} = -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\mathrm{d}\omega\,\mathrm{d}\Omega\,M(\omega)N(\Omega)\epsilon(\omega+\Omega)\mathrm{e}^{-\mathrm{i}(\omega+\Omega)t}\}. \tag{C.38}$$

Now, under the assumptions of the theorem, there exists a cut-off W, such that $M(\omega) = 0$ for $\omega > W$ and $N(\Omega) = 0$ for $\Omega < W$. Under these conditions, the signum function reduces $\epsilon(\omega + \Omega) = \epsilon(\Omega)$ and integral (C.38) factors:

$$\mathcal{H}\{m(t)n(t)\} = \left\{\int_{-\infty}^{\infty} \mathrm{d}\omega\, M(\omega)\mathrm{e}^{-\mathrm{i}\omega t}\right\}\left\{-\int_{-\infty}^{\infty} \mathrm{d}\Omega\, N(\Omega)\epsilon(\Omega)\mathrm{e}^{-\mathrm{i}\Omega t}\right\}. \tag{C.39}$$

The result (C.35) follows immediately.

Appendix D

Frequency domain representations of $\delta(t)$ and $\epsilon(t)$

Fourier's theorem for the Fourier transform states

$$g(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{d}\omega\, \mathrm{e}^{\mathrm{i}\omega t}\left(\int_{-\infty}^{\infty} \mathrm{d}\tau\, \mathrm{e}^{-\mathrm{i}\omega\tau} g(\tau)\right) \tag{D.1}$$

or, rearranging

$$g(t) = \int_{-\infty}^{\infty} \mathrm{d}\tau\, g(\tau)\left(\frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{d}\omega\, \mathrm{e}^{\mathrm{i}\omega(t-\tau)}\right). \tag{D.2}$$

Now the defining property of the Dirac δ-function is the *projection property*

$$g(t) = \int_{-\infty}^{\infty} \mathrm{d}\tau\, g(\tau)\delta(t-\tau) \tag{D.3}$$

so (D.2) allows the identification

$$\delta(t-\tau) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{d}\omega\, \mathrm{e}^{\mathrm{i}\omega(t-\tau)} \tag{D.4}$$

or, equally well,

$$\delta(t-\tau) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{d}\omega\, \mathrm{e}^{-\mathrm{i}\omega(t-\tau)}. \tag{D.5}$$

Now, consider the integral

$$I(t) = \int_{-\infty}^{\infty} \mathrm{d}\omega\, \frac{\mathrm{e}^{\mathrm{i}\omega t}}{\omega} = 2\mathrm{i}\int_{0}^{\infty} \mathrm{d}\omega\, \frac{\sin(\omega t)}{\omega}, \quad t > 0. \tag{D.6}$$

Taking the one-sided Laplace transform of both sides yields

$$\mathcal{L}[I(t)] = \tilde{I}(p) = 2\mathrm{i}\int_{0}^{\infty} \mathrm{d}t\, \mathrm{e}^{-pt}\int_{0}^{\infty} \mathrm{d}\omega\, \frac{\sin(\omega t)}{\omega} \tag{D.7}$$

and assuming that one can interchange the order of integration, this becomes

$$\tilde{I}(p) = 2\mathrm{i}\int_0^\infty \mathrm{d}\omega \left(\int_0^\infty \mathrm{d}t\, \mathrm{e}^{-pt}\sin(\omega t)\right)\frac{1}{\omega} \tag{D.8}$$

and, using standard tables of Laplace transforms, this is

$$\tilde{I}(p) = 2\mathrm{i}\int_0^\infty \mathrm{d}\omega\, \frac{1}{p^2+\omega^2} = \frac{\mathrm{i}\pi}{p}. \tag{D.9}$$

Taking the inverse transform gives $I(t) = \mathrm{i}\pi$ if $t > 0$. A simple change of variables in the original integral gives $I(t) = -\mathrm{i}\pi$ if $t < 0$ and it follows that $I(t) = \mathrm{i}\pi\epsilon(t)$, or

$$\epsilon(t) = \frac{1}{\mathrm{i}\pi}\int_{-\infty}^\infty \mathrm{d}\omega\, \frac{\mathrm{e}^{\mathrm{i}\omega t}}{\omega} \tag{D.10}$$

or in $\mathcal{F}_+$

$$\epsilon(t) = \frac{1}{\mathrm{i}\pi}\int_{-\infty}^\infty \mathrm{d}\omega\, \frac{\mathrm{e}^{-\mathrm{i}\omega t}}{\omega}. \tag{D.11}$$

A simple application of the shift theorem for the Fourier transform gives (in $\mathcal{F}_-$)

$$\frac{1}{\mathrm{i}\pi}\int_{-\infty}^\infty \mathrm{d}\omega\, \frac{\mathrm{e}^{\mathrm{i}\omega t}}{\omega-\Omega} = \mathrm{e}^{\mathrm{i}\Omega t}\epsilon(t). \tag{D.12}$$

Appendix E

Advanced least-squares techniques

Chapter 6 discussed the solution of least-squares (LS) problems by the use of the normal equations, but indicated that more sophisticated techniques exist which give the user more control for ill-conditioned problems. Two such methods are the subject of this appendix.

E.1 Orthogonal least squares

As discussed in chapter 6, there are more robust and informative means of solving LS problems than the normal equations. The next two sections describe two of the most widely used. The first is the *orthogonal* approach. Although the basic Gram–Schmidt technique has been used for many years and is described in the classic text [159], the technique has received only limited use for system identification until comparatively recently. Since the early 1980s, orthogonal methods have been strengthened and generalized by Billings and his co-workers who have used them to great effect in the NARMAX nonlinear modelling approach [149, 60] which is described in detail in chapter 6. The discussion here follows [32] closely. In order to make a framework suitable for generalizing to nonlinear systems, the analysis will be for the model form

$$y_i = \sum_{i=1}^{N_p} \beta_i \theta_i \tag{E.1}$$

where the θ_i are the model terms or *basis* and the a_i are the associated parameters. In the linear ARX case the θ's are either lagged y's or x's (ignoring noise for the moment). This model structure generates the usual LS system

$$[A]\{\beta\} = \{Y\} \tag{E.2}$$

(repeated here for convenience). However, it will be useful to rewrite these equations in a different form, namely

$$\beta_1\{\theta_1\} + \beta_2\{\theta_2\} + \cdots + \beta_{N_p}\{\theta_{N_p}\} = \{Y\} \tag{E.3}$$

where the vectors $\{\theta_i\}$ are the ith columns of $[A]$. Each column consists of a given model term evaluated for each time. So $[A]$ is the vector of vectors $(\{\theta_1\},\ldots,\{\theta_{N_p}\})$. In geometrical terms, $\{Y\}$ is decomposed into the linear combination of basis vectors $\{\theta_i\}$. Now, as there are only N_p vectors $\{\theta_i\}$ they can only generate a N_p-dimensional subspace of the N-dimensional space in which $\{Y\}$ sits and in general $N_p \ll N$; this subspace is called the *range* of the model basis. Clearly, $\{Y\}$ need not lie in the range (and in general because of measurement noise, it will not). In this situation, the system of equations (E.2) does not have a solution. However, as a next best case, one can find the parameters $\{\hat{\beta}\}$, for the *closest* point in the range to $\{Y\}$. This is the geometrical content of the LS method. This picture immediately shows why correlated model terms produce problems. If two model terms are the same up to a constant multiple, then the corresponding vectors $\{\theta\}$ will be parallel and therefore indistinguishable from the point of view of the algorithm. The contribution can be shared between the coefficients arbitrarily and the model will not be unique. The same situation arises if the set of $\{\theta\}$ vectors is linearly dependent. The orthogonal LS algorithm allows the identification of linear dependencies and the offending vectors can be removed.

The method assumes the existence of a square matrix $[T]$ with the following properties:

(1) $[T]$ is invertible.
(2)

$$[W] = [A][T]^{-1} \tag{E.4}$$

is column orthogonal, i.e. if $[W] = (\{W_1\},\ldots,\{W_{N_p}\})$, then

$$\langle\{W_i\},\{W_j\}\rangle = ||\{W_i\}||^2\delta_{ij} \tag{E.5}$$

where $\langle\ ,\ \rangle$ is the standard scalar product defined by

$$\langle\{u\},\{v\}\rangle = \langle\{v\},\{u\}\rangle = \sum_{i=1}^{N_p} u_i v_i \tag{E.6}$$

and $\|\ \|$ is the standard Euclidean norm $\|\{u\}\|^2 = \langle\{u\},\{u\}\rangle$.

If such a matrix exists, one can define the *auxiliary* parameters $\{g\}$ by

$$\{g\} = [T]\{\beta\} \tag{E.7}$$

and these are the solution of the original problem with respect to the new basis $\{W_i\}$, $i = 1,\ldots,N_p$, i.e.

$$[W]\{g\} = [A][T]^{-1}[T]\{\beta\} = [A]\{\beta\} = \{Y\} \tag{E.8}$$

or, in terms of the column vectors

$$g_1\{W_1\} + g_2\{W_2\} + \cdots + g_{N_p}\{W_{N_p}\} = \{Y\}. \tag{E.9}$$

The advantage of the coordinate transformation is that parameter estimation is almost trivial in the new basis. Taking the scalar product of this equation with the vector $\{W_i\}$, leads to

$$\begin{aligned} g_1\langle\{W_1\},\{W_j\}\rangle + g_2\langle\{W_2\},\{W_j\}\rangle + \cdots + g_{N_p}\langle\{W_{N_p}\},\{W_j\}\rangle \\ = \langle\{Y\},\{W_j\}\rangle \end{aligned} \tag{E.10}$$

and the orthogonality relation (E.5) immediately gives

$$g_j = \frac{\langle\{Y\},\{W_j\}\rangle}{\langle\{W_j\},\{W_j\}\rangle} = \frac{\langle\{Y\},\{W_j\}\rangle}{\|\{W_j\}\|^2} \tag{E.11}$$

so the auxiliary parameters can be obtained one at a time, unlike the situation in the physical basis where they must be estimated *en bloc*. Before discussing why this turns out to be important, the question of constructing an appropriate $[T]$ must be answered. If this turned out to be impossible the properties of the orthogonal basis would be irrelevant. In fact, it is a well-known problem in linear algebra with an equally well-known solution. The first step is to obtain the orthogonal basis $(\{W_1\},\ldots,\{W_{N_p}\})$ from the physical basis $(\{\theta_1\},\ldots,\{\theta_{N_p}\})$. The method is iterative and starts from the initial condition

$$\{W_1\} = \{\theta_1\}. \tag{E.12}$$

The *Gram–Schmidt* procedure now generates $\{W_2\}$ by subtracting the component of $\{\theta_2\}$ parallel to $\{W_1\}$, i.e.

$$\{W_2\} = \{\theta_2\} - \frac{\langle\{W_1\},\{\theta_2\}\rangle}{\langle\{W_1\},\{W_1\}\rangle}\{W_1\} \tag{E.13}$$

and $\{W_2\}$ and $\{W_1\}$ are orthogonal by construction. In the next step, $\{W_3\}$ is obtained by subtracting from $\{\theta_3\}$, components parallel to $\{W_2\}$ *and* $\{W_1\}$. After $N_p - 1$ iterations, the result is an orthogonal set. In matrix form, the procedure is generated by

$$[W] = [A] - [W][\alpha] \tag{E.14}$$

where the matrix $[\alpha]$ is defined by

$$[\alpha] = \begin{pmatrix} 0 & \alpha_{12} & \alpha_{13} & \ldots & \alpha_{1N_p} \\ 0 & 0 & \alpha_{23} & \ldots & \alpha_{2N_p} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{E.15}$$

and the $\alpha_{ij} = \langle\{W_i\},\{\theta_j\}\rangle/\langle\{W_i\},\{W_i\}\rangle$ must be evaluated from the top line down. A trivial rearrangement of (E.14)

$$[A] = [W](I + [\alpha]) \tag{E.16}$$

gives, by comparison with (E.4)

$$[T] = I + [\alpha] \tag{E.17}$$

or

$$[T] = \begin{pmatrix} 1 & \alpha_{12} & \alpha_{13} & \dots & \alpha_{1N_p} \\ 0 & 1 & \alpha_{23} & \dots & \alpha_{2N_p} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}. \tag{E.18}$$

Obtaining $[T]^{-1}$ is straightforward as the representation of $[T]$ in (E.18) is upper-triangular. One simply carries out the back-substitution part of the Gaussian elimination algorithm [102]. If the elements of $[T]^{-1}$ are labelled t_{ij}, then the jth column is calculated by back-substitution from

$$\begin{pmatrix} 1 & \alpha_{12} & \alpha_{13} & \dots & \alpha_{1N_p} \\ 0 & 1 & \alpha_{23} & \dots & \alpha_{2N_p} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix} \begin{pmatrix} t_{1j} \\ t_{2j} \\ \vdots \\ t_{N_p j} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \tag{E.19}$$

where the unit is in the jth position. The algorithm gives

$$t_{ij} = \begin{cases} 0, & \text{if } i > j \\ 1, & \text{if } i = j \\ -\displaystyle\sum_{k=i+1}^{j} \alpha_{jk} t_{kj}, & \text{if } i < j. \end{cases} \tag{E.20}$$

Having estimated the set of auxiliary parameters, $[T]^{-1}$ is used to recover the physical parameters, i.e.

$$\{\hat{\beta}\} = [T]^{-1}\{\hat{g}\}. \tag{E.21}$$

So, it is possible to solve the LS problem in the orthogonal basis and work back to the physical parameters. What then are the advantages of working in the orthogonal basis. There are essentially two fundamental advantages. The first relates to the fact that the model terms are orthogonal and the parameters can be obtained one at a time. Because of this, the model is expanded easily to include extra terms; previous terms need not be re-estimated. The second is related to the conditioning of the problem. Recall that the normal-equations approach fails if the columns of $[A]$ are linearly dependent. The orthogonal estimator is able to diagnose this problem. Suppose in the physical basis, $\{\theta_j\}$ is linearly dependent on $\{\theta_1\}, \dots, \{\theta_{j-1}\}$. As the subspace spanned by $\{\theta_1\}, \dots, \{\theta_{j-1}\}$ is identical to that spanned by $\{W_1\}, \dots, \{W_{j-1}\}$ by construction, then $\{\theta_j\}$ is in the latter subspace. This means that at the step in the Gram–Schmidt process where one

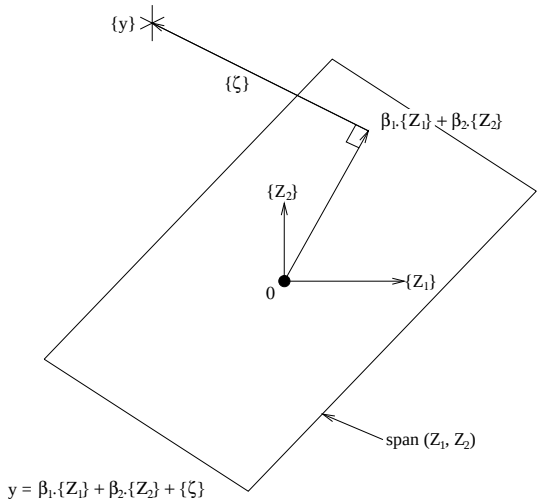


Figure E.1. The geometrical interpretation of LS estimation.

subtracts off components parallel to earlier vectors, $\{W_j\}$ will turn out to be the zero vector and $\|\{W_j\}\| = 0$. So if the algorithm generates a zero-length vector at any point, the corresponding physical basis term θ_j should be removed from the regression problem. If the vector is allowed to remain, there will be a division by zero at the next stage. In practice, problems are caused by vectors which are nearly parallel, rather than parallel. In fact measurement noise will ensure that vectors are never exactly parallel. The strategy in this case is to remove vectors which generate orthogonal terms with $\|\{W_j\}\| < \epsilon$ with epsilon some small constant. (There is a parallel here with the singular value decomposition method discussed in the next section.) In the normal-equations approach, one can monitor the conditioning number of the matrix $([A]^{\mathrm{T}}[A])$ and this will indicate when problems are likely to occur; however, this is a diagnosis only and cannot lead to a cure. Having established that the orthogonal estimator has useful properties, it remains to show why it is an LS estimator

Consider again the equations

$$\{Y\} = [A]\{\beta\} + \{\zeta\}. \tag{E.22}$$

If the correct model terms are present and the measurement noise ζ is zero, then the left-hand side vector $\{Y\}$ will lie inside the range of $[A]$ and (E.22) will have a clear solution. If the measurements are noisy, the vector $\{\zeta\}$ pushes the right-hand side vector outside the range (figure E.1) and there will only be a least-squares solution, i.e. the point in the range *nearest* $\{Y\}$. Now, the shortest distance between the range and the point $\{Y\}$ is the perpendicular distance. So the LS condition is met if $\{\zeta\}$ is perpendicular to the range. It is sufficient for this that

$\{\zeta\}$ is perpendicular to all the basis vectors of the range, i.e.

$$\langle\{\zeta\},\{W_j\}\rangle = 0, \quad \forall j \tag{E.23}$$

so

$$\langle\{Y\} - \sum_{i=1}^{N_p} g_i\{W_i\},\{W_j\}\rangle = 0, \quad \forall j \tag{E.24}$$

or

$$\langle\{Y\}\{W_j\}\rangle - \sum_{i=1}^{N_p} g_i\langle\{W_i\},\{W_j\}\rangle = 0 \tag{E.25}$$

and on using orthogonality (E.5), one recovers (E.11). This shows that the orthogonal estimator satisfies the LS condition. As a matter of fact, this approach applies just as well in the physical basis. In this case, the LS condition is simply

$$\langle\{\zeta\},\{\theta_j\}\rangle = 0, \quad \forall j \tag{E.26}$$

so

$$\langle\{Y\} - \sum_{i=1}^{N_p} \beta_i\{\theta_i\},\{\theta_j\}\rangle = 0, \quad \forall j \tag{E.27}$$

and

$$\langle\{Y\}\{\theta_j\}\rangle = \sum_{i=1}^{N_p} \beta_i\langle\{\theta_i\},\{\theta_j\}\rangle = 0. \tag{E.28}$$

Writing the $\{\theta\}$ vectors as components of $[A]$ and expanding the scalar products gives

$$\sum_{k=1}^{N} y_k A_{ki} = \sum_{j=1}^{N_p} \beta_j \left(\sum_{k=1}^{N} A_{ik} A_{ij} \right) \tag{E.29}$$

which are the normal equations as expected.

The final task is the evaluation of the covariance matrix for the parameter uncertainties. This is available for very little effort. As $[W]$ is column-orthogonal, $([W]^{\mathrm{T}}[W])$ is diagonal with ith element $\|\{W_i\}\|^2$ and these quantities have already been computed during the diagonalization procedure. This means that $([W]^{\mathrm{T}}[W])^{-1}$ is diagonal with ith-element $\|\{W_i\}\|^{-2}$ and is readily available. The covariance matrix in the auxiliary basis is simply

$$[\Sigma]_g = \sigma_\zeta^2 \,\mathrm{diag}(\|\{W_i\}\|^{-2}) \tag{E.30}$$

and the covariance matrix in the physical basis is obtained from standard statistical theory as [149]

$$[\Sigma] = [T]^{-1}[\Sigma]_g[T]. \tag{E.31}$$

In order to carry out effective structure detection, the significance factors (6.31) can be evaluated in the orthogonal basis, Because the model terms are

uncorrelated in the auxiliary basis, the error variance is reduced when a term $g_i\{W_i\}$ is added to the model by the variance of the term. By construction, the $\{W_i\}$ sequences are zero-mean[1] so the variance of a given term is $\sum_{j=1}^{N} g_i^2 W_{ij}^2$ and the significance factor for the ith model term is simply

$$s_{W_i} = 100 \frac{\sum_{j=1}^{N} g_i^2 W_{ij}^2}{\sigma_y^2}. \tag{E.32}$$

In the literature relating to the NARMAX model, notably [149], this quantity is referred to as the *error reduction ratio* or ERR. To stress the point, because the model terms are uncorrelated in the auxiliary basis, low significance terms will necessarily have low ERRs and as such will all be detected.

A noise model is incorporated into the estimator by fitting parameters, estimating the prediction errors, fitting the noise model and then iterating to convergence. In the extended orthogonal basis which incorporates noise terms, all model terms are uncorrelated so the estimator is guaranteed to be free of bias.

E.2 Singular value decomposition

The subject of this section is the second of the robust LS procedures alluded to earlier. Although the algorithm is arguably more demanding than any of the others discussed, it is also the most foolproof. The theoretical bases for the results presented here are actually quite deep and nothing more than a cursory summary will be presented here. In fact, this is the one situation where the use of a 'canned' routine is recommended. The SVDCMP routine from [209] is recommended as excellent. If more theoretical detail is required, the reader is referred to [159] and [101].

Suppose one is presented with a square matrix $[A]$. It is a well-known fact that there is almost always a matrix $[U]$ which converts $[A]$ to a diagonal form by the similarity transformation[2]

$$[S] = [U]^{\mathrm{T}}[A][U] \tag{E.33}$$

where the diagonal elements s_i are the eigenvalues of $[A]$ and the ith column $\{u_i\}$ of $[U]$ is the eigenvector of $[A]$ belonging to s_i. $[U]$ is an orthogonal matrix. An alternative way of regarding this fact is to say that any matrix $[A]$ admits a decomposition

$$[A] = [U][S][U]^{\mathrm{T}} \tag{E.34}$$

with $[S]$ diagonal containing the eigenvalues and $[U]$ orthogonal containing the eigenvectors. Now, it is a non-trivial fact that this decomposition also extends to

[1] Except at most one if $y(t)$ has a finite mean. In this case, the Gram–Schmidt process can be initialized with $\{W_0\}$ equal to a constant term.

[2] Mathematically, the diagonalizable matrices are *dense* in the space of matrices. This means if a matrix fails to be diagonalizable, it can be approximated arbitrarily closely by one that is.

rectangular matrices $[A]$ which are $M \times N$ with $M > N$. In this case

$$[A] = [U][S][V]^{\mathrm{T}} \tag{E.35}$$

where $[U]$ is a $M \times N$ column-orthogonal matrix, i.e. $[U]^{\mathrm{T}}[U] = I$, $[S]$ is a $N \times N$ diagonal matrix and $[V]$ is a $N \times N$ column-orthogonal matrix, i.e. $[V]^{\mathrm{T}}[V] = I$. Because $[V]$ is square, it is also row-orthogonal, i.e. $[V][V]^{\mathrm{T}} = I$.

If $[A]$ is square and invertible, then $[V] = [U]$ and the inverse is given by

$$[A]^{-1} = [U][S]^{-1}[U]^{\mathrm{T}}. \tag{E.36}$$

If $[A]$ is $M \times N$ with $M > N$, then the quantity

$$[A]^{\dagger} = [U][S]^{-1}[V]^{\mathrm{T}} \tag{E.37}$$

is referred to as the *pseudo-inverse* because

$$[A]^{\dagger}[A] = I. \tag{E.38}$$

(Note that $[A][A]^{\dagger} \neq I$ because $[U]$ is not row orthogonal.)

Now $[S]^{-1}$ is the diagonal matrix with entries s_i^{-1} and it is clear that a square matrix $[A]$ can only be singular if one of the *singular values* s_i is zero. The number of non-zero singular values is the *rank* of the matrix., i.e. if $[A]$ has only $r < N$ linearly independent columns, then the rank is r and $N - r$ singular values are zero.

Consider, the familiar system of equations

$$[A]\{\beta\} = \{Y\} \tag{E.39}$$

and suppose that $[A]$ is square and invertible. ($\{Y\}$ is guaranteed to be in the range of $[A]$ in this case.) In this case the solution of the equation is simply

$$\{\hat{\beta}\} = [A]^{-1}\{Y\} = [U][S]^{-1}[U]^{\mathrm{T}}\{Y\} \tag{E.40}$$

and there are no surprises.

The next most complicated solution is that $\{Y\}$ is in the range of $[A]$, but $[A]$ is not invertible. The solution in this case is not unique. The reason is as follows: if $[A]$ is singular, there exist vectors $\{n\}$, such that

$$[A]\{n\} = \{0\}. \tag{E.41}$$

In fact, if the rank of the matrix $[A]$ is $r < N$, there are $N - r$ linearly independent vectors which satisfy condition (E.41). These vectors span a space called the *nullspace* of $[A]$. Now suppose $\{\hat{\beta}\}$ is a solution of (E.39), then so is $\{\hat{\beta}\} + \{n\}$ where $\{n\}$ is any vector in the nullspace, because

$$[A](\{\hat{\beta}\} + \{n\}) = [A]\{\hat{\beta}\} + [A]\{n\} = \{Y\} + \{0\} = \{Y\}. \tag{E.42}$$

Now if the matrix $[S]^{-1}$ has all elements corresponding to zero singular values replaced by zeroes to form the matrix $[S_{\mathrm{d}}]^{-1}$, then the remarkable fact is that the solution

$$\{\hat{\beta}\} = [U][S_{\mathrm{d}}]^{-1}[U]^{\mathrm{T}}\{Y\} \tag{E.43}$$

is the one with smallest norm, i.e. with smallest $||\{\hat{\beta}\}||$. The reason for this is that the columns of $[V]$ corresponding to zero singular values span the nullspace. Taking this prescription for $[S_{\mathrm{d}}]^{-1}$ means that there is no nullspace component in the solution. This simple replacement $[S_{\mathrm{d}}]$ for $[S]$, automatically kills any linearly dependent vectors in $[A]$. Now, recall that linear dependence in $[A]$ is a problem with the LS methods previously discussed in chapter 6.

It transpires that in the case of interest for system identification, where $\{Y\}$ is *not* in the range of $[A]$, whether it is singular or not, then (E.43) actually furnishes the LS solution. This remarkable fact means that the singular value decomposition provides a LS estimator which automatically circumnavigates the problem of linear dependence in $[A]$. The proofs of the various facts asserted here can be found in the references cited at the beginning of this section.

In practice, because of measurement noise, the singular values will not be exactly zero. In this case, one defines a tolerance ϵ, and deletes any singular values less than this threshold. The number of singular values less than ϵ is referred to as the *effective nullity* n_ϵ and the *effective rank* is $N - n_\epsilon$. The critical fact about the singular-value-decomposition estimator is that one *must* delete the near-zero singular values; hence the method is foolproof. If this is neglected, the method is no better than using the normal equations.

A similar derivation to the one given in section (6.3.2) suffices to establish the covariance matrix for the estimator

$$[\Sigma] = \sigma_\zeta^2[V][S_{\mathrm{d}}]^{-2}[V]^{\mathrm{T}}. \tag{E.44}$$

E.3 Comparison of LS methods

As standard, the operation counts in the following discussions refer to multiplications only; additions are considered to be negligible.

E.3.1 Normal equations

As the inverse covariance matrix $[A]^{\mathrm{T}}[A]$ (sometimes called the *information matrix*) is symmetric, it can be formed with $\frac{1}{2}P(P+1)N \approx \frac{1}{2}P^2N$ operations. If the inversion is carried out using LU decomposition as recommended [209], the operation count is P^3. (If the covariance matrix is not needed, the LU decomposition can be used to solve the normal equations without an inversion; in which case the operation count is $\frac{1}{3}P^3$.) Back-substitution generates another $PN + P^2$ operation, so to leading order the operation count is $P^3 + \frac{1}{2}P^2N$.

Questions of speed aside, the normal equations have the advantage of simplicity. In order to implement the solution, the most complicated operations

are matrix inversion and multiplication. A problem occurs if the information matrix is singular and the parameters do not exist. More seriously, the matrix may be near-singular so that the parameters cannot be trusted. The determinant or condition number of $[A]^{\mathrm{T}}[A]$ gives an indication of possible problems, but cannot suggest a solution, i.e. which columns of $[A]$ are correlated and should be removed.

E.3.2 Orthogonal least squares

Computing the $[T]$ matrix requires $\frac{1}{2}P(P-1)N \approx \frac{1}{2}P^2N$ operations. Generating the auxiliary data requires exactly the same number. Generating the auxiliary parameters costs $2PN$. A little elementary algebra suffices to show that inverting the $[T]$ matrix needs

$$\tfrac{1}{12}P(P+1)(2P+1) - \tfrac{1}{4}P(P+1) \approx \tfrac{1}{6}P^3 \tag{E.45}$$

operations. Finally, generating the true parameters requires $\frac{1}{2}P(P-1)$ multiplications. To leading order, the overall operations count is $\frac{1}{6}P^3 + P^2N$. This count is only smaller than that for the normal equations if $N < \frac{5}{3}P$ which is rather unlikely. As a consequence, this method is a little slower than using the normal equations.

The orthogonal estimator has a number of advantages over the normal equations. The most important one relates to the fact that linear dependence in the $[A]$ matrix can be identified and the problem can be removed. Another useful property is that the parameters are identified one at a time, so if the model is enlarged the parameters which already exist need not be re-estimated.

E.3.3 Singular value decomposition

The 'black-box' routine recommended in this case is SVDCMP from [209]. The routine is divided into two steps. First a householder reduction to bidiagonal form is needed with an operation count of between $\frac{2}{3}P^3$ and $\frac{4}{3}P^3$. The second step is a QR step with an operation count of roughly $3\,P^3$. This gives an overall count of about $4P^3$. This suggests that singular value decomposition is one of the slower routines.

The advantage that singular value decomposition has over the two previous routines is that it is foolproof as long as small singular values are deleted. Coping with linear dependence in $[A]$ is part of the algorithm.

E.3.4 Recursive least squares

Calculating the $\{K\}$ matrix at each step requires $2P^2 + P$ operations. The $[P]$ matrix and $\{\beta\}$ vector require $2P^2$ and $2P$ respectively. Assuming one complete pass through the data, the overall count is $\approx 4P^2N$ and as N is usually much greater than P, this is the slowest of the routines. One can speed things up in two

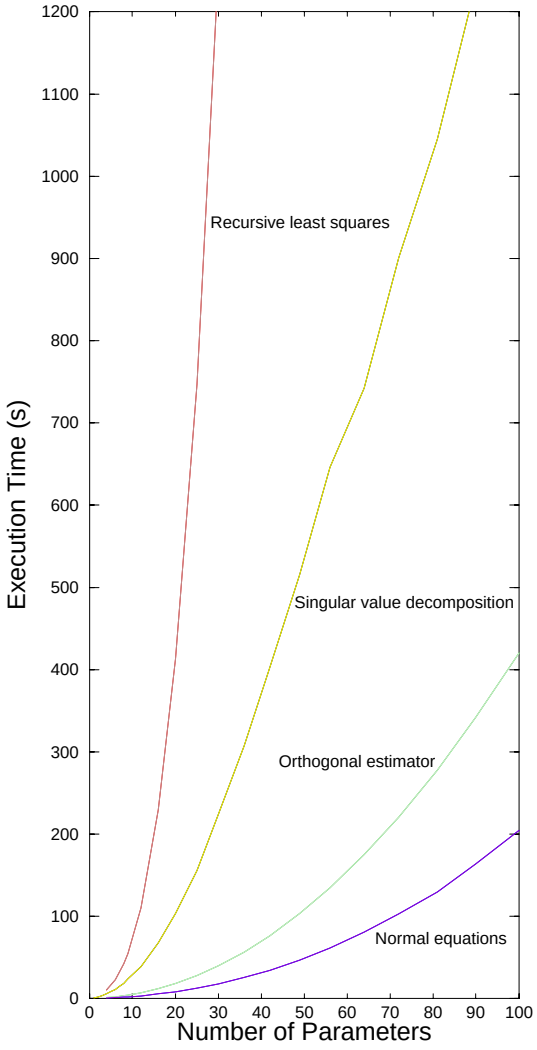


Figure E.2. Comparison of execution times for LS methods.

ways. First, the algorithm can be terminated once the parameters have stabilized to an appropriate tolerance. Secondly, a so-called 'fast' recursive LS scheme can be applied where only the diagonals of the covariance matrix are updated [286].

The recursive LS scheme should not be used if the processes are stationary and the system is time-invariant because of the overheads. If it is necessary to track time-variation of the parameters there may be no alternative.

In order to check these conclusions, all the methods were implemented with

a common input–output interface and the times for execution were evaluated for a number of models with up to 100 parameters. Figure E.2 shows the results, which more or less confirm the operation counts given earlier.

Appendix F

Neural networks

Artificial neural networks are applied in this book only to the problem of system identification. In fact, the historical development of the subject was mainly in terms of pattern recognition. For those readers interested in the history of the subject, this appendix provides a precis. Readers only interested in the structures used in system identification may skip directly to sections F.4 and F.5 where the relevant network paradigms—multi-layer perceptrons (MLP) and radial basis function networks (RBF)—are discussed. Any readers with an interest in software implementation of these networks should consult the following appendix.

F.1 Biological neural networks

Advanced as contemporary computers are, none has the capability of carrying out certain tasks—notably pattern recognition—as effectively as the human brain (or mammalian brain for that matter). In recent years a considerable effort has been expended in pursuing the question of why this should be.

There are essential differences in the way in which the brain and standard serial machines compute. A conventional Von Neumann computer operates by passing instructions sequentially to a single processor. The processor is able to carry out moderately complex instructions very quickly. As an example, at one point many IBM-compatible personal computers were based on the Intel 80486 microprocessor. This chip operates with a clock cycle of 66 MHz, and is capable of carrying out approximately 60 distinct operations (if different address modes are considered, this number is closer to 500). Averaging over long and short instructions, the chip is capable of performing about 25 million instructions per second (MIPs). (There is little point in describing the performance of a more modern processor as it will without doubt be obsolete by the time this book is published.) State-of-the-art vector processors may make use of tens or hundreds of processors.

In contrast, neurons—the processing units of the brain—can essentially carry out only a single instruction. Further, the delay between instructions is of the

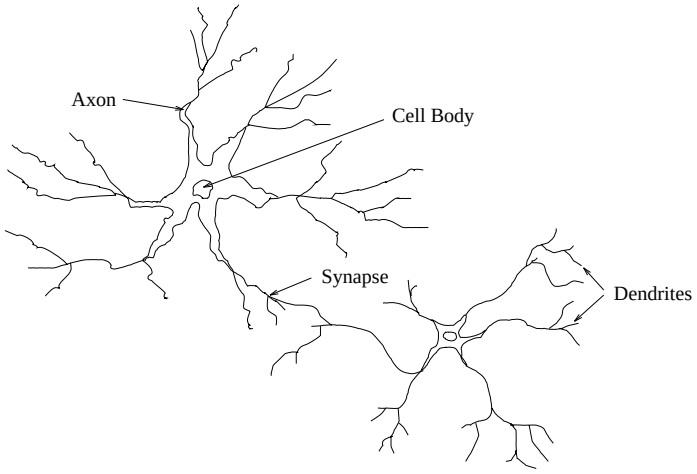


Figure F.1. Structure of the biological neuron.

order of milliseconds; the neuron operates at approximately 0.001 MIPs. The essential difference with an electronic computer is that the brain comprises a densely interconnected network of about 10^{10} processors operating in parallel.

It is clear that any superiority that the brain enjoys over electronic computers can only be due to its massively parallel nature; the individual processing units are considerably more limited. (In tasks where an algorithm is serial by nature, the brain cannot compete).

The construction of *artificial neural networks* (ANNs) has been an active field of research since the mid-1940s. In the first case, it was hoped that theoretical and computational models would shed light on the properties of the brain. Secondly, it was hoped that a new paradigm for a computer would emerge which would prove more powerful than a Von Neumann serial computer when presented with certain tasks.

Before proceeding to a study of artificial neural networks, it is useful to discuss the construction and behaviour of biological neurons in order to understand the properties which have been incorporated into model neurons.

F.1.1 The biological neuron

As discussed earlier, the basic processing unit of the brain is the nerve cell or *neuron*; the structure and operation of the neuron is the subject of this section. In brief, the neuron acts by summing stimuli from connected neurons. If the total stimulus or *activation* exceeds a certain threshold, the neuron 'fires', i.e. it generates a stimulus which is passed on into the network. The essential components of the neuron are shown in the schematic figure F.1.

The cell body, which contains the cell nucleus, carries out those biochemical

reactions which are necessary for sustained functioning of the neuron. Two main types of neuron are found in the cortex (the part of the brain associated with the higher reasoning capabilities); they are distinguished by the shape of the cell body. The predominant type have a pyramid-shaped body and are usually referred to as *pyramidal* neurons. Most of the remaining nerve cells have star-shaped bodies and are referred to as *stellate* neurons. The cell bodies are typically a few micrometres in diameter. The fine tendrils surrounding the cell body are the *dendrites*; they typically branch profusely in the neighbourhood of the cell and extend for a few hundred micrometres. The *nerve fibre* or *axon* is usually much longer than the dendrites, sometimes extending for up to a metre. The axon only branches at its extremity where it makes connections with other cells.

The dendrites and axon serve to conduct signals to and from the cell body. In general, input signals to the cell are conducted along the dendrites, while the cell output is directed along the axon. Signals propagate along the fibres as electrical impulses. Connections between neurons, called *synapses*, are usually made between axons and dendrites although they can occur between dendrites, between axons and between an axon and a cell body.

Synapses operate as follows: the arrival of an electrical nerve impulse at the end of an axon say, causes the release of a chemical—a *neurotransmitter* into the synaptic gap (the region of the synapse, typically 0.01 μm). The neurotransmitter then binds itself to specific sites—*neuroreceptors* usually in the dendrites of the target neuron. There are distinct types of neurotransmitters: *excitatory* transmitters, which trigger the generation of a new electrical impulse at the receptor site; and *inhibitory* transmitters, which act to prevent the generation of new impulses. A discussion of the underlying biochemistry of this behaviour is outside the scope of this appendix; however, those interested can consult the brief discussion in [73], or the more detailed treatment in [1].

Table F.1 (reproduced from [1]) gives the typical properties of neurons within the cerebral cortex (the term remote sources refers to sources outside the cortex).

The operation of the neuron is very simple. The cell body carries out a summation of all the incoming electrical impulses directed inwards along the dendrite. The elements of the summation are individually weighted by the strength of the connection or synapse. If the value of this summation—the *activation* of the neuron—exceeds a certain threshold, the neuron fires and directs an electrical impulse outwards via its axon. From synapses with the axon, the signal is communicated to other neurons. If the activation is less than the threshold, the neuron remains dormant.

A mathematical model of the neuron, exhibiting most of the essential features of the biological neuron, was developed as early as 1943 by McCulloch and Pitts [181]. This model forms the subject of the next section; the remainder of this section is concerned with those properties of the brain which emerge as a result of its massively parallel nature.

Table F.1. Properties of the cortical neural network

Variable	Value
Neuronal density	40 000 mm^{-3}
Neuronal composition:	
Pyramidal	75%
Stellate	25%
Synaptic density	8×10^8 mm^{-3}
Axonal length density	3200 m mm^{-3}
Dendritic length density	400 m mm^{-3}
Synapses per neuron	20 000
Inhibitory synapses per neuron	2000
Excitatory synapses from remote sources per neuron	9000
Excitatory synapses from local sources per neuron	9000
Dendritic length per neuron	10 mm

F.1.2 Memory

The previous discussion was concerned with the neuron, the basic processor of the brain. An equally important component of any computer is its memory. In an electronic computer, regardless of the particular memory device in use, data are stored at specific physical locations within the device from which they can be retrieved and directed to the processor. The question now arises of how knowledge can be stored in a neural network, i.e. a massively connected network of nominally identical processing elements.

It seems clear that the only place where information can be stored is in the network connectivity and the strengths of the connections or synapses between neurons. In this case, knowledge is stored as a distributed quantity throughout the entire network. The act of retrieving information from such a memory is rather different from that for an electronic computer. In order to access data on a PC say, the processor is informed of the relevant address in memory, and it retrieves it from that location. In a neural network, a stimulus is presented (i.e. a number of selected neurons receive an external input) and the required data are encoded in the subsequent pattern of neuronal activations. Potentially, recovery of the pattern is dependent on the entire distribution of connection weights or synaptic strengths.

One advantage of this type of memory retrieval system is that it has a much greater resistance to damage. If the surface of a PC hard disk is damaged, all data at the affected locations may be irreversibly corrupted. In a neural network, because the knowledge is encoded in a distributed fashion, local damage to a portion of the network may have little effect on the retrieval of a pattern when a stimulus is applied.

F.1.3 Learning

According to the argument in the previous section, knowledge is encoded in the connection strengths between the neurons in the brain. The question arises of how a given distributed representation of data is obtained. There appear to be only two ways: it can be present from birth or it can be learned. The first type of knowledge is common in the more basic animals; genetic 'programming' provides an initial encoding of information which is likely to prove vital to the survival of the organism. As an example, the complex mating rituals of certain insects are certainly not learnt, the creatures having undergone their initial development in isolation.

The second type of knowledge is more interesting: the initial state of the brain at birth is gradually modified as a result of its interaction with the environment. This development is thought to occur as an evolution in the connection strengths between neurons as different patterns of stimulus and appropriate response are activated in the brain as a result of signals from the sense organs.

The first explanation of learning in terms of the evolution of synaptic connections was given by Hebb in 1949 [131]. Following [73], a general statement of Hebb's principle is:

> When a cell A excites cell B by its axon and when in a repetitive and persistent manner it participates in the firing of B, a process of growth or of changing metabolism takes place in one or both cells such that the effectiveness of A in stimulating and impulsing cell B is increased with respect to all other cells which can have this effect.

If some similar mechanism could be established for computational models of neural networks, there would be the attractive possibility of 'programming' these systems simply by presenting them with a sequence of stimulus-response pairs so that the network can learn the appropriate relationship by reinforcing some of its internal connections. Fortunately, Hebb's rule proves to be quite simple to implement for artificial networks (although in the following discussions, more general learning algorithms will be applied).

F.2 The McCulloch–Pitts neuron

Having found a description of a biological neural network, the first stage in deriving a computational model was to represent mathematically the behaviour of a single neuron. This step was carried out in 1943 by the neurophysiologist Warren McCulloch and the logician Walter Pitts [181].

The McCulloch–Pitts model (MCP model) constitutes the simplest possible neural network model. Because of its simplicity it is possible without too much effort to obtain mathematically rigorous statements regarding its range

of application; the major disadvantage of the model is that this range is very limited. The object of this section is to demonstrate which input–output systems or functions allow representation as an MCP model. In doing this, a number of techniques which are generally applicable to more complex network paradigms are encountered.

F.2.1 Boolean functions

For a fruitful discussion, limits must be placed upon the range of systems or functions which the MCP model will be asked to represent; the output of a nonlinear dynamical system, for example, can be represented as a nonlinear *functional* of the whole input history. This is much too general to allow a simple analysis. For this reason, the objects of study here are the class of multi-input–single-output (MISO) systems which have a representation as a *function* of the instantaneous input values, i.e.

$$y = f(x_1, x_2, \ldots, x_n) \tag{F.1}$$

y being the output and $x_1, \ldots, x_n$ being the inputs. A further constraint is imposed which will be justified in the next section; namely, the variables y and $x_1, \ldots, x_n$ are only allowed to take the values 0 and 1. If this set of values is denoted $\{0, 1\}$, the functions of interest have the form

$$f : \{0, 1\}^n \longrightarrow \{0, 1\}. \tag{F.2}$$

Functions of this type are call *Boolean*. They arise naturally in symbolic logic where the value 1 is taken to indicate truth of a proposition while 0 indicates falsity (depending on which notation is in use, $1 = T =$ .true. and $0 = F =$ .false.). In the following, curly brackets shall be used to represent those Boolean functions which are represented by logical propositions, e.g. the function

$$f(x_1, x_2) = \{x_1 = x_2\}. \tag{F.3}$$

Given the inputs to this function, the output is evaluated as follows

$$\begin{aligned} f(0, 0) &= \{0 = 0\} = \text{.true.} = 1 \\ f(0, 1) &= \{0 = 1\} = \text{.false.} = 0 \\ f(1, 0) &= \{1 = 0\} = \text{.false.} = 0 \\ f(1, 1) &= \{1 = 1\} = \text{.true.} = 1. \end{aligned}$$

A Boolean function which is traditionally of great importance in neural network theory is the exclusive-or function XOR(x_1, x_2) which is true if one, but not both, of its arguments is true. It is represented by the Boolean,

	0	1
0	0	1
1	1	0

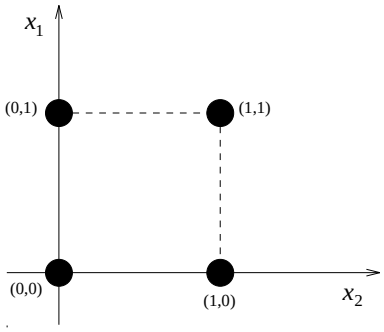


Figure F.2. Domain of Boolean function with two inputs.

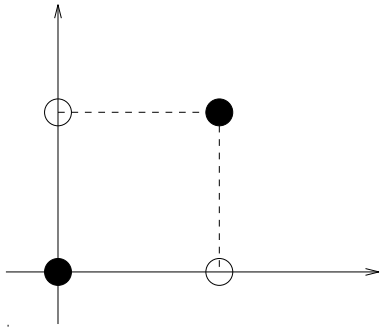


Figure F.3. Pictorial representation of the Exclusive-Or.

Note that this function also has a representation as the proposition $\{x_1 \neq x_2\}$.

There is a very useful pictorial representation of the Boolean functions with two arguments $f : \{0,1\}^2 \longrightarrow \{0,1\}$. The possible combinations of input values can be represented as the vertices of the unit square in the Cartesian plane (figure F.2).

This set of possible inputs is called the *domain* of the function. Each Boolean function on this domain is now specified by assigning the value 0 or 1 to each point in the domain. If a point on which the function is true is represented by a white circle, and a point on which the function is false by a black circle, one obtains the promised pictorial representation. As an example, the XOR function has the representation shown in figure F.3.

For the general Boolean function with n inputs, the domain is the set of vertices of the unit hypercube in n dimensions. It is also possible to use the black and white dots to represent functions in three dimensions, but clearly not in four or more.

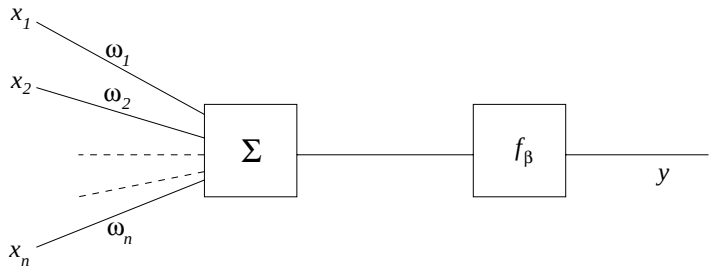


Figure F.4. McCulloch–Pitts neuron.

F.2.2 The MCP model neuron

In the MCP model, each input to a neuron is assumed to come from a connected neuron; the only information considered to be important is whether the connected neuron has fired or not (all neurons are assumed to fire with the same intensity). This allows a restriction of the possible input values to 0 and 1. On the basis of this information, the neuron will either fire or not fire, so the output values are restricted to be 0 or 1 also. This means that a given neuron can be identified with some Boolean function. The MCP model must therefore be able to represent an arbitrary Boolean function. The MCP neuron can be illustrated as in figure F.4.

The input values $x_i \in \{0, 1\}$ are weighted by a factor w_i before they are passed to the body of the MCP neuron (this allows the specification of a strength for the connection). The weighted inputs are then summed and the MCP neuron fires if the weighted sum exceeds some predetermined threshold β. So the model fires if

$$\sum_{i=1}^{n} w_i x_i > \beta \tag{F.4}$$

and does not fire if

$$\sum_{i=1}^{n} w_i x_i \leq \beta. \tag{F.5}$$

Consequently, the MCP neuron has a representation as the proposition

$$\left\{ \sum_{i=1}^{n} w_i x_i > \beta \right\} \tag{F.6}$$

which is clearly a Boolean function. As a real neuron could correspond to an arbitrary Boolean function, there are two fundamental questions which can be asked:

(1) Can a MCP model of the form (F.6) represent an *arbitrary* Boolean function $f(x_1, \ldots, x_n)$? That is, do there exist values for $w_1, \ldots, w_n$ and β such that $f(x_1, \ldots, x_n) = \{\sum_{i=1}^{n} w_i x_i > \beta\}$?

(2) If a MCP model exists, how can the weights and thresholds be determined? In keeping with the spirit of neural network studies one would like a *training* algorithm which would allow the MCP model to learn the correct parameters by presenting it with a *finite* number of input–output pairs. Does such an algorithm exist?

Question (2) can be answered in the affirmative, but the discussion is rather technical and the resulting training algorithm is given in the next section in the slightly generalized form used for *perceptrons*, or networks of MCP neurons. The remainder of this section is concerned with question (1) and, in short, the answer is *no*.

Consider the function $f : \{0,1\}^2 \longrightarrow \{0,1\}$,

$$f(x_1, x_2) = \{x_1 = x_2\}. \tag{F.7}$$

Suppose an MCP model exists, i.e. there exist parameters $w_1, \ldots, w_n, \beta$ such that

$$\left\{\sum_{i=1}^{n} w_i x_i > \beta\right\} = \{x_1 = x_2\} \tag{F.8}$$

or, in this two-dimensional case,

$$\{w_1 x_1 + w_2 x_2 > \beta\} = \{x_1 = x_2\}. \tag{F.9}$$

Considering all possible values of x_1, x_2 leads to

x_1	x_2	$\{x_1 = x_2\}$	$\Rightarrow$	$\{w_1 x_1 + w_2 x_2 > \beta\}$	
0	0	1	$\Rightarrow$	$0 > \beta$	(a)
0	1	0	$\Rightarrow$	$w_2 \leq \beta$	(b)
1	0	0	$\Rightarrow$	$w_1 \leq \beta$	(c)
1	1	1	$\Rightarrow$	$w_1 + w_2 > \beta$	(d)

Now, (a) and (b) $\Rightarrow$

$$0 > \beta \geq w_2. \tag{F.10}$$

Therefore w_2 is strictly negative, so

$$w_1 + w_2 < w_1. \tag{F.11}$$

However, according to (c) $w_1 \leq \beta$, therefore

$$w_1 + w_2 < \beta \tag{F.12}$$

which contradicts (d). This contradiction shows that the initial hypothesis was false, i.e. no MCP exists. This proof largely follows [202].

The simplest way to determine the limitations of the class of MCP models is to consider the geometry of the situation. In n dimensions the equation

$$\sum_{i=1}^{n} w_i z_i = \beta \tag{F.13}$$

represents a hyperplane which separates two regions of the n-dimensional input space. One region $\mathcal{U}$ consists of all those points $(z_1, \ldots, z_n)$ (where the z_i can take any real values) such that

$$\sum_{i=1}^{n} w_i z_i > \beta. \tag{F.14}$$

The other region $\mathcal{L}$ contains all those points such that

$$\sum_{i=1}^{n} w_i z_i < \beta. \tag{F.15}$$

The proof of this is straightforward. Consider the function

$$l(z_1, z_2) = \sum_{i=1}^{n} w_i z_i - \beta \tag{F.16}$$

By equations (F.14) and (F.15), this function is strictly positive on $\mathcal{U}$, strictly negative on $\mathcal{L}$ and zero *only* on the hyperplane [1].

This means that each MCP model (F.6) specifies a plane which divides the input space into two regions $\mathcal{U}$ and $\mathcal{L}$ (where $\mathcal{L}$ is now defined to include the plane itself). Further, by (F.14) and (F.15) the MCP model takes the values 0 on $\mathcal{L}$ and 1 on $\mathcal{U}$. This means that if one is to represent an arbitrary Boolean function f by an MCP model, there must exist a plane which splits off the points on which $f = 1$ from the points on which $f = 0$. Using the pictorial representation of section F.1, such a plane should separate the white dots from the black dots. It is now obvious why there is no MCP model for the Boolean $\{x_1 = x_2\}$, no such plane exists (figure F.5).

The XOR function of figure F.3 is a further example. In fact, these are the only two-input Boolean functions which do not have an MCP model. It is quite simple to determine how many n-input Boolean functions are possible, there are clearly 2^n points in the domain of such a function (consider figure F.2), and the

[1] Suppose that the hyperplane does not partition the space in the manner described earlier. In that case there will be a point of $\mathcal{U}$ say P, and a point of $\mathcal{L}$ say Q on the same side of the hyperplane with neither lying on it. Now, f is positive at P and negative at Q and because it is continuous it must, by the intermediate value theorem have a zero somewhere on the straight line segment between P and Q. However, no point of this line segment is on the hyperplane so there is a zero of f off the hyperplane. This establishes a contradiction, so the hyperplane (F.13) must partition the space in the way described above.

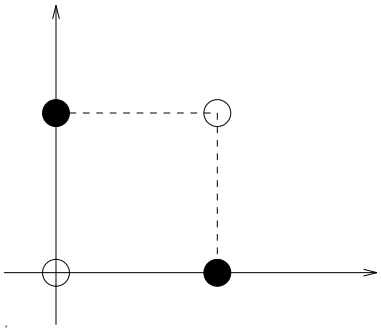


Figure F.5. Pictorial representation of $\{x_1 = x_2\}$.

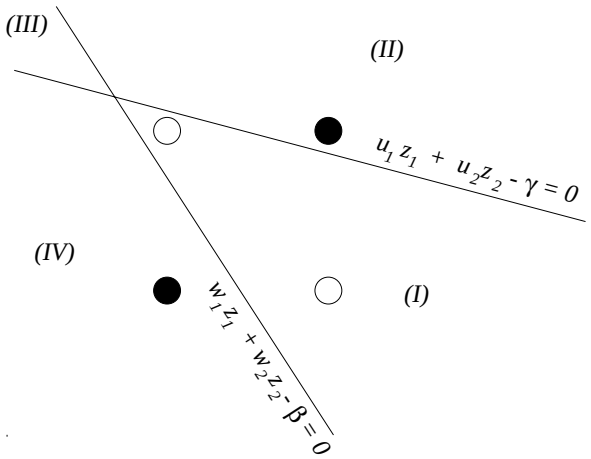


Figure F.6. Division of space into four regions.

value of the function can be freely chosen to be 0 or 1 at each of these points. Consequently, the number of possible functions is 2^{2^n}.

There are therefore 16 two-input Boolean functions, of which only two are not representable by an MCP model, i.e. 87.5% of them are. This is not such a bad result. However, the percentage of functions which can be represented falls off rapidly with increasing number of inputs.

The solution to the problem is to use more than one MCP unit to represent a MISO system. For example, if two MCP units are used in the two-input case, one can partition the plane of the XOR(x_1, x_2) function into four regions as follows, and thereby solve the problem.

Consider the two lines in figure F.6. The parameters of the first line $w_1 z_1 + w_2 z_2 = \beta$ define an MCP model MCP$_\beta$, the parameters of the second $u_1 z_1 + u_2 z_2 = \gamma$ define a model MCP$_\gamma$. This configuration of lines separates the white dots (in region I) from the black dots as required. The points where the

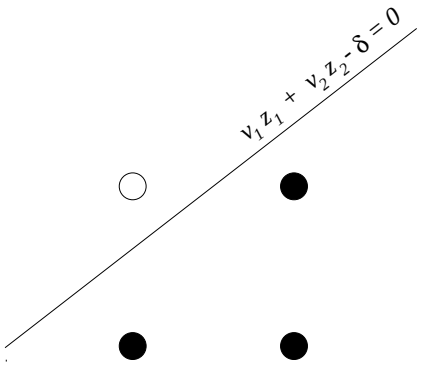


Figure F.7. Pictorial representation of MCP_δ.

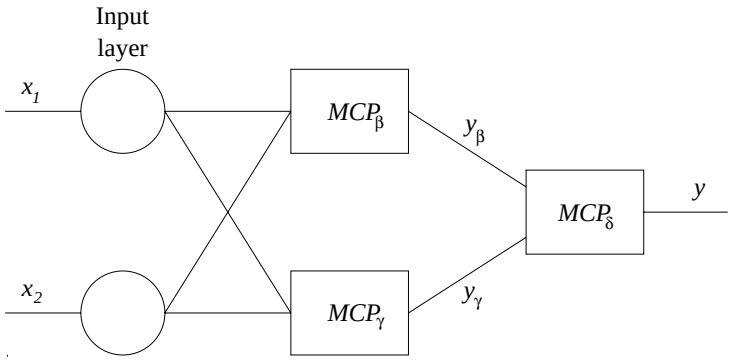


Figure F.8. Network of MCP neurons representing XOR.

XOR function is 1 are in the region I where the outputs y_β and y_γ from MCP_β and MCP_γ are 1 and 0 respectively, all other pairs of outputs indicate regions where XOR is false. It is possible to define a Boolean function $f(y_\beta, y_\gamma)$ whose output is 1, if and only if $(y_\beta, y_\gamma) = (1, 0)$. The pictorial representation of this Boolean is shown in figure F.7.

It is clear from the figure that this function has an MCP model, say MCP_δ (with weights v_1, v_2 and threshold δ). Considering the network of MCP models shown in figure F.8, it is clear that the final output is 1 if, and only if, the input point (x_1, x_2) is in region I in figure F.6. Consequently, the network provides a representation of the XOR function.

There are an infinite number of possible MCP models representing the Boolean in figure F.7, each one corresponding to a line which splits off the white dots. There are also infinitely many pairs of lines which can be used to define region I in figure F.6.

This three-neuron network (with two trivial input neurons whose only purpose is to distribute unchanged the inputs to the first MCP layer) actually gives

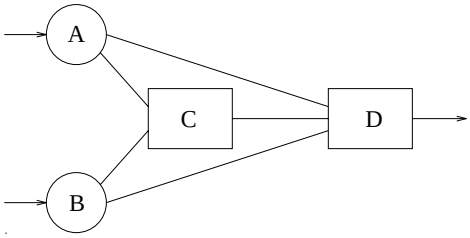


Figure F.9. Minimal network of MCP neurons representing XOR.

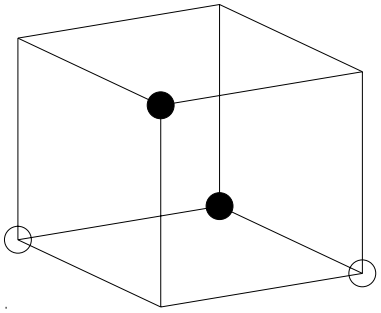


Figure F.10. Minimal network of MCP neurons representing XOR.

the minimal representation of the XOR function *if* the network is layered as before and neurons only communicate with neurons in adjacent layers.

If a fully connected network is allowed, the representation in figure F.9 is minimal, containing only four neurons (A and B are input neurons which do not compute and C and D are MCP neurons). The geometrical explanation of this fact is quite informative [73]. Suppose neuron C in figure F.9 computes the value of (y_{A} .and. y_{B}). The resulting inputs and outputs to neuron D are summarized in the following table:

Inputs			Output
y_{A}	y_{B}	y_{A} .and. y_{B}	y_{D}
0	0	0	0
0	1	0	1
1	0	0	1
1	1	1	0

The geometrical representation of this three-input Boolean is given in figure F.10.

It is immediately obvious that the function permits a representation as a MCP model. This verifies that the network in figure 1.9 is sufficient. As no

three-neuron (two inputs and one MCP neuron) network is possible, it must also be minimal. Note that the end result in both of these cases, is a heterogeneous network structure in which there are three types of neurons:

Input neurons which communicate directly with the outside world, but serve no computational purpose beyond distributing the input signals to all of the first layer of computing neurons.

Hidden neurons which do not communicate with the outside world; compute.

Output neurons which do Communicate with the outside world; compute.

Signals pass forward through the input layer and hidden layers and emerge from the output layer. Such a network is called a *feed-forward* network. Conversely, if signals are communicated backwards in a feedback fashion, the network is termed a *feed-back* or *recurrent* network.

The constructions that have been presented suggest why the MCP model proves to be of interest; by passing to networks of MCP neurons, it can be shown fairly easily that any Boolean function can be represented by an appropriate network; this forms the subject of the next section. Furthermore, a training algorithm exists for such networks [202], which terminates in a *finite* time.

F.3 Perceptrons

In the previous section, it was shown how the failure of the MCP model to represent certain simple functions led to the construction of simple networks of MCP neurons which could overcome the problems. The first serious study of such networks was carried out by Rosenblatt and is documented in his 1962 book [216]. Rosenblatt's *perceptron* networks are composed of an input layer and two layers of MCP neurons as shown in figure F.11. The hidden layer is referred to as the *associative* layer while the output layer is termed the *decision* layer. Only the connections between the decision nodes and associative nodes are adjustable in strength; those between the input nodes and associative nodes are preset before training takes place[2].

The neurons operate as threshold devices exactly as described in the previous section: if the weighted summation of inputs to a neuron exceeds the threshold, the neuron output is unity, otherwise it is zero.

In the following, the output of node i in the associate layer will be denoted by $y_i^{(a)}$ and the corresponding output in the decision layer by $y_i^{(d)}$; the connection weight between decision node i and associative node j will be denoted by w_{ij}. The thresholds will be labelled $\beta_i^{(a)}$ and $\beta_i^{(d)}$.

[2] This means that the associative nodes can actually represent any Boolean function of the inputs which is representable by an MCP model. The fact that there is only one layer of trainable weights means that the perceptron by-passes the credit-assignment problem—more about this later.

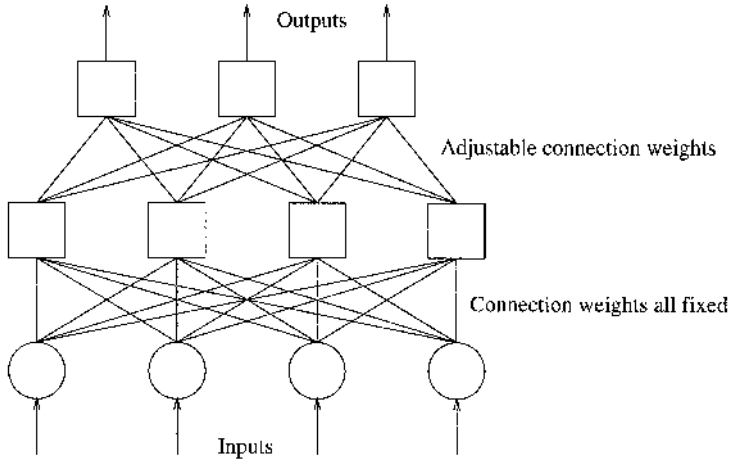


Figure F.11. Structure of Rosenblatt's perceptron.

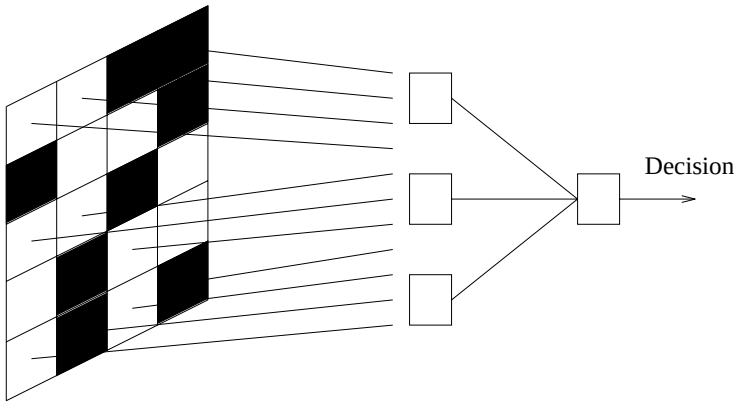


Figure F.12. Perceptron as pattern recognizer.

It is immediately apparent that the perceptrons have applications in pattern recognition. For example, the input layer could be associated with a screen or retina as in figure F.12, such that an input of 0 corresponded to a white pixel and 1 to a black pixel. The network could then be trained to respond at the decision layer only if certain patterns appeared on the screen.

This pattern recognition problem is clearly inaccessible to an MCP model since there are no restrictions on the form of Boolean function which could arise. However, it is possible to show that *any* Boolean function can be represented by a perceptron network; a partial proof is presented here. The reason is that *any*

Boolean $f(x_1, \ldots, x_{\mathrm{N}})$ can be represented by a sum of products

$$f(x_1, \ldots, x_{\mathrm{N}}) = \sum_{p=0}^{N} \sum_{\text{distinct } i_j} a^{i_1 \ldots i_p} x_{i_1} \ldots x_{i_p} \tag{F.17}$$

where the coefficients $a^{i_1 \ldots i_p}$ are integers. (For a proof see [186] or [202].) For example, the XOR function has a representation

$$\mathrm{XOR}(x_1, x_2) = x_1 + x_2 - 2x_1 x_2. \tag{F.18}$$

In a perceptron, the associative units can be used to compute the products and the decision units to make the final linear combination. The products are computed as follows.

Suppose it is required to compute the term $x_1 x_3 x_6 x_9$ say. First, the connections from the associative node to the inputs x_1, x_3, x_6 and x_9 are set to unity while all other input connections are zeroed. The activation of the node is therefore $x_1 + x_3 + x_6 + x_9$. Now, the product $x_1 x_3 x_6 x_9$ is 1, if and only if $x_1 = x_3 = x_6 = x_9 = 1$, in which case the activation of the unit is 4. If any of the inputs is equal to zero, the activation is clearly ≤ 3. Therefore if the threshold β for that node is set at 3.5

$$\sum w_{ij}^{(i)} x_j \geq \beta \Longrightarrow y^{(a)} = 1, \quad \text{if and only if } x_1 x_3 x_6 x_9 = 1$$

where the $w_{ij}^{(i)}$ are the connections to the input layer. This concludes the proof.

Note that it may be necessary to use *all* possible products of inputs in forming the linear combination (F.17). In this case, for N inputs, 2^N products and hence 2^N associative nodes are required. The XOR function here is an example of this type of network.

F.3.1 The perceptron learning rule

Having established that any Boolean function can be computed using a perceptron, the next problem is to establish a training algorithm which will lead to the correct connection weights. A successful learning rule was obtained by Rosenblatt [216]. This rule corrects the weights after comparing the network outputs for a given input with a set of desired outputs; the approach is therefore one of *supervised learning*.

At each step, a set of inputs $x_1^{(0)}, \ldots, x_n^{(0)}$ are presented and the outputs $x_1^{(d)}, \ldots x_{\mathrm{M}}^{(d)}$ are obtained from the decision layer. As the desired outputs $y_1, \ldots, y_{\mathrm{M}}$ are known, the errors $\delta_i = y_i - x_i^{(d)}$ are computable. The connections to decision node i are then updated according to the *delta rule*

$$w_{ij} \longrightarrow w_{ij} + \eta \delta_i x_j^{(a)} \tag{F.19}$$

where $\eta > 0$ is the *learning coefficient* which mediates the extent of the change.

This is a form of Hebb's rule. If δ_i is high, i.e. the output should be high and the prediction is low, the connections between decision node i and those associative nodes j which are active should be strengthened. In order to see exactly what is happening, consider the four possibilities,

	y_i	$x_i^{(d)}$	δ_i
(i)	0	0	0
(ii)	0	1	-1
(iii)	1	0	1
(iv)	1	1	0

In cases (i) and (iv) the network response is correct and no adjustments to the weights are needed. In case (ii) the weighted sum is too high and the network fires even though it is not required to. In this case, the adjustment according to (F.19) changes the weighted sum at the neuron i as follows.

$$\sum w_{ij} x_j^{(a)} \longrightarrow \sum w_{ij} x_j^{(a)} - \eta (x_j^{(a)})^2 \tag{F.20}$$

and therefore leads to the desired reduction in the activation at i. In case (iii) the neuron does not fire when required; the delta rule modification then leads to

$$\sum w_{ij} x_j^{(a)} \longrightarrow \sum w_{ij} x_j^{(a)} + \eta (x_j^{(a)})^2 \tag{F.21}$$

and the activation is higher next time the inputs are presented.

Originally, it was hoped that repeated application of the learning rule would lead to convergence of the weights to the correct values. In fact, the perceptron convergence theorem [186, 202] showed that this is guaranteed within finite time. However, the situation is similar to that for training an MCP neuron in that the theorem gives no indication of how long this time will be.

F.3.2 Limitations of perceptrons

The results that have been presented indicate why perceptrons were initially received with enthusiasm. They can represent a Boolean function of arbitrary complexity and are provided with a training algorithm which is guaranteed to converge in finite time. The problem is that in representing a function with N arguments, the perceptron may need 2^N elements in the associative layers; the networks grow exponentially in complexity with the dimension of the problem.

A possible way of avoiding this problem was seen to be to restrict the number of connections between the input layer and associative layer, so that each associative node connects to a (hopefully) small subset of the inputs. A perceptron with this restriction is called a *diameter-limited* perceptron. The justification for such perceptrons is that the set of Booleans which require full connections might consist of a small set of uninteresting functions. Unfortunately, this has proved not to be the case.

In 1969, Minsky and Papert published the book *Perceptrons* [186]. It constituted a completely rigorous investigation into the capabilities of perceptrons. Unfortunately for neural network research, it concluded that perceptrons were of limited use. For example, one result stated that a perceptron pattern recognizer of the type shown in figure F.12 cannot even establish if a pattern is connected (i.e. if it is composed of one or more disjoint pieces), if it is diameter-limited. A further example, much quoted in the literature, is the parity function F. $F(x_1, \ldots, x_n) = 1$ if, and only if, an odd number of the inputs x_i are high. Minsky and Papert showed that this function cannot be computed by a diameter-limited perceptron.

Another possible escape route was the use of perceptrons with several hidden layers in the hope that the more complex organization would avoid the exponential growth in the number of neurons. The problem here is that the adjustment of connection weights to a node by the delta rule requires an estimate of the output error at that node. However, only the errors at the output layer are given, and at the time there was no means of assigning meaningful errors to the internal nodes. This was referred to as the *credit-assignment problem*. The problem remained unsolved until 1974 [264]. Unfortunately, Minsky and Papert's book resulted in the almost complete abandonment of neural network research until Hopfield's paper [132A] of 1982 brought about a resurgence of interest. As a result of this, Werbos' 1974 solution [264] of the credit-assignment problem was overlooked until after Rumelhart *et al* independently arrived at the solution in 1985 [218]. The new paradigm the latter introduced—the multi-layer perceptron (MLP)—is probably the most widely-used neural network so far.

F.4 Multi-layer perceptrons

The network is a natural generalization of the perceptrons described in the previous section. The main references for this discussion are [37] or the seminal work [218]. A detailed analysis of the network structure and learning algorithm is given in appendix G, but a brief discussion is given here if the reader is prepared to take the theory on trust.

The MLP is a feedforward network with the neurons arranged in layers (figure F.13). Signal values pass into the *input layer* nodes, progress forward through the network *hidden layers*, and the result finally emerges from the *output layer*. Each node i is connected to each node j in the preceding and following layers through a connection of weight w_{ij}. Signals pass through the nodes as follows: in layer k a weighted sum is performed at each node i of all the signals $x_j^{(k-1)}$ from the preceding layer $k-1$, giving the excitation $z_i^{(k)}$ of the node; this is then passed through a nonlinear *activation function* f to emerge as the output

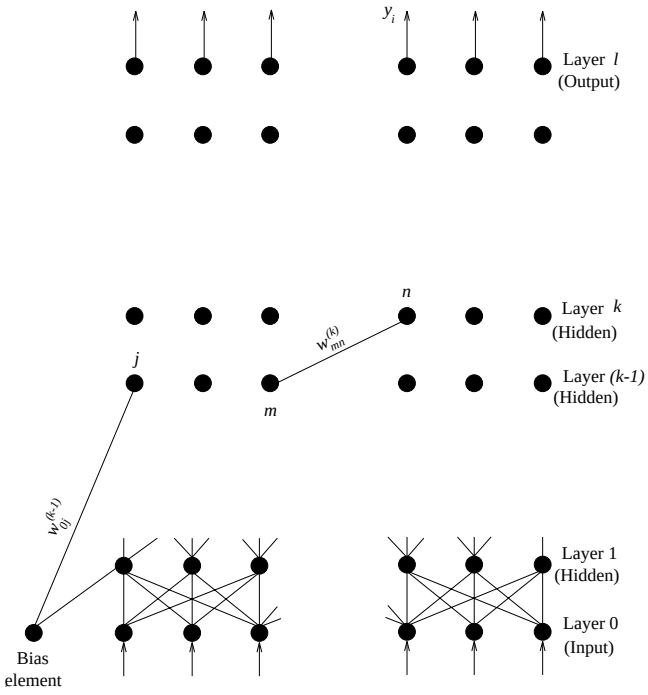


Figure F.13. The general multi-layer perceptron network.

of the node $x_i^{(k)}$ to the next layer, i.e.

$$x_i^{(k)} = f(z_i^{(k)}) = f\left(\sum_j w_{ij}^{(k)} x_j^{(k-1)}\right). \tag{F.22}$$

Various choices for the function f are possible, the one adopted here is the hyperbolic tangent function $f(x) = \tanh(x)$. (Note that the hard threshold of the MCP neurons is *not* allowed. The validity of the learning algorithm depends critically on the differentiability of f. The reason for this is discussed in appendix G.) A novel feature of this network is that the neuron outputs can take any values in the interval $[-1, 1]$. There are also no explicit threshold values associated with the neurons. One node of the network, the *bias* node, is special in that it is connected to all other nodes in the hidden and output layers; the output of the bias node is held fixed throughout in order to allow constant offsets in the excitations z_i of each node.

The first stage of using a network is to establish the appropriate values for the connection weights w_{ij}, i.e. the *training* phase. The type of training usually used is a form of *supervised* learning and makes use of a set of network inputs for which the desired network outputs are known. At each training step a set of

inputs is passed forward through the network yielding trial outputs which can be compared with the desired outputs. If the comparison error is considered small enough, the weights are not adjusted. If, however, a significant error is obtained, the error is passed *backwards* through the net and the training algorithm uses the error to adjust the connection weights so that the error is reduced. The learning algorithm used is usually referred to as the *back-propagation* algorithm, and can be summarized as follows. For each presentation of a training set, a measure J of the network error is evaluated where

$$J(t) = \frac{1}{2}\sum_{j=1}^{n^{(l)}}(y_i(t) - \hat{y}_i(t))^2 \tag{F.23}$$

and $n^{(l)}$ is the number of output layer nodes. J is implicitly a function of the network parameters $J = J(\theta_1, \ldots, \theta_n)$ where the θ_i are the connection weights, ordered in some way. The integer t labels the presentation order of the training sets. After presentation of a training set, the standard steepest descent algorithm requires an adjustment of the parameters according to

$$\triangle\theta_i = -\eta\frac{\partial J}{\partial \theta_i} = -\eta\nabla_i J \tag{F.24}$$

where ∇_i is the gradient operator in the parameter space. The parameter η determines how large a step is made in the direction of steepest descent and therefore how quickly the optimum parameters are obtained. For this reason η is called the *learning coefficient*. Detailed analysis (appendix G) gives the update rule after the presentation of a training set

$$w_{ij}^{(m)}(t) = w_{ij}^{(m)}(t-1) + \eta\delta_i^{(m)}(t)x_j^{(m-1)}(t) \tag{F.25}$$

where $\delta_i^{(m)}$ is the error in the output of the ith node in layer m. This error is not known *a priori* but must be constructed from the known errors $\delta_i^{(l)} = y_i - \hat{y}_i$ at the output layer l. This is the source of the name back-propagation, the weights must be adjusted layer by layer, moving backwards from the output layer.

There is little guidance in the literature as to what the learning coefficient η should be; if it is taken too small, convergence to the correct parameters may take an extremely long time. However, if η is made large, learning is much more rapid but the parameters may diverge or oscillate. One way around this problem is to introduce a *momentum* term into the update rule so that previous updates persist for a while, i.e.

$$\triangle w_{ij}^{(m)}(t) = \eta\delta_i^{(m)}(t)x_j^{(m-1)}(t) + \alpha\triangle w_{ij}^{(m)}(t-1) \tag{F.26}$$

where α is termed the *momentum coefficient*. The effect of this additional term is to damp out high-frequency variations in the back-propagated error signal. This is the form of the algorithm used throughout the case studies in chapter 6.

Once the comparison error is reduced to an acceptable level over the whole training set, the training phase ends and the network is established.

F.5 Problems with MLPs and (partial) solutions

This section addresses some questions regarding the desirability of solving problems using MLPs and training them using back-propagation.

F.5.1 Existence of solutions

Before advocating the use of neural networks in representing functions and processes, it is important to establish what they are capable of. As described earlier, artificial neural networks were all but abandoned as a subject of study following Minsky and Papert's book [186] which showed that perceptrons were incapable of modelling very simple logical functions. In fact, recent years have seen a number of rigorous results [72, 106, 133], which show that an MLP network is capable of approximating a given function with arbitrary accuracy, even if possessed of only a single hidden layer. Unfortunately, the proofs are not constructive and offer no guidelines as to the complexity of network required for a given function. A single hidden layer may be sufficient but might require many more neurons than if two hidden layers were used.

F.5.2 Convergence to solutions

In this case the situation is even more depressing. There is currently no proof that back-propagation results in convergence to a solution even if restricted conditions are adopted [73]. The situation here contrasts interestingly with that for a perceptron network. In that case, a solution to a given problem is rather unlikely, yet if it does exist, the perceptron learning algorithm is guaranteed to converge to it in finite time [186, 202]. Note that the question for MLP networks is whether it converges at all.

F.5.3 Uniqueness of solutions

This is the problem of local minima again. The error function for an MLP network is an extremely complex object. Given a converged MLP network, there is no way of establishing if it has arrived at the global minimum. Present attempts to avoid the problem are centred around the association of a temperature with the learning schedule. Roughly speaking, at each training cycle the network may randomly be given enough 'energy' to escape from a local minimum. The probable energy is calculated from a network temperature function which decreases with time. Recall that molecules of a solid at high temperature escape the energy minimum which specifies their position in the lattice. An alternative approach is to seek network paradigms with less severe problems, e.g. radial basis function networks [62]. Having said all this, problem local minima do not seem to appear in practice with the monotonous regularity with which they appear in cautionary texts. Davalo and Naïm [73] have it that they most often appear in the construction of pathological functions in the mathematical literature.

F.5.4 Optimal training schedules

There is little guidance in the literature as to which values of momentum and learning coefficients should be used. Time-varying training coefficients are occasionally useful, where initially high values are used to induce large steps in the parameter space. Later, the values can be reduced to allow 'fine-tuning'.

Another question is to do with the order of presentation of the training data to the network; it is almost certain that some strategies for presentation will slow down convergence. Current 'best practice' appears to be to present the training sets randomly.

The question of when to update the connection weights remains open. It is almost certainly better to update only after several training cycles have passed (an *epoch*). However, again there are no rigorous results.

The question of *overtraining* of networks is often raised. This is the failure of networks to generalize as a result of spending too long learning a specific set of training examples. This subject is most easily discussed in the context of neural network pattern recognition and is therefore not discussed here.

F.6 Radial basis functions

The use of radial basis functions (RBF) stems from the fact that functions can be approximated arbitrarily closely by superpositions of the form

$$f(\{x\}) \approx \sum_{i=1}^{N_\mathrm{b}} a_i \varphi_i(\|\{x\} - \{c_i\}\|) \tag{F.27}$$

where φ_i is typically a Gaussian

$$\varphi_i(u) = \exp\left(-\frac{u^2}{2r_i^2}\right) \tag{F.28}$$

and r_i is a radius parameter. The vector $\{c_i\}$ is called the *centre* of the ith basis function.

Their use seems to date from applications by Powell [208]. Current interest in the neural network community stems from the observation by Broomhead and Lowe [47] that they can be implemented as a network (figure F.14).

The weights a_i can be obtained by off-line LS methods as discussed in chapter 6 or by back-propagation using the iterative formula

$$\triangle a_i(t) = \eta\delta(t)x_j^{(\mathrm{h})}(t) + \alpha\triangle a_i(t-1) \tag{F.29}$$

(where $x_j^{(\mathrm{h})}$ is the output from the jth hidden node and δ is the output error) once the centres and radii for the Gaussians are established. This can be accomplished using clustering methods in an initial phase to place the centres at regions of high

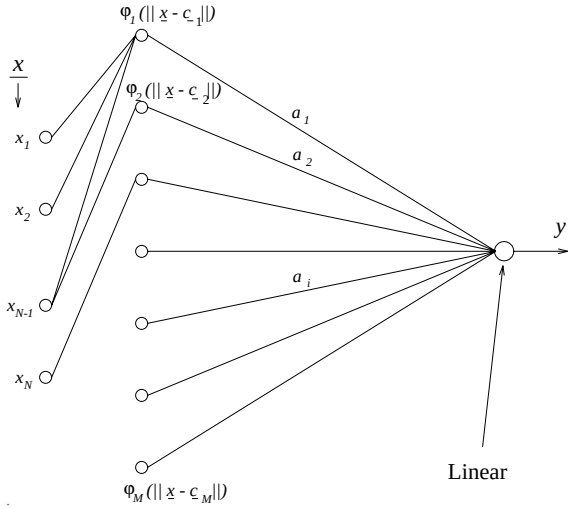


Figure F.14. Single-output radial basis function network.

data density [190]. When a pattern $\boldsymbol{x}(t)$ is presented at the input layer, the position of the *nearest* centre, say $\{c_j\}$, is adjusted according to the rule

$$\{c_j\}(t+1) = \{c_j\}(t) + \beta[\{x\}(t) - \{c_j\}(t)] \tag{F.30}$$

where β is the *clustering gain*. β is usually taken as a function of time $\beta(t)$, large in the initial stages of clustering and small in the later stages.

The radii are set using a simple nearest-neighbour rule; r_j is set to be the distance to the nearest-neighbouring cluster.

The RBF network generalizes trivially to the multi-output case (figure F.15). The representation becomes

$$f_j(\{x\}) \approx \sum_{i=1}^{N_b} a_{ij}\varphi_i(\|\{x\} - \{c_i\}\|) \tag{F.31}$$

and the output weights are trained using a simple variant of (F.29).

RBF networks differ from MLPs in a number of respects.

(1) Radial basis function networks have a single hidden layer while MLP networks have potentially more (although they theoretically only need one to approximate any function [72, 106, 235]). With this structure, RBF networks can approximate any function [203].
(2) All the neurons in the MLP network are usually of the same form (although heterogeneous networks will be constructed later for system identification purposes). In the RBF network, the hidden layer nodes are quite different from the linear output nodes.

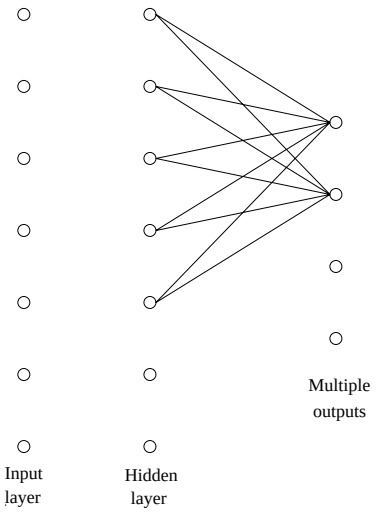


Figure F.15. Multi-output radial basis function network.

(3) The activation in the RBF comes from a Euclidean distance; in the MLP it arises from an inner product.
(4) Most importantly, the RBF networks give a representation of a function as a sum of local processing units, i.e. Gaussians $\Rightarrow$ local approximations. They therefore have reduced sensitivity to the training data
MLP networks construct a *global* approximation to mappings. This allows them potentially to generalize to regions of pattern space distant from the data.

Note that the RBF network can be trained in a single phase so that all parameters are iterated at the same time. This is essentially back-propagation [243].

Appendix G

Gradient descent and back-propagation

The back-propagation procedure for training multi-layer neural networks was initially developed by Paul Werbos and makes its first appearance in his doctoral thesis in 1974 [264]. Unfortunately, it languished there until the mid-eighties when it was discovered independently by Rumelhart *et al* [218]. This is possibly due to the period of dormancy that neural network research underwent following the publication of Minsky and Paperts' book [186] on the limitations of perceptron networks.

Before deriving the algorithm, it will prove beneficial to consider a number of simpler optimization problems as warm-up exercises; the back-propagation scheme will eventually appear as a (hopefully) natural generalization.

G.1 Minimization of a function of one variable

For the sake of simplicity, a function with a single minimum is assumed. The effect of relaxing this restriction will be discussed in a little while.

Consider the problem of minimizing the function $f(x)$ shown in figure G.1. If an analytical form for the function is known, elementary calculus provides the means of solution. In general, such an expression may not be available. However, if some means of determining the function and its first derivative at a point x is known, the solution can be obtained by the iterative scheme described below.

Suppose the iterative scheme begins with guessing or estimating a trial position x_0 for the minimum at x^m. The next estimate x_1 is obtained by adding a small amount δx to x_0. Clearly, in order to move nearer the minimum, δx should be positive if $x_0 < x^m$, and negative otherwise. It appears that the answer is needed before the next step can be carried out. However, note that

$$\frac{\mathrm{d}f}{\mathrm{d}x} < 0, \quad \text{if } x_0 < x^m \tag{G.1}$$

$$\frac{\mathrm{d}f}{\mathrm{d}x} > 0, \quad \text{if } x_0 > x^m. \tag{G.2}$$

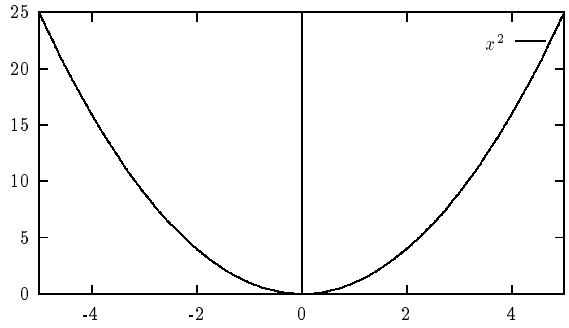


Figure G.1. A simple function of one variable.

So, *in the vicinity of the minimum*, the update rule,

$$x_1 - x_0 = \delta x = +\eta, \quad \text{if } \frac{\mathrm{d}f}{\mathrm{d}x} < 0 \tag{G.3}$$

$$x_1 - x_0 = \delta x = -\eta, \quad \text{if } \frac{\mathrm{d}f}{\mathrm{d}x} > 0 \tag{G.4}$$

with η a small positive constant, moves the iteration closer to the minimum. In a simple problem of this sort, η would just be called the *step-size*, it is essentially the *learning coefficient* in the terminology of neural networks. Clearly, η should be small in order to avoid overshooting the minimum. In a more compact notation

$$\delta x = -\eta \ \mathrm{sgn}\left(\frac{\mathrm{d}f}{\mathrm{d}x}\right). \tag{G.5}$$

Note that $|\frac{\mathrm{d}f}{\mathrm{d}x}|$ actually increases with distance from the minimum x^m. This means that the update rule

$$\delta x = -\eta \frac{\mathrm{d}f}{\mathrm{d}x} \tag{G.6}$$

also encodes the fact that large steps are desirable when the iterate is far from the minimum. In an ideal world, iteration of this update rule would lead to convergence to the desired minimum. Unfortunately, a number of problems can occur; the two most serious are now discussed.

G.1.1 Oscillation

Suppose that the function is $f(x) = (x - x^m)^2$. (This is not an unreasonable assumption as Taylor's theorem shows that most functions are approximated by a quadratic in the neighbourhood of a minimum.)

As mentioned earlier, if η is too large the iterate x_{i+1} may be on the opposite side of the minimum to x_i (figure G.2). A particularly ill-chosen value of η, η_c say, leads to x_{i+1} and x_i being *equidistant* from x^m. In this case, the iterate

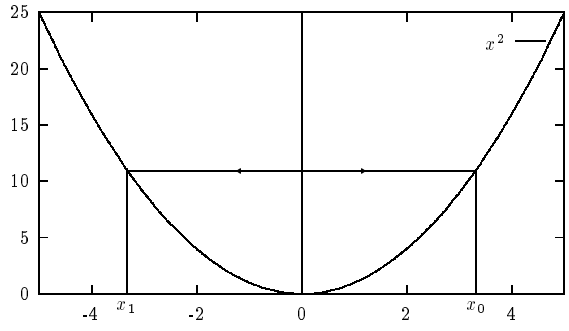


Figure G.2. The problem of oscillation.

will oscillate about the minimum *ad infinitum* as a result of the symmetry of the function. It could be argued that choosing $\eta = \eta_c$ would be extremely unlucky; however, any values of η slightly smaller than η_c will cause damped oscillations of the iterate about the point x^m. Such oscillations delay convergence, possibly substantially.

Fortunately, there is a solution to this problem. Note that the updates δ_i and δ_{i-1} will have opposite signs and similar magnitudes at the onset of oscillation. This means that they will cancel to a large extent, and updating at step i with $\delta_i + \delta_{i-1}$ would provide more stable iteration. If the iteration is not close to oscillation, the addition of the last-but-one update produces no quantitative difference. This circumstance leads to a modified update rule

$$\delta x_i = -\eta \frac{\mathrm{d}f(x_i)}{\mathrm{d}x} + \alpha \delta x_{i-1}. \tag{G.7}$$

The new coefficient α is termed the *momentum* coefficient, a sensible choice of this can lead to much better convergence properties for the iteration. Unfortunately, the next problem with the procedure is not dealt with so easily.

G.1.2 Local minima

Consider the function shown in figure G.3, this illustrates a feature—a local minimum—which can cause serious problems for the iterative minimization scheme. Although x^m is the global minimum of the function, it is clear that starting the iteration at any x_0 to the right of the local minimum at x^{lm} will very likely lead to convergence to x^{lm}. There is no simple solution to this problem.

G.2 Minimizing a function of several variables

For this section it is sufficient to consider functions of two variables, i.e. $f(x, y)$; no new features appear on generalizing to higher dimensions. Consider the

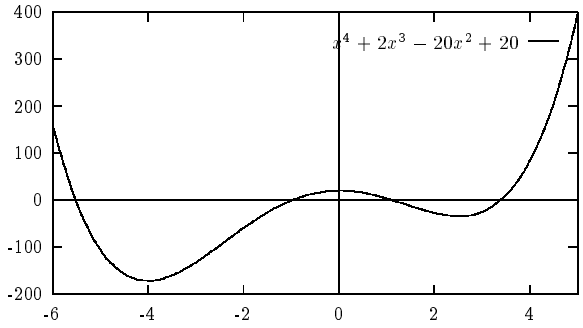


Figure G.3. The problem of local minima.

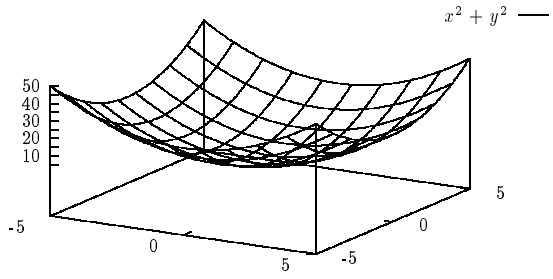


Figure G.4. Minimizing a function over the plane.

function in figure G.4. The position of the minimum is now specified by a point in the (x, y)-plane. Any iterative procedure will require the update of both x and y. An analogue of equation (G.6) is required. The most simple generalization would be to update x and y separately using partial derivatives, e.g.,

$$\delta x = -\eta \frac{\partial f}{\partial x} \tag{G.8}$$

which would cause a decrease in the function by moving the iterate along a line of constant y, and

$$\delta y = -\eta \frac{\partial f}{\partial y} \tag{G.9}$$

which would achieve the same with movement along a line of constant x. In fact, this update rule proves to be an excellent choice. In vector notation, which shall be used for the remainder of this section, the coordinates are given by $\{x\} = (x_1, x_2)$ and the update rule is

$$\{\delta x\} = (\delta x_1, \delta x_2) = -\eta \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2} \right) = -\eta \{\nabla\} f \tag{G.10}$$

where ∇ is the gradient operator

$$\{\nabla\} f = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2} \right). \tag{G.11}$$

With the choices (G.8) and (G.9) for the update rules, this approach to optimization is often referred to as the *method of gradient descent*.

A problem which did not occur previously is that of choosing the direction for the iteration, i.e. the *search direction*. For a function of one variable, only two directions are possible, one of which leads to an increase in the function. In two or more dimensions, a continuum of search directions is available and the possibility of optimally choosing the direction arises.

Fortunately, this problem admits a fairly straightforward solution. (The following discussion follows closely to that in [66].) Suppose the current position of the iterate is $\{x\}_0$. The next step should be in the direction which produces the greatest decrease in f, given a fixed step-length. Without loss of generality, the step-length can be taken as unity; the update vector, $\{u\} = (u_1, u_2)$ is therefore a unit vector. The problem is to maximize δf, where

$$\delta f = \frac{\partial f(\{x\}_0)}{\partial x_1} u_1 + \frac{\partial f(\{x\}_0)}{\partial x_2} u_2 \tag{G.12}$$

subject to the constraint on the step-length

$$u_1^2 + u_2^2 = 1. \tag{G.13}$$

Incorporating the length constraint into the problem via a Lagrange multiplier λ [233] leads to $F(u_1, u_2, \lambda)$ as the function to be maximized, where

$$F(u_1, u_2, \lambda) = \frac{\partial f(\{x\}_0)}{\partial x_1} u_1 + \frac{\partial f(\{x\}_0)}{\partial x_2} u_2 + \lambda(u_1^2 + u_2^2 - 1). \tag{G.14}$$

Zeroing the derivatives with respect to the variables leads to the equations for the optimal u_1, u_2 and λ.

$$\frac{\partial F}{\partial u_1} = 0 = \frac{\partial f(\{x\}_0)}{\partial x_1} - 2\lambda u_1 \Rightarrow u_1 = \frac{1}{2\lambda} \frac{\partial f(\{x\}_0)}{\partial x_1} \tag{G.15}$$

$$\frac{\partial F}{\partial u_2} = 0 = \frac{\partial f(\{x\}_0)}{\partial x_2} - 2\lambda u_2 \Rightarrow u_2 = \frac{1}{2\lambda} \frac{\partial f(\{x\}_0)}{\partial x_2} \tag{G.16}$$

$$\frac{\partial F}{\partial \lambda} = 0 = 1 - u_1^2 - u_2^2. \tag{G.17}$$

Substituting (G.15) and (G.16) into (G.17) gives

$$1 - \frac{1}{4\lambda^2} \left\{ \left(\frac{\partial f(\{x\}_0)}{\partial x_1} \right)^2 + \left(\frac{\partial f(\{x\}_0)}{\partial x_2} \right)^2 \right\} = 1 - \frac{1}{4\lambda^2} \| \{\nabla\} f(\{x\}_0) \|^2 = 0 \tag{G.18}$$

$$\Rightarrow \lambda = \pm | \{\nabla\} f(\{x\}_0) |. \tag{G.19}$$

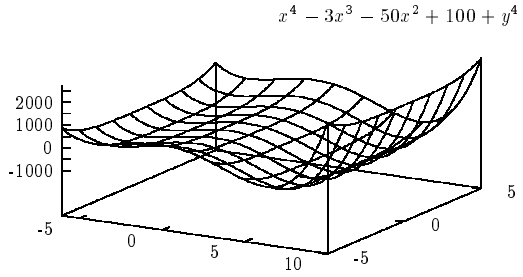


Figure G.5. Local minimum in a function over the plane.

Substituting this result into (G.15) and (G.16) gives

$$u_1 = \pm \frac{1}{|\{\nabla\} f(\{x\}_0)|} \frac{\partial f(\{x\}_0)}{\partial x_1} \tag{G.20}$$

$$u_2 = \pm \frac{1}{|\{\nabla\} f(\{x\}_0)|} \frac{\partial f(\{x\}_0)}{\partial x_2} \tag{G.21}$$

or

$$\{u\} = \pm \frac{\{\nabla\} f(\{x\}_0)}{|\{\nabla\} f(\{x\}_0)|}. \tag{G.22}$$

A consideration of the second derivatives reveals that the $+$ sign gives a vector in the direction of maximum increase of f, while the $-$ sign gives a vector in the direction of maximum decrease. This shows that the gradient descent rule

$$\{\delta x\}_{i+1} = -\eta \{\nabla\} f(\{x\}_i) \tag{G.23}$$

is actually the *best possible*. For this reason, the approach is most often referred to as the *method of steepest descent*.

Minimization of functions of several variables by steepest descent is subject to all the problems associated with the simple iterative method of the previous section. The problem of oscillation certainly occurs, but can be alleviated by the addition of a momentum term. The modified update rule is then

$$\{\delta x\}_{i+1} = -\eta \{\nabla\} f(\{x\}_i) + \alpha \delta \{x\}_i. \tag{G.24}$$

The problems presented by local minima are, if anything, more severe in higher dimensions. An example of a troublesome function is given in figure G.5.

In addition to stalling in local minima, the iteration can be directed out to infinity along valleys.

G.3 Training a neural network

The relevant tools have been developed and this section is concerned with deriving a learning rule for training a multi-layer perceptron (MLP) network. The method

of steepest descent is directly applicable; the function to be minimized is a measure of the network error in representing a desired input–output process. Steepest-descent is used because there is no analytical relationship between the network parameters and the prediction error of the network. However, at each iteration, when an input signal is presented to the network, the error is known because the desired outputs for a given input are assumed known. Steepest-descent is therefore a method based on *supervised learning*. It will be shown later that applying the steepest-descent algorithm results in update rules coinciding with the back-propagation rules which were stated without proof in appendix E. This establishes that back-propagation has a rigorous basis unlike some of the more *ad hoc* learning schemes. The analysis here closely follows that of Billings *et al* [37].

A short review of earlier material will be given first to re-establish the appropriate notation. The MLP network neurons are assembled into layers and only communicate with neurons in the adjacent layers; intra-layer connections are forbidden (see figure E.13). Each node j in layer m is connected to each node i in the following layer $m+1$ by connections of *weight* $w_{ij}^{(m+1)}$. The network has $l+1$ layers, layer 0 being the input layer and layer l the output. Signals are passed through each node in layer $m+1$ as follows: a weighted sum is performed at i of all outputs $x_j^{(m)}$ from the preceding layer, this gives the *excitation* $z_i^{(m+1)}$ of the node

$$z_i^{(m+1)} = \sum_{j=0}^{n^{(m)}} w_{ij}^{(m+1)} x_j^{(m)} \tag{G.25}$$

where $n^{(m)}$ is the number of nodes in layer m. (The summation index starts from zero in order to accommodate the bias node.) The excitation signal is then passed through a nonlinear *activation function* f to emerge as the output $x_i^{(m+1)}$ of the node to the next layer

$$x_i^{(m+1)} = f(z_i^{(m+1)}) = f\left(\sum_{j=0}^{n^{(m)}} w_{ij}^{(m+1)} x_j^{(m)}\right). \tag{G.26}$$

Various choices for f are possible, in fact, the only restrictions on f are that it should be differentiable and monotonically increasing [219]. The hyperbolic tangent function $f(x) = \tanh(x)$ is used throughout this work, although the sigmoid $f(x) = (1 - \mathrm{e}^{-x})^{-1}$ is also very popular. The input layer nodes do not have nonlinear activation functions as their purpose is simply to distribute the network inputs to the nodes in the first hidden layer. The signals propagate only forward through the layers so the network is of the *feedforward* type.

An exception to the rule stated earlier, forbidding connections between layers which are not adjacent, is provided by the *bias* node which passes signals to all other nodes except those in the input layer. The output of the bias node is held constant at unity in order to allow constant offsets in the excitations. This is an

alternative to associating a threshold $\beta_i^{(m)}$ with each node so that the excitation is calculated from

$$z_i^{(m+1)} = \sum_{j=1}^{n^{(m)}} w_{ij}^{(m+1)} x_j^{(m)} + \beta_i^{(m+1)}. \tag{G.27}$$

The bias node is considered to be the 0th node in each layer.

As mentioned, training of the MLP requires sets of network inputs for which the desired network outputs are known. At each training step, a set of network inputs is passed forward through the layers yielding finally a set of trial outputs $\hat{y}_i$, $i = 1, \ldots, n^{(l)}$. These are compared with the desired outputs y_i. If the comparison errors $\delta_i^{(l)} = y_i - \hat{y}_i$ are considered small enough, the network weights are not adjusted. However, if a significant error is obtained, the error is passed backwards through the layers and the weights are updated as the error signal propagates back through the connections. This is the source of the name back-propagation.

For each presentation of a training set, a measure J of the network error is evaluated where

$$J(t) = \frac{1}{2} \sum_{j=1}^{n^{(l)}} (y_i(t) - \hat{y}_i(t))^2 \tag{G.28}$$

and J is implicitly a function of the network parameters $J = J(\theta_1, \ldots, \theta_n)$ where the θ_i are the connection weights ordered in some way. The integer t labels the presentation order of the training sets (the index t is suppressed in most of the following theory as a single presentation is considered). After a presentation of a training set, the steepest-descent algorithm requires an adjustment of the parameters

$$\triangle\theta_i = -\eta \frac{\partial J}{\partial \theta_i} = -\eta \nabla_i J \tag{G.29}$$

where ∇_i is the gradient operator in the parameter space. As before, the *learning coefficient* η determines the step-size in the direction of steepest descent. Because only the errors for the output layer are known, it is necessary to construct effective errors for each of the hidden layers by propagating back the error from the output layer. For the output (lth) layer of the network an application of the chain rule of partial differentiation [233] yields

$$\frac{\partial J}{\partial w_{ij}^{(l)}} = \frac{\partial J}{\partial \hat{y}_i} \cdot \frac{\partial \hat{y}_i}{\partial w_{ij}^{(l)}}. \tag{G.30}$$

Now

$$\frac{\partial J}{\partial \hat{y}_i} = -(y_i - \hat{y}_i) = -\delta_i^{(l)} \tag{G.31}$$

and as

$$\hat{y}_i = f\left(\sum_{j=0}^{n^{(l-1)}} w_{ij}^{(l)} x_j^{(l-1)} \right) \tag{G.32}$$

a further application of the chain rule

$$\frac{\partial \hat{y}_i}{\partial w_{ij}^{(l)}} = \frac{\partial f}{\partial z^{(l)}} \frac{\partial z_i^{(l)}}{\partial w_{ij}^{(l)}} \tag{G.33}$$

where z is defined as in (G.25), yields

$$\frac{\partial \hat{y}_i}{\partial w_{ij}^{(l)}} = f'\left(\sum_{j=0}^{n^{(l-1)}} w_{ij}^{(l)} x_j^{(l-1)} \right) x_j^{(l-1)} = f'(z_i^{(l)}) x_j^{(l-1)}. \tag{G.34}$$

So substituting this equation and (G.31) into (G.30) gives

$$\frac{\partial J}{\partial w_{ij}^{(l)}} = -f'\left(\sum_{j=0}^{n^{(l-1)}} w_{ij}^{(l)} x_j^{(l-1)} \right) x_j^{(l-1)} \delta_i^{(l)} \tag{G.35}$$

and the update rule for connections to the output layer is obtained from (G.29) as

$$\triangle w_{ij}^{(l)} = \eta f'\left(\sum_{j=0}^{n^{(l-1)}} w_{ij}^{(l)} x_j^{(l-1)} \right) x_j^{(l-1)} \delta_i^{(l)} = \eta f'(z_i^{(l)}) x_j^{(l-1)} \delta_i^{(l)} \tag{G.36}$$

where

$$f'(z) = (1 + f(z))(1 - f(z)) \tag{G.37}$$

if f is the hyperbolic tangent function, and

$$f'(z) = f(z)(1 + f(z)) \tag{G.38}$$

if f is the sigmoid. Note that the whole optimization hinges critically on the fact that the transfer function f is differentiable. The existence of f' is crucial to the propagation of errors to the hidden layers and to their subsequent training. This is the reason why perceptrons could not have hidden layers and were consequently so limited. The use of discontinuous 'threshold' functions as transfer functions meant that hidden layers could not be trained.

Updating of the parameters is essentially the same for the hidden layers except that an explicit error $\delta_i^{(m)}$ is not available. The errors for the hidden layer nodes must be constructed.

Considering the $(l-1)$th layer and applying the chain rule once more gives

$$\frac{\partial J}{\partial w_{ij}^{(l-1)}} = \sum_{k=1}^{n^{(l)}} \frac{\partial J}{\partial \hat{y}_k} \frac{\partial \hat{y}_k}{\partial x_i^{(l-1)}} \frac{\partial x_i^{(l-1)}}{\partial z_i^{(l-1)}} \frac{\partial z_i^{(l-1)}}{\partial w_{ij}^{(l-1)}}. \tag{G.39}$$

Now

$$\frac{\partial \hat{y}_k}{\partial x_i^{(l-1)}} = f'\left(\sum_{j=0}^{n^{(l-1)}} w_{kj}^{(l)} x_j^{(l-1)} \right) w_{ki}^{(l)} \tag{G.40}$$

$$\frac{\partial x_i^{(l-1)}}{\partial z_i^{(l-1)}} = f'(z_i^{(l-1)}) = f'\left(\sum_{j=0}^{n^{(l-2)}} w_{ij}^{(l-1)} x_j^{(l-2)} \right) \tag{G.41}$$

and

$$\frac{\partial z_i^{(l-1)}}{\partial w_{ij}^{(l-1)}} = x_j^{(l-2)} \tag{G.42}$$

so (G.39) becomes

$$\frac{\partial J}{\partial w_{ij}^{(l-1)}} = -\sum_{k=1}^{n^{(l)}} \delta_k^{(l)} f'\left(\sum_{j=0}^{n^{(l-1)}} w_{kj}^{(l)} x_j^{(l-1)}\right) w_{ki}^{(l)} f'\left(\sum_{j=0}^{n^{(l-2)}} w_{ij}^{(l-1)} x_j^{(l-2)}\right) x_j^{(l-2)}. \tag{G.43}$$

If the errors for the ith neuron of the $(l-1)$th layer are now *defined* as

$$\delta_i^{(l-1)} = f'\left(\sum_{j=0}^{n^{(l-2)}} w_{ij}^{(l-1)} x_j^{(l-2)}\right) \sum_{k=1}^{n^{(l)}} f'\left(\sum_{j=0}^{n^{(l-1)}} w_{kj}^{(l)} x_j^{(l-1)}\right) w_{ki}^{(l)} \delta_k^{(l)} \tag{G.44}$$

or

$$\delta_i^{(l-1)} = f'(z_i^{(l-1)}) \sum_{k=1}^{n^{(l)}} f'(z_k^{(l)}) w_{ki}^{(l)} \delta_k^{(l)} \tag{G.45}$$

then equation (G.43) takes the simple form

$$\frac{\partial J}{\partial w_{ij}^{(l-1)}} = -\delta_i^{(l-1)} x_j^{(l-2)}. \tag{G.46}$$

On carrying out this argument for all hidden layers $m \in l-1, l-2, \ldots, 1$ the general rules

$$\delta_i^{(m-1)}(t) = f'\left(\sum_{j=0}^{n^{(m-2)}} w_{ij}^{(m-1)}(t-1) x_j^{(m-2)}(t)\right) \sum_{k=1}^{n^{(m)}} \delta_k^{(m)}(t) w_{ki}^{(m)}(t-1) \tag{G.47}$$

or

$$\delta_i^{(m-1)}(t) = f'\left(z_i^{(m-1)}(t)\right) \sum_{k=1}^{n^{(m)}} \delta_k^{(m)}(t) w_{ki}^{(m)}(t-1) \tag{G.48}$$

and

$$\frac{\partial J}{\partial w_{ij}^{(m-1)}}(t) = -\delta_i^{(m-1)}(t) x_j^{(m-2)}(t) \tag{G.49}$$

are obtained (on restoring the t index which labels the presentation of the training set). Hence the name back-propagation.

Finally, the update rule for all the connection weights of the hidden layers can be given as

$$w_{ij}^{(m)}(t) = w_{ij}^{(m)}(t-1) + \triangle w_{ij}^{(m)}(t) \tag{G.50}$$

where

$$\triangle w_{ij}^{(m)}(t) = \eta \delta_i^{(m)}(t) x_j^{(m-1)}(t) \tag{G.51}$$

for each presentation of a training set.

There is little guidance in the literature as to what the learning coefficient η should be; if it is taken too small, convergence to the correct parameters may take an extremely long time. However, if η is made large, learning is much more rapid but the parameters may diverge or oscillate in the fashion described in earlier sections. One way around this problem is to introduce a *momentum* term into the update rule as before:

$$\triangle w_{ij}^{(m)}(t) = \eta \delta_i^{(m)}(t) x_j^{(m-1)}(t) + \alpha \triangle w_{ij}^{(m)}(t-1) \tag{G.52}$$

where α is the momentum coefficient. The additional term essentially damps out high-frequency variations in the error surface.

As usual with steepest-descent methods, back-propagation only guarantees convergence to a *local* minimum of the error function. In fact the MLP is highly nonlinear in the parameters and the error surface will consequently have many minima. Various methods of overcoming this problem have been proposed, none has met with total success.

Appendix H

Properties of Chebyshev polynomials

The basic properties are now fairly well known [103, 209]; however, for the sake of completeness they are described here along with one or two less well-known results.

H.1 Definitions and orthogonality relations

The definition of the Chebyshev polynomial of order n is

$$\begin{aligned} T_n(x) &= \cos(n\cos^{-1}(x)), \quad |x| \leq 1 \\ T_n(x) &= \cosh(n\cosh^{-1}(x)), \quad |x| \geq 1. \end{aligned} \tag{H.1}$$

It is not immediately obvious that this *is* a polynomial. That it is follows from applications of De Moivre's theorem. For example

$$T_3(x) = \cos(3\cos^{-1}(x)) = 4\cos^3(\cos^{-1}(x)) - 3\cos(\cos^{-1}(x)) = 4x^3 - 3x. \tag{H.2}$$

The Chebyshev polynomials are orthogonal on the interval $[-1, 1]$ with weighting factor $w(x) = (1 - x^2)^{\frac{1}{2}}$ which means that

$$\int_{-1}^{1} \mathrm{d}x\, w(x) T_n(x) T_m(x) = \frac{\pi}{2}(1 + \delta_{n0})\delta_{nm} \tag{H.3}$$

where δ_{nm} is the Kronecker delta.

The proof of this presents no problems: first the substitution $y = \cos^{-1}(x)$ is made; second, making use of the definition (H.1) changes the integral (H.3) to

$$\int_0^{\pi} \mathrm{d}y\, \cos(my)\cos(ny) \tag{H.4}$$

and this integral forms the basis of much of Fourier analysis. In fact, Chebyshev expansion is entirely equivalent to the more usual Fourier sine and cosine

expansions. Returning to the integral, one has

$$\int_0^\pi \mathrm{d}y\ \cos(my)\cos(ny) = \begin{cases} 0, & \text{if } m \neq n \\ \pi, & \text{if } m = n = 0 \\ \dfrac{\pi}{2}, & \text{if } m = n \neq 0. \end{cases} \tag{H.5}$$

With the help of the orthogonality relation (H.3) it is possible to expand any given function in terms of a series of Chebyshev polynomials, i.e.

$$f(x) = \sum_{i=0}^{m} a_i T_i(x). \tag{H.6}$$

Multiplying through by $w(x)T_j(x)$ and using the relation (H.3) gives for the coefficients

$$a_i = X_i \int_{-1}^{1} \mathrm{d}x\, w(x) T_i(x) f(x) \tag{H.7}$$

where $X_i = 1/\pi$ if $i \neq 0$ and $X_i = 2/\pi$ if $i = 0$.

The extension to a double series is fairly straightforward. If an expansion is needed of the form

$$f(x,y) = \sum_{i=0}^{m}\sum_{j=0}^{n} C_{ij} T_i(x) T_j(y) \tag{H.8}$$

then

$$C_{ij} = X_i X_j \int_{-1}^{+1}\int_{-1}^{+1} \mathrm{d}x\,\mathrm{d}y\, w(x)w(y) T_i(x) T_j(y) f(x,y). \tag{H.9}$$

The orthogonality relations can also be used to show that the Chebyshev expansion of order n is unique. If

$$f(x) = \sum_{i=0}^{m} a_i T_i(x) = \sum_{i=0}^{m} b_i T_i(x) \tag{H.10}$$

then multiplying by $w(x)T_j(x)$ and using the relation (H.3) gives $a_i = b_i$.

H.2 Recurrence relations and Clenshaw's algorithm

Like all orthogonal polynomials, the Chebyshev polynomials satisfy a number of recursion relations. Probably the most useful is

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x). \tag{H.11}$$

The proof is elementary. If $y = \cos^{-1}(x)$ then

$$\begin{aligned} T_{n+1}(x) &= \cos((n+1)y) = \cos(ny)\cos(y) - \sin(ny)\sin(y) \\ T_{n-1}(x) &= \cos((n+1)y) = \cos(ny)\cos(y) + \sin(ny)\sin(y) \end{aligned} \tag{H.12}$$

and adding gives

$$T_{n+1}(x) + T_{n-1}(x) = 2\cos(ny)\cos(y) = 2xT_n(x) \tag{H.13}$$

as required.

It is clear that if the recurrence begins with $T_0(x) = 1$ and $T_1(x) = x$, equation (H.11) will yield values of $T_n(x)$ for any n. This is the preferred means of evaluating $T_n(x)$ numerically as it avoids the computation of polynomials.

In order to evaluate how good a Chebyshev approximation is, one compares the true function to the approximation over a testing set. This means that one is potentially faced with mamy summations of the form (H.6). Although current computers are arguably powerful enough to allow a brute force approach, there is in fact a much more economical means of computing (H.6) than evaluating the polynomials and summing the series. The method uses *Clenshaw's recurrence formula*. In fact this can be used for any polynomial which uses a recursion relation although the version here is specific to the Chebyshev series. The general result is given in [209].

First define a sequence by

$$y_{n+2} = y_{n+1} = 0, \quad y_i = 2xy_{i-1} - y_i + a_i. \tag{H.14}$$

Then

$$\begin{aligned} f(x) = {} & [y_n - 2xy_{n+1} + y_{n+2}]T_n(x) + \cdots + [y_i - 2xy_{i+1} + y_{i+2}]T_i(x) \\ & + \cdots + [a_0 - y_2 + y_2]T_0(x) \end{aligned} \tag{H.15}$$

after adding and subtracting $y_2T_0(x)$. In the middle of this summation one has

$$\begin{aligned} & + \cdots + [y_{i+1} - 2xy_{i+2} + y_{i+3}]T_{i+1}(x) + [y_i - 2xy_{i+1} + y_{i+2}]T_i(x) \\ & + [y_{i-1} - 2xy_i + y_{i+1}]T_{i-1}(x) \end{aligned} \tag{H.16}$$

so the coefficient of y_{i+1} is

$$T_{n+1}(x) - 2xT_n(x) + T_{n-1}(x) \tag{H.17}$$

which vanishes by virtue of the recurrence relation (H.11). Similarly all the coefficients vanish down to y_2. All that remains is the end of the summation which is found to be

$$f(x) = a_0 + xy_1 - y_2. \tag{H.18}$$

Therefore to evaluate $f(x)$ for each x, one simply passes downwards through the recurrence (H.14) to obtain y_1 and y_2 and then evaluates the linear expression (H.18). Unfortunately there is no obvious analogue of Clenshaw's result for two-dimensional expansions of the form (H.8). This means that in evaluating a double series, one can only use the recurrence if the function $f(x, y)$ splits into single-variable functions, i.e. $f(x, y) = g(x) + h(y)$. Of all the examples considered in

chapter 7, only the Van der Pol oscillator fails to satisfy this condition, although it would be unlikely to hold in practice.

Clenshaw's algorithm can also be used algebraically in order to turn Chebyshev expansions into ordinary polynomials. However, one should be aware that this is not always a good idea [209].

H.3 Chebyshev coefficients for a class of simple functions

In chapter 7, the Chebyshev expansion for the restoring force $f(y,\dot{y})$ is estimated for a number of simple systems. In order to form an opinion of the accuracy of these estimates, one needs to know the exact values of the coefficients. A function sufficiently general to include the examples of chapter 7 is

$$f(x,y) = ax^3 + bx^2 + cx + dy + ey^2 + fx^2y. \tag{H.19}$$

The x and y are subjected to a linear transformation

$$\begin{aligned} x \longrightarrow \zeta(x) &= \overline{x} = \frac{x-\alpha_2}{\alpha_1} \\ y \longrightarrow \dot{\zeta}(y) &= \overline{y} = \frac{y-\beta_2}{\beta_1} \end{aligned} \tag{H.20}$$

where

$$\begin{aligned} \alpha_1 &= \tfrac{1}{2}(x_{\max} - x_{\min}), \quad \alpha_2 = \tfrac{1}{2}(x_{\max} + x_{\min}) \\ \beta_1 &= \tfrac{1}{2}(y_{\max} - y_{\min}), \quad \beta_2 = \tfrac{1}{2}(y_{\max} + y_{\min}). \end{aligned} \tag{H.21}$$

The form of f in the $(\overline{x},\overline{y})$ coordinate system is given by

$$\overline{f}(\overline{x},\overline{y}) = f(x,y) = f(\zeta^{-1}(\overline{x}),\dot{\zeta}^{-1}(\overline{y})) = f(\alpha_1\overline{x}+\alpha_2,\beta_1\overline{y}+\beta_2). \tag{H.22}$$

A little algebra produces the result

$$\overline{f}(\overline{x},\overline{y}) = \overline{a}\overline{x}^3 + \overline{b}\overline{x}^2 + \overline{c}\overline{x} + \overline{d}\overline{y} + \overline{e}\overline{y}^2 + \overline{f}\overline{x}^2\overline{y} + \overline{g}\overline{x}\overline{y} + \overline{h} \tag{H.23}$$

where

$$\begin{aligned} \overline{a} &= a\alpha_1^3 \\ \overline{b} &= 3a\alpha_1^2\alpha_2 + b\alpha_1^2 + f\alpha_1\beta_2 \\ \overline{c} &= 3a\alpha_1\alpha_2^2 + 2b\alpha_1\alpha_2 + 2f\alpha_1\alpha_2\beta_2 \\ \overline{d} &= d\beta_1 + 2e\beta_1\beta_2 + f\alpha_2^2\beta_1 \\ \overline{e} &= e\beta_1^2 \\ \overline{f} &= f\alpha_1^2\beta_1 \\ \overline{g} &= 2f\alpha_1\alpha_2\beta_2 \\ \overline{h} &= a\alpha_2^3 + b\alpha_2^2 + c\alpha_2 + d\beta_2 + e\beta_2^2 + f\alpha_2^2\beta_2. \end{aligned} \tag{H.24}$$

One can now expand this function as a double Chebyshev series of the form

$$\overline{f}(\overline{x},\overline{y}) = \sum_{i=0}^{m}\sum_{j=0}^{n} C_{ij}T_i(\overline{x})T_j(\overline{y}) \tag{H.25}$$

either by using the orthogonality relation (H.9) or by direct substitution. The exact coefficients for $\overline{f}(\overline{x},\overline{y})$ are found to be

$$\begin{aligned}
C_{00} &= \overline{h} + \tfrac{1}{2}(\overline{b}+\overline{e}) \\
C_{01} &= \overline{d} + \tfrac{1}{2}\overline{f} \\
C_{02} &= \tfrac{1}{2}\overline{e} \\
C_{10} &= \tfrac{1}{2}\overline{a} + \overline{c} \\
C_{11} &= \overline{g} \\
C_{12} &= 0 \\
C_{20} &= \tfrac{1}{2}\overline{b} \\
C_{21} &= \tfrac{1}{2}\overline{f} \\
C_{22} &= 0 \\
C_{30} &= \tfrac{1}{2}\overline{a}.
\end{aligned} \tag{H.26}$$

H.4 Least-squares analysis and Chebyshev series

It has already been noted in chapter 7 that Chebyshev polynomials are remarkably good approximating polynomials. In fact, fitting a Chebyshev series to data is entirely equivalent to fitting a LS model. With a little extra effort one can show that this is the case for any orthogonal polynomials as follows [88]:

Let $\{\psi_i(x), i = 1,\ldots,\infty\}$ be a set of polynomials orthonormal on the interval $[a,b]$ with weighting function $w(x)$, i.e.

$$\int_{\mathrm{a}}^{b} \mathrm{d}x\, w(x)\psi_i(x)\psi_j(x) = \delta_{ij}. \tag{H.27}$$

(The Chebyshev polynomials used in this work are not orthonormal. However, the set $\psi_0(x) = \pi^{-\frac{1}{2}}$ and $\psi_i(x) = (2/\pi)^{\frac{1}{2}}$ are.) Suppose one wishes to approximate a function $f(x)$ by a summation of the form

$$\hat{f}(x) = \sum_{i=0}^{n} c_i\psi_i(x). \tag{H.28}$$

A least-squared error functional can be defined by

$$\begin{aligned}
I_n[c_i] &= \int_{\mathrm{a}}^{b} \mathrm{d}x\, w(x)|f(x) - \hat{f}(x)|^2 \\
&= \int_{\mathrm{a}}^{b} \mathrm{d}x\, w(x)|f(x) - \sum_{i=0}^{n} c_i\psi_i(x)|^2
\end{aligned} \tag{H.29}$$

and expanding this expression gives

$$I_n[c_i] = \int_{\mathrm{a}}^{b} \mathrm{d}x\, w(x) f(x)^2 + 2\sum_{i=0}^{n} c_i \int_{\mathrm{a}}^{b} dx\, w(x) f(x) \psi_i(x) + \sum_{i=0}^{n}\sum_{j=0}^{n} c_i c_j \int_{\mathrm{a}}^{b} \mathrm{d}x\, w(x) \psi_i(x) \psi_j(x). \tag{H.30}$$

Now, the Fourier coefficients for a_i for an expansion are defined by

$$a_i = \int_{\mathrm{a}}^{b} \mathrm{d}x\, w(x) f(x) \psi_i(x) \tag{H.31}$$

so using this and the orthogonality relation (H.27) gives

$$I_n[c_i] = \int_{\mathrm{a}}^{b} \mathrm{d}x\, w(x) f(x)^2 - 2\sum_{i=0}^{n} a_i c_i + \sum_{i=0}^{n} c_i^2 \tag{H.32}$$

and finally completing the square gives

$$I_n[c_i] = \int_{\mathrm{a}}^{b} \mathrm{d}x\, w(x) f(x)^2 - \sum_{i=0}^{n} a_i^2 + \sum_{i=0}^{n} (c_i - a_i)^2. \tag{H.33}$$

Now, the first two terms of this expression are fixed by the function $f(x)$ and the Fourier coefficients, so minimizing the error functional by varying c_i is simply a matter of minimizing the last term. This is only zero if $a_i = c_i$. This shows clearly that using a Fourier expansion of orthogonal functions is an LS procedure. The only point which needs clearing up is that the usual LS error functional is

$$I_n[c_i] = \int_{\mathrm{a}}^{b} \mathrm{d}x\, |f(x) - \hat{f}(x)|^2 \tag{H.34}$$

without the weighting function. In fact, for the Chebyshev expansion changing the variables from x to $y = \cos(x)$ changes (H.28) to

$$I_n[c_i] = \int_{0}^{\pi} \mathrm{d}y\, |f(\cos^{-1}(y)) - \hat{f}(\cos^{-1}(y))|^2 \tag{H.35}$$

which is the usual functional over a different interval.

Appendix I

Integration and differentiation of measured time data

The contents of chapter 7 illustrate the power of the restoring force surface methods in extracting the equations of motion of real systems. The main problem, common to the Masri–Caughey and direct parameter estimation methods, is that displacement, velocity and acceleration data are all needed simultaneously at each sampling instant. In practical terms, this would require a prohibitive amount of instrumentation, particularly for MDOF systems. For each degree-of-freedom, four transducers are needed, each with its associated amplification etc. It is also necessary to sample and store the data. A truly pragmatic approach to the problem demands that only one signal should be measured and the other two estimated from it, for each DOF. The object of this appendix is to discuss which signal should be measured and how the remaining signals should be estimated. There are essentially two options:

(1) measure $\ddot{y}(t)$ and numerically integrate the signal to obtain $\dot{y}(t)$ and $y(t)$; and
(2) measure $y(t)$ and numerically differentiate to obtain $\dot{y}(t)$ and $\ddot{y}(t)$.

There are of course other strategies, Crawley and O'Donnell [71] measure displacement and acceleration and then form the velocity using an optimization scheme. Here, options 1 and 2 are regarded as the basic strategies. Note that analogue integration is possible [134], but tends to suffer from most of the same problems as digital or numerical integration.

Integration and differentiation methods fall into two distinct categories—the time domain and the frequency domain—and they will be dealt with separately here. It is assumed that the data are sampled with high enough frequency to eliminate aliasing problems. In any case, it will be shown that Shannon's rule of sampling at twice the highest frequency of interest is inadequate for accurate implementation of certain integration rules.

I.1 Time-domain integration

There are two main problems associated with numerical integration: the introduction of spurious low-frequency components into the integrated signal and the introduction of high-frequency pollution. In order to illustrate the former problem, one can consider the trapezium rule as it will be shown to have no high-frequency problems. In all cases, the arguments will be carried by example and the system used will be the simple SDOF oscillator with equation of motion

$$\ddot{y} + 40\dot{y} + 10^4 y = x(t) \tag{I.1}$$

with $x(t)$ a Gaussian white-noise sequence band-limited onto the range $[0, 200]$ Hz. The sampling frequency for the system is 1 kHz. The undamped natural frequency of the system is 15.92 Hz, so the sampling is carried out at 30 times the resonance. This is to avoid any problems of smoothness of the signal for now. As y, $\dot{y}$ and $\ddot{y}$ are available noise-free from the simulation, an LS fit to the simulated data will generate benchmark values for the parameter estimates later. The estimates are found to be

$$m = 1.000, \quad c = 40.00, \quad k = 10\,000.0, \quad \alpha = 0.0$$

where α is added as a constant offset to the restoring force. The model MSE was essentially zero.

I.1.1 Low-frequency problems

The first attempt at signal processing considered here used the trapezium rule

$$v_i = v_{i-1} + \frac{\Delta t}{2}(u_i + u_{i-1}) \tag{I.2}$$

where $v(t)$ is the estimated integral with respect to time of $u(t)$. This rule was applied in turn to the $\ddot{y}_i$ signal from the simulation and then to the resulting $\dot{y}_i$. Each step introduces an unknown constant of integration, so that

$$\dot{y}(t) = \int \mathrm{d}t\, \ddot{y}(t) + A \tag{I.3}$$

and

$$y(t) = \left(\int \mathrm{d}t \right)^2 \ddot{y}(t) + At + B. \tag{I.4}$$

The spurious mean level A is clearly visible in $\dot{y}$ (figure I.1) and the linear drift component $At + B$ can be seen in $y(t)$ (figure I.2). In the frequency domain, the effect manifests itself as a spurious low-frequency component as the unwanted A and B affect the d.c. line of $\dot{y}$ and y respectively. The effect on the system identification is, as one might expect, severe; the estimated parameters are:

$$m = 0.755, \quad c = 38.8, \quad k = -27.6, \quad \alpha = 0.008$$

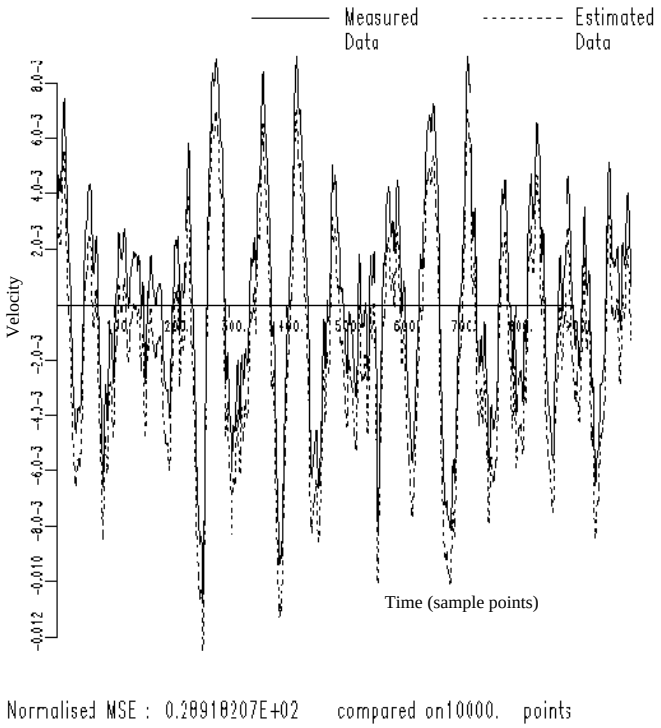


Figure I.1. Comparison of exact and estimated velocity data—no low-frequency correction.

and the MSE is raised to 24%. The restoring force surface computed from this data is shown in figure I.3, it is impossible to infer linearity of the system.

Usually, the constant of integration is fixed by say, initial data $\dot{y}(0)$. Unfortunately when dealing with a stream of time-data, this information is not available. However, all is not lost. Under certain conditions ($x(t)$ is a zero-mean sequence and the nonlinear restoring force $f(y, \dot{y})$ is an odd function of its arguments) it can be assumed that $\dot{y}(t)$ and $y(t)$ are zero-mean signals. This means that A and B can be set to the appropriate values by removing the mean level from $\dot{y}$ and a linear drift component from $y(t)$. (Note that the only problem here is with $y(t)$; in any laboratory experiment, $\dot{y}$ must be zero-mean in order for the apparatus to remain confined to the bench!) If these operations are applied to the data obtained earlier, there is a considerable improvement in the displacement and velocity estimates as shown in the comparisons in figures I.4 and I.5. Although the signal estimates are now excellent, this is not a general result as in the generic case, higher-order polynomial trends remain at a sufficient level to corrupt the displacement. An LS fit to these data produces the parameter

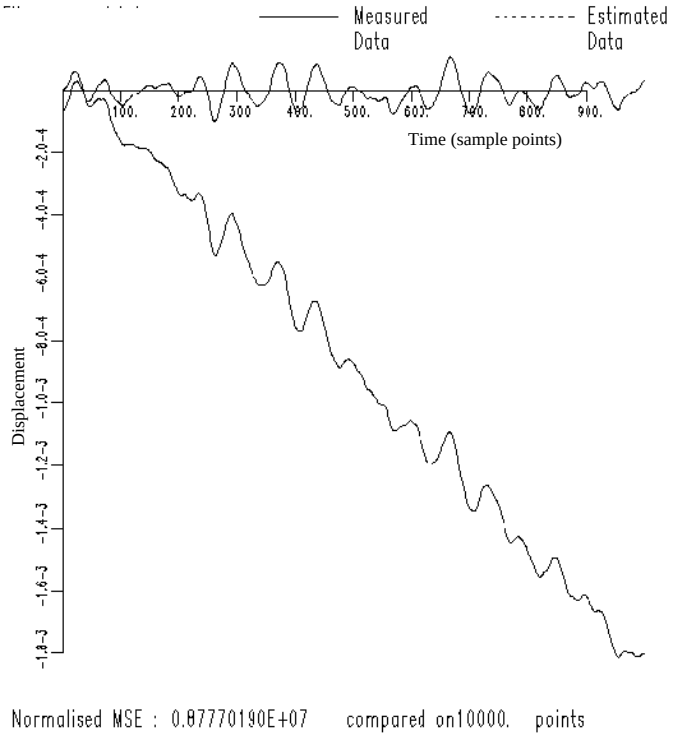


Figure I.2. Comparison of exact and estimated displacement data—no low-frequency correction.

estimates:

$$m = 0.999, \quad c = 40.07, \quad k = 10\,008.3, \quad \alpha = 0.000$$

which is a vast improvement on the untreated case. The MSE for the model is 5.7×10^{-9}.

It is worth noting at this point that the experimenter is not powerless to change matters. The form of the input signal is after all under his or her control. Suppose no energy is supplied to the system at the contentious low frequencies, one would then be justified in removing any low-frequency components in the estimated velocity and displacement by drift removal or filtration. In order to examine this possibility, the system (I.1) was simulated with $x(t)$ band-limited onto the interval $[5, 40]$ Hz. The acceleration data were integrated as before using the trapezium rule and LS parameter estimates were obtained. The estimates were identical to those obtained earlier; however, one should recall that the results for broadband excitation were surprisingly good. In general a band-limited signal is recommended in order to have a robust state estimation procedure.

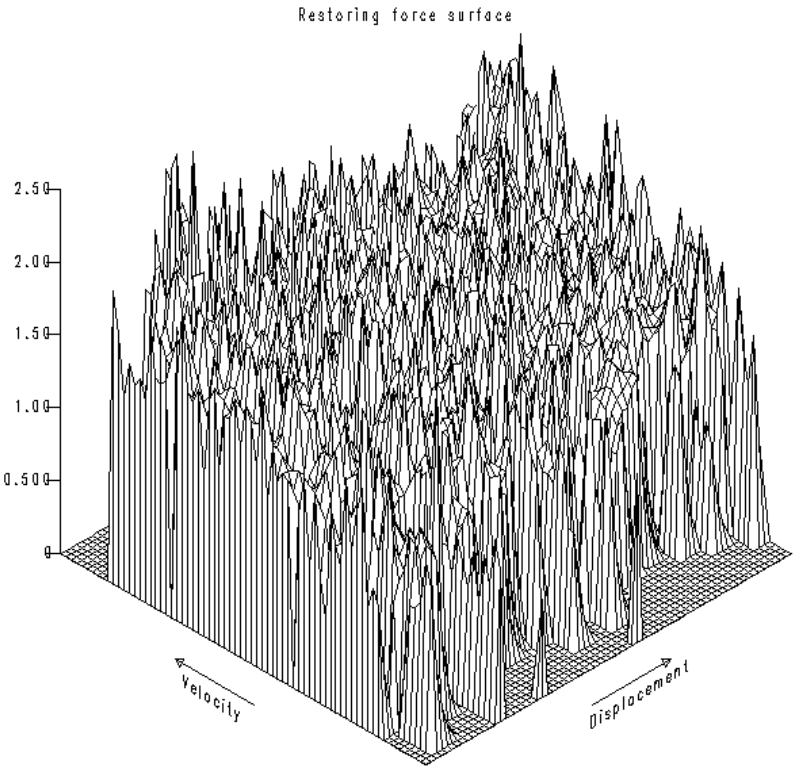


Figure I.3. Restoring force surface constructed from the data from figures I.1 and I.2.

The resulting restoring force surface is shown in figure I.6; the linear form is very clear as expected given the accuracy of the parameters. This small aside raises a significant question: How far one can proceed in the definition of optimal experimental strategies? This will not be discussed in any detail here, the reader is referred to [271B] for a catalogue of simple excitation types with a discussion of their effectiveness and also to [84] in which optimal excitations are derived from a rigorous viewpoint.

Regarding low-frequency pollution, one other caveat is worthy of mention. It is possible that d.c. components can be introduced into the acceleration signal before integration. This should strictly *not* be removed. The reason is as follows. Although $\ddot{y}$ is constrained to be zero-mean, any *finite* sample of acceleration data will necessarily have a non-zero mean $\ddot{y}_{\mathrm{s}}$, subtracting this gives a signal $\ddot{y}(t) - \ddot{y}_{\mathrm{s}}$

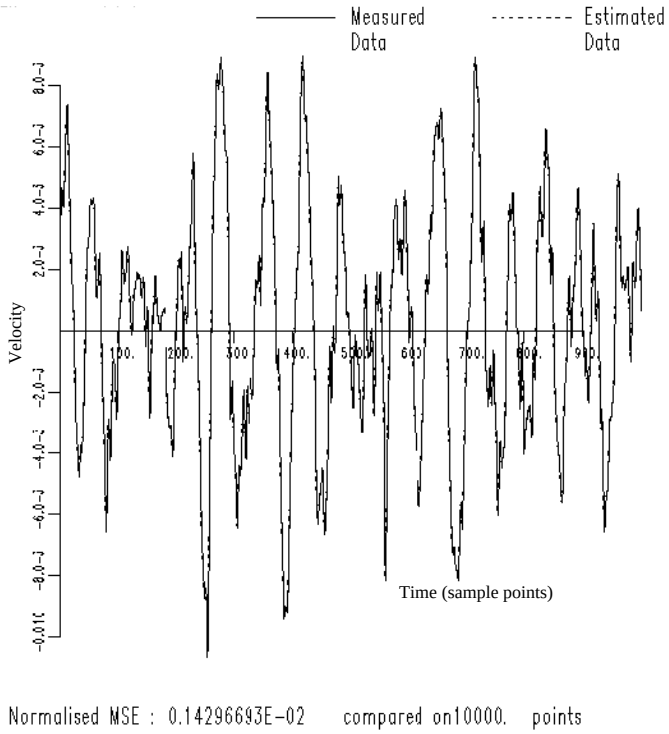


Figure I.4. Comparison of exact and estimated velocity data—mean removed.

which is not asymptotically zero-mean. Integration then gives

$$\dot{y}(t) = \int \mathrm{d}t\, \ddot{y}(t) - \ddot{y}_{\mathrm{s}} t + A \tag{I.5}$$

and

$$y(t) = \left(\int \mathrm{d}t \right)^2 \ddot{y}(t) - \frac{1}{2}\ddot{y}_{\mathrm{s}} t^2 + At + B \tag{I.6}$$

and it becomes necessary to remove a linear trend from the velocity and a *quadratic* trend from the displacement. The rather dramatic result of removing the mean acceleration initially is shown in figure I.7.

It is clear from these examples that linear trend removal is not sufficient to clean the displacement signal totally. This can be achieved by two means: first there is filtering as discussed earlier but note that if the signal *should* have a component below the low cut-off of the filter, this will be removed too. The

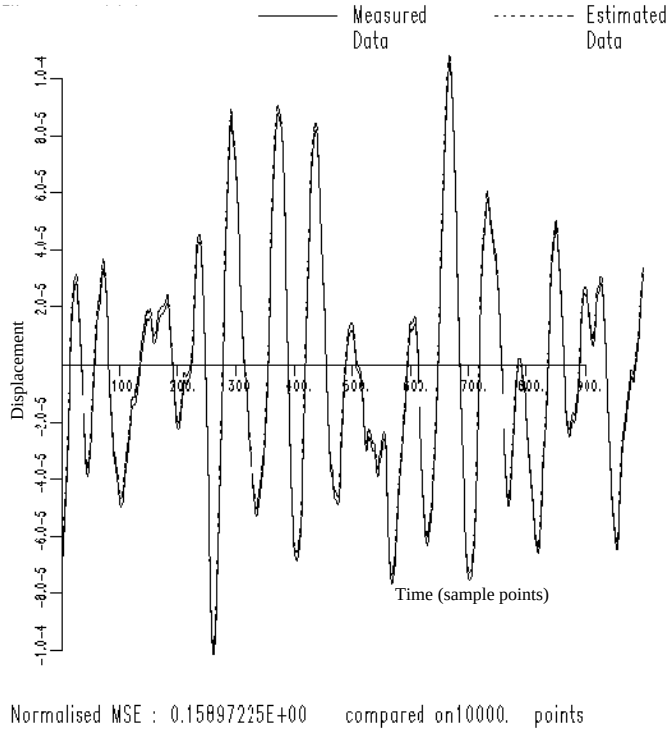


Figure I.5. Comparison of exact and estimated displacement data—linear drift removed.

second approach is to remove polynomial trends, i.e. a model of the form

$$y(t) = \sum_{i=0}^{i_{\max}} a_i t^i \tag{I.7}$$

is fitted and removed using LS. As with filtering, if $i_{\max}$ is too large, the procedure will remove low-frequency data which should be there. In fact, the two methods are largely equivalent and the choice will be dictated by convenience. Suppose the data comprise a record of T s, sampled at Δt. Fitting a polynomial of order n will account for up to n zero-crossings within the record. As there are two zero-crossings per harmonic cycle, this accounts for up to $n/2$ cycles. So removing a polynomial trend of order n is equivalent to high-pass filtering with cut-off $n/(2T)$.

Note that data must be passed through any filter in both the forward and backward directions in order to zero the phase lags introduced by the filter. Any such phase lags will destroy simultaneity of the signals and will have a disastrous effect on the estimated force surface.

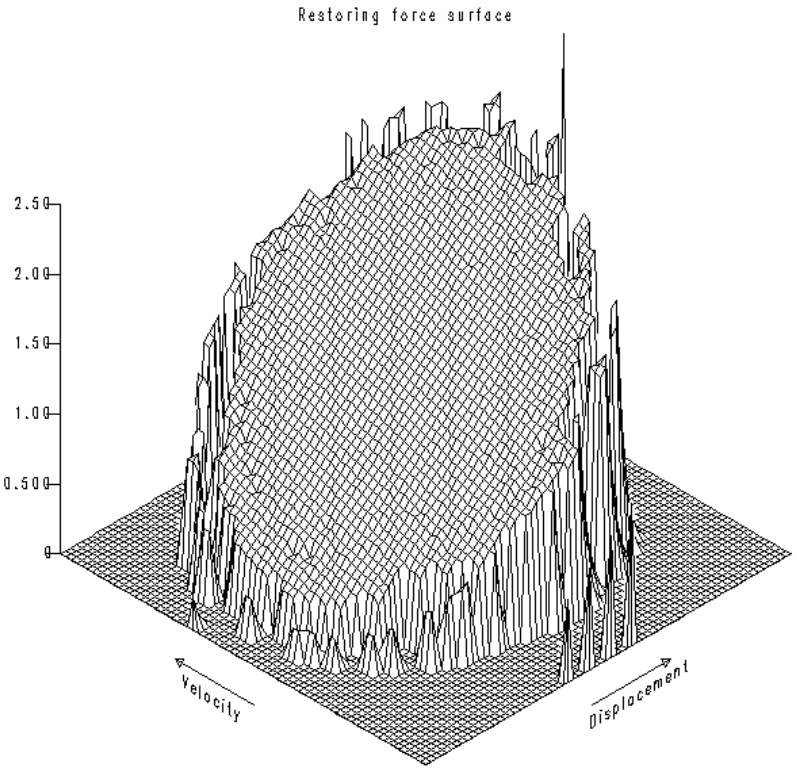


Figure I.6. Restoring force surface constructed from the data from figures I.4 and I.5.

I.1.2 High-frequency problems

It will be shown in the next section that the trapezium rule only suffers from low-frequency problems. However, it is not a particularly accurate integration rule and, unfortunately, in passing to rules with higher accuracy, the possibility of high-frequency problems arises in addition to the omnipresent low-frequency distortion. If an integration routine is unstable at high frequencies, any integrated signals must be band-pass-filtered rather than simply high-passed.

The two rules considered here are Simpson's rule

$$v_{i+1} = v_{i-1} + \frac{\Delta t}{3}(u_{i+1} + 4u_i + u_{i-1}) \tag{I.8}$$

and Tick's rule (or one of them)

$$v_{i+1} = v_{i-1} + \Delta t(0.3854u_{i+1} + 1.2832u_i + 0.3854u_{i-1}). \tag{I.9}$$

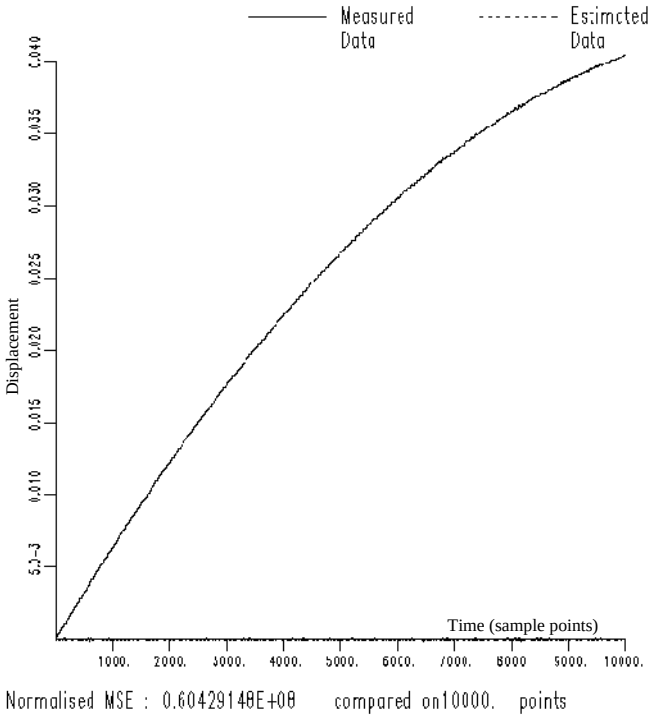


Figure I.7. Comparison of exact displacement data and that estimated after the acceleration mean level was removed. The quadratic trend in the estimate is shown.

The latter algorithm is more accurate than Simpson's rule over low frequencies but suffers more over high.

In order to illustrate the high-frequency problem, the acceleration data from the previous simulation with $x(t)$ band-limited between 5 and 40 Hz (i.e. with no appreciable high-frequency component) was integrated using Tick's rule. The resulting displacement signal is shown in figure I.8; an enormous high-frequency component has been introduced.

These simulations lead us to the conclusion that careful design of the experiment may well allow the use of simpler routines, with a consequent reduction in the post-integration processing requirements. Integration can be thought of as a solution of the simplest type of differential equation; this means that routines for integrating differential equations could be used. A comparison of six methods is given in [28], namely centred difference, Runge–Kutta, Houbolt's method, Newmark's method, the Wilson theta method and the harmonic acceleration method. With the exception of centred difference, all

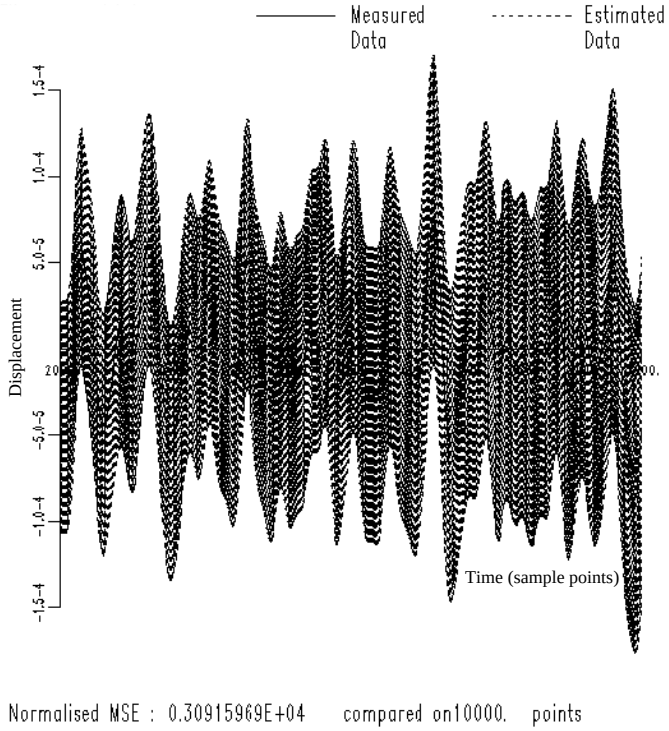


Figure I.8. Comparison of exact and estimated displacement data showing the large high-frequency component introduced into the estimate if Tick's rule is used.

the methods are more complex and time-consuming than the simple routines discussed here and the possible increase in accuracy does not justify their use.

I.2 Frequency characteristics of integration formulae

The previous discussion has made a number of statements without justification; it is time now to provide the framework for this. It is possible to determine the frequency-domain behaviour of the integration and differentiation rules by considering them as digital filters or, alternatively, as ARMA models. The basic ideas are taken from [129]. Throughout this section, a time scale is used such that $\Delta t = 1$. This means that the sampling frequency is also unity and the Nyquist frequency is 0.5: the angular Nyquist frequency is π.

The simplest integration rule considered here is the trapezium rule (I.2), which is, with the conventions described earlier

$$v_i = v_{i-1} + \tfrac{1}{2}(u_i + u_{i-1}). \tag{I.10}$$

This is little more than an ARMA model as described in chapter 1. The only difference is the presence of the present input u_i. It can be written in terms of the backward shift operator $\triangle$ as follows:

$$(1 - \triangle)v_i = \tfrac{1}{2}(1 + \triangle). \tag{I.11}$$

Now, applying the approach of section 1.6, setting $u_i = \mathrm{e}^{\mathrm{i}\omega t}$ and $v_i = H(\omega)\mathrm{e}^{\mathrm{i}\omega t}$, where $H(\omega)$ is the FRF of the process $u \longrightarrow v$, it is a simple matter to obtain (using (1.91) and (1.92))

$$H(\omega) = \frac{1}{2}\frac{1 + \mathrm{e}^{\mathrm{i}\omega}}{1 - \mathrm{e}^{\mathrm{i}\omega}} = \frac{\cos(\frac{\omega}{2})}{2\mathrm{i}\sin(\frac{\omega}{2})}. \tag{I.12}$$

Now, following [129], one can introduce an alternative FRF $H_{\mathrm{a}}(\omega)$, which is a useful measure of the accuracy of the formula. It is defined by

$$H_{\mathrm{a}}(\omega) = \frac{\text{Spectrum of estimated result}}{\text{Spectrum of true result}}. \tag{I.13}$$

Now, if $u(t) = \mathrm{e}^{\mathrm{i}\omega t}$, the true integral—without approximation—is $v(t) = \mathrm{e}^{\mathrm{i}\omega t}/i\omega$. For the trapezium rule, it follows from (I.12) that the estimate of $v(t)$ is

$$\hat{v}(t) = \frac{\cos(\frac{\omega}{2})}{2\mathrm{i}\sin(\frac{\omega}{2})}\mathrm{e}^{\mathrm{i}\omega t} \tag{I.14}$$

so for the trapezium rule

$$H_{\mathrm{a}}(\omega) = \cos\left(\frac{\omega}{2}\right)\left(\frac{\frac{\omega}{2}}{\sin(\frac{\omega}{2})}\right). \tag{I.15}$$

This function is equal to unity at $\omega = 0$ and decreases monotonically to zero at $\omega = \pi$—the Nyquist frequency. This means that the trapezium rule can only integrate constant signals without error. It underestimates the integral $v(t)$ at all other frequencies.

In the units of this section, Simpson's rule (I.8) becomes

$$v_{i+1} = v_{i-1} + \tfrac{1}{3}(u_{i+1} + 4u_i + u_{i-1}). \tag{I.16}$$

Application of this procedure to formula (I.16) gives

$$H(\omega) = \frac{\mathrm{e}^{\mathrm{i}\omega} + 4 + \mathrm{e}^{-\mathrm{i}\omega}}{3(\mathrm{e}^{\mathrm{i}\omega} - \mathrm{e}^{-\mathrm{i}\omega})} = \frac{2 + \cos(\omega)}{3\sin(\omega)} \tag{I.17}$$

and

$$H_{\mathrm{a}}(\omega) = \frac{2 + \cos(\omega)}{3}\left(\frac{\omega}{\sin(\omega)}\right). \tag{I.18}$$

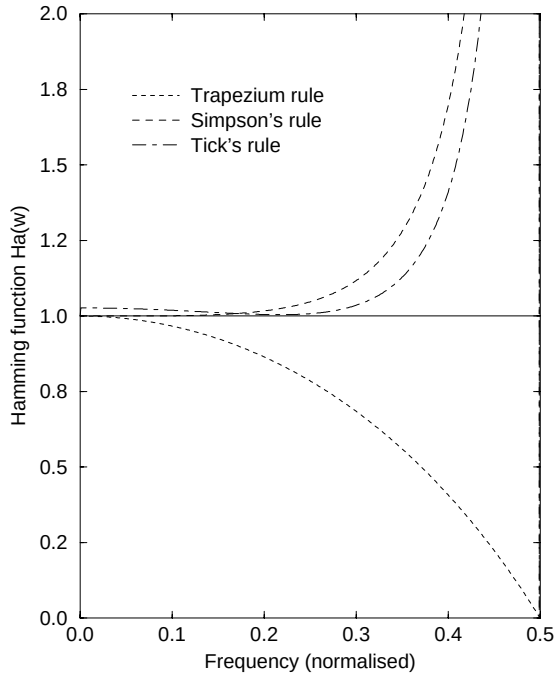


Figure I.9. FRFs for various time-domain integration procedures.

It follows that $H_{\mathrm{a}}(\omega)$ tends to unity at $\omega = 0$ in the same way as for the trapezium rule. However, unlike the simpler integrator, $H_{\mathrm{a}}(\omega)$ for Simpson's rule tends to infinity as ω approaches the Nyquist frequency, indicating instability at high frequencies. Figure I.9 shows $H_{\mathrm{a}}(\omega)$ for the three integration rules discussed here. It shows that they all have the same low-frequency behaviour, but Simpson's rule and Tick's rule blow up at high frequencies. It also substantiates the statement that Tick's rule is superior to Simpson at low frequencies but has worse high-frequency behaviour. In fact, there are a whole family of Tick rules; the one shown here has been designed to be flat over the first half of the Nyquist interval, which explains its superior performance there. The penalty for the flat response is the faster blow-up towards the Nyquist frequency.

It remains to show how low-frequency problems arise. For simplicity, it is assumed that the trapezium rule is used. The implication of figure I.9 is that all the rules are perfect as $\omega \longrightarrow 0$. Unfortunately, the analysis reckons without measurement noise. If the sampled u_i have a measurement error ζ_i or even a truncation error for simulated data, then the spectrum of the estimated integral is given by

$$V(\omega) = H(\omega)(U(\omega) + Z(\omega)) \tag{I.19}$$

where $V(\omega)$ (respectively $U(\omega)$, $Z(\omega)$) is the spectrum of v_i, (respectively u_i,

ζ_i). The spectrum of the error in the integral $E(\omega)$ is straightforwardly obtained as

$$E(\omega) = H(\omega)Z(\omega) = \frac{1}{2\mathrm{i}} \cot\left(\frac{\omega}{2}\right) Z(\omega) \tag{I.20}$$

and as $H(\omega)$ tends to infinity as ω tends to zero, *any* low-frequency input noise is magnified greatly by the integration process. The integration is unstable under small perturbations. As all of the methods have the same low-frequency $H(\omega)$ behaviour, they all suffer from the same problem.

In the numerical simulations considered in the previous section, the highest frequency of interest was about 50 Hz where the band-limited input was used, the Nyquist frequency was 500 Hz. This gives a normalized value of 0.05 for the highest frequency in the input. Figure I.9 shows that the three integration rules are indistinguishable in accuracy at this frequency; one is therefore justified in using the simplest rule to integrate.

If frequencies are present up to 0.25 (a quarter of the Nyquist limit), Tick's rule should be used. At this upper limit, figure I.9 shows that $H_{\mathrm{a}}(\omega)$ for the trapezium rule is less than 0.8 so integrating the acceleration data twice using the trapezium rule would only yield 60% of the displacement data at this frequency. If Simpson's rule were used, $H_{\mathrm{a}}(\omega)$ is approximately 1.1, so integrating twice would give an overestimate of 20%. Tick's rule has a unit gain up to 0.25 exactly as it was designed to do. Diagrams like figure I.9 can be of considerable use in choosing an appropriate integration formula.

I.3 Frequency-domain integration

The theoretical basis of this approach is simple: if

$$Y_{\mathrm{a}}(\omega) = \int_{-\infty}^{\infty} \mathrm{d}t\, \mathrm{e}^{-\mathrm{i}\omega t} \ddot{y}(t) \tag{I.21}$$

is the Fourier transform of the acceleration $\ddot{y}(t)$, then $Y_v(\omega) = Y_{\mathrm{a}}(\omega)/i\omega$ is the corresponding transform of the velocity $\dot{y}(t)$ and $Y(\omega) = Y_{\mathrm{a}}(\omega)/-\omega^2$ is the transform of the displacement. So in the frequency domain, division by $\mathrm{i}\omega$ is equivalent to time integration in the time domain. In practice, for sampled data, the discrete or fast Fourier transform is used, but the principle is the same. Mean removal is accomplished by setting the $\omega = 0$ line to 0 (in any case, one cannot carry out the division here).

At first sight this appears to be a very attractive way of looking at the problem. However, on closer inspection, it turns out to have more or less the same problems as time-domain integration and also to have a few of its own.

The first problem to arise concerns the acceleration signal. If the excitation is random, the signal will not be periodic over the Fourier window and will consequently have leakage problems [23]. Figure I.10 shows (a) the spectrum of a sine wave which was periodic over the range of data transformed and (b)

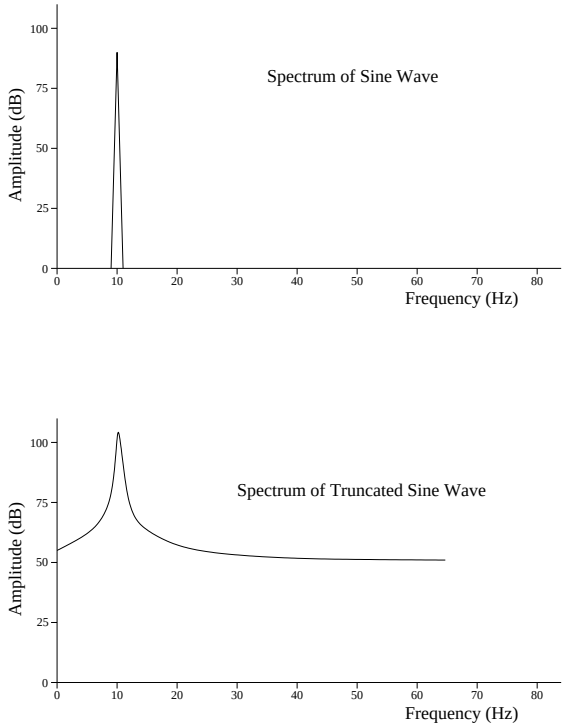


Figure I.10. The effect of leakage on the spectrum of a sine wave.

the spectrum of a sine wave which was not. In the latter case, energy has 'leaked' out into neighbouring bins. More importantly, it has been transferred to low frequencies where it will be greatly magnified by the integration procedure. Figure I.11 shows a twice-integrated sine wave which was not periodic over the Fourier window; the low-frequency drift due to leakage is evident.

The traditional approach to avoiding leakage is to window the data and this is applied here using a standard cosine Hanning window [129]. Because the Hanning window is only close to unity in its centre, only the integrated data in this region are reliable. To overcome this problem, the Fourier windows over the data record should overlap to a large extent. The effect of the multiple windowing is a small amplitude modulation over the data record as one can see from figure I.12 (which shows the double integral of the same data as figure I.11 except with windowed data). The modulation can be suppressed by discarding a higher proportion of the window for each transform, at the expense of extended processing time. Other windows like the flat-top window can sometimes be used with greater efficiency.

To illustrate the procedure, the data from the simulation of equation (I.1) with

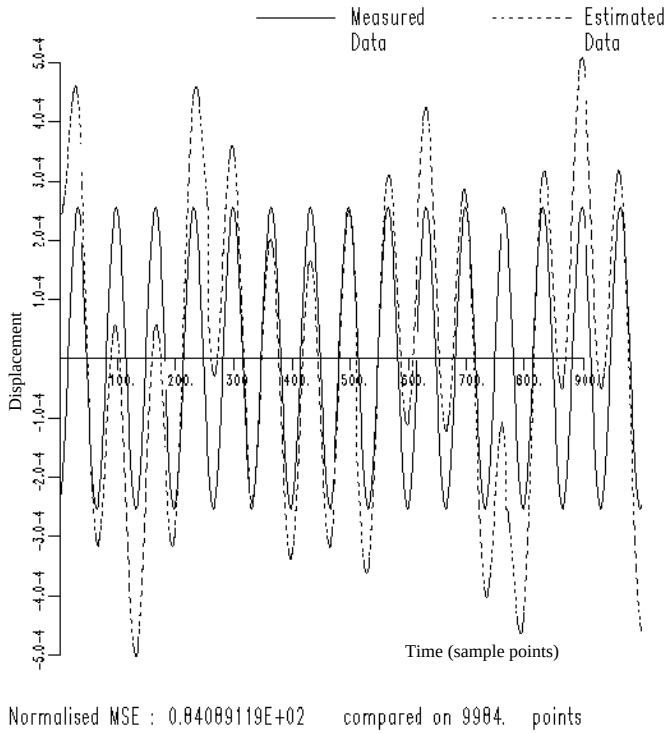


Figure I.11. Comparison of exact and estimated displacement data when the system excitation is a sine wave not periodic over the FFT window. Rectangular window used.

band-limited input were integrated using a Hanning window and an overlap which discarded 80% of the data. The band-limited force was used for the same reason as that discussed earlier—to eliminate low-frequency noise amplification. The mechanism for noise gain is much more transparent here: because the spectrum is divided by $-\omega^2$, the noise is magnified with the signal. There will generally be noise in the spectrum at low frequencies, either from leakage or from fold-back of high-frequency signal and noise caused by aliasing. An LS curve-fit generated the parameter estimates:

$$m = 0.864, \quad c = 39.64, \quad k = 7643.0, \quad \alpha = 0.004$$

and the model MSE was 4.0%. The force surface is shown in figure I.13, the linearity of the system is clearly shown despite the poor estimates.

Note that the division by $-\omega^2$ means that the Fourier transform method does not suffer from high-frequency problems.

Where frequency-domain methods come into their own is where the forcing signal is initially *designed* in the frequency domain as in the work in [9] and

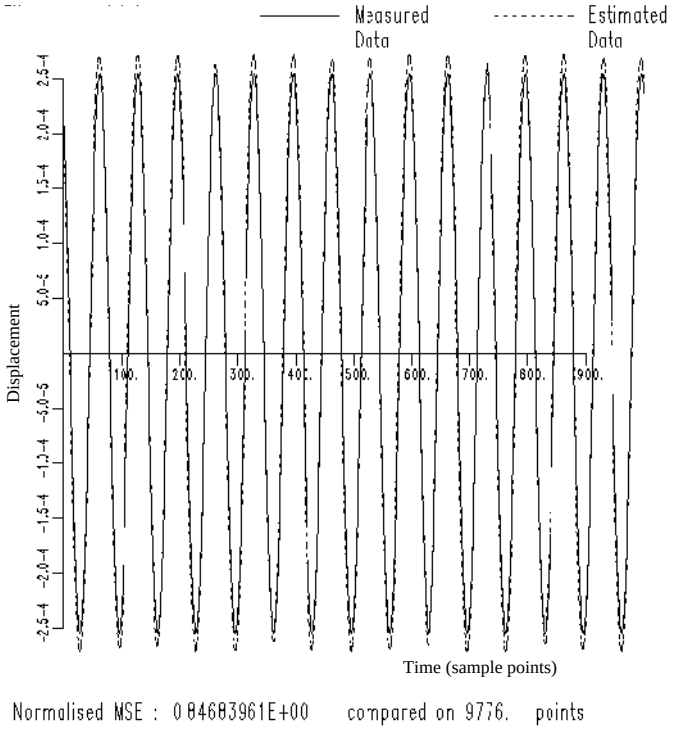


Figure I.12. Comparison of exact and estimated displacement data when the system excitation is a sine wave not periodic over the FFT window. Hanning window used.

[84]. There a periodic pseudo-random waveform is defined as a spectrum and then inverse-transformed (without leakage) into the time domain for exciting the system. As long as subharmonics are not generated, the system response will be periodic over the same window length and can be Fourier transformed with a rectangular window with no leakage.

I.4 Differentiation of measured time data

Because differentiation is defined in terms of a limit, it is notoriously difficult to carry out. The approximation

$$\frac{\mathrm{d}y}{\mathrm{d}t} = \lim_{\delta t \longrightarrow 0} \frac{\delta y}{\delta t} \approx \frac{\Delta y}{\Delta t} \tag{I.22}$$

will clearly become better as Δt is decreased. Unfortunately, this is the sort of operation which will produce significant round-off errors when performed on a digital computer. (Some languages like ADA and packages like Mathematica

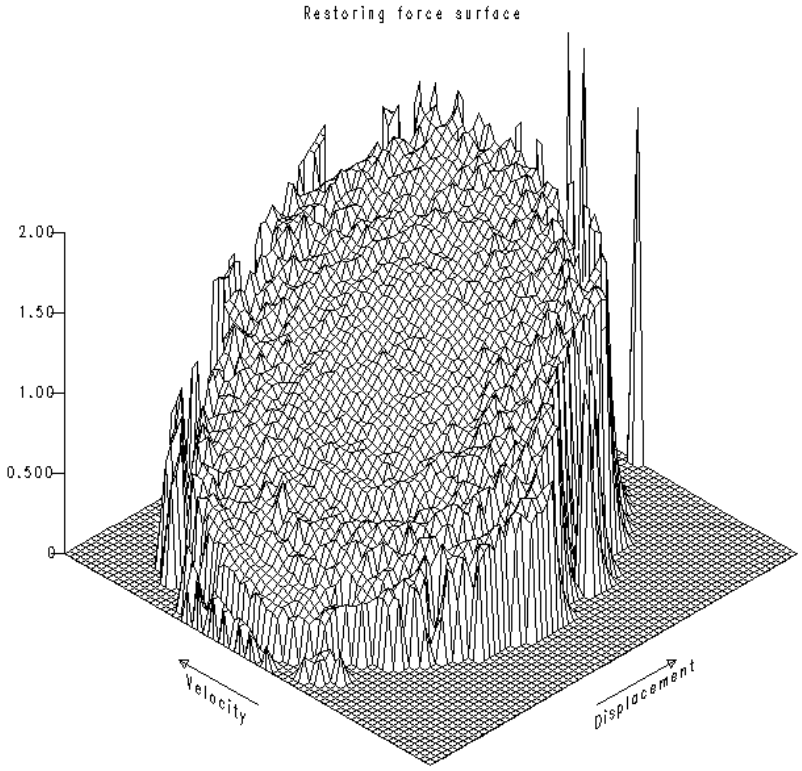


Figure I.13. Force surface obtained using velocity and displacement data from frequency-domain integration of acceleration data. The system excitation is band-limited.

offer accuracy to an arbitrary number of places. However, this is irrelevant. This could only prove of any use for simulation, any laboratory equipment used for data acquisition will be limited in precision.) Numerical differentiation requires a trade-off between approximation errors and round-off errors and optimization is needed in any particular situation. For this reason, numerical differentiation is not recommended except when it is unavoidable. It may be necessary in some cases, e.g. if the restoring force methods are applied to rotor systems as in [48], in order to estimate bearing coefficients. It is usually displacement which is measured for rotating systems, because of the possibility of using non-contact sensors. For this reason, some methods of numerical differentiation are considered.

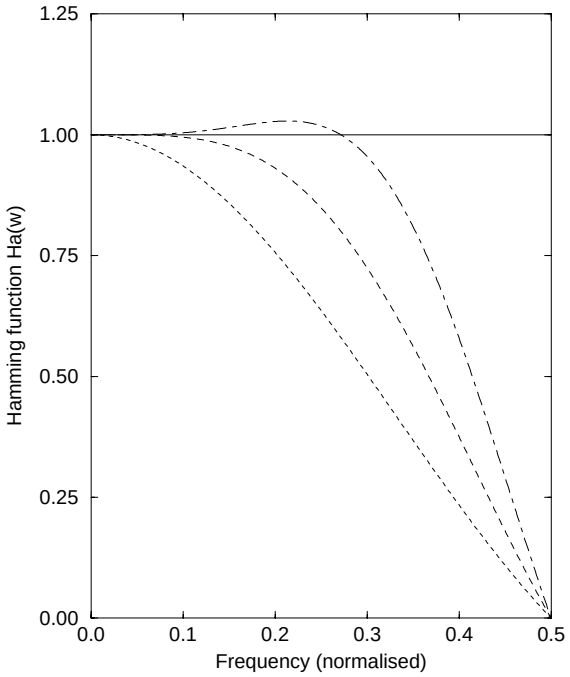


Figure I.14. FRFs for various time-domain differentiation procedures.

I.5 Time-domain differentiation

The most common means of numerical differentiation is by a difference formula. Only the centred differences will be considered here as they are the most stable. The three-, five- and seven-point formulae are given by

$$u_i = \frac{1}{2\Delta t}(v_{i+1} - v_{i-1}) \tag{I.23}$$

$$u_i = \frac{1}{12\Delta t}(-y_{i+2} + 8y_{i+1} - 8y_{i-1} + y_{i+2}) \tag{I.24}$$

$$u_i = \frac{1}{60\Delta t}(2y_{i+3} - 13y_{i+2} + 50y_{i+1} - 50y_{i-1} + 13y_{i-2} - 2y_{i-3}). \tag{I.25}$$

In principle, the formulae using more lags are accurate to a higher order in the step-size Δt. In practice, it is a good idea to keep an eye on the remainder terms for the formulae. The five-point formulae offers a fairly good result in most situations.

Frequency-domain analysis of the formula is possible in the same way as for the integration formula. For example, setting $\Delta t = 1$, the three-point formula becomes

$$u_i = \tfrac{1}{2}(v_{i+1} - v_{i-1}). \tag{I.26}$$

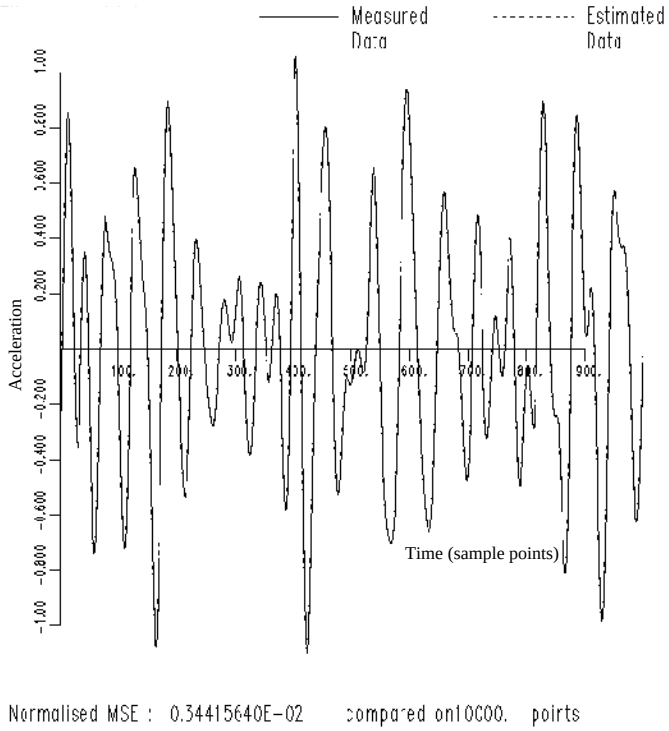


Figure I.15. Comparison of exact and estimated acceleration data obtained by using the five-point centred-difference formula twice on displacement data.

The FRF of the process is obtained as

$$H(\omega) = \mathrm{i}\sin(\omega). \tag{I.27}$$

As the true derivative for $v(t) = \mathrm{e}^{\mathrm{i}\omega t}$ is $u(t) = \mathrm{i}\omega\mathrm{e}^{\mathrm{i}\omega t}$

$$H_{\mathrm{a}}(\omega) = \frac{\sin(\omega)}{\omega}. \tag{I.28}$$

This differentiator is only accurate at $\omega = 0$ and underestimates at all higher frequencies. At $\omega = \pi/2$, i.e. half-way up the Nyquist interval, $H_{\mathrm{a}}(\omega) = 0.63$, so the formula only reproduces 40% of the acceleration at this frequency if it is applied twice to the displacement.

The normalized five-point rule is

$$u_i = \tfrac{1}{12}(-y_{i+2} + 8y_{i+1} - 8y_{i-1} + y_{i+2}) \tag{I.29}$$

and the corresponding $H_{\mathrm{a}}(\omega)$ is

$$H_{\mathrm{a}}(\omega) = \frac{8\sin(\omega) - \sin(2\omega)}{6\omega}. \tag{I.30}$$

The $H_{\mathrm{a}}(\omega)$ functions for the three rules are shown in figure I.14.

The use of the five-point formula will be illustrated by example; the acceleration data were taken from the simulation of (I.1) with the input between 0 and 200 Hz. The displacement was differentiated twice to give velocity and acceleration, the comparison errors were 1.7×10^{-5} and 3.4×10^{-3} (figure I.15). This is remarkably good. However, there is evidence that this is not a general result. Differentiation can sometimes produce inexplicable phase shifts in the data which result in very poor parameter estimates [271A]. The parameters obtained from an LS fit are

$$m = 1.00, \quad c = 40.2, \quad k = 10\,018.0, \quad \alpha = 0.0$$

and the model MSE is 0.03%.

I.6 Frequency-domain differentiation

The basis of this method is the same as for integration except that differentiation in the frequency domain is implemented by *multiplying* the Fourier transform by $\mathrm{i}\omega$. So if the displacement spectrum $Y(\omega)$ is given as follows:

$$Y_{\mathrm{d}}(\omega) = \mathrm{i}\omega Y(\omega) \tag{I.31}$$

and

$$Y_{\mathrm{a}}(\omega) = -\omega^2 Y(\omega) \tag{I.32}$$

The frequency-domain formulation shows clearly that differentiation amplifies high-frequency noise.

The leakage problem is dealt with in exactly the same way as for integration.

Appendix J

Volterra kernels from perturbation analysis

The method of harmonic probing was discussed in chapter 8 as a means of calculating HFRFs from equations of motion. Historically, calculations with the Volterra series began with methods of estimating the kernels themselves. This section will illustrate one of the methods of extracting kernels—the perturbation method—by considering a simple example: the ubiquitous Duffing oscillator

$$m\ddot{y} + c\dot{y} + ky + k_3 y^3 = x(t). \tag{J.1}$$

Now, assume that k_3 is small enough to act as an expansion parameter and the solution of the equation (J.1) can be expressed as an infinite series

$$y(t) = y^{(0)}(t) + k_3 y^{(1)}(t) + k_3^2 y^{(2)}(t) + \cdots \tag{J.2}$$

(where, in this section, the superscript labels the perturbation order and *not* the Volterra term order). Once (J.2) is substituted into (J.1), the coefficients of each k_3^i can be projected out to yield equations for the y_i. To order k_3^2, one has

$$\begin{aligned} k_3^0 &: m\ddot{y}^{(0)} + c\dot{y}^{(0)} + ky^{(0)} = x(t) \\ k_3^1 &: m\ddot{y}^{(1)} + c\dot{y}^{(1)} + ky^{(1)} + y^{(0)3} = 0 \\ k_3^2 &: m\ddot{y}^{(2)} + c\dot{y}^{(2)} + ky^{(2)} + 3y^{(0)2}y^{(1)} = 0. \end{aligned} \tag{J.3}$$

The solution method is iterative. The first step is to solve the order k_3^0 equation. This is the standard SDOF linear equation and the solution is simply

$$y^{(0)}(t) = \int_{-\infty}^{\infty} \mathrm{d}\tau \, h_1(\tau) x(t-\tau) \tag{J.4}$$

where $h_1(\tau)$ is the impulse response of the underlying linear system.

The frequency content of the expansion is summed by

$$Y(\omega) = Y^{(0)}(\omega) + k_3 Y^{(1)}(\omega) + k_3^2 Y^{(2)}(\omega) + \cdots \tag{J.5}$$

and so to order k_3^0, one has

$$Y^{(0)}(\omega) = H_1(\omega)X(\omega) = Y_1(\omega). \tag{J.6}$$

The next equation is to order k_3^1 from (8.23). Note that the nonlinear term y_0^3 is actually known from the k_3^0 calculation, and the equation has a forced linear SDOF form

$$m\ddot{y}^{(1)} + c\dot{y}^{(1)} + ky^{(1)} = -y^{(0)3} \tag{J.7}$$

and this has solution

$$y^{(1)}(t) = -\int_{-\infty}^{\infty} \mathrm{d}\tau\, h_1(t-\tau)y^{(0)3}(\tau). \tag{J.8}$$

Substituting (J.4) yields

$$\begin{aligned} y^{(1)}(t) = & -\int_{-\infty}^{\infty} \mathrm{d}\tau\, h_1(t-\tau)\left(\int_{-\infty}^{\infty} \mathrm{d}\tau_1\, h_1(\tau-\tau_1)x(\tau_1)\right) \\ & \times \left(\int_{-\infty}^{\infty} \mathrm{d}\tau_2\, h_1(\tau-\tau_2)x(\tau_2)\right)\left(\int_{-\infty}^{\infty} \mathrm{d}\tau_3\, h_1(\tau-\tau_3)x(\tau_3)\right) \end{aligned} \tag{J.9}$$

or

$$\begin{aligned} y^{(1)}(t) = & -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{d}\tau_1\, \mathrm{d}\tau_2\, \mathrm{d}\tau_3 \left(\int_{-\infty}^{\infty} \mathrm{d}\tau\, h_1(t-\tau)h_1(\tau-\tau_1)\right. \\ & \left. \times h_1(\tau-\tau_2)h_1(\tau-\tau_3)\right)x(\tau_1)x(\tau_2)x(\tau_3). \end{aligned} \tag{J.10}$$

Comparing this with equation (8.6),

$$\begin{aligned} & h_3(t-\tau_1, t-\tau_2, t-\tau_3) \\ & \quad = -\int_{-\infty}^{\infty} \mathrm{d}\tau\, h_1(t-\tau)h_1(\tau-\tau_1)h_1(\tau-\tau_2)h_1(\tau-\tau_3) \end{aligned} \tag{J.11}$$

setting $t = 0$,

$$h_3(-\tau_1, -\tau_2, -\tau_3) = -\int_{-\infty}^{\infty} \mathrm{d}\tau\, h_1(-\tau)h_1(\tau-\tau_1)h_1(\tau-\tau_2)h_1(\tau-\tau_3) \tag{J.12}$$

and finally letting $t_i = -\tau_i$; $i = 1, \ldots, 3$, one obtains

$$h_3(t_1, t_2, t_3) = -\int_{-\infty}^{\infty} \mathrm{d}\tau\, h_1(-\tau)h_1(\tau+t_1)h_1(\tau+t_2)h_1(\tau+t_3). \tag{J.13}$$

Note that this, the third kernel, factors into a functional of the the first kernel h_1. This behaviour will be repeated at higher order. Before proceeding though,

it is instructive to consider what is happening in the frequency domain. The appropriate transformation is (from (8.13))

$$H_3(\omega_1,\omega_2,\omega_3)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\mathrm{d}t_1\,\mathrm{d}t_2\,\mathrm{d}t_3\,\mathrm{e}^{-\mathrm{i}(\omega_1t_1+\omega_2t_2+\omega_3t_3)}h_3(t_1,t_2,t_3) \tag{J.14}$$

or

$$\begin{aligned}H_3(\omega_1,\omega_2,\omega_3)=&-\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\mathrm{d}t_1\,\mathrm{d}t_2\,\mathrm{d}t_3\,\mathrm{e}^{-\mathrm{i}(\omega_1t_1+\omega_2t_2+\omega_3t_3)}\\&\times\left(\int_{-\infty}^{\infty}\mathrm{d}\tau\,h_1(-\tau)h_1(\tau+t_1)h_1(\tau+t_2)h_1(\tau+t_3)\right).\end{aligned} \tag{J.15}$$

Now, this expression factors:

$$\begin{aligned}H_3(\omega_1,\omega_2,\omega_3)=&-\int_{-\infty}^{\infty}\mathrm{d}\tau\,h_1(-\tau)\left(\int_{-\infty}^{\infty}\mathrm{d}t_1\mathrm{e}^{-i\omega_1t_1}h_1(t_1+\tau)\right)\\&\times\left(\int_{-\infty}^{\infty}\mathrm{d}t_2\,\mathrm{e}^{-\mathrm{i}\omega_2t_2}h_1(t_2+\tau)\right)\\&\times\left(\int_{-\infty}^{\infty}\mathrm{d}t_3\,\mathrm{e}^{-\mathrm{i}\omega_3t_3}h_1(t_3+\tau)\right).\end{aligned} \tag{J.16}$$

According to the shift theorem for the Fourier transform

$$\int_{-\infty}^{\infty}\mathrm{d}t_1\mathrm{e}^{-\mathrm{i}\omega_1t_1}h_1(t_1+\tau)=\mathrm{e}^{\mathrm{i}\omega_1\tau}H_1(\omega_1) \tag{J.17}$$

and using this result three times in (J.16) yields

$$\begin{aligned}H_3(\omega_1,\omega_2,\omega_3)&=-\int_{-\infty}^{\infty}\mathrm{d}\tau\,h_1(-\tau)\mathrm{e}^{\mathrm{i}\omega_1\tau}H_1(\omega_1)\mathrm{e}^{\mathrm{i}\omega_2\tau}H_1(\omega_2)\mathrm{e}^{\mathrm{i}\omega_3\tau}H_1(\omega_3)\\&=-H_1(\omega_1)H_1(\omega_2)H_1(\omega_3)\int_{-\infty}^{\infty}\mathrm{d}\tau\,\mathrm{e}^{\mathrm{i}(\omega_1+\omega_2+\omega_3)\tau}h_1(-\tau)\end{aligned} \tag{J.18}$$

and a final change of variables $\tau\longrightarrow-\tau$ gives

$$\begin{aligned}H_3(\omega_1,\omega_2,\omega_3)&=-H_1(\omega_1)H_1(\omega_2)H_1(\omega_3)\int_{-\infty}^{\infty}\mathrm{d}\tau\,\mathrm{e}^{-\mathrm{i}(\omega_1+\omega_2+\omega_3)\tau)}h_1(\tau)\\&=-H_1(\omega_1)H_1(\omega_2)H_1(\omega_3)H_1(\omega_1+\omega_2+\omega_3).\end{aligned} \tag{J.19}$$

It will be shown later that this is the sole contribution to y_3, the third-order Volterra functional, i.e.

$$y_3(t)=k_3y^{(1)}(t) \tag{J.20}$$

which agrees with (8.87) from harmonic probing as it should.

To drive the method home, the next term is computed from

$$m\ddot{y}^{(2)} + c\dot{y}^{(2)} + ky^{(2)} = -3y^{(0)2}y^{(1)} \tag{J.21}$$

so

$$y^{(1)}(t) = -\int_{-\infty}^{\infty} \mathrm{d}\tau\, h_1(t-\tau)y^{(0)2}(\tau)y^{(1)}(\tau). \tag{J.22}$$

Very similar calculations to those used earlier lead to the Volterra kernel

$$\begin{aligned} h_5(t_1,t_2,t_3,t_4,t_5) = & -3\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{d}\tau\, \mathrm{d}\tau'\, h_1(-\tau)h_1(\tau-\tau')h_1(\tau+t_1) \\ & \times h(\tau+t_2)h_1(\tau'+t_3)h_1(\tau'+t_4)h_1(\tau'+t_5) \end{aligned} \tag{J.23}$$

and kernel transform

$$\begin{aligned} H_1(\omega_1,\omega_2,\omega_3,\omega_4,\omega_5) = & -3H_1(\omega_1)H_1(\omega_2)H_1(\omega_3)H_1(\omega_4)H_1(\omega_5) \\ & \times H_1(\omega_3+\omega_4+\omega_5)H_1(\omega_1+\omega_2+\omega_3+\omega_4+\omega_5). \end{aligned} \tag{J.24}$$

Note that these expressions are not symmetric in their arguments. In fact, any non-symmetric kernel can be replaced by a symmetric version with impunity, as discussed in section 8.1.

Appendix K

Further results on random vibration

The purpose of this appendix is to expand on the analysis given in section 8.7. Further results on the Volterra series analysis of randomly excited systems are presented.

K.1 Random vibration of an asymmetric Duffing oscillator

This is simply a Duffing oscillator with a non-zero k_2 as in (8.49), (or a symmetric oscillator with $x(t)$ with a non-zero mean).

The expression for the $\Lambda_{\mathrm{r}}(\omega)$ expansion remains as given in equation (8.131). However, due to the increased complexity of the HFRF expressions with the introduction of the k_2 term, only the first three terms will be considered here. The required HFRFs can be calculated by harmonic probing, the H_3 needed is given by (8.54), and $H_3(\omega_1, -\omega_1, \omega)$ is given by

$$\begin{aligned} H_3(\omega_1, -\omega_1, \omega) &= H_1(\omega)^2 |H_1(\omega_1)|^2 \\ &\quad \times \{\tfrac{2}{3} k_2^2 [H_1(0) + H_1(\omega + \omega_1) + H_1(\omega - \omega_1)] - k_3\}. \end{aligned} \tag{K.1}$$

The expression for $H_5(\omega_1, -\omega_1, \omega_2, -\omega_2, \omega)$ in terms of k_2, k_3 and H_1 is composed of 220 terms and will therefore not be given here.

Substituting equation (K.1) into the $\frac{S_{y_3 x}(\omega)}{S_{xx}(\omega)}$ term of equation (8.131) gives

$$\begin{aligned} \frac{S_{y_3 x}(\omega)}{S_{xx}(\omega)} = \frac{P k_2^2 H_1(\omega)^2}{\pi} \Bigg\{ & H_1(0) \int_{-\infty}^{+\infty} \mathrm{d}\omega_1 \, |H_1(\omega_1)|^2 \\ & + \int_{-\infty}^{+\infty} \mathrm{d}\omega_1 \, H_1(\omega + \omega_1) |H_1(\omega_1)|^2 \\ & + \int_{-\infty}^{+\infty} \mathrm{d}\omega_1 \, H_1(\omega - \omega_1) |H_1(\omega_1)|^2 \Bigg\} \\ & - \frac{3 P k_3 H_1(\omega)^2}{2\pi} \int_{-\infty}^{+\infty} \mathrm{d}\omega_1 \, |H_1(\omega_1)|^2. \end{aligned} \tag{K.2}$$

As before, simplifications are possible. The first and last integrals are the same and changing coordinates from ω_1 for $-\omega_1$ in the second integral gives the same expression as the third integral. The simplified form of the equation is

$$\frac{S_{y_3x}(\omega)}{S_{xx}(\omega)} = \left\{ \frac{Pk_2^2 H_1(\omega)^2 H_1(0)}{\pi} - \frac{3Pk_3 H_1(\omega)^2}{2\pi} \right\} \int_{-\infty}^{+\infty} \mathrm{d}\omega_1 \, |H_1(\omega_1)|^2 + \frac{2Pk_2^2 H_1(\omega)^2}{\pi} \int_{-\infty}^{+\infty} \mathrm{d}\omega_1 \, H_1(\omega + \omega_1) |H_1(\omega_1)|^2. \tag{K.3}$$

Both of these integrals follow from results in chapter 8. The first integral is identical to that in equation (8.134). The second integral is equal to the second part of the integral on the right-hand side of equation (8.147) with ω_2 set to zero. Substituting the expressions from equations (8.136) and (8.150) into this equation and setting $H_1(0) = \frac{1}{k_1}$ results in

$$\frac{S_{y_3x}(\omega)}{S_{xx}(\omega)} = \frac{Pk_2^2 H_1(\omega)^2}{ck_1^2} - \frac{3Pk_3 H_1(\omega)^2}{2ck_1} - \frac{2Pk_2^2 H_1(\omega)^2 (\omega - 4\mathrm{i}\zeta\omega_n)}{mck_1(\omega - 2\mathrm{i}\zeta\omega_n)(\omega + 2\omega_\mathrm{d} - 2\mathrm{i}\zeta\omega_n)(\omega - 2\omega_\mathrm{d} - 2\mathrm{i}\zeta\omega_n)}. \tag{K.4}$$

Whereas the $\frac{S_{y_3x}(\omega)}{S_{xx}(\omega)}$ term for the classical Duffing oscillator did not affect the position of the poles, the same term for the asymmetric case results in new poles being introduced at

$$2\omega_\mathrm{d} + 2\mathrm{i}\zeta\omega_n; \quad -2\omega_\mathrm{d} + 2\mathrm{i}\zeta\omega_n; \quad 2\mathrm{i}\zeta\omega_n \tag{K.5}$$

as well as creating double poles at the linear system pole locations.

As stated earlier, the $H_5(\omega_1, -\omega_1, \omega_2, -\omega_2, \omega)$ expression for this system consists of 220 H_1 terms. Even when the procedure of combining identical integrals is used, there are still 38 double integrals to evaluate. These integrals will not be given here but they have been solved, again with the aid of a symbolic manipulation package. The $\frac{S_{y_5x}(\omega)}{S_{xx}(\omega)}$ term in the $\Lambda_\mathrm{r}(\omega)$ expansion was found to generate new poles

$$\omega_\mathrm{d} + 3\mathrm{i}\zeta\omega_n; \quad -\omega_\mathrm{d} + 3\mathrm{i}\zeta\omega_n; \quad 3\omega_\mathrm{d} + 3\mathrm{i}\zeta\omega_n; \quad -3\omega_\mathrm{d} + 3\mathrm{i}\zeta\omega_n. \tag{K.6}$$

Note that these poles arise not only due to the k_3 term but also in integrals which depend only upon k_2. This suggests that even nonlinear terms result in poles in the composite FRF at all locations $a\omega_\mathrm{d} + b\mathrm{i}\zeta\omega_n$ where $a \leq b$ are both odd integers or both even. Also, the poles at the locations given in (K.5) became double poles whilst triple poles were found to occur at the positions of the linear system poles.

The pole structure of the first three terms of $\Lambda_\mathrm{r}(\omega)$ for this system is shown in figure K.1. As in the case of the symmetric Duffing oscillator, the poles are

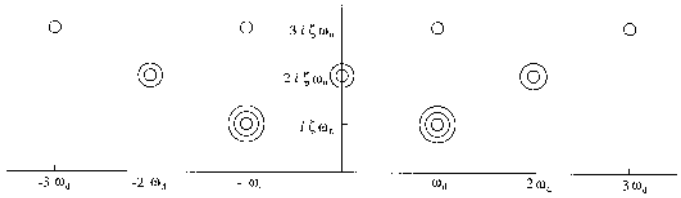


Figure K.1. Pole structure of the first three terms of $\Lambda_{\mathrm{r}}(\omega)$ for the asymmetric Duffing oscillator.

all located in the upper-half of the ω-plane. As discussed in chapter 8 this would cause a Hilbert transform analysis to label the system as linear.

As in the classical Duffing oscillator case, it is expected that the inclusion of all terms in the $\Lambda_{\mathrm{r}}(\omega)$ expansion will result in an infinite array of poles, positioned at $a\omega_{\mathrm{d}} + b\mathrm{i}\zeta\omega_n$ where $a \leq b$ are both odd integers or both even.

K.2 Random vibrations of a simple MDOF system

K.2.1 The MDOF system

The system investigated here is a simple 2DOF nonlinear system with lumped-mass characteristics. The equations of motion are

$$m\ddot{y}_1 + 2c\dot{y}_1 - c\dot{y}_2 + 2ky_1 - k_2y_2 + k_3y_1^3 = x(t) \tag{K.7}$$

$$m\ddot{y}_2 + 2c\dot{y}_2 - c\dot{y}_1 + 2ky_2 - k_1y_1 = 0. \tag{K.8}$$

This system has been discussed before in chapter 3, but the salient facts will be repeated here for convenience. The underlying linear system is symmetrical but the nonlinearity breaks the symmetry and shows itself in both modes. If the FRFs for the processes $x(t) \longrightarrow y_1(t)$ and $x(t) \longrightarrow y_2(t)$ are denoted $H_1^{(1)}(\omega)$ and $H_1^{(2)}(\omega)$, then it can be shown that

$$H_1^{(1)}(\omega) = R_1(\omega) + R_2(\omega) \tag{K.9}$$

$$H_1^{(2)}(\omega) = R_1(\omega) - R_2(\omega) \tag{K.10}$$

and the R_1 and R_2 are (up to a multiplicative constant) the FRFs of the individual modes:

$$R_1(\omega) = \frac{1}{2}\frac{1}{-m(\omega^2 - \omega_{n1}^2) + 2i\zeta_1\omega_{n1}\omega} = \frac{-1}{2m}\frac{1}{(\omega - p_1)(\omega - p_2)} \tag{K.11}$$

$$R_2(\omega) = \frac{1}{2}\frac{1}{-m(\omega^2 - \omega_{n2}^2) + 2i\zeta_2\omega_{n2}\omega} = \frac{-1}{2m}\frac{1}{(\omega - q_1)(\omega - q_2)} \tag{K.12}$$

where $\omega_{\mathrm{n}1}$ and $\omega_{\mathrm{n}2}$ are the first and second undamped natural frequencies and ζ_1 and ζ_2 are the corresponding dampings. p_1 and p_2 are the poles of the first mode and q_1 and q_2 are the poles of the second mode

$$\begin{aligned} p_1, p_2 &= \pm\omega_{\mathrm{d}1} + \mathrm{i}\zeta_1\omega_{\mathrm{n}1} \\ q_1, q_2 &= \pm\omega_{\mathrm{d}2} + \mathrm{i}\zeta_2\omega_{\mathrm{n}2} \end{aligned} \tag{K.13}$$

where $\omega_{\mathrm{d}1}$ and $\omega_{\mathrm{d}2}$ are the first and second damped natural frequncies.

From this point on, the calculation will concentrate on the FRF $H_1^{(1)}(\omega)$ and the identifying superscript will be omitted, the expressions are always for the process $x(t) \longrightarrow y_1(t)$.

In order to calculate the FRF up to order $\mathrm{O}(P^2)$ it is necessary to evaluate equation (8.130), restated here as

$$\begin{aligned} \Lambda_{\mathrm{r}}(\omega) = H_1(\omega) &+ \frac{3P}{2\pi}\int_{-\infty}^{+\infty} \mathrm{d}\omega_1\, H_3(\omega_1, -\omega_1, \omega) \\ &+ \frac{15P^2}{(2\pi)^2}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \mathrm{d}\omega_1\, \mathrm{d}\omega_2\, H_5(\omega_1, -\omega_1, \omega_2, -\omega_2, \omega) \\ &+ \mathrm{O}(P^3). \end{aligned} \tag{K.14}$$

The simple geometry chosen here results in an identical functional form for $\Lambda_{\mathrm{r}}(\omega)$ in terms of $H_1(\omega)$ as that obtained in section 8.7.2. The relevant equations for H_3 and H_5 are given in (8.132) and (8.133). The critical difference is now that $H_1(\omega)$ corresponds to a multi-mode system, and this complicates the integrals in (8.131) a little.

K.2.2 The pole structure of the composite FRF

The first integral which requires evaluation in (8.131) is the order P term

$$I_1 = -\frac{3k_3P}{2\pi}H_1(\omega)^2\int_{-\infty}^{\infty}\mathrm{d}\omega_1\,|H_1(\omega_1)|^2. \tag{K.15}$$

However, as the integral does not involve the parameter ω, it evaluates to a constant and the order P term does not introduce any new poles into the FRF but raises the order of the linear system poles.

The order P^2 term requires more effort, this takes the form

$$\begin{aligned} \frac{S_{y_5x}(\omega)}{S_{xx}(\omega)} = &\frac{9P^2k_3^2H_1(\omega)^3}{4\pi^2}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\mathrm{d}\omega_1\,\mathrm{d}\omega_2\,|H_1(\omega_1)|^2|H_1(\omega_2)|^2 \\ &+ \frac{9P^2k_3^2H_1(\omega)^2}{2\pi^2} \\ &\times\left\{\mathrm{Re}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\mathrm{d}\omega_1\,\mathrm{d}\omega_2\,H_1(\omega_1)|H_1(\omega_1)|^2|H_1(\omega_2)|^2\right. \end{aligned}$$

$$+\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \mathrm{d}\omega_1\,\mathrm{d}\omega_2\, H_1(\omega_1+\omega_2+\omega)|H_1(\omega_1)|^2|H_1(\omega_2)|^2\Big\}. \tag{K.16}$$

The first and second integrals may be dispensed with as they also contain integrals which do not involve ω, and there is no need to give the explicit solution here; no new poles are introduced. The terms simply raise the order of the linear system poles to three again.

The third term in (K.16) is the most complicated. However, it is routinely expressed in terms of 32 integrals I_{jklmn} where

$$I_{jklmn} = \frac{9P^2k_3^2H_1(\omega)^2}{2\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{d}\omega_1\,\mathrm{d}\omega_2\, R_j(\omega_1+\omega_2+\omega) \times R_k^*(\omega_1)R_l(\omega_1)R_{\mathrm{m}}^*(\omega_2)R_n(\omega_2). \tag{K.17}$$

In fact, because of the manifest symmetry in ω_1 and ω_2, it follows that

$$I_{jklmn} = I_{jmnkl} \tag{K.18}$$

and this reduces the number of independent integrals to 20. A little further thought reveals the relation

$$I_{jklmn} = S[I_{\mathrm{s}(j)s(k)s(l)s(m)s(n)}] \tag{K.19}$$

where the s operator changes the value of the index from 1 to 2 and vice-versa and the S operator exchanges the subscripts on the constants, i.e. $\omega_{\mathrm{d}1} \longleftrightarrow \omega_{\mathrm{d}2}$ etc. This reduces the number of integrals to 10. It is sufficient to evaluate the following: I_{11111}, I_{11112}, I_{11121}, I_{11122}, I_{11212}, I_{11221}, I_{11222}, I_{12121}, I_{12122} and I_{12222}. Evaluation of the integral is an exercise in the calculus of residues which requires some help from computer algebra. The expression for the integral is rather large and will not be given here, the important point is that the term I_{jklmn} is found to have poles in the positions

$$\pm\omega_{\mathrm{d}k} \pm \omega_{\mathrm{d}l} \pm \omega_{\mathrm{d}m} + \mathrm{i}(\omega_{nk}\zeta_k + \omega_{nl}\zeta_l + \omega_{nm}\zeta_{\mathrm{m}}). \tag{K.20}$$

It transpires that as a result of pole–zero cancellation, the number of poles varies for each of the independent integrals. I_{11111} and I_{11112} have simple poles at:

$$\omega_{\mathrm{d}1}+3\mathrm{i}\omega_{\mathrm{n}1}\zeta_1, \quad -\omega_{\mathrm{d}1}+3\mathrm{i}\omega_{\mathrm{n}1}\zeta_1, \quad 3\omega_{\mathrm{d}1}+3\mathrm{i}\omega_{\mathrm{n}1}\zeta_1, \quad -3\omega_{\mathrm{d}1}+3\mathrm{i}\omega_{\mathrm{n}1}\zeta_1 \tag{K.21}$$

so by the symmetries described above, I_{12222}, amongst others, has poles at:

$$\omega_{\mathrm{d}2} + 3\mathrm{i}\omega_{\mathrm{n}2}\zeta_2, \quad -\omega_{\mathrm{d}2} + 3\mathrm{i}\omega_{\mathrm{n}2}\zeta_2, \quad 3\omega_{\mathrm{d}2} + 3\mathrm{i}\omega_{\mathrm{n}2}\zeta_2, \quad -3\omega_{\mathrm{d}2} + 3\mathrm{i}\omega_{\mathrm{n}2}\zeta_2. \tag{K.22}$$

I_{11121}, I_{11122} and I_{11212} have simple poles at:

$$\begin{aligned} \omega_{\mathrm{d}2} + \mathrm{i}(2\omega_{\mathrm{n}1}\zeta_1 + \omega_{\mathrm{n}2}\zeta_2), \quad & -\omega_{\mathrm{d}2} + \mathrm{i}(2\omega_{\mathrm{n}1}\zeta_1 + \omega_{\mathrm{n}2}\zeta_2) \\ 2\omega_{\mathrm{d}1} + \omega_{\mathrm{d}2} + \mathrm{i}(2\omega_{\mathrm{n}1}\zeta_1 + \omega_{\mathrm{n}2}\zeta_2), \quad & -2\omega_{\mathrm{d}1} + \omega_{\mathrm{d}2} + \mathrm{i}(2\omega_{\mathrm{n}1}\zeta_1 + \omega_{\mathrm{n}2}\zeta_2) \\ 2\omega_{\mathrm{d}1} - \omega_{\mathrm{d}2} + \mathrm{i}(2\omega_{\mathrm{n}1}\zeta_1 + \omega_{\mathrm{n}2}\zeta_2), \quad & -2\omega_{\mathrm{d}1} - \omega_{\mathrm{d}2} + \mathrm{i}(2\omega_{\mathrm{n}1}\zeta_1 + \omega_{\mathrm{n}2}\zeta_2) \end{aligned} \tag{K.23}$$

and finally I_{11221}, I_{11222}, I_{12121} and I_{12122} have poles at:

$$\begin{aligned} \omega_{\mathrm{d}1} + \mathrm{i}(2\omega_{\mathrm{n}2}\zeta_2 + \omega_{\mathrm{n}1}\zeta_1), \quad & -\omega_{\mathrm{d}1} + \mathrm{i}(2\omega_{\mathrm{n}2}\zeta_2 + \omega_{\mathrm{n}1}\zeta_1) \\ 2\omega_{\mathrm{d}2} + \omega_{\mathrm{d}1} + \mathrm{i}(2\omega_{\mathrm{n}2}\zeta_2 + \omega_{\mathrm{n}1}\zeta_1), \quad & -2\omega_{\mathrm{d}2} + \omega_{\mathrm{d}1} + \mathrm{i}(2\omega_{\mathrm{n}2}\zeta_2 + \omega_{\mathrm{n}1}\zeta_1) \\ 2\omega_{\mathrm{d}2} - \omega_{\mathrm{d}1} + \mathrm{i}(2\omega_{\mathrm{n}2}\zeta_2 + \omega_{\mathrm{n}1}\zeta_1), \quad & -2\omega_{\mathrm{d}2} - \omega_{\mathrm{d}1} + \mathrm{i}(2\omega_{\mathrm{n}2}\zeta_2 + \omega_{\mathrm{n}1}\zeta_1) \end{aligned} \tag{K.24}$$

and this exhausts all the possibilities.

This calculation motivates the following conjecture. In an MDOF system, the composite FRF from random excitation has poles at all the combination frequencies of the single-mode resonances. This is a pleasing result; there are echoes of the fact that a two-tone periodically excited nonlinear MDOF system has output components at all the combinations of the input frequencies (see chapter 3). A further observation is that all of the poles are in the upper half-plane. This means that the Hilbert transform test will fail to diagnose nonlinearity from the FRF (chapter 5). It was observed in section 8.7.2 that, in the SDOF system, each new order in P produced higher multiplicities for the poles leading to the conjecture that the poles are actually isolated essential singularities. It has not been possible to pursue the calculation here to higher orders. These results do show, however, that the multiplicity of the linear system poles appears to be increasing with the order of P in much the same way as for the SDOF case.

Earlier in this appendix, the case of a Duffing oscillator with an additional quadratic nonlinearity was considered and it was found that poles occurred at even multiples of the fundamental. It is conjectured on the basis of these results that an even nonlinearity in an MDOF system will generate poles at all the even sums and differences. (This is partially supported by the simulation which follows.)

K.2.3 Validation

The validation of these results will be carried out using data from numerical simulation. Consider the linear mass–damper–spring system of figure K.2 which is a simplified version of (K.7) and (K.8). The equations of motion are

$$m\ddot{y}_1 + c\dot{y}_1 + k(2y_1 - y_2) = x_1(t) \tag{K.25}$$

$$m\ddot{y}_2 + c\dot{y}_2 + k(2y_2 - y_1) = x_2(t). \tag{K.26}$$

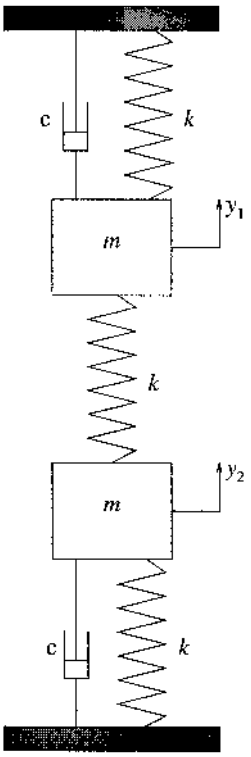


Figure K.2. Basic 2DOF linear system.

The system clearly possesses a certain symmetry. Eigenvalue analysis reveals that the two modes are $(1,1)^{\mathrm{T}}$ and $(1,-1)^{\mathrm{T}}$. Suppose a cubic nonlinearity is added between the two masses, the equations are modified to

$$m\ddot{y}_1 + c\dot{y}_1 + k(2y_1 - y_2) + k_3(y_1 - y_2)^3 = x_1(t) \tag{K.27}$$

$$m\ddot{y}_2 + c\dot{y}_2 + k(2y_2 - y_1) + k_3(y_2 - y_1)^3 = x_2(t) \tag{K.28}$$

and the nonlinearity couples the two equations. In modal space, the situation is a little different. Changing to normal coordinates yields

$$m\ddot{u}_1 + c\dot{u}_1 + ku_1 = \frac{1}{\sqrt{2}}(x_1 + x_2) = p_1 \tag{K.29}$$

$$m\ddot{u}_2 + c\dot{u}_2 + 3ku_2 + 4\sqrt{2}k_3u_2^2 = \frac{1}{\sqrt{2}}(x_1 - x_2) = p_2. \tag{K.30}$$

The system decouples into two SDOF systems, one linear and one nonlinear. This is due to the fact that in the first mode, masses 1 and 2 are moving in phase

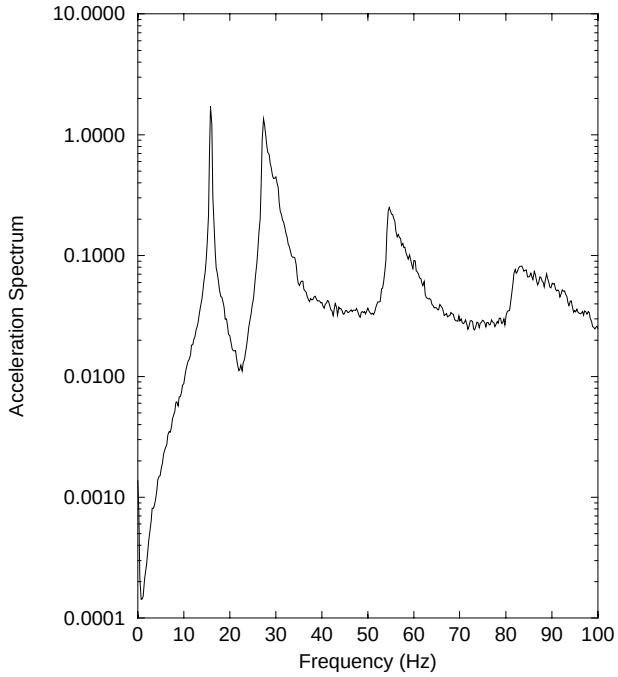


Figure K.3. Spectrum from 2DOF system with nonlinear spring centred.

with constant separation. As a result, the nonlinear spring is never exercised and the mode is linear.

Suppose the nonlinearity were between the first mass and ground. The equations of motion in physical space would then be

$$m\ddot{y}_1 + c\dot{y}_1 + k(2y_1 - y_2) + k_3 y_1^3 = x_1(t) \tag{K.31}$$

$$m\ddot{y}_2 + c\dot{y}_2 + k(2y_2 - y_1) = x_2(t) \tag{K.32}$$

and in modal coordinates would be

$$m\ddot{u}_1 + c\dot{u}_1 + ku_1 + \frac{k_3}{2}(u_1 + u_2)^3 = p_1 \tag{K.33}$$

$$m\ddot{u}_2 + c\dot{u}_2 + 3ku_2 - \frac{k_3}{2}(u_1 + u_2)^3 = p_2 \tag{K.34}$$

and the two modes are coupled by the nonlinearity.

Both nonlinear systems above were simulated using fourth-order Runge-Kutta with a slight modification, a quadratic nonlinearity was added to the cubic of the form $k_2 y_1^2$ or $k_2(y_1 - y_2)^2$. The values of the parameters were $m = 1$, $c = 2$, $k = 10^4$, $k_2 = 10^7$ and $k_3 = 5 \times 10^9$. The excitation x_2 was zero and x_1 initially had rms 2.0, but this was low-pass filtered into the interval 0–100 Hz.

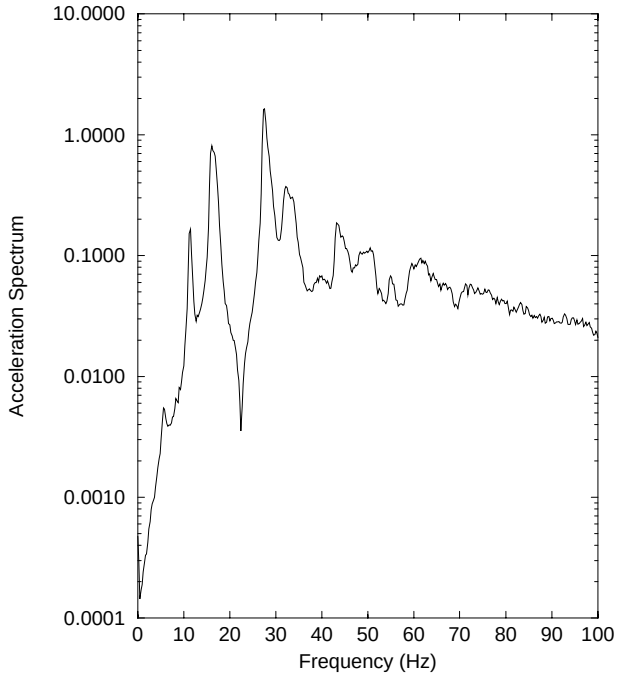


Figure K.4. Spectrum from 2DOF system with nonlinear spring grounded.

With these parameter values the undamped natural frequencies were 15.92 and 27.57 Hz. The sampling frequency was 500 Hz. Using the acceleration response data $\ddot{y}_1$, the output spectra were computed, a 2048-point FFT was used and 100 averages were taken.

Figure K.3 shows the output spectrum for the uncoupled system. As only the second mode is nonlinear, the only additional poles above those for the linear system occur at multiples of the second natural frequency. The presence of the poles is clearly indicated by the peaks in the spectrum at twice and thrice the fundamental.

Figure K.4 shows the output spectrum for the coupled system. Both modes are nonlinear and as in the analysis earlier, poles occur at the sum and differences between the modes. Among the peaks present are: $2f_1 \approx 31.84$ Hz, $2f_2 \approx 55.14$, $f_2 - f_1 \approx 11.65$, $f_2 + f_1 \approx 55.14$, $3f_1 \approx 47.76$, $2f_1 - f_2 \approx 4.27$. The approximate nature of the positions is due to the fact that the peaks move as result of the interactions between the poles as discussed in section 8.7.2.

The conclusions from this section are very simple. The poles for a nonlinear system composite FRF appear to occur at well-defined combinations of the natural frequencies of the underlying nonlinear system. As in the SDOF case, frequency shifts in the FRF peaks at higher excitations can be explained in terms of the

presence of the higher-order poles. Because of the nature of the singularities as previously conjectured, the implications for curve-fitting are not particularly hopeful unless the series solution can be truncated meaningfully at some finite order of P. These results also shed further light on the experimental fact that the Hilbert transform test for nonlinearity fails on FRFs obtained using random excitation.

Bibliography

[1] Abeles M 1991 *Corticonics—Neural Circuits of the Cerebral Cortex* (Cambridge: Cambridge University Press)

[2] Adams D E and Allemang R J 1998 Survey of nonlinear detection and identification techniques for experimental vibrations. *Proc. ISMA 23: Noise and Vibration Engineering* (Leuven: Catholic University) pp 269–81

[3] Adams D E and Allemang R J 1999 A new derivation of the frequency response function matrix for vibrating nonlinear systems *Preprint* Structural Dynamics Research Laboratory, University of Cincinnati

[4] Agneni A and Balis-Crema L 1989 Damping measurements from truncated signals via the Hilbert transform *Mech. Syst. Signal Process.* **3** 1–13

[5] Agneni A and Balis-Crema L A time domain identification approach for low natural frequencies and light damping structures *Preprint* Dipartimento Aerospaziale, Universitá di Roma 'La Sapienza'

[6] Ahlfors L V 1966 *Complex Analysis: Second Edition* (New York: McGraw-Hill)

[7] Ahmed I 1987 Developments in Hilbert transform procedures with applications to linear and non-linear structures *PhD Thesis* Department of Engineering, Victoria University of Manchester.

[8] Al-Hadid M A 1989 Identification of nonlinear dynamic systems using the force-state mapping technique *PhD Thesis* University of London

[9] Al-Hadid M A and Wright J R 1989 Developments in the force-state mapping technique for non-linear systems and the extension to the location of non-linear elements in a lumped-parameter system *Mech. Syst. Signal Process.* **3** 269–90

[10] Al-Hadid M A and Wright J R 1990 Application of the force-state mapping approach to the identification of non-linear systems *Mech. Syst. Signal Process.* **4** 463–82

[11] Al-Hadid M A and Wright J R 1992 Estimation of mass and modal mass in the identification of nonlinear single and multi DOF systems using the force-state mapping approach *Mech. Syst. Signal Process.* **6** 383–401

[12] Arrowsmith D K and Place C M 1990 *An Introduction to Dynamical Systems* (Cambridge: Cambridge University Press)

[13] Arthurs A M 1973 *Probability Theory* (London: Routledge)

[14] Astrom K J 1969 On the choice of sampling rates in parameter identification of time series *Inform. Sci.* **1** 273–87

[15] Atkinson J D 1970 Eigenfunction expansions for randomly excited nonlinear systems *J. Sound Vibration* **30** 153–72
[16] Audenino A, Belingardi G and Garibaldi L 1990 An application of the restoring force mapping method for the diagnostic of vehicular shock absorbers dynamic behaviour *Preprint* Dipartimento di Meccanica del Politecnico di Torino
[17] Barlow R J 1989 *Statistics—A Guide to the Use of Statistical Methods in the Physical Sciences* (Chichester: Wiley)
[18] Baumeister J 1987 *Stable Solution of Inverse Problems (Vieweg Advanced Lectures in Mathematics)* (Braunschweig: Vieweg)
[19] Belingardi G and Campanile P 1990 Improvement of the shock absorber dynamic simulation by the restoring force mapping method *Proc. 15th Int. Seminar in Modal Analysis and Structural Dynamics* (Leuven: Catholic University)
[20] Barrett J F 1963 The use of functionals in the analysis of nonlinear systems *J. Electron. Control* **15** 567–615
[21] Barrett J F 1965 The use of Volterra series to find the region of stability of a Non-linear differential equation *Int. J. Control* **1** 209–16
[22] Bedrosian E and Rice S O 1971 The output properties of Volterra systems driven by harmonic and Gaussian inputs *Proc. IEEE* **59** 1688–707
[23] Bendat J S and Piersol A C 1971 *Random Data: Analysis and Measurement* (New York: Wiley–Interscience)
[24] Bendat J S 1985 *The Hilbert Transform and Applications to Correlation Measurements* (Bruel and Kjaer)
[25] Bendat J S 1990 *Non-Linear System Analysis and Identification* (Chichester: Wiley)
[26] Bendat J S 1998 *Nonlinear Systems Techniques and Applications* (New York: Wiley–Interscience)
[27] Benedettini F, Capecchi D and Vestroni F 1991 Nonparametric models in identification of hysteretic oscillators *Report* DISAT N.4190, Dipartimento di Ingegneria delle Strutture, Universita' dell'Aquila, Italy
[28] Bert C W and Stricklin J D 1988 Comparitive evaluation of six different integration methods for non-linear dynamic systems *J. Sound Vibration* **127** 221–9
[29] Billings S A and Voon W S F 1983 Structure detection and model validity tests in the identification of nonlinear systems *IEE Proc.* **130** 193–9
[30] Billings S A 1985 Parameter estimation *Lecture Notes* Department of Automatic Control and Systems Engineering, University of Sheffield, unpublished
[31] Billings S A and Fadzil M B 1985 The practical identification of systems with nonlinearities *Proc. IFAC Symp. on System Identification and Parameter Estimation (York)*
[32] Billings S A and Tsang K M 1989 Spectral analysis for non-linear systems, part I: parameteric non-linear spectral analysis *Mech. Syst. Signal Process.* **3** 319–39
[33] Billings S A and Chen S 1989 Extended model set, global data and threshold model identification of severely non-linear systems *Int. J. Control* **50** 1897–923
[34] Billings S A, Chen S and Backhouse R J 1989 Identification of linear and nonlinear models of a turbocharged automotive diesel engine *Mech. Syst. Signal Process.* **3** 123–42
[35] Billings S A and Tsang K M 1990 Spectral analysis of block-structured non-linear systems *Mech. Syst. Signal Process.* **4** 117–30

[36] Billings S A, Jamaluddin H B and Chen S 1991 Properties of neural networks with applications to modelling non-linear dynamical systems *Int. J. Control* **55** 193–224

[37] Billings S A, Jamaluddin H B and Chen S 1991 A comparison of the backpropagation and recursive prediction error algorithms for training neural networks *Mech. Syst. Signal Process.* **5** 233–55

[38] Birkhoff G and Maclane S 1977 *A Survey of Modern Algebra* 4th edn (New York: Macmillan)

[39] Bishop J R 1979 Aspects of large scale wave force experiments and some early results from Christchurch Bay *National Maritime Institute Report* no NMI R57

[40] Bishop C M 1996 *Neural Networks for Pattern Recognition* (Oxford: Oxford University Press)

[41] Blaquiere A 1966 *Nonlinear System Analysis* (London: Academic)

[42] Blevins R D 1979 *Formulas for Natural Frequency and Mode Shape* (Krieger)

[43] Bode H W 1945 *Network Analysis and Feedback Amplifier Design* (New York: Van Nostrand Reinhold)

[44] Bouc R 1967 Forced vibration of mechanical system with hysteresis *Proc. 4th Conf. on Nonlinear Oscillation (Prague)*

[45] Boyd S, Tang Y S and Chua L O 1983 Measuring Volterra kernels *IEEE Trans. CAS* **30** 571–7

[46] Box G E P and Jenkins G M 1970 *Time Series Analysis, Forecasting and Control* (San Francisco, CA: Holden-Day)

[47] Broomhead D S and Lowe D 1988 Multivariable functional interpolation and adaptive networks *Complex Systems* **2** 321–55

[48] Brown R D, Wilkinson P, Ismail M and Worden K 1996 Identification of dynamic coefficients in fluid films with reference to journal bearings and annular seals *Proc. Int. Conf. on Identification in Engineering Systems (Swansea)* pp 771–82

[49] *Bruel and Kjaer Technical Review* 1983 System analysis and time delay, Part I and Part II

[50] Cafferty S 1996 Characterisation of automotive shock absorbers using time and frequency domain techniques *PhD Thesis* School of Engineering, University of Manchester

[51] Cafferty S and Tomlinson G R 1997 Characterisation of automotive dampers using higher order frequency response functions *Proc. I. Mech. E., Part D—J. Automobile Eng.* **211** 181–203

[52] Cai G Q and Lin Y K 1995 *Probabilistic Structural Mechanics: Advanced Theory and Applications* (New York: McGraw-Hill)

[53] Cai G Q and Lin Y K 1997 Response Spectral densities of strongly nonlinear systems under random excitation *Probab. Eng. Mech.* **12** 41–7

[54] Caughey T K 1963 Equivalent linearisation techniques *J. Acoust. Soc. Am.* **35** 1706–11

[55] Caughey T K 1971 Nonlinear theory of random vibrations *Adv. Appl. Mech.* **11** 209–53

[56] Chance J E 1996 Structural fault detection employing linear and nonlinear dynamic characteristics *PhD Thesis* School of Engineering, University of Manchester

[57] Chance J E, Worden K and Tomlinson G R 1998 Frequency domain analysis of NARX neural networks *J. Sound Vibration* **213** 915–41

[58] Chen Q and Tomlinson G R 1994 A new type of time series model for the identification of nonlinear dynamical systems *Mech. Syst. Signal Process.* **8** 531–49
[59] Chen S and Billings S A 1989 Representations of non-linear systems: the NARMAX model *Int. J. Control* **49** 1013–32
[60] Chen S, Billings S A and Luo W 1989 Orthogonal least squares methods and their application to non-linear system identification *Int. J. Control* **50** 1873–96
[61] Chen S, Billings S A, Cowan C F N and Grant P M 1990 Practical identification of NARMAX models using radial basis functions *Int. J. Control* **52** 1327–50
[62] Chen S, Billings S A, Cowan C F N and Grant P M 1990 Non-linear systems identification using radial basis functions *Int. J. Syst. Sci.* **21** 2513–39
[63] Christensen G S 1968 On the convergence of Volterra series *IEEE Trans. Automatic Control* **13** 736–7
[64] Chu S R, Shoureshi R and Tenorio M 1990 Neural networks for system identification *IEEE Control Syst. Mag.* **10** 36–43
[65] Cizek V 1970 Discrete Hilbert transform *IEEE Trans. Audio Electron. Acoust.* **AU-18** 340–3
[66] Cooper L and Steinberg D 1970 *Introduction to Methods of Optimisation* (Philadelphia, PA: Saunders)
[67] Cooper J E 1990 Identification of time varying modal parameters *Aeronaut. J.* **94** 271–8
[68] Crandall S H 1963 Perturbation techniques for random vibration of nonlinear systems *J. Acoust. Soc. Am.* **36** 1700–5
[69] Crandall S H The role of damping in vibration theory *J. Sound Vibration* **11** 3–18
[70] Crawley E F and Aubert A C 1986 Identification of nonlinear structural elements by force-state mapping *AIAA J.* **24** 155–62
[71] Crawley E F and O'Donnell K J 1986 Identification of nonlinear system parameters in joints using the force-state mapping technique *AIAA Paper 86-1013* pp 659–67
[72] Cybenko G 1989 Approximations by superpositions of a sigmoidal function *Math. Control, Signals Syst.* **2** 303–14
[73] Davalo E and Naïm P 1991 *Neural Networks* (Macmillan)
[74] Bourcier De Carbon C 1950 Theorie matématique et réalisation practique de la suspension amortie des vehicles terrestres *Atti Congresso SIA, Parigi*
[75] Dienes J K 1961 Some applications of the theory of continuous Markoff processes to random oscillation problems *PhD Thesis* California Institute of Technology
[76] Dinca F and Teodosiu C 1973 *Nonlinear and Random Vibrations* (London: Academic)
[77] Ditchburn 1991 *Light* (New York: Dover)
[78] Donley M G and Spanos P D 1990 *Dynamic Analysis of Non-linear Structures by the Method of Statistical Quadratization (Lecture Notes in Engineering vol 57)* (Berlin: Springer)
[79] Drazin P G 1992 *Nonlinear Systems* (Cambridge: Cambridge University Press)
[80] Duffing G 1918 *Erzwungene Schwingungen bei Veränderlicher Eigenfrequenz (Forced Oscillations in the Presence of Variable Eigenfrequencies)* (Braunschweig: Vieweg)
[81] Dugundji J 1958 Envelopes and pre-envelopes of real waveforms *Trans. IRE* **IT-4** 53–7

[82] Duym S, Stiens R and Reybrouck K 1996 Fast parametric and nonparametric identification of shock absorbers *Proc. 21st Int. Seminar on Modal Analysis (Leuven)* pp 1157–69
[83] Duym S, Schoukens J and Guillaume P 1996 A local restoring force surface *Int. J. Anal. Exp. Modal Anal.* **5**
[84] Duym S and Schoukens J 1996 Selection of an optimal force-state map *Mech. Syst. Signal Process.* **10** 683–95
[85] Eatock Taylor R (ed) 1990 Predictions of loads on floating production systems *Environmental Forces on Offshore Structures and their Prediction* (Dordrecht: Kluwer Academic) pp 323–49
[86] Emmett P R 1994 Methods of analysis for flight flutter data *PhD Thesis* Department of Mechanical Engineering, Victoria University of Manchester
[87] Ewins D J 1984 *Modal Testing: Theory and Practice* (Chichester: Research Studies Press)
[88] Erdelyi A, Magnus W, Oberhettinger F and Tricomi F G 1953 The Bateman manuscript project *Higher Transcendental Functions* vol II (New York: McGraw-Hill)
[89] Erdelyi A, Magnus W, Oberhettinger F and Tricomi F G 1954 *Tables of Integral Transforms* vol II (New York: McGraw-Hill) pp 243–62
[90] Ewen E J and Wiener D D 1980 Identification of weakly nonlinear systems using input and output measurements *IEEE Trans. CAS* **27** 1255–61
[91] Fei B J 1984 Transformées de Hilbert numeriques *Rapport de Stage de Fin d'Etudes, ISMCM (St Ouen, Paris)*
[92] Feldman M 1985 Investigation of the natural vibrations of machine elements using the Hilbert transform *Soviet Machine Sci.* **2** 44–7
[93] Feldman M 1994 Non-linear system vibration analysis using the Hilbert transform—I. Free vibration analysis method 'FREEVIB' *Mech. Syst. Signal Process.* **8** 119–27
[94] Feldman M 1994 Non-linear system vibration analysis using the Hilbert transform—I. Forced vibration analysis method 'FORCEVIB' *Mech. Syst. Signal Process.* **8** 309–18
[95] Feldman M and Braun S 1995 Analysis of typical non-linear vibration systems by use of the Hilbert transform *Proc. 11th Int. Modal Analysis Conf. (Florida)* pp 799–805
[96] Feldman M and Braun S 1995 Processing for instantaneous frequency of 2-component signal: the use of the Hilbert transform *Proc. 12th Int. Modal Analysis Conf.* pp 776–81
[97] Feller W 1968 *An Introduction to Probability Theory and its Applications* vol 1, 3rd edn (New York: Wiley)
[98] Ferry J D 1961 *Viscoelastic Properties of Polymers* (New York: Wiley)
[99] Fletcher R 1987 *Practical Methods of Optimization* 2nd edn (Chichester: Wiley)
[100] Fonseca C M, Mendes E M, Fleming P J and Billings S A 1993 Non-linear model term selection with genetic algorithms *IEEE/IEE Workshop on Natural Algorithms in Signal Processing* pp 27/1–27/8
[101] Forsyth G E, Malcolm M A and Moler C B 1972 *Computer Methods for Mathematical Computations* (Englewood Cliffs, NJ: Prentice-Hall)
[102] Fox L 1964 *An Introduction to Numerical Linear Algebra (Monographs on Numerical Analysis)* (Oxford: Clarendon)

[103] Fox L and Parker I 1968 *Chebyshev Polynomials in Numerical Analysis* (Oxford: Oxford University Press)
[104] Friswell M and Penny J E T 1994 The accuracy of jump frequencies in series solutions of the response of a Duffing oscillator *J. Sound Vibration* **169** 261–9
[105] Frolich 1958 *Theory of Dielectrics* (Oxford: Clarendon)
[106] Funahashi K 1989 On the approximate realization of continuous mappings by neural networks *Neural Networks* **2** 183–92
[107] Gardner M 1966 *New Mathematical Diversions from Scientific American* (Pelican)
[108] Genta G and Campanile P 1989 An approximated approach to the study of motor vehicle suspensions with nonlinear shock absorbers *Meccanica* **24** 47–57
[109] Giacomin J 1991 Neural network simulation of an automotive shock absorber *Eng. Appl. Artificial Intell.* **4** 59–64
[110] Gifford S J 1989 Volterra series analysis of nonlinear structures *PhD Thesis* Department of Mechanical Engineering, Heriot-Watt University
[111] Gifford S J and Tomlinson G R 1989 Recent advances in the application of functional series to non-linear structures *J. Sound Vibration* **135** 289–317
[112] Gifford S J 1990 Detection of nonlinearity *Private Communication*
[113] Gifford S J 1993 Estimation of second and third order frequency response function using truncated models *Mech. Syst. Signal Process.* **7** 145–60
[114] Gill P E, Murray W and Wright M H 1981 *Practical Optimisation* (London: Academic)
[115] Gillespie T D 1992 *Fundamentals of Vehicle Dynamics* (Society of Automotive Engineers)
[116] Gold B, Oppenheim A V and Rader C M 1970 Theory and implementation of the discrete Hilbert transform *Symposium on Computer Processing in Communications* vol 19 (New York: Polytechnic)
[117] Goldberg D E. 1989 *Genetic Algorithms in Search, Machine Learning and Optimisation* (Reading, MA: Addison-Wesley)
[118] Goldhaber M Dispersion relations *Theorie de la Particules Elementaire* (Paris: Hermann)
[119] Goodwin G C and Payne R L 1977 *Dynamic System Identification: Experiment Design and Data Analysis* (London: Academic)
[120] Gottileb O, Feldman M and Yim S C S 1996 Parameter identification of nonlinear ocean mooring systems using the Hilbert transform *J. Offshore Mech. Arctic Eng.* **118** 29–36
[121] Goyder H G D 1976 Structural modelling by the curve fitting of measured response data *Institute of Sound and Vibration Research Technical Report* no 87
[122] Goyder H G D 1984 Some theory and applications of the relationship between the real and imaginary parts of a frequency response function provided by the Hilbert transform *Proc. 2nd Int. Conf. on Recent Advances in Structural Dynamics (Southampton)* (Institute of Sound and Vibration Research) pp 89–97
[123] Gradshteyn I S and Ryzhik I M 1980 *Tables of Integrals, Series and Products* (London: Academic)
[124] Grimmett G R and Stirzaker D R 1992 *Probability and Random Processes* (Oxford: Clarendon)
[125] Guckenheimer J and Holmes P 1983 *Nonlinear Oscillators, Dynamical Systems and Bifurcations of Vector Fields* (Berlin: Springer)

[126] Guillemin E A 1963 *Theory of Linear Physical Systems* (New York: Wiley)
[127] Hagedorn P and Wallaschek J 1987 On equivalent harmonic and stochastic linearisation for nonlinear shock-absorbers *Non-Linear Stochastic Dynamic Engineering Systems* ed F Ziegler and G I Schueller (Berlin: Springer) pp 23–32
[128] Hall B B and Gill K F 1987 Performance evaluation of motor vehicle active suspension systems *Proc. I.Mech.E., Part D: J. Automobile Eng.* **201** 135–48
[129] Hamming R W 1989 *Digital Filters* 3rd edn (Englewood Cliffs, NJ: Prentice-Hall)
[130] Haoui A 1984 Transformées de Hilbert et applications aux systèmes non linéaires *Thèse de Docteur Ingénieur, ISMCM (St Ouen, Paris)*
[131] Hebb D O 1949 *The Organisation of Behaviour* (New York: Wiley)
[132] Holland J H 1975 *Adaption in Natural and Artificial Systems* (Ann Arbor: University of Michigan Press)
[132A] Hopfield J J 1982 Neural networks and physical systems with emergent collective computational facilities *Proc. Natl Acad. Sci.* **79** 2554–8
[133] Hornik K, Stinchcombe M and White H 1990 Universal approximation of an unknown mapping and its derivatives using multilayer feedforward networks *Neural Networks* **3** 551–60
[134] Hunter N, Paez T and Gregory D L 1989 Force-state mapping using experimental data *Proc. 7th Int. Modal Analysis Conf. (Los Angeles, CA)* (Society for Experimental Mechanics) pp 843–69
[135] Inman D J 1994 *Engineering Vibration* (Englewood Cliffs, NJ: Prentice-Hall)
[136] Isidori A 1995 *Nonlinear Control Systems* 3rd edn (Berlin: Springer)
[137] Johnson J P and Scott R A 1979 Extension of eigenfunction-expansion solutions of a Fokker-Planck equation—I. First order system *Int. J. Non-Linear Mech.* **14** 315
[138] Johnson J P and Scott R A 1980 Extension of eigenfunction-expansion solutions of a Fokker–Planck equation—II. Second order system *Int. J. Non-Linear Mech.* **15** 41–56
[139] Kennedy C C and Pancu C D P 1947 Use of vectors and vibration and measurement analysis *J. Aeronaut. Sci.* **14** 603–25
[140] Kennedy J B and Neville A M 1986 *Basic Statistical Methods for Engineers and Scientists* (New York: Harper and Row)
[141] Kim W-J and Park Y-S 1993 Non-linearity identification and quantification using an inverse Fourier transform *Mech. Syst. Signal Process.* **7** 239–55
[142] King N E 1994 Detection of structural nonlinearity using Hilbert transform procedures *PhD Thesis* Department of Engineering, Victoria University of Manchester
[143] King N E and Worden K An expansion technique for calculating Hilbert transforms *Proc. 5th Int. Conf. on Recent Advances in Structural Dynamics (Southampton)* (Institute of Sound and Vibration Research) pp 1056–65
[144] King N E and Worden K 1997 A rational polynomial technique for calculating Hilbert transforms *Proc. 6th Conf. on Recent Advances in Structural Dynamics (Southampton)* (Institute of Sound and Vibration Research)
[145] Kirk N 1985 The modal analysis of nonlinear structures employing the Hilbert transform *PhD Thesis* Department of Engineering, Victoria University of Manchester

[146] Kirkegaard P H 1992 Optimal selection of the sampling interval for estimation of modal parameters by an ARMA-model *Preprint* Department of Building Technology and Structural Engineering, Aalborg University, Denmark

[147] Khabbaz G R 1965 Power spectral density of the response of a non-linear system *J. Acoust. Soc. Am.* **38** 847–50

[148] Klein F 1877 *Lectures on the Icosahedron and the Solution of Equations of the Fifth Degree* (New York: Dover)

[149] Korenberg M J, Billings S A and Liu Y P 1988 An orthogonal parameter estimation algorithm for nonlinear stochastic systems *Int. J. Control* **48** 193–210

[150] Korenberg M J and Hunter I W 1990 The identification of nonlinear biological systems: Wiener kernel approaches *Ann. Biomed. Eng.* **18** 629–54

[151] Koshigoe S and Tubis A 1982 Implications of causality, time-translation invariance, and minimum-phase behaviour for basilar-membrance response *J. Acoust. Soc. Am.* **71** 1194–200

[152] Kozin F and Natke H G 1986 System identification techniques *Structural Safety* **3** 269–316

[153] Kreyszig E 1983 *Advanced Engineering Mathematics* 5th edn (New York: Wiley)

[154] Kronig R de L 1926 On the theory of dispersion of x-rays *J. Opt. Soc. Am.* **12** 547

[155] Krylov N N and Bogoliubov N N 1947 *Introduction to Nonlinear Mechanics* (Princeton: Princeton University Press)

[156] Ku Y H and Wolf A A 1966 Volterra–Wiener functionals for the analysis of nonlinear systems *J. Franklyn Inst.* **281** 9–26

[157] Lang H H 1977 A study of the characteristics of automotive hydraulic dampers at high stroking frequencies *PhD Dissertation* Department of Mechanical Engineering, University of Michigan

[158] Laning J H and Battin R H 1956 *Random Processes in Automatic Control* (New York: McGraw-Hill)

[159] Lawson C L and Hanson R J 1974 *Solving Least Squares Problems (Prentice-Hall Series in Automatic Computation)* (Englewood Cliffs, NJ: Prentice-Hall)

[160] Lawson C L 1977 Software for C^1 surface interpolation *Mathematical Software* vol III (London: Academic)

[161] Leontaritis I J and Billings S A 1985 Input–output parametric models for nonlinear systems, part I: deterministic nonlinear systems *Int. J. Control* **41** 303–28

[162] Leontaritis I J and Billings S A 1985 Input–output parametric models for nonlinear systems, part II: stochastic nonlinear systems *Int. J. Control* **41** 329–44

[163] Leuridan J M 1984 Some direct parameter modal identification methods applicable for multiple input modal analysis *PhD Thesis* Department of Mechanical and Industrial Engineering, University of Cincinnati

[164] Leuridan J M 1986 Time domain parameter identification methods for linear modal analysis: A unifying approach *J. Vibration, Acoust., Stress Reliability Des.* **108** 1–8

[165] Liang Y C and Cooper J 1992 Physical parameter identification of distributed systems *Proc. 10th Int. Modal Analysis Conf. (San Diego, CA)* (Society for Experimental Mechanics) pp 1334–40

[166] Lighthill M J *Fourier Series and Generalised Functions* (Cambridge: Cambridge University Press)

[167] Ljung L 1987 *System Identification: Theory for the User* (Englewood Cliffs, NJ: Prentice-Hall)
[168] Ljung L and Söderstrom T 1983 *Theory and Practice of Recursive Identification* (Cambridge, MA: MIT Press)
[169] Lo H R and Hammond J K 1988 Identification of a class of nonlinear systems *Preprint* Institute of Sound and Vibration Research, Southampton, England
[170] Low H S 1989 Identification of non-linearity in vibration testing *BEng Honours Project* Department of Mechanical Engineering, Heriot-Watt University
[171] Marmarelis P K and Naka K I 1974 Identification of multi-input biological systems *IEEE Trans. Biomed. Eng.* **21** 88–101
[172] Marmarelis V Z and Zhao X 1994 On the relation between Volterra models and feedforward artificial neural networks *Advanced Methods of System Modelling* vol 3 (New York: Plenum) pp 243–59
[173] Manson G 1996 Analysis of nonlinear mechanical systems using the Volterra series *PhD Thesis* School of Engineering, University of Manchester
[174] Masri S F and Caughey T K 1979 A nonparametric identification technique for nonlinear dynamic problems *J. Appl. Mech.* **46** 433–47
[175] Masri S F, Sassi H and Caughey T K 1982 Nonparametric identification of nearly arbitrary nonlinear systems *J. Appl. Mech.* **49** 619–28
[176] Masri S F, Miller R K, Saud A F and Caughey T K 1987 Identification of nonlinear vibrating structures: part I—formalism *J. Appl. Mech.* **54** 918–22
[177] Masri S F, Miller R K, Saud A F and Caughey T K 1987 Identification of nonlinear vibrating structures: part II—applications *J. Appl. Mech.* **54** 923–9
[178] Masri S F, Smyth A and Chassiakos A G 1995 Adaptive identification for the control of systems incorporating hysteretic elements *Proc. Int. Symp. on Microsystems, Intelligent Materials and Robots (Sendai)* pp 419–22
[179] Masri S F, Chassiakos A G and Caughey T K 1993 Identification of nonlinear dynamic systems using neural networks *J. Appl. Mech.* **60** 123–33
[180] Masri S F, Chassiakos A G and Caughey T K 1992 Structure-unknown non-linear dynamic systems: identification through neural networks *Smart Mater. Struct.* **1** 45–56
[181] McCulloch W S and Pitts W 1943 A logical calculus of the ideas imminent in nervous activity *Bull. Math. Biophys.* **5** 115–33
[182] McIain D M 1978 Two-dimensional interpolation from random data *Computer J.* **21** 168
[183] McMillan A J 1997 A non-linear friction model for self-excited oscillations *J. Sound Vibration* **205** 323–35
[184] Miles R N 1989 An approximate solution for the spectral response of Duffing's oscillator with random input *J. Sound Vibration* **132** 43–9
[185] Milne H K The impulse response function of a single degree of freedom system with hysteretic damping *J. Sound Vibration* **100** 590–3
[186] Minsky M L and Papert S A 1988 *Perceptrons (Expanded Edition)* (Cambridge, MA: MIT Press)
[187] Mohammad K S and Tomlinson G R 1989 A simple method of accurately determining the apparent damping in non-linear structures *Proc. 7th Int. Modal Analysis Conf. (Las Vegas)* (Society for Experimental Mechanics)
[188] Mohammad K S 1990 Identification of the characteristics of non-linear structures *PhD Thesis* Department of Mechanical Engineering, Heriot-Watt University

[189] Mohammad K S, Worden K and Tomlinson G R 1991 Direct parameter estimation for linear and nonlinear structures *J. Sound Vibration* **152** 471–99
[190] Moody J and Darken C J 1989 Fast learning in networks of locally-tuned processing units *Neural Comput.* **1** 281–94
[191] Moore B C 1981 Principal component analysis in linear systems: controllability, observability and model reduction *IEEE Trans. Automatic Control* **26** 17–32
[192] Morison J R, O'Brien M P, Johnson J W and Schaf S A 1950 The force exerted by surface waves on piles *Petroleum Trans.* **189** 149–57
[193] Muirhead H *The Physics of Elementary Particles* (Oxford: Pergamon)
[194] Narendra K S and Parthasarathy K 1990 Identification and control of dynamical systems using neural networks *IEEE Trans. Neural Networks* **1** 4–27
[195] Natke H G 1994 The progress of engineering in the field of inverse problems *Inverse Problems in Engineering Mechanics* ed H D Bui *et al* (Rotterdam: Balkema) pp 439–44
[196] Nayfeh A H and Mook D T 1979 *Nonlinear Oscillations* (New York: Wiley–Interscience)
[197] Nayfeh A H 1973 *Perturbation Methods* (New York: Wiley)
[198] Newland D E 1993 *An Introduction to Random Vibrations, Spectral and Wavelet Analysis* (New York: Longman)
[199] Obasaju E D, Bearman P W and Graham J M R 1988 A study of forces, circulation and vortex patterns around a circular cylinder in oscillating flow *J. Fluid Mech.* **196** 467–94
[200] Palm G and Poggio T 1977 The Volterra representation and the Wiener expansion: validity and pitfalls *SIAM J. Appl. Math.* **33** 195–216
[201] Palm G and Pöpel B 1985 Volterra representation and Wiener-like identification of nonlinear systems: scope and limitations *Q. Rev. Biophys.* **18** 135–64
[202] Paris J B 1991 *Machines* unpublished lecture notes, Department of Mathematics, University of Manchester.
[203] Park J and Sandberg I W 1991 Universal approximation using radial basis function networks *Neural Comput.* **3** 246–57
[204] Peters J M H 1995 A beginner's guide to the Hilbert transform *Int. J. Math. Education Sci. Technol.* **1** 89–106
[205] Peyton Jones J C and Billings S A 1989 Recursive algorithm for computing the frequency response of a class of non-linear difference equation models *Int. J. Control* **50** 1925–40
[206] Poggio T and Girosi F 1990 Network for approximation and learning *Proc. IEEE* **78** 1481–97
[207] Porter B 1969 *Synthesis of Dynamical Systems* (Nelson)
[208] Powell M J D 1985 Radial basis functions for multivariable interpolation *Technical Report* DAMPT 1985/NA12, Department of Applied Mathematics and Theoretical Physics, University of Cambridge
[209] Press W H, Flannery B P, Teukolsky S A and Vetterling W T 1986 *Numerical Recipes—The Art of Scientific Computing* (Cambridge: Cambridge University Press)
[210] Rabiner L R and Schafer T W 1974 On the behaviour of minimax FIR digital Hilbert transformers *Bell Syst. J.* **53** 361–88
[211] Rabiner L R and Gold B 1975 *Theory and Applications of Digital Signal Processing* (Englewood Cliffs, NJ: Prentice-Hall)

[212] Rades M 1976 Methods for the analysis of structural frequency-response measurement data *Shock and Vibration Digest* **8** 73–88
[213] Rauch A 1992 Corehence: a powerful estimator of nonlinearity theory and application *Proc. 10th Int. Modal Analysis Conf. (San Diego, CA)* (Society for Experimental Mechanics)
[214] Richards C M and Singh R 1998 Identification of multi-degree-of-freedom nonlinear systems under random excitation by the 'reverse-path' spectral method *J. Sound Vibration* **213** 673–708
[215] Rodeman R 1988 Hilbert transform implications for modal analysis *Proc. 6th Int. Modal Analysis Conf. (Kissimee, FL)* (Society for Experimental Mechanics) pp 37–40
[216] Rosenblatt F 1962 *Principles of Neurodynamics* (New York: Spartan)
[217] Rugh W J 1981 *Nonlinear System Theory: The Volterra/Wiener Approach* (Johns Hopkins University Press)
[218] Rumelhart D E, Hinton G E and Williams R J Learning representations by back propagating errors 1986 *Nature* **323** 533–6
[219] Rumelhart D E and McClelland J L 1988 *Parallel Distributed Processing: Explorations in the Microstructure of Cognition* (two volumes) (Cambridge, MA: MIT Press)
[220] Sauer G 1992 A numerical and experimental investigation into the dynamic response of a uniform beam with a simulated crack *Internal Report* Department of Engineering, University of Manchester
[221] Schetzen M 1980 *The Volterra and Wiener Theories of Nonlinear Systems* (New York: Wiley–Interscience)
[222] Schmidt G and Tondl A 1986 *Non-Linear Vibrations* (Cambridge: Cambridge University Press)
[223] Segel L and Lang H H 1981 The mechanics of automotive hydraulic dampers at high stroking frequency *Proc. 7th IAVSD Symp. on the Dynamics of Vehicles (Cambridge)*
[224] Sharma S 1996 *Applied Multivariate Techniques* (Chichester: Wiley)
[225] Sibson R 1981 A brief description of natural neighbour interpolation *Interpreting Multivariate Data* ed V Barnett (Chichester: Wiley)
[226] Sibson R 1981 *TILE4: A Users Manual* Department of Mathematics and Statistics, University of Bath
[227] Simmons G F 1974 *Differential Equations* (New York: McGraw-Hill)
[228] Simmons G F 1963 *Topology and Modern Analysis* (New York: McGraw-Hill)
[229] Simon M 1983 Developments in the modal analysis of linear and non-linear structures *PhD Thesis* Department of Engineering, Victoria University of Manchester
[230] Simon M and Tomlinson G R 1984 Application of the Hilbert transform in modal analysis of linear and non-linear structures *J. Sound Vibration* **90** 275–82
[231] Söderstrom T and Stoica P 1988 *System Identification* (London: Prentice-Hall)
[232] Sperling L and Wahl F 1996 The frequency response estimation for weakly nonlinear systems *Proc. Int. Conf. on Identification of Engineering Systems (Swansea)* ed J E Mottershead and M I Friswell
[233] Stephenson G 1973 *Mathematical Methods for Science Students* 2nd edn (London: Longman)

[234] Stewart I and Tall D 1983 *Complex Analysis* (Cambridge: Cambridge University Press)

[235] Stinchcombe M and White H 1989 Multilayer feedforward networks are universal approximators *Neural Networks* **2** 359–66

[236] Storer D M 1991 An explanation of the cause of the distortion in the transfer function of the Duffing oscillator subject to sine excitation *Proc. European Conf. on Modal Analysis (Florence)* pp 271–9

[237] Storer D M 1991 Dynamic analysis of non-linear structures using higher-order frequency response functions *PhD Thesis* School of Engineering, University of Manchester

[238] Storer D M and Tomlinson G R 1993 Recent developments in the measurement and interpretation of higher order transfer functions from non-linear structures *Mech. Syst. Signal Process.* **7** 173–89

[239] Surace C, Worden K and Tomlinson G R 1992 On the nonlinear characteristics of automotive shock absorbers *Proc. I.Mech.E., Part D: J. Automobile Eng.*

[240] Surace C, Storer D and Tomlinson G R 1992 Characterising an automotive shock absorber and the dependence on temperature *Proc. 10th Int. Modal Analysis Conf. (San Diego, CA)* (Society for Experimental Mechanics) pp 1317–26

[241] Tan K C, Li Y, Murray-Smith D J and Sharman K C 1995 System identification and linearisation using genetic algorithms with simulated annealing *Genetic Algorithms in Engineering Systems: Innovations and Applications (Sheffield)* pp 164–9

[242] Tanaka M and Bui H D (ed) 1992 *Inverse Problems in Engineering Dynamics* (Berlin: Springer)

[243] Tarassenko L and Roberts S 1994 Supervised and unsupervised learning in radial basis function classifiers *IEE Proc.—Vis. Image Process.* **141** 210–16

[244] Tao Q H 1992 Modelling and prediction of non-linear time-series *PhD Thesis* Department of Automatic Control and Systems Engineering, University of Sheffield

[245] Thrane N 1984 The Hilbert transform *Bruel and Kjaer Technical Review* no 3

[246] Tikhonov A N and Arsenin V Y 1977 *Solution of Ill-Posed Problems* (New York: Wiley)

[247] Titchmarsh E C 1937 *Introduction to the Fourier Integral* (Oxford: Oxford University Press)

[248] Thompson J M T and Stewart H B 1986 *Nonlinear Dynamics and Chaos* (Chichester: Wiley)

[249] Thompson W T 1965 *Mechanical Vibrations with Applications* (George Allen and Unwin)

[250] Tognarelli M A, Zhao J, Baliji Rao K and Kareem A 1997 Equivalent statistical quadratization and cubicization for nonlinear systems *J. Eng. Mech.* **123** 512–23

[251] Tomlinson G R 1979 Forced distortion in resonance testing of structures with electrodynamic shakers *J. Sound Vibration* **63** 337–50

[252] Tomlinson G R and Lam J 1984 Frequency response characteristics of structures with single and multiple clearance-type non-linearity *J. Sound Vibration* **96** 111–25

[253] Tomlinson G R and Storer D M 1994 Reply to a note on higher order transfer functions *Mech. Syst. Signal Process.* **8** 113–16

[254] Tomlinson G R, Manson G and Lee G M 1996 A simple criterion for establishing an upper limit of the harmonic excitation level to the Duffing oscillator using the Volterra series *J. Sound Vibration* **190** 751–62

[255] Tricomi F G 1951 *Q. J. Math.* **2** 199–211

[256] Tsang K M and Billings S A 1992 Reconstruction of linear and non-linear continuous time models from discrete time sampled-data systems *Mech. Syst. Signal Process.* **6** 69–84

[257] Vakakis A F, Manevitch L I, Mikhlin Y V, Pilipchuk V N and Zevin A A 1996 *Normal Modes and Localization in Nonlinear Systems* (New York: Wiley–Interscience)

[258] Vakakis A F 1997 Non-linear normal modes (NNMs) and their applications in vibration theory: an overview *Mech. Syst. Signal Process.* **11** 3–22

[259] Vidyasagar M 1993 *Nonlinear Systems Analysis* 2nd edn (Englewood Cliffs, NJ: Prentice-Hall)

[260] Vihn T, Fei B J and Haoui A 1986 Transformees de Hilbert numeriques rapides *Session de Perfectionnement: Dynamique Non Lineaire des Structures, Institut Superieur des Materiaux et de la Construction Mecanique (Saint Ouen)*

[261] Volterra V 1959 *Theory of Functionals and Integral equations* (New York: Dover)

[262] Wallaschek J 1990 Dynamics of nonlinear automotive shock absorbers *Int. J. Non-Linear Mech.* **25** 299–308

[263] Wen Y K 1976 Method for random vibration of hysteretic systems *J. Eng. Mechanics Division, Proc. Am. Soc. of Civil Engineers* **102** 249–63

[264] Werbos P J 1974 Beyond regression: new tools for prediction and analysis in the behavioural sciences *Doctoral Dissertation* Applied Mathematics, Harvard University

[265] White R G and Pinnington R J 1982 Practical application of the rapid frequency sweep technique for structural frequency response measurement *Aeronaut. J. R. Aeronaut. Soc.* **86** 179–99

[266] Worden K and Tomlinson G R 1988 Identification of linear/nonlinear restoring force surfaces in single- and multi-mode systems *Proc. 3rd Int. Conf. on Recent Advances in Structural Dynamics (Southampton)* (Institute of Sound and Vibration Research) pp 299–308

[267] Worden K 1989 Parametric and nonparametric identification of nonlinearity in structural dynamics *PhD Thesis* Department of Mechanical Engineering, Heriot-Watt University

[268] Worden K and Tomlinson G R 1989 Application of the restoring force method to nonlinear elements *Proc. 7th Int. Modal Analysis Conf. (Las Vegas)* (Society for Experimental Mechanics)

[269] Worden K and Tomlinson G R 1990 The high-frequency behaviour of frequency response functions and its effect on their Hilbert transforms *Proc. 7th Int. Modal Analysis Conf. (Florida)* (Society for Experimental Mechanics)

[270] Worden K, Billings S A, Stansby P K and Tomlinson G R 1990 Parametric modelling of fluid loading forces II *Technical Report to DoE* School of Engineering, University of Manchester

[271A] Worden K 1990 Data processing and experiment design for the restoring force surface method, Part I: integration and differentiation of measured time data *Mech. Syst. Signal Process.* **4** 295–321

[271B] Worden K 1990 Data processing and experiment design for the restoring force surface method, Part II: choice of excitation signal *Mech. Syst. Signal Process.* **4** 321–44

[272] Worden K and Tomlinson G R 1991 An experimental study of a number of nonlinear SDOF systems using the restoring force surface method *Proc. 9th Int. Modal Analysis Conf. (Florence)* (Society for Experimental Mechanics)

[273] Worden K and Tomlinson G R 1991 Restoring force identification of shock absorbers *Technical Report to Centro Ricerche FIAT, Torino, Italy* Department of Mechanical Engineering, University of Manchester

[274] Worden K and Tomlinson G R 1992 Parametric and nonparametric identification of automotive shock absorbers *Proc. of 10th Int. Modal Analysis Conf. (San Diego, CA)* (Society for Experimental Mechanics) pp 764–5

[275] Worden K and Tomlinson G R 1993 Modelling and classification of nonlinear systems using neural networks. Part I: simulation *Mech. Syst. Signal Process.* **8** 319–56

[276] Worden K, Billings S A, Stansby P K and Tomlinson G R 1994 Identification of nonlinear wave forces *J. Fluids Struct.* **8** 18–71

[277] Worden K 1995 On the over-sampling of data for system identification *Mech. Syst. Signal Process.* **9** 287–97

[278] Worden K 1996 On jump frequencies in the response of the Duffing oscillator *J. Sound Vibration* **198** 522–5

[279] Worden K, Manson G and Tomlinson G R 1997 A harmonic probing algorithm for the multi-input Volterra series *J. Sound Vibration* **201** 67–84

[280] Wray J and Green G G R 1994 Calculation of the Volterra kernels of nonlinear dynamic systems using an artificial neural network *Biol. Cybernet.* **71** 187–95

[281] Wright J R and Al-Hadid M A 1991 Sensitivity of the force-state mapping approach to measurement errors *Int. J. Anal. Exp. Modal Anal.* **6** 89–103

[282] Wright M and Hammond J K 1990 The convergence of Volterra series solutions of nonlinear differential equations *Proc. 4th Conf. on Recent Advances in Structural Dynamics* (Institute of Sound and Vibration Research) pp 422–31

[283] Yang Y and Ibrahim S R 1985 A nonparametric identification technique for a variety of discrete nonlinear vibrating systems *Trans. ASME, J. Vibration, Acoust., Stress, Reliability Des.* **107** 60–6

[284] Yar M and Hammond J K 1986 Spectral analysis of a randomly excited Duffing system *Proc. 4th Int. Modal Analysis Conf. (Los Angeles, CA)* (Society for Experimental Mechanics)

[285] Yar M and Hammond J K 1987 Parameter estimation for hysteretic systems *J. Sound Vibration* **117** 161–72

[286] Young P C 1984 *Recursive Estimation and Time-Series Analysis* (Berlin: Springer)

[287] Young P C 1996 Identification, estimation and control of continuous-time and delta operator systems *Proc. Identification in Engineering Systems (Swansea)* pp 1–17

[288] Zarrop M B 1979 *Optimal Experiment Design for Dynamic System Identification (Lecture Notes in Control and Information Sciences vol 21)* (Berlin: Springer)

[289] Ziemer R E and Tranter W H 1976 *Principles of Communications: Systems, Modulation and Noise* (Houghton Mifflin)

Index