DETERMINING MODELS OF STRUCTURES FROM EARTHQUAKE RECORDS

Thesis by

James Leslie Beck

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1979

(Submitted June 26, 1978)

ACKNOWLEDGMENTS

My advisor, Professor P. C. Jennings, offered many valuable suggestions during my research and during the preparation of this thesis. I am particularly appreciative of his readiness to interrupt his work whenever I wanted to discuss my research. Discussions of system identification with Professor T. K. Caughey during coffee break in Room 210 Thomas were always interesting, as were the discussions with fellow graduate student Graeme McVerry. Many others in Thomas Building and elsewhere on campus have also contributed to making my stay at Caltech both pleasant and stimulating.

I wish to acknowledge the Fellowhip I received during my graduate study from the National Research Advisory Council of New Zealand and also the support received from personnel of the New Zealand Department of Scientific and Industrial Research, particularly Dr. M. C. Probine, during my application and tenure of this Fellowship.

Sue Berkley did a great job with the typing of this manuscript, working long hours to finish on time. The help given by Cecilia Lin in drawing the figures is also much appreciated.

I especially want to mention my family who have shared the ups and downs of my studying and research over the last four years. My

children, Diane, Craig and Michael, have provided me with many hours of relaxation at home and on outings, while my wife Silvia has unselfishly supported me for many years as I have pursued the goals of my career. Her understanding and companionship have made these goals that much easier to attain.

ABSTRACT

The problem of determining linear models of structures from seismic response data is studied using ideas from the theory of system identification. The investigation employs a general formulation called the output-error approach, in which optimal estimates of the model parameters are obtained by minimizing a selected measure-of-fit between the responses of the structure and the model. The question of whether the parameters can be determined uniquely and reliably in this way is studied for a general class of linear structural models. Because earthquake records are normally available from only a small number of locations in a structure, and because of measurement noise, it is shown that it is necessary in practice to estimate parameters of the dominant modes in the records, rather than the stiffness and damping matrices.

Two output-error techniques are investigated. Tests of the first, an optimal filter method, show that its advantages are offset by weaknesses which make it unsatisfactory for application to seismic response. A new technique, called the modal minimization method, is developed to overcome these difficulties. It is a reliable and efficient method to determine the optimal estimates of modal parameters for linear structural models.

The modal minimization method is applied to two multi-story buildings that experienced the 1971 San Fernando earthquake. New information is obtained concerning the properties of the higher modes of the taller building and more reliable estimates of the properties of the fundamental modes of both structures are found. The time-varying

character of the equivalent linear parameters is also studied for both buildings. It is shown for the two buildings examined that the optimal, time-invariant, linear models with a small number of modes can reproduce the strong-motion records much better than had been supposed from previous work using less systematic techniques.

TABLE OF CONTENTS

TABLE OF CONTENTS (CONTINUED)

TABLE OF CONTENTS (CONTINUED)

TABLE OF CONTENTS (CONTINUED)

I. INTRODUCTION

1.1 Structural Identification, Introduction and Previous Work

Broadly speaking, system identification is the process of trying to deduce a model of a real system from its output and possibly its input. In this definition, a model is any mathematical representation of the system which allows a good approximation to its output to be computed.* An important aspect of system identification is to allow for the fact that measurements made on the system are inevitably contaminated by noise. Some survey articles on system identification in general are those by Cuenod and Sage (1968), Bekey (1970), Nieman, Fisher and Seborg (1971), Åström and Eykhoff (1971), Bowles and Straeter (1972), and Sage (1972).** The book by Eykhoff (1974) also has an extensive bibliography.

This dissertation is concerned with the application of system identification ideas to structural systems such as buildings, bridges and dams. In this context, the output of the system refers to the histories of response quantities measured at points within the structure. These quantities could be the displacement or its time derivatives, velocity and acceleration, or even the stress or strain. However, it is rare for the latter to be measured in structures and

---

* The terms input and output are used here in a technical sense to describe the observed portions of the excitation and response, respectively. They need not correspond to the complete excitation and response of the system.

** References are given at the end of each chapter.

and the term "response at a point" will be used to refer to the displacement or its derivatives unless otherwise specified. The input to the system refers to the measured portion of the excitation producing the structural response.

Some survey articles on different areas of structural identification are those by Schiff (1972), Collins, Young and Kiefling (1972), Rodeman and Yao (1973) and Hart and Yao (1977). Hudson (1977) has reviewed the means by which structural response data can be produced. We shall first consider two widely-used sources, steady-state harmonic tests and ambient vibrations, and then concentrate on the major concern of this dissertation, which is the use of seismic response records.

1.1.1 Steady-State Harmonic Tests

Steady-state harmonic tests are performed by shaking a structure with special mechanical vibrators which effectively exert a sinusoidal point-force on the structure. In this area, structural identification with linear models has been applied quite extensively. The basic approach, as in most frequency-domain methods, is to estimate the amplitude and phase components of the transfer function, $H(i\omega)$, between the location of the response measurement and the location of the excitation. Because of both the steady-state character and the monochromatic frequency content of the input and output, these functions can be evaluated directly from the amplitude and phase of the response, relative to the exciting force, for each frequency of

excitation.

In practice, the phase information is often ignored and the modal parameters are estimated from resonant peaks of $|\hat{H}(i\omega)|$, the amplitude of the estimated transfer function. The modal frequencies are estimated from the location of these peaks; the modal damping factors are estimated from their half-power bandwidth; and the (unscaled) modeshape values are estimated from their heights. Some consideration of the phase is required to determine the correct sign of each modeshape value. In some cases, the parameter estimates obtained by this resonant-peak technique are strongly affected by modal interference, that is, by the contribution to the response of other modes in the neighborhood of a given modal frequency. This can make the damping estimates particularly unreliable. Hoerner and Jennings (1969) have investigated a particular case of modal interference.

A deficiency of the resonant-peak technique is that only a small number of points of $|\hat{H}(i\omega)|$ are used to estimate the modal parameters, so much of the data is ignored or, at best, used only qualitatively. This makes the estimates sensitive to measurement noise and to model error, where the latter refers to errors arising because the structure is not a time-invariant linear system with uncoupled modes as assumed in the model. Nevertheless, the approach has proved successful with low-amplitude forced vibration tests because the noise levels involved are small for the lower modes of vibration. However, as discussed later, the same technique applied to $|\hat{H}(i\omega)|$

estimated from seismic response records leads to unreliable parameter estimates because the noise levels and model error are much greater.

Ibáñez (1972) has pointed out the above deficiency in the context of steady-state harmonic tests and proposes a technique which uses all the frequency-domain data. This procedure, which he calls YFIT, estimates the parameters by minimizing an output-error functional. It is essentially a frequency-domain version of the general system identification approach adopted in this work.

The above discussion has been concerned with identification using linear models. The identification of structures using nonlinear models and steady-state harmonic data has been investigated by several authors, including Ibáñez (1972), Jennings (1967) and Novak (1971).

1.1.2 Ambient Vibrations

Structural identification has also been carried out by utilizing ultra low-level ambient vibrations induced by wind and microtremors. Techniques for this application generally assume that the system is linear, the excitation is (band-limited) white-noise and that the response is an ergodic random process. The stochastic hypotheses are necessary because the actual excitation, which is spatially-distributed, is not recorded.

By treating each individual modal response as that of a single degree-of-freedom oscillator, it is possible to determine an effective transfer

function using an approach based on the equation:

$$P_o(\omega) = |H(i\omega)|^2 P_i(\omega)$$

where P_o and P_i are the power spectral density of the output and input respectively of a linear system. For ambient vibrations, P_o must be estimated from the recorded response whereas P_i is unknown but, by hypothesis, assumed constant. In practice, because only records of finite length are used, the stochastic hypotheses above have the same effect as making the deterministic assumption that the average over the records of the Fourier amplitude spectrum of each point-excitation does not vary greatly with frequency. The average spectrum of the records can then be used directly as an estimate of $|H(i\omega)|$ in the neighborhood of each modal frequency. This allows the efficient Cooley-Tukey FFT algorithm to be used.

Once $|H(i\omega)|$ has been estimated, the parameters can be determined by the resonant-peak technique discussed in the previous section. Again, difficulties arise because of modal interference and the use of only a few data points, which are accentuated in this application because of the more variable character of the estimated transfer function. In addition, the assumption that the average spectrum of the excitation is approximately constant is often violated. This can be caused by strong wind gusts for example. A further consideration relates to the frequency resolution. In steady-state harmonic tests the frequency resolution depends on the frequency control of the shaker. With modern equipment, a frequency

resolution of 0.01 Hz or less can be achieved. However, when the transfer function is estimated from ambient data the frequency resolution is given by $1/T$ where T is the record length, so that very long records are required to adequately define the resonant peaks of the low modes.

Schiff and his colleagues (1972, 1973) give a discussion of the difficulties which arise when the modal parameters are estimated from an estimate of the transfer function made under the assumption of white-noise excitation. Schiff proposes applying a parametric curve-fitting method to $|\hat{H}(i\omega)|$ which considers all the frequency-domain information in the neighborhood of a modal frequency in order to get more reliable estimates of the corresponding modal parameters. In the second paper, the authors carry out some tests by applying this technique and Vanmarcke's method of moments (1970) to simulated data. They were interested in investigating whether these techniques could successfully estimate the damping from short-duration records so that they could be used with seismic response data. The results indicate that for a single-degree-of-freedom linear oscillator at least ten cycles from a stationary response are required to get reasonable damping estimates from either of the methods mentioned above. Furthermore, nonstationarity of the response has a strong influence on the accuracy of the damping estimates. By way of comparison, one of the time-domain techniques discussed later gives nearly exact results in an analogous situation, even when only half of a cycle of nonstationary response, together with the corresponding

nonstationary excitation, are used in the identification process. This illustrates the importance of using input records if they are available.

The discussion so far has concentrated on frequency-domain identification methods for ambient vibration data and the attendant difficulties. Gersch and his colleagues (1974, 1976) have developed a time-domain technique which is based on an auto-regressive moving-average model of a discrete time-series. This technique appears to be a promising one for ambient vibration applications, particularly since it gives some idea of the accuracy of the computed estimates of the parameters.

1.1.3 Seismic Response Data

It has long been recognized that an earthquake can be viewed as a full-scale, large-amplitude experiment on a structure, and that if the structural motion is recorded, it offers an opportunity to make a quantitative study of the behavior of the structure at dynamic force and deflection levels directly relevant to earthquake-resistant design. However, the time and location of a strong-motion earthquake can not be predicted with confidence so the acquisition of such data requires an extensive deployment of dedicated instrumentation, which must be capable of remaining operational over long periods of time. For these reasons, response data of good quality were not readily available until recently, so there was little motivation to develop systematic techniques for structural identification from earthquake records.

The 1971 San Fernando earthquake in California dramatically changed this situation. Seismic response records from about 50 buildings in the Los Angeles area were obtained (Jennings, 1971; California Institute of Technology, 1971-1974). None of the instrumented buildings was heavily damaged but the peak acceleration response in some buildings approached $\frac{1}{2}$ g and many of the buildings exhibited nonlinear behavior, at least to the extent of lengthening fundamental periods.

To date, the ideas of system identification have not been fully utilized in the interpretation of these records. A common approach has been to compare the recorded response of a building with the response of a synthesized linear model subjected to the recorded base excitation. This comparison has been followed by some trial-and-error adjustment of the model parameters to achieve better visual matching of the theoretical and recorded response. (Wood, 1972; Blume and Associates, 301-443, Gates, 445-574, Martin and Associates, 575-596, in Murphy, 1973). Such an approach can be viewed as a rudimentary scheme for estimating parameters in the time domain. One of the aims of this work is to investigate systematic versions of this procedure which give the best possible response matching in a well-defined sense.

Systematic techniques for structural identification from earthquake records must contend with the transient nature of the excitation and response records. However, in contrast to ambient vibrations excited by wind and microtremors, most of the excitation

can be recorded.

If a building is supported solely by a rigid foundation then the excitation would be completely specified by recording the motion in the six rigid-body degrees of freedom of the base. In the past, only the three translational components at one point on the base have been recorded so that it is difficult to separate the rocking and twisting components of the base motion from the translational components. Nevertheless, in the absence of strong soil-structure interaction, the dominant contribution to the lateral response of the structure will arise from the horizontal motion of the base. It is also often assumed that the building axes define two orthogonal horizontal directions in which the total horizontal response can be decomposed so that the component in each direction is due only to the base motion in that direction. This leads to the commonly assumed planar structural models. One limitation of these models is that they do not treat properly any torsional response of the structure.

Several authors have applied frequency-domain identification to data from the San Fernando earthquake (Hart, 597-607, in Murphy, 1973; Udwadia and Trifunac, 1974; Hart et al, 1975; Hart and Vasudevan, 1975). For a planar linear model the response history y (acceleration, velocity or displacement) at any point is related in the frequency domain to the base acceleration history $\ddot{z}$ by the transformed Duhamel equation:

$$Y(\omega) = H(i\omega)\,\ddot{Z}(\omega) \qquad (1.1.1)$$

where $H(i\omega)$ is the appropriate transfer function. In theory, this relation could be applied to estimate $H(i\omega)$ and then any of the relevant techniques discussed in the previous sections could be applied to estimate the modal parameters, although the half-power band-width method for estimating damping is generally replaced by an approach based on the height of the resonant peak. This approach requires a prior estimate of the corresponding modeshape so that the participation factor can be evaluated.

In practice, difficulties arise because the estimated transfer functions are characterized by extreme variability with numerous peaks which appear to be a function of measurement noise and model error and are not related to resonant peaks. Smoothing of $|Y(\omega)|$ and $|\ddot{Z}(\omega)|$ before taking their quotient, or smoothing of $|\hat{H}(i\omega)|$ after division, can reduce the variability and therefore make the resonant peaks more apparent, but this leads to a loss of information which can result in the damping being overestimated. Generally, past work suggests that the only modal parameters which can be reliably estimated from $|\hat{H}(i\omega)|$ by current techniques are the frequencies of the first few modes and possibly the damping factor of the fundamental translational modes.

A further complication in any frequency-domain approach arises from the typical short duration of earthquake records. This leads to a frequency resolution which is inadequate for long-period structures when the Cooley-Tukey FFT algorithm is used to determine the Fourier spectra of the base motion and structural response.

Spectral ordinates can be calculated at intermediate frequency points either by adding zeros to the digitized time-history data or by evaluating the Fourier transform integral at selected frequencies. This will produce valid estimates of the true Fourier spectrum only if the major portion of the complete excitation and response histories are used in the spectral analysis. This same requirement is also necessary of course for the transformed Duhamel equation (1.1.1) to be a valid approximation, unless it is modified to include nonzero initial and final conditions. Because of these considerations, there are difficulties in any frequency-domain approach which must be overcome if short time segments of the full response are to be used.

Some of the above difficulties can be avoided by using the time-domain version of the nonparametric identification procedure based on equation (1.1.1), that is, the impulse response function $h(t)$ is estimated from the Duhamel or superposition integral equation:

$$y(t) = \int_0^t h(\tau)\ddot{z}(t-\tau)d\tau \qquad (1.1.2)$$

where it is assumed that there is no motion until time $t = 0$. The modal parameters are then estimated from the computed $h(t)$.

Torkamani and Hart (1975) have estimated the impulse response function by discretizing equation (1.1.2), which leads to a set of ill-conditioned linear equations. They apply a smoothing criterion during the estimation of $h(t)$ to help overcome this problem. Udwadia and Marmarelis (1976) have estimated $h(t)$ from equation (1.1.2) by

using a correlation technique based on the assumption that the base motion is white noise. They used the basement and roof records produced in the Millikan Library building at the California Institute of Technology during the San Fernando earthquake. In a companion paper, these authors have also applied the correlation technique to determine the second-order Wiener kernel of the Millikan Library from this earthquake data to attempt to gain some insight into the nonlinear processes which occurred in the building (Marmarelis and Udwadia, 1976). The Wiener kernels give a nonlinear, nonparametric model which is based on a representation of the response as a sum of integral terms that is valid for a general class of nonlinear systems. The first-order kernel is analogous to the impulse response function because the corresponding integral term in the representation of the response has the same form as the right-hand side of equation (1.1.2). One problem in identifying Wiener models from earthquake records is that the excitation is not band-limited white noise and it is difficult to determine the effect of this on the estimated kernels.

In the cited papers by Udwadia and Marmarelis, the authors point out the nature of the compromise that must be made in the selection of the record length to be analyzed. On one hand this should be long so that the statistical variability of the estimates is reduced but on the other hand it should be short enough that the structural properties can be considered stationary. This is a major difficulty for non-parametric identification of structural systems because many cycles of response are required to give reliable estimates.

Iemura and Jennings (1973) have developed a novel nonparametric technique based on the general form of the equation of motion for a single degree-of-freedom oscillator. They used their approach to estimate the global hysteresis loops from the roof response of the Millikan Library during the San Fernando earthquake. The same approach has recently been applied to some seismic response records for an earth dam (Abdel-Ghaffar et al, 1977). The hysteresis loops identified by this technique appear to be contaminated by considerable noise unless the original data are severely band-pass filtered about the fundamental frequency of the structure.

In general, the performance of nonparametric identification methods when applied to earthquake records has not been completely satisfactory. It is felt that these difficulties may stem from the lack of model constraints during identification in the presence of high levels of measurement noise and model error, particularly the latter. For example, in the nonparametric procedures based on equation (1.1.1) or (1.1.2), the only assumptions made about the structural model is that it is linear and time-invariant. Much useful information, such as the fact that the dynamics satisfy Newton's Second Law, is ignored. A parametric model is imposed only after the transfer function $H(i\omega)$ or the impulse response function $h(t)$ is estimated, so the prior information contained in this model is not used in the critical first stage of the identification where it would facilitate the extraction of the signal information from the noise. It would appear to be advantageous to impose the parametric model right from the

start when such a model is available. Prior knowledge can then be utilized more efficiently to reduce statistical variability and hence to enable the structural parameters to be estimated more reliably from short records.

Most of the work in structural identification using parametric models has been based on variations of the response-matching idea mentioned earlier. This is referred to as the output-error approach to parameter estimation and it is the basis of the techniques investigated in this dissertation. The model parameters are estimated by minimizing an integral (continuous data) or sum (discrete data) of the squared response error. Although past work has favored the time domain, linear parametric models can also be determined in the frequency domain by applying the output-error approach using the square of the transformed response error. If the complete records are used, then by Parseval's identity the parameter estimates should be equal to those obtained by minimizing in the time domain.

Many authors have tested identification techniques for parametric models by employing simulated seismic response data. Distefano and Rath (1974) have applied two output-error techniques, one based on an optimal filter and the other on a Gauss-Newton procedure (which is also known as the modified Newton-Raphson method). They use these techniques to estimate the parameters of some single degree-of-freedom nonlinear models from simulated data. The same optimal filter approach is used in this work with

linear models. Beliveau (1975) has also used the Gauss-Newton method to estimate the parameters of a single-mass linear soil-structure system and a single degree-of-freedom nonlinear system on a rigid foundation. Udwadia and Shah (1975) estimated the stiffness distribution of a continuous shear beam. They found it necessary for this continuous case to add derivative terms to the integral squared response error to provide smoothing constraints during minimization, which was done by a mixed gradient technique (steepest descent followed by conjugate gradient near the minimum). Finally, a discrete equation-error technique has been developed by Caravani et al (1977) to estimate the stiffness and damping matrices for a linear chain model. In contrast to output-error techniques, this technique requires a response record for each degree of freedom and so it has limited potential as far as seismic data is concerned.

Several authors have applied time-domain techniques to determine parametric models from both simulated and real data. Raggett (1974) has employed an output-error approach to estimate modal parameters. He uses the simulated seismic response of a three degree-of-freedom linear chain system and response data from a real structure. His technique is described more fully in Chapter 5 because the technique discussed there has several similar features. Distefano and Pena-Pardo (1976) have used the optimal filter technique to estimate the parameters of a linear three degree-of-freedom chain model and the same model with cubic softening added. They tested the algorithm with simulated data and then applied it to records

obtained from a three-story steel-frame structure, which was shaken by simulated earthquakes on the large shaking-table at the Richmond Field Station, University of California, Berkeley. This facility has also been used by Matzen and McNiven (1976, 1977) to generate "seismic" response records for a single-story steel-frame. They then use these recorded data, after some prior testing with simulated data, to estimate the parameters of a single degree-of-freedom model with a Ramberg-Osgood hysteresis law. They employed a Gauss-Newton procedure to minimize the integral squared response error. Finally, Beck and Jennings (1977) tested an optimal filter algorithm on a single degree-of-freedom linear oscillator and then applied this algorithm to short time-segments of the response to investigate the changes in the equivalent linear parameters of the fundamental mode of the Union Bank building during the San Fernando earthquake.

1.2. Outline of This Work

The principal aim of this work was to devise a practical approach which would allow the best estimates of parameters of linear structural models to be determined systematically from records of base motion and response during an earthquake.

Linear models were chosen partly because they are a natural starting point for identification of structures and partly because they are easily formulated. In addition, the identification of time-invariant linear models is of practical importance because these are the

models commonly used in dynamic design. This is either through their use in the response spectrum approach (Hudson, 1956; Housner, 1959), which is based on the modal decomposition of linear structural models, or through the use of synthesized models and particular ground motion records to compute full response histories. One of the aims of this work was to investigate how well time-invariant linear models can reproduce the strong-motion response of a building.

The general features of parametric and nonparametric models for structural identification are discussed in Chapter 2, and it is concluded that the former models are more useful in earthquake engineering. It is noted that empirical parametric models obtained by the identification of existing structures can be used to evaluate the accuracy of techniques for synthesizing models from structural plans. In addition, empirical models can be used to estimate parameters, such as those describing structural damping, which are difficult to determine by synthesis.

These remarks in Chapter 2 are followed by the formulation of what is termed the output-error approach to parameter estimation. This approach is based on the idea of estimating the parameters by calculating those values which optimize the match between the recorded and model responses. It is noted that any technique which implements this formulation will not only provide a means for determining the optimal estimates of the parameters of specified models, but in the case where measurement noise is known to be small, it

will also allow the mathematical form of the model to be evaluated. The remainder of Chapter 2 contains a discussion of the reliability of the estimates of the parameters obtained by an output-error method.

In Chapter 3, the question of identifiability of a general class of linear structural models is examined. This involves an investigation into whether the values of the parameters of the model are specified uniquely by its input and output, which is a necessary condition for uniqueness of the optimal estimates given by an output-error method. An investigation of identifiability is particularly important when the measured output from a system does not correspond to the history of the complete state of the model used in the identification. This is the situation when earthquake records are used in structural identification because on one hand, the response is typically measured at only a small number of locations in the structure, while on the other hand, it is desirable to have a large number of degrees of freedom to model adequately the distribution of stiffness.

Two results of importance are proved in Chapter 3 relating to the identifiability of the class of linear structural models considered. The first shows which parameters are specified uniquely by the input and output of a model. These are the modal periods, damping factors and effective participation factors. The second result shows that to determine the stiffness and damping matrices uniquely within the general class of linear models with N degrees

of freedom, it is necessary to measure the response at no less than $\frac{1}{2}N$ of the degrees of freedom. This assumes that sufficiently good prior information about these parameters is available so that the appropriate values can be chosen from a finite number of possible values. If this is not the case, uniqueness is strictly guaranteed only if the response is measured at every degree of freedom.

The class of models may be further restricted to ensure identifiability. However, it is concluded that even if the models are identifiable, the stiffness and damping matrices generally cannot be estimated reliably in applications because of noise in the records. A practical strategy is then suggested for structural identification using linear models and earthquake records, in which the parameters of the dominant modes are estimated by performing a series of identifications.

An investigation is made of two output-error techniques to estimate modal parameters of linear models from seismic records. The first, described in Chapter 4, is an optimal filter method which was adapted from the literature on state estimation. This technique processes the data sequentially and leads to sequential estimates of the parameters. The second method, described in Chapter 5, is an iterative approach which uses all the data at each iteration. It is referred to as the modal minimization method and it was developed in this work to provide a reliable technique to estimate the modal parameters after certain weaknesses of the optimal filter technique

became apparent when it was applied to seismic records. Both methods were initially tested by using simulated seismic response records.

The modal minimization method was applied to seismic records from two multi-story buildings and the results are reported in Chapter 6. Optimal estimates of the parameters of the first few dominant modes are presented and their reliability is discussed. It is shown that the optimal time-invariant linear models for the buildings can reproduce their strong-motion response remarkably well. In addition, time-varying linear models are used to examine changes in the structural properties of the buildings during their earthquake response.

REFERENCES

Abdel-Ghaffar, A. M., R. F. Scott, G. W. Housner and J. B. Berrill (1977). An investigation of the dynamic characteristics of an earth dam from the analysis of two recorded earthquake motions. Report No. EERL 77-04, Calif. Inst. of Tech., Pasadena, California.

Åström, K. J. and P. Eykhoff (1971). System identification — a survey. Automatica 7, 123-162.

Beck, J. L. and P. C. Jennings (1977). An optimal filter approach to identification in structural dynamics. Proc. Symp. on Applns of Computer Methods in Engineering, 1, 251-260, Univ. of Southern California, Los Angeles, California.

Bekey, G. A. (1970). System identification — an introduction and a survey. Simulation 15, 151-166.

Beliveau, J. G. (1975). Structural identification during an earthquake. Paper No. 12, Proc. 2nd Canadian Conf. on Earthq. Engng, McMaster Univ., Hamilton, Ontario.

Bowles, R. L. and T. A. Straeter (1972). System identification computational considerations, in System Identification of Vibrating Structures, W. D. Pilkey and R. Cohen (eds). ASME, New York.

California Institute of Technology (1971-1974). Analyses of Strong Motion Earthquake Accelerograms, Parts C To S of EERL reports. Index volume, Report No. EERL 76-02, Calif. Inst. of Tech., Pasadena, California.

Caravani, P., M. L. Watson, W. T. Thomson (1977). Recursive least-squares time domain identification of structural parameters. J. Appl. Mech., ASME 44, 135-140.

Collins, J. D., J. P. Young and L. Kiefling (1972). Methods and application of system identification in shock and vibration, in System Identification of Vibrating Structures, W. D. Pilkey and R. Cohen (eds). ASME, New York.

Cuenod, M. and A. P. Sage (1968). Comparison of some methods used for process identification. Automatica 4, 235-269.

REFERENCES

Distefano, N. and B. Pena-Pardo (1976). System identification of frames under seismic loads. J. Engng Mech. Div., ASCE 102, EM2, 313-330.

Distefano, N. and A. Rath (1974). Modeling and Identification in Nonlinear Structural Dynamics. Report No. EERC 74-15, Univ. of Calif., Berkeley, California.

Eykhoff, P. (1974). System Identification. Wiley and Sons, New York.

Gersch, W. and D. A. Foutch (1974). Least squares estimates of structural system parameters using covariance function data. IEEE Trans. Autom. Control, AC-19, 898-903.

Gersch, W., G. T. Taoka and R. Liu (1976). Structural system parameter estimation by two-stage least-squares method. J. Engng Mech. Div., ASCE 102, EM5, 883-899.

Hart, G. C., R. M. DiJulio and M. Lew (1975). Torsional response of high-rise buildings. J. Struc. Div., ASCE 101, ST2, 397-416.

Hart, G. C. and R. Vasudevan (1975). Earthquake design of buildings: damping. J. Struc. Div., ASCE 101, ST1, 11-29.

Hart G. C. and J. T. P. Yao (1976). System identification in structural dynamics. J. Engng Mech. Div., ASCE 103, EM6, 1089-1104.

Hoerner, J. B. and P. C. Jennings (1969). Modal interference in vibration tests. J. Engng. Mech. Div., ASCE 95, EM4, 827-839.

Housner, G. W. (1959). Behavior of structures during earthquakes. J. Engng Mech. Div., ASCE 85, EM4, 109-129.

Hudson, D. E. (1956). Response spectrum techniques in engineering seismology. Proc. 1st Wld. Conf. Earthq. Engng, Paper No. 4-1, Berkeley, California.

Hudson, D. E. (1977). Dynamic tests of full-scale structures. J. Engng Mech. Div., ASCE 103, EM6, 1141-1157.

Ibáñez, P. (1972). Identification of Dynamic Structural Models from Experimental Data. Report No. UCLA-ENG-7225, Univ. of Calif., Los Angeles, California.

REFERENCES

Iemura, H. and P. C. Jennings (1973). Hysteretic Response of a Nine-story Reinforced Concrete Building during the San Fernando Earthquake. Report No. EERL 73-07, Calif. Inst. of Tech., Pasadena, California.

Jennings, P. C. (1967). Force-deflection relations from dynamic tests. J. Engng Mech. Div., ASCE 93, EM2, 115-129.

Jennings, P. C., ed. (1971). Engineering Features of the San Fernando Earthquake, February 9, 1971. Report No. EERL 71-02, Calif. Inst. of Tech., Pasadena, California.

Marmarelis, P. Z. and F. E. Udwadia (1976). The identification of building structural systems. II. The nonlinear case. Bull. Seism. Soc. Am. 66, 153-171.

Matzen, V. C. and H. D. McNiven (1976). Investigation of the Inelastic Characteristics of a Single Story Steel Structure Using System Identification and Shaking Table Experiments. Report No. EERC 76-20, Univ. of Calif., Berkeley, California.

Matzen, V. C. and H. D. McNiven (1977). Formulation of a two-phase mathematical model. Proc. ASCE/EMD Specialty Conf, 423-426, N. Carolina State Univ., Raleigh, North Carolina.

Murphy, L. M., ed. (1973). San Fernando, California, Earthquake of February 9, 1971, 1. U. S. Dept. of Commerce, Washington, D. C.

Nieman, R. E., D. G. Fisher and D. E. Seborg (1971). A review of process identification and parameter estimation techniques. Int. J. Control 13, 209-264.

Novak, M. (1971). Data reduction from nonlinear response curves. J. Engng Mech. Div., ASCE 97, EM4, 1187-1204.

Raggett, J. D. (1974). Time domain analysis of structural motions. Preprint 2209, ASCE Nat. Struct. Engng Meeting, Cincinnati, Ohio.

Rodeman, R. and J. T. P. Yao (1973). Structural Identification — Literature Review. Report No. CE-STR-73-3, Purdue Univ., West Lafayette, Indiana.

REFERENCES

Sage, A. P. (1972). System identification history, methodology, future prospects, in System Identification of Vibrating Structures, W. D. Pilkey and R. Cohen (eds). ASME, New York.

Schiff, A. J. (1972). Identification of large structures using data from ambient and low level excitations, in System Identification of Vibrating Structures, W. D. Pilkey and R. Cohen (eds). ASME, New York.

Schiff, A. J., P. J. Feil and J. L. Bogdanoff (1973). Estimating structural parameters from response data. Proc. 5th Wld Conf. Earthq. Engng, 2558-2567, Rome.

Torkamani, M. A. and G. C. Hart (1975). Building System Identification Using Earthquake Data. Report No. UCLA-ENG-7507, Univ. of Calif., Los Angeles, California.

Udwadia, F. E. and P. Z. Marmarelis (1976). The identification of building structural systems. I. The linear case. Bull. Seism. Soc. Am. 66, 125-151.

Udwadia, F. E. and P. C. Shah (1975). Identification of structures through records obtained during strong earthquake ground motion. J. Engng for Industry, ASME 98, 1347-1362.

Udwadia, F. E. and M. D. Trifunac (1974). Time and amplitude dependent response of structures. Int. J. Earthq. Engng & Struc. Dyn., 2, 359-378.

Vanmarcke, E. H. (1970). Parameters of the Spectral Density Function: Their Significance in the Time and Frequency Domains. Report No. R70-58, Civil Engng, Mass. Inst. of Tech., Cambridge, Massachusetts.

Wood, J. H. (1972). Analysis of the Earthquake Response of a Nine-story Steel Frame Building during the San Fernando Earthquake. Report No. EERL 72-04, Calif. Inst. of Tech., Pasadena, California.

II. IDENTIFICATION USING PARAMETRIC MODELS

Some of the features of two principal categories of models, parametric and nonparametric, are discussed in this Chapter and the advantages derived from using the former in earthquake engineering are given. The output-error approach to parameter estimation is then formulated and several associated problems are examined, including the reliability of the estimates of the parameters. Except for §2.1.1 and parts of §2.1.2, the discussion in this Chapter has general applicability in system identification. In later Chapters, several aspects will be specialized to linear structural models.

2.1. Parametric and Nonparametric Models in System Identification

A model is defined here to be any mathematical representation which approximates the relation between the input and output of a system. The models employed in system identification can be classified into two principal categories:

(a) Parametric models. Here a particular mathematical form is chosen to describe the essential features of the input-output relation of the system under study, but certain parameters must be assigned values before the model is completely specified. Often prior information is available to assist in this step, but in general some of the parameters must be estimated from the input and output of the system. As an example, a single degree-of-freedom model could be represented in the time domain by the differential equation:

$$\ddot{x} + f(t, x, \dot{x}; \underline{a}) = z(t) \qquad (2.1.1)$$

where $x(t)$ is the output of the model; $z(t)$ is the input to the model; the restoring force f is a prescribed function or functional and $\underline{a}$ is a vector of unknown parameters to be estimated. If the model is linear and time-invariant, an equivalent representation in the frequency domain (Fourier-transform space) is:

$$X(\omega) = H(i\omega; \underline{a})Z(\omega) \qquad (2.1.2)$$

where H is a prescribed function of ω containing unknown parameters $\underline{a}$ to be estimated.

(b) Nonparametric models. Here the unknown parts of the model are functions rather than parameters, and so they are like infinite-dimensional "parameters" for identification. The only assumptions that need be made about the system are that it has finite memory and is time-invariant, although linearity is also often assumed. The system is treated as a "black box" since the aim is to determine a functional relationship between the input and output without recourse to any prior information about the internal structure of the system. For example, a time-invariant linear model with a single input and a single output could be characterized by the impulse response function $h(t)$ and the associated input-output relation:

$$x(t) = \int_0^\infty h(\tau)z(t - \tau)d\tau \qquad (2.1.3)$$

The corresponding model in the frequency domain would be given by the transfer function $H(i\omega)$, the Fourier transform of $h(t)$, and the input-output relation:

$$X(\omega) = H(i\omega)Z(\omega) \quad . \qquad (2.1.4)$$

Note that in the nonparametric formulation, $h(t)$ and $H(i\omega)$ are arbitrary functions to be estimated from the input and output, whereas in (a) these functions are of a prescribed form but with unknown parameters. Different identification procedures are therefore required in the two cases.

In view of the preceding discussion of models, system identification can be considered as the process of:

1) specifying the mathematical form (input-output relation) of the model for the system under study,

2) estimating the unknown parameters for a parametric model, or the unknown functions for a nonparametric model, using input and output data from the system,

3) evaluating the capability of the selected model to describe the essential features of the system.

2.1.1. Parametric and Nonparametric Models in Earthquake Engineering

The prime motivation to engage in system identification research in earthquake engineering is to provide the design engineer with more accurate models with which to predict the seismic response

of a proposed structure during its design. Nonparametric models suffer from several disadvantages for this application which stem from the fact that they neglect prior knowledge about the system.

Firstly, the identification of nonparametric models for structural systems is inevitably followed by a parametric interpretation, whenever this is possible. For example, the estimated transfer function $\hat{H}(i\omega)$ completely characterizes a linear nonparametric model for a structure, but without imposing some parametric model it is difficult to give a physical interpretation of the information in this model. This is the reason for the common practice of subsequently estimating the parameters of a linear parametric model from the estimated transfer function. It has already been pointed out in Chapter 1 that if parametric models based on prior information are available, it would be better from the point of view of reducing statistical variability to use these models from the beginning of the identification process. This is particularly the case when linearity is assumed because the parametric form of linear structural models is well known.

Possibly the greatest disadvantage of truly nonparametric models in earthquake engineering is that they are empirical models which cannot be constructed by synthesis. Successful identification from records at a number of points in a structure leads to a relation between the excitation and response at only those points. The behavior at other points in the structure, or the seismic response of different structures, cannot be predicted from a purely nonparametric

model. In particular, these models are not useful in the process of designing for earthquake resistance wherein the seismic behavior of a proposed structure must be predicted.

These disadvantages of nonparametric models can be avoided by using empirical and synthesized parametric models, so that the latter models are more useful in earthquake engineering. In the remainder of this dissertation, the emphasis will therefore shift to parametric models and the adjective "parametric" will often be omitted.

2.1.2. Empirical and Synthesized Parametric Models

To predict realistically the seismic response of a structure during its design, theoretical models are required for which the parameters can be estimated from the properties of the structural subcomponents and their interactions. The resulting parametric models will be called synthesized models to distinguish them from empirical models for which the parameters are estimated from records of the structural response. Synthesized models are sometimes called theoretical models, but in this work a theoretical model will mean a general mathematical form describing the internal structure of a system, without specification of values of the parameters.

To illustrate these definitions, consider the equation

$$M\ddot{\underline{x}} + C\dot{\underline{x}} + K\underline{x} = \underline{f}(t) \qquad (2.1.5)$$

which will be used later as a theoretical structural model. Its

mathematical form is based on Newton's Second Law and the constitutive laws for a linear viscoelastic solid. If the unknown parameters in this theoretical model are estimated by synthesis using the plans of a structure, it becomes a synthesized model, whereas if the unknown parameters are estimated from structural records, it becomes an empirical model. A useful interpretation of a theoretical model is that it is a generic form defining a whole class of models. Each model has the same mathematical form and is given by a particular set of values for the parameters.

Despite recent advances, which include the development of the finite element method and great improvements in computer technology, synthesis of structural models has only met with partial success. One of the reasons for this is that it is extremely difficult to estimate system damping from the damping of each subcomponent. Raggett (1975) has made a contribution in this area. However, even the values of the significant modal periods for linear models are often not predicted well (Wood, 1972; Murphy, 1973), and these are the most important parameters in predicting the seismic response to a given ground-motion history.

The lack of complete success with structural synthesis could be due to a number of factors which include the uncertainties associated with the properties of the structural and nonstructural components, the

simplification necessary to ensure the model is computationally feasible, and the difficulties in selecting a theoretical model which is capable of realistically modelling the physics of the strong-motion response of a structure. As a consequence of these problems, it becomes necessary to complement the a priori knowledge used in a synthesized model by the a posteriori knowledge derived from empirical models.

Empirical models of existing structures have some intrinsic value of their own. However, it is the interplay between synthesized and empirical models based on the same theoretical model which is of greatest value in earthquake engineering. Generally, the empirical model will be a reduced form of the full theoretical model because it must be identifiable from records at only a few positions in the structure (§2.4.1) and because of the limited resolution of the parameters in the presence of noise (§2.4.1). For example, an empirical structural model corresponding to a linear theoretical model should be based on parameters of the dominant modes and not upon the equation of motion in physical coordinates which involves all the parameters of the mass, stiffness and damping matrices (see Chapter 3). Thus, when identification is performed on a structure in the field, the empirical model cannot be expected to give the same level of detail as a synthesized model, but it will impose constraints on that model.

These empirically-determined constraints can be used to modify a synthesized model of the structure to ensure that such a model is consistent with the observed behavior. For a linear model, the modification could be as simple as scaling the synthesized stiffness matrix to match the observed fundamental frequency or it could be the smallest possible change in the elements of the stiffness matrix necessary to give the observed values for all of the related modal quantities.

To lead to improvements in the earthquake-resistant design process, the identification of an empirical model for an existing structure may best be viewed as having three functions:

1) The estimated parameters may be used to evaluate the accuracy of the techniques used to synthesize the parameters for a corresponding theoretical model. For example, for linear models the accuracy of the modal periods and participation factors obtained by synthesis can be determined.

2) For those parameters of a theoretical model which cannot be reliably estimated by synthesis, the corresponding estimated parameters of the empirical model can be used to determine typical values for a given type of structure. For example, modal damping factors determined empirically can be used with linear theoretical models during design.

3) Some evaluation can be made of the mathematical form of

the theoretical model from the degree to which the empirical model matches the structural response. For example, the ability of linear structural models to describe the response of structures to strong ground motion can be examined in this way.

One of the fundamental problems arising during the identification of an empirical model is whether the parameter estimates are reliable. Different aspects of this problem are discussed after the output-error approach to system identification is introduced in the next section.

2.2. Output-error Approach to System Identification

The output-error approach (Bekey, 1970; Bowles and Straeter, 1972) to the estimation of parameters of dynamic models is used in this dissertation. The equation-error approach (Bowles and Straeter, 1972; Distefano and Rath, p. 16 and 51, 1974) was investigated for linear single degree-of-freedom models but its accuracy in several cases was found to be inferior to the output-error approach. Furthermore, it is well known that this approach is not useful for multi-degree-of-freedom models because it either requires measurements at every degree of freedom, or measurement at one degree of freedom of each modal contribution if a modal approach is taken.

The idea behind the output-error method, illustrated in Fig. 2.1,

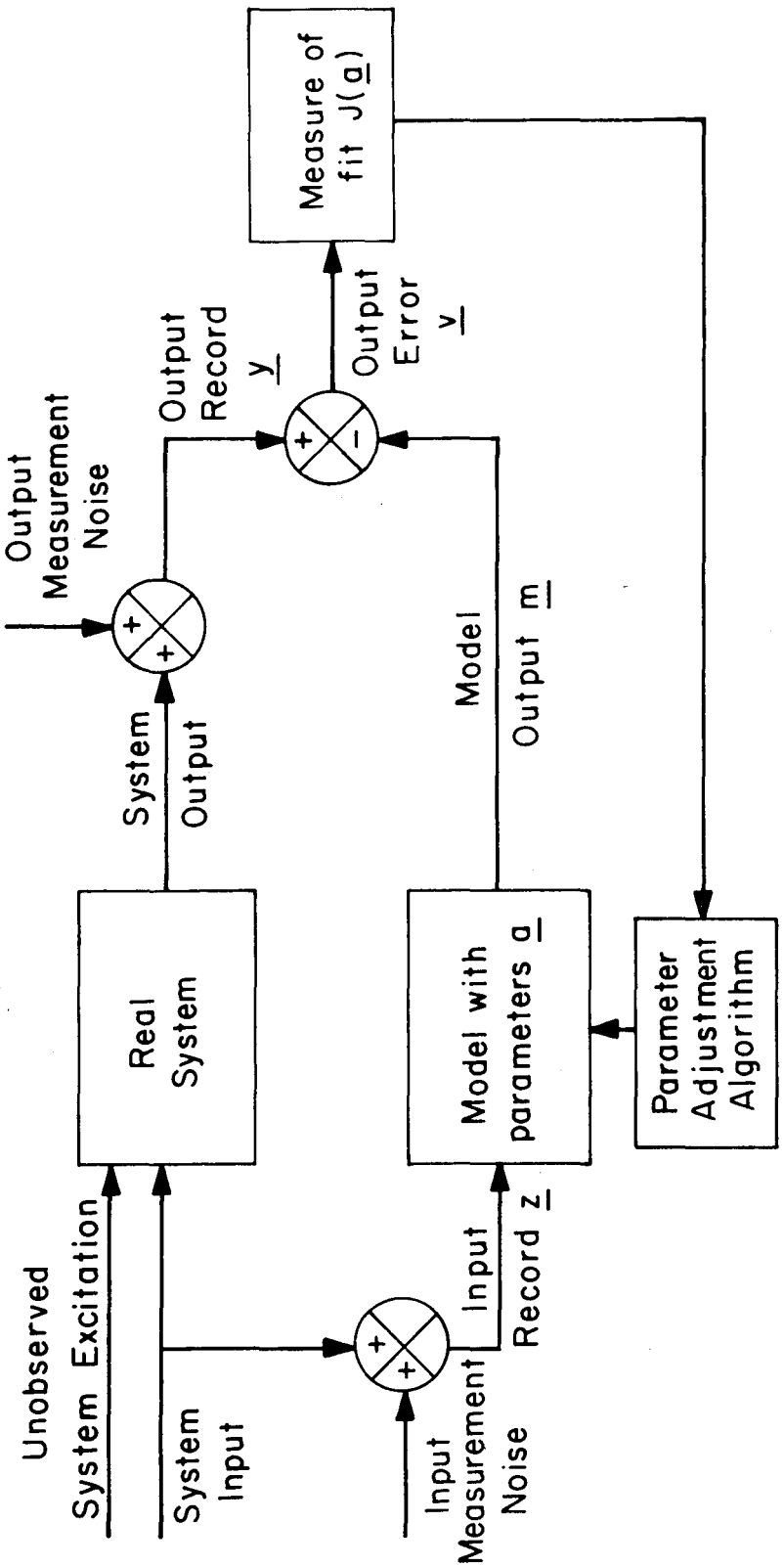


Figure 2.1. Block diagram illustrating the output-error approach

is to estimate the parameters of a model by determining those values which give an optimal match of the output of the model and the output of the real system, when both are subjected to nominally the same input. The quality of the output match is determined by some scalar measure-of-fit, J, which is a positive-definite function of the output-error. Either a continuous form or a discrete form can be chosen for the measure-of-fit. In the applications in this dissertation, continuous records are used which are obtained by linear interpolation between discrete data points, and so an integral mean-square output-error is chosen for J. Finally, the purpose of the parameter-adjustment algorithm, shown in a schematic way in Fig. 2.1, is to select the optimal parameter values by minimizing the measure-of-fit J in a systematic manner. Appropriate algorithms are discussed later in §2.2.4.

It is convenient to formulate the output-error approach in four parts: state equation, output equation, criterion for optimality, and minimization (or parameter-adjustment) algorithm.

2.2.1. State Equation

It is assumed that a theoretical model is available which is spatially discretized, so that its dynamics may be described by a state equation expressed in the general first-order form:

$$\dot{\underline{x}}(t) = \underline{f}(\underline{x}, \underline{z}, t; \underline{\alpha}) \qquad (2.2.1)$$

Here $\underline{x}$ is the state vector of the model. For structural models, it will consist of the generalized displacements and velocities for every degree of freedom of the model. It is not necessary for the state vector to correspond to physical coordinates; for example, it may correspond to modal coordinates if a linear model is used. The vector function or functional $\underline{f}$ describes the mathematical form of the theoretical model and its argument $\underline{z}$ represents the input history to the model. The vector $\underline{\alpha}$ consists of the parameters of the model.

Notice that the history of the state is not uniquely defined by Eq. (2.2.1) unless the initial state, $\underline{x}(T_i)$, is prescribed. However, this is likely to be unknown in many applications. For example, when using seismic records to identify structures, it is generally not possible to take advantage of the fact that the structure starts from rest. The reason for this is that the initial start-up motion is usually lost because a certain threshold motion is required before recording occurs. If the time interval used in the identification is only a portion of the full history of the response, $\underline{x}(T_i)$ is still likely to be unknown because observations of the state will be contaminated by noise. Furthermore, the complete state is typically not observed anyway.

The value chosen for the initial state will influence the estimated values for the model parameters. Both sets of unknown quantities are therefore combined into one vector $\underline{a}$ and all the components are

treated as "model parameters" which are to be estimated from available data, that is,

$$\underline{a} = \begin{bmatrix} \underline{x}(T_i) \\ \underline{\alpha} \end{bmatrix}$$

It should also be noted that what is to be considered as the input $\underline{z}$ is model-dependent and that this model input may not include all the excitation of the real system, as indicated in Fig. 2.1. For example, for planar structural models the seismic input corresponds to one component of the horizontal acceleration at one point on the base of the structure, whereas the real structural motion parallel to a vertical plane may also be caused partly by out-of-plane excitation and rotation of the base.

2.2.2. Output Equation and Output-error

The output equation describes how the output of the model is related to the state of the model. It is sufficient for most purposes to take a linear relation between the model output $\underline{m}$ and the state and its rate of change, so that:

$$\underline{m} = \Gamma_1 \underline{x} + \Gamma_2 \dot{\underline{x}} \qquad (2.2.2)$$

where Γ_1 and Γ_2 are constant rectangular matrices. The elements

of these matrices might be chosen, for example, to be either zero or unity in such a way that they select those components of $\underline{x}$ and $\dot{\underline{x}}$ which contribute to the output.

The output-error $\underline{v}$ is the difference between the output measurements $\underline{y}$ of the real system and the model output $\underline{m}$, that is,

$$\underline{v}(t;\underline{a}) = \underline{y}(t) - \underline{m}(t;\underline{a},\underline{z}) \qquad (2.2.3)$$

where the implicit dependence of the model output on the parameters of the model and the input to the model has been shown. There are two contributions to $\underline{v}$, measurement noise and model error, which are discussed in §2.4.3. Also, the dimension of $\underline{y}$, $\underline{m}$ and $\underline{v}$ will in general be smaller than the dimension of $\underline{x}$ because the number of output records will be less than the desired number of degrees of freedom in the model.

In structural identification, the output vector $\underline{y}$ will be the recorded response (displacement, velocity or acceleration) at various points in the structure. The term $\Gamma_2\dot{\underline{x}}$ is included in (2.2.2) so that it is possible to use acceleration records. Although the acceleration may be integrated to provide displacement and velocity histories, this process accentuates the long-period errors in the digitized data, which in some cases may cause difficulties in the identification. It also lowers the signal-to-noise ratio at high frequencies, which can be an

advantage when determining the properties of the lower modes but not if the higher modes are of interest.

2.2.3. Optimality Criterion

For a given recorded input $\underline{z}$ and recorded output $\underline{y}$ over a time interval $[T_i, T_f]$, the optimal estimates of the parameters are defined to be the values which minimize the measure-of-fit:

$$J(\underline{a}) = \int_{T_i}^{T_f} \|\underline{v}(t;\underline{a})\|^2_{V(t)} \, dt + \|\underline{a} - \underline{\hat{a}}_0\|^2_A \qquad (2.2.4)$$

subject to the constraints of Eqs. (2.2.1), (2.2.2) and (2.2.3). The vector of optimal estimates is denoted by $\underline{\hat{a}}$. It is assumed for the present that $\underline{\hat{a}}$ is defined uniquely by the minimization.

In Eq. (2.2.4), $\underline{\hat{a}}_0$ is an a priori estimate of the parameters, and A and V(t) are prescribed symmetric positive semi-definite and positive definite matrices respectively, which allow weighting of the parameters and output-error based on prior knowledge. Some judgment is required in selecting these quantities. The norms in Eq. (2.2.4) are the weighted Euclidean norms:

$$\|\underline{v}\|^2_{V(t)} = \sum_i \sum_j V_{ij}(t) v_i v_j$$

and (2.2.5)

$$\|\underline{a} - \underline{\hat{a}}_0\|^2_A = \sum_i \sum_j A_{ij}(a_i - \hat{a}_{0,i})(a_j - \hat{a}_{0,j})$$

The weighting matrices are commonly taken diagonal so that, for example, $\|\underline{v}\|^2_{V(t)}$ reduces to $\sum_i V_{ii}(t) v_i^2$.

Instead of viewing the output-error approach as estimating the parameters of a theoretical model, it is often useful to take an alternate point of view: a class of models is defined, then the recorded input and output from the system under study are used to determine the optimal model within the class. The class is defined by the theoretical model chosen to represent the system, together with the output equation. Each model in the class is given by assigning values to the parameters of the theoretical model from within a set of allowable values; the optimal model being given by $\underline{\hat{a}}$.

The optimal model is essentially that model with the smallest weighted integral-squared output-error but with some constraints, governed by the size of the elements of A, which prevent too large a departure from the prior estimates $\underline{\hat{a}}_0$. For example, if A_{ii} is relatively large, a_i will be constrained to remain close to $\hat{a}_{0,i}$ during the minimization of J. It is desirable in many structural

applications to set A equal to zero so that the parameters are not constrained by prior estimates. However, for reasons explained later, this cannot be done with one group of output-error techniques, the filter methods.

It is apparent that in the case $A = 0$ the output-error approach allows the chosen theoretical model to be evaluated, when there is prior information available which indicates that measurement noise has only a small influence on the optimal output-error. In this case, since the mean-square output-error is minimized, if the agreement between the response of the real system and the optimal model is not satisfactory, then the theoretical model must be at fault.

The optimality criterion has been given in a deterministic setting where the presence of noise in the data is acknowledged but no statistical assumptions are made about its form. It is possible to give a stochastic interpretation of the optimality criterion, because the same minimization problem can be derived by assuming the output-error $\underline{v}$ is Gaussian white noise with zero mean and covariance matrix $V^{-1}(t)$. In this case, if $A = 0$, $\hat{\underline{a}}$ is the maximum likelihood estimate of $\underline{a}$. On the other hand, if the parameters are assumed to be Gaussian random variables with mean $\hat{\underline{a}}_0$ and covariance A^{-1}, then $\hat{\underline{a}}$ is the Bayesian maximum probability estimate. These ideas for a discrete measure-of-fit are discussed in Bowles and Straeter (1972), while

Jazwinski (p. 150, 1970) treats both discrete and continuous cases.

2.2.4. Minimization Algorithms

The problem of identifying the optimal model from system data has been reduced to minimizing the function $J(\underline{a})$ in Eq. (2.2.4) where $\underline{v}$ is subjected to the constraints of Eqs. (2.2.3), (2.2.2) and (2.2.1). This minimization could be tackled by directly solving the condition for the stationarity of J with respect to $\underline{a}$:

$$\nabla J \Big|_{\underline{a}=\hat{\underline{a}}} = \underline{0} \qquad (2.2.6)$$

although this usually leads to a set of simultaneous nonlinear algebraic equations in $\underline{a}$ which cannot be solved analytically. The nonlinearity arises because the model response is almost always a nonlinear function of the parameters, even if the model itself is linear in the state and linear in the parameters (Eykhoff, p. 113 and p. 446; 1974). Most techniques actually carry out the minimization by other means although the Gauss-Newton minimization method is equivalent to applying to Eq. (2.2.6) a modification of the classical Newton-Raphson method for finding the zeros of a multi-variable vector function.

Two major groups of methods for determining the minimum of J can be distinguished and these will be considered briefly. A number of authors, including Bekey (1970), Bowles and Straeter (1972) and

Eykhoff (p. 151, 1974), have given a more extensive review of minimization techniques.

(a) Filtering methods: These are based on state estimation theory (assuming no "process" or "plant" noise) and the minimization is achieved in an indirect manner by solving an initial-value problem. Either a deterministic setting (invariant-imbedding filter) or a stochastic setting (extended Kalman filter) can be used but the final equations to be solved are formally equivalent.

A characteristic feature of these methods is that they process the data sequentially and give rise to sequential estimates of both the parameters and the state. One drawback of these methods for parameter estimation is that they give only an approximation to the optimal estimates.

The invariant-imbedding filter is discussed in more detail in Chapter 4.

(b) Descent methods: These are iterative methods which use all the data over a given time segment at each iteration. They may be interpreted geometrically as finding the minimum by a search in the multi-dimensional space represented by the allowable values of the parameter vector $\underline{a}$. An initial estimate $\hat{\underline{a}}_0$ is required to start the algorithm, even if A is zero in (2.2.4).

Some techniques in this category which have been used in

structural identification have been given in §1.1.3 and include the Gauss-Newton method (also called the modified Newton-Raphson method), the method of steepest descent and the conjugate gradient method. The first procedure is a modification of the classical Newton-Raphson method. The Hessian matrix $\frac{1}{2}\nabla\nabla J$ [S in (2.3.8)]is modified by neglecting the term containing the second derivatives of the model response with respect to the parameters (Matzen and McNiven, p. 17, 1976; Distefano and Rath, p. 16, 1974). Bard (1970) has compared several descent methods for their application to parameter estimation.

A new descent method called the modal minimization method is introduced in Chapter 5. This was specifically developed to provide a reliable technique for the identification of linear multi-degree-of-freedom models.

2.3. Some Useful Definitions and Shorthand Notation

It is convenient to introduce the scalar product $<\cdot, \cdot>$ defined on the space of continuous vector functions by:

$$<\underline{b}, \underline{c}> = \int_{T_i}^{T_f} (\underline{b}(t), V(t)\underline{c}(t))dt \qquad (2.3.1)$$

where V is a prescribed continuous matrix function which is symmetric and positive definite and $\underline{b}, \underline{c}$ are any continuous vector functions

defined on the given time interval $[T_i, T_f]$. The notation $(\cdot, \cdot)$ refers to the usual Euclidean (or vector) scalar product, so that:

$$(\underline{b}(t), V(t)\underline{c}(t)) = \sum_i \sum_j V_{ij}(t) b_i(t) c_j(t) \tag{2.3.2}$$

and
$$(\underline{a} - \hat{\underline{a}}_0, A(\underline{a} - \hat{\underline{a}}_0)) = \sum_i \sum_j A_{ij}(a_i - \hat{a}_{0,i})(a_j - \hat{a}_{0,j})$$

$$= \|\underline{a} - \hat{\underline{a}}_0\|_A^2 , \tag{2.3.3}$$

from Eq. (2.2.5). It is easy to show from (2.3.7) that $\langle \cdot, \cdot \rangle$ satisfies the required properties (symmetry, linearity and positive definiteness) to make it a scalar product.

With this shorthand notation, Eqs. (2.2.4) and (2.2.3) may be written as:

$$J(\underline{a}) = \langle \underline{v}, \underline{v} \rangle + (\underline{a} - \hat{\underline{a}}_0, A(\underline{a} - \hat{\underline{a}}_0)) \tag{2.3.4}$$

and
$$\underline{v}(\underline{a}) = \underline{y} - \underline{m}(\underline{a}, \underline{z}) \tag{2.3.5}$$

It is also useful for later work to define:

$$J_0(\underline{a}) = \langle \underline{v}, \underline{v} \rangle . \tag{2.3.6}$$

From the properties of a scalar product, it can be shown that Eqs. (2.3.4) and (2.3.5) imply:

$$[\nabla J(\underline{a})]_k \triangleq \frac{\partial J}{\partial a_k}$$

$$= -2 \langle \underline{v}, \frac{\partial \underline{m}}{\partial a_k} \rangle + 2(\underline{i}_k, A(\underline{a} - \hat{\underline{a}}_0)) \tag{2.3.7}$$

where $\underline{i}_k$ is the unit vector $(\underline{i}_k)_j = \delta_{jk}$, and:

$$S_{jk}(\underline{a}) \triangleq \left[\frac{1}{2}\nabla\nabla J(\underline{a})\right]_{jk} \triangleq \frac{1}{2}\frac{\partial^2 J}{\partial a_j \partial a_k}$$
$$= <\frac{\partial \underline{m}}{\partial a_j}, \frac{\partial \underline{m}}{\partial a_k}> + A_{jk} - <\underline{v}, \frac{\partial^2 \underline{m}}{\partial a_j \partial a_k}> \qquad (2.3.8)$$

The symmetric matrix $\hat{S} = S(\hat{\underline{a}})$ is called the sensitivity matrix. It plays an important role in the application of any output-error approach and it will be discussed in more detail later. It is also convenient to introduce a reduced sensitivity matrix-function $\tilde{S}(\underline{a})$ defined by:

$$\tilde{S}_{jk}(\underline{a}) = <\frac{\partial \underline{m}}{\partial a_j}, \frac{\partial \underline{m}}{\partial a_k}> \qquad (2.3.9)$$

The matrix $\tilde{S}(\underline{a})$ is symmetric and at least positive semi-definite. It can easily be shown that $\tilde{S}$ is positive definite if and only if the $\frac{\partial \underline{m}}{\partial a_j}$ (the sensitivity coefficients) are linearly independent over the time interval $[T_i, T_f]$ (see Appendix A).

Several technical points are to be noted in relation to these definitions. First, the output-error $\underline{v}(\underline{a})$, and hence $J(\underline{a})$ and $S(\underline{a})$, are also functions of the input and output records, $\underline{z}$ and $\underline{y}$ respectively, while $\tilde{S}(\underline{a})$ depends on $\underline{m}(\underline{a}, \underline{z})$ and so it is a function of $\underline{z}$ (but not $\underline{y}$). These arguments have been omitted in the above notation and the time dependence of $\underline{y}$, $\underline{z}$, $\underline{m}$ and $\underline{v}$ is also not denoted explicitly. Second, it is assumed that there is sufficient continuity and differentiability for all the quantities involved to be meaningful and for

the manipulations carried out on these quantities to be valid.

The definitions above are useful for two reasons. The first reason is that they are a convenient shorthand which makes the analysis in subsequent discussions more economical. The second and most important reason is that they give wide generality to the arguments developed in the remainder of this chapter. Thus, the discussion need not be restricted to the particular measure-of-fit J defined in (2.2.4). It will apply to any measure-of-fit which has the form of Eq. (2.3.4), where $<\cdot, \cdot>$ is now to be interpreted as an arbitrary scalar product and the time interval $[T_i, T_f]$ is to be interpreted as the appropriate data interval. The discussion which follows can therefore be applied to measures-of-fit which are integrals (continuous data) or sums (discrete data) in either the time domain or frequency domain. The model does not even have to be dynamic; it could be a "static" model, that is, $\underline{m}(\underline{a}, \underline{z})$ could be simply an algebraic relation between the "output" $\underline{m}$ and "input" $\underline{z}$ involving unknown parameters $\underline{a}$. With a suitable interpretation, the output-error approach and the discussion in the following sections are therefore applicable to structural identification using the data from steady-state harmonic tests.

2.4. Reliability of Optimal Estimates of Parameters

There are a number of questions relating to the reliability of the optimal estimates of the parameters which should be considered when applying an output-error algorithm to a theoretical model of a

system.

One important question is whether the values of the parameters of the theoretical model can be expected to be defined uniquely by the input and output for the system. The first step is to examine whether the corresponding class of models is identifiable. This concept is discussed in §2.4.1 and its relation to the resolution of the internal structure of the system is given.

Identifiability of the models is necessary for meaningful results but it does not ensure that the optimal estimates of the parameters are unique. Conditions for uniqueness are considered in §2.4.2 along with the question of convergence of the algorithm, that is, whether the values of the parameters returned by the algorithm actually give the global minimum of J.

The next question considered is how the accuracy of the optimal estimates is affected by measurement noise. A fundamental difficulty is that there are no true or exact values for the parameters because every theoretical model gives only an approximation to the physical processes occurring in the real system. This problem is considered in §2.4.3 where the concept of an ideal model is introduced to act as a basis for judging the accuracy of the optimal model. In §2.4.4, a deterministic error analysis is carried out to investigate the accuracy of the optimal estimates of the parameters with respect to the ideal values. Only limited results can be obtained unless quantitative assumptions are made about the level of measurement

noise and ideal model error. In §2.4.5, some properties of the sensitivity matrix desirable for good accuracy are discussed and a geometrical interpretation of these properties is mentioned.

In §2.4.6, attention is drawn to the fact that the optimal estimates of the parameters can be expected to change as different portions of the data from a system are used because of limitations of the theoretical model. This is followed by a section containing some final remarks on the problem of assessing the reliability of the parameter estimates.

2.4.1. Identifiability and Resolution:

Let $\mathbb{m}$ denote the class of models corresponding to a theoretical model to be used in the identification of a system. The first question considered in this subsection is whether knowledge of the input and output of any model in $\mathbb{m}$ gives sufficient information to allow the values of the parameters for that model to be determined. To show that this need not be the case, an example is given which is based on some work by Udwadia and Sharma (1978), but given from a slightly different point of view.

Consider a theoretical structural model which is a linear chain model with two degrees of freedom (Fig. 2.2). To begin with, suppose that the output m corresponds to the response of the top mass m_2, so that a class of models is defined by the state equations:

$$\begin{aligned} m_2\ddot{x}_2 + k_2x_2 - k_2x_1 &= -m_2\ddot{z} \\ m_1\ddot{x}_1 + (k_1+k_2)x_1 - k_2x_2 &= -m_1\ddot{z} \end{aligned} \qquad (2.4.1)$$

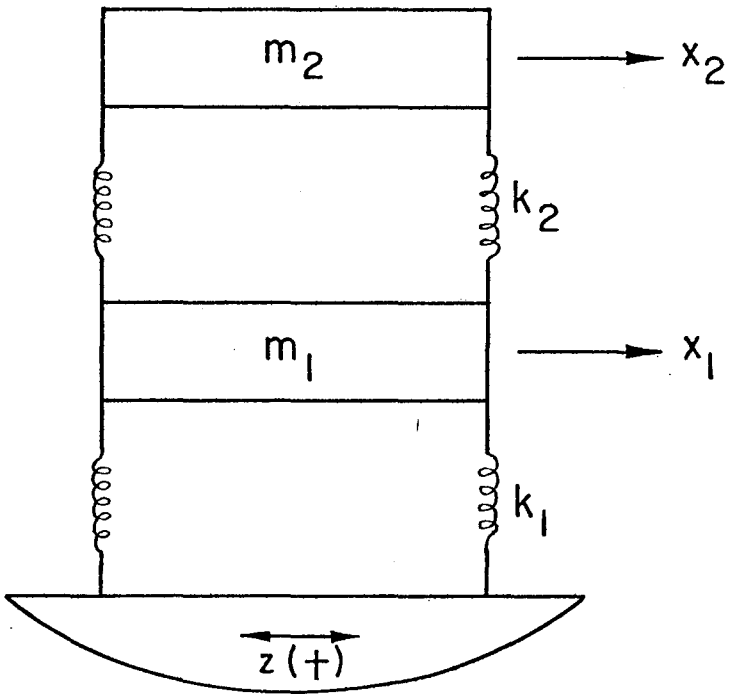


Figure 2.2. Linear chain model with two degrees of freedom

and the output equation:

$$m = x_2 \tag{2.4.2}$$

where the parameters k_1 and k_2 can have only positive values. The masses m_1 and m_2 are assumed to be known and equal, and are denoted by m_0. It is also assumed that the model is initially at rest so that the initial conditions $x_i(0) = 0 = \dot{x}_i(0)$, $i = 1,2$ are known. Thus, the vector of parameters is $\underline{a} = [k_1, k_2]^t$ and the set of allowable values $\mathcal{A} = \{\underline{a}: k_1 > 0, k_2 > 0\}$.

The input-output relation for the models can be given in an explicit form, rather than as a differential equation, by using Duhamel's integral:

$$m(t;\underline{a}, z) = -\sum_{r=1}^{2} \frac{\alpha_r}{\omega_r} \int_0^t \sin \omega_r (t - \tau) \ddot{z}(\tau) d\tau \tag{2.4.3}$$

Here the modal frequencies ω_1 and $\omega_2 (\omega_1 < \omega_2)$ are given by the positive roots of:

$$m_0^2 (\omega^2)^2 - m_0 (k_1 + 2k_2)\omega^2 + k_1 k_2 = 0 \tag{2.4.4}$$

and the modal participation factors α_r, $r = 1,2$, are given by:

$$\alpha_r = \frac{1 + \varphi_1^{(r)}}{1 + [\varphi_1^{(r)}]^2} \tag{2.4.5}$$

where $$\varphi_1^{(r)} = 1 - \frac{m_0}{k_2}\omega_r^2 \qquad (2.4.6)$$

The model output m is controlled by the four <u>derived parameters</u> ω_1, ω_2, α_1 and α_2 which are functions of the model parameters k_1 and k_2. It may be shown by substitution that these derived parameters have the same values for the two models given by:

$$k_1 = k_1^* > 0 \quad , \quad k_2 = k_2^* > 0$$
$$\text{and} \qquad k_1 = 2k_2^* \quad , \quad k_2 = \frac{1}{2}k_1^* \qquad (2.4.7)$$

(The algebra is shortened if the numerator of the difference $\alpha_r(k_1^*, k_2^*) - \alpha_r(2k_2^*, \frac{1}{2}k_1^*)$ is shown to vanish.) Thus, these two models will have the same output regardless of the input z and so the values of the parameters for any model in the prescribed class of models are not specified uniquely by the input and output unless $k_1 = 2k_2$. It should be noted, however, that the response of mass m_1 will be different in the two models.

Suppose the output Eq. (2.4.2) was changed to:

$$m = x_1 \qquad (2.4.8)$$

so that the output now corresponds to the response of the bottom mass m_1. For this new class of models, it turns out that the parameters are defined uniquely by any input-output pair if the input z has finite duration. This is because in this case the four derived parameters

controlling the output, ω_1, ω_2, $\varphi_1^{(1)}\alpha_1$ and $\varphi_1^{(2)}\alpha_2$ are specified uniquely (see Chapter 3), and these in turn specify k_1 and k_2 uniquely.

Definitions of Identifiability

Consider a class of models $\mathbb{m}$ and a class of inputs $\mathbb{C}$, then:

A model in $\mathbb{m}$ is globally identifiable for $\mathbb{C}$ if the values of its parameters are specified uniquely by each input in $\mathbb{C}$ and the corresponding output.

A weaker property can be defined which is implied by the above but which is motivated by the fact that local uniqueness may be all that is required if sufficiently good prior knowledge of the parameters is available:

If the values of the parameters are specified uniquely by each input and output only in some neighborhood of the actual parameter values, the model will be said to be locally identifiable for $\mathbb{C}$.

Prior to identification, it is not known which particular model in a class will be determined by the input and output so it is useful to investigate the identifiability of the whole class:

The class $\mathbb{m}$ is globally (or locally) identifiable if each model in $\mathbb{m}$ is globally (or locally) identifiable. The adjective "globally" will sometimes be omitted.

With this terminology, the first class of models in the example, which used the response of the top mass as output, is not globally identifiable for any input, although it is locally identifiable for any

input of finite duration. The second class of models, which used the response of the bottom mass as output, is globally identifiable for the class of inputs with finite duration.

Other definitions of identifiability have appeared in the literature. It is shown in Appendix A that the definitions used here are equivalent to the concepts of global and local identifiability introduced by Bellman and Åström (1970), except for an essential change in their definition of global identifiability. In addition, the definitions used here have been generalized from the delta-function input used by Bellman and Åström to a prescribed class of inputs, so that the inputs expected in applications can be included. Another definition of identifiability has been given by Beck and Arnold (1977) in their recently published book on parameter estimation. It is shown to be a stronger form of local identifiability in Appendix A.

It is emphasized that identifiability as defined here relates to the unique determination of the parameters of a model from the input and output of the model. An obvious question to be asked is what happens when input and output records from a real system are used to determine the optimal model within an identifiable class of models. The situation is now complicated by noise in the records and the limitations of the class of models in describing the behavior of the system, and minimizing J might not lead to unique optimal estimates of the parameters. However, it is easily seen that global and local identifiability of an optimal model are necessary conditions for global

and local uniqueness, respectively, of the corresponding optimal estimates of the parameters based on minimizing J_0. The difficulty in finding sufficient conditions for uniqueness is that the output-error is unknown prior to identification. A partial result is given in §2.4.2 where it is shown that linear independence of the sensitivity coefficients $\frac{\partial \underline{m}}{\partial a_j}$ is sufficient for local uniqueness of the optimal estimates, provided the optimal output-error is sufficiently small.

When a class of models based on some theoretical model is used in the identification of a system, unique determination of the optimal estimates of the model parameters may be viewed as resolving the internal structure of the system, as it is portrayed by the theoretical model. If too much detail is asked for, the class of models may not be identifiable and the desired resolution will be unattainable.

Even if the desired resolution is attained, some of the model parameters might be estimated inaccurately because of noise. The accuracy of the parameter estimates is governed by the sensitivity of the model output to each parameter and by the characteristics of the noise in the records from the system. In general, as the resolution is increased by refining the models, the optimal model becomes more sensitive to the particular noise content of the records used. A compromise must therefore be made between the amount of resolution asked for and the variance of the optimal estimates of the parameters. This trade-off between resolution and variance is well-known in the literature relating to the geophysical inverse problem (see, for

example, Jackson, 1972).

In Chapter 3, the ideas in this subsection are applied to a class of linear structural models. It is shown that these models are typically not identifiable if the unknown parameters of the theoretical model are the elements of the stiffness and damping matrices, the mass matrix being assumed known. It is also shown that certain parameters of each mode give all the information about the stiffness and damping distributions that is contained in the input and output. Furthermore, as a compromise between the resolution of this information and the accuracy of the estimates, only the parameters of the dominant modes should be estimated.

2.4.2. Convergence and Uniqueness of the Optimal Estimates

Recall that an optimal parameter vector gives the global minimum of $J(\underline{a})$ in Eq. (2.3.4), subject to the constraints of (2.3.5) and the input-output relation (state equation and output equation) for the class of models being used. Two questions which should be considered are whether the minimization algorithm has converged to the global minimum of J and whether this minimum defines unique optimal estimates. Convergence cannot be confirmed simply by examining the output because the effect of lack of convergence on the output-error cannot be distinguished from the effects of measurement noise and model error (§2.4.4). Also, uniqueness is not implied, of course, by the existence of the global minimum.

A technical point requires clarification. Recall that a class

of models is defined by a theoretical model, an output equation and a set G of allowable values for the parameters of the theoretical model. The allowable set G will usually be determined on physical grounds and, for example, might correspond simply to each parameter being positive. The global minimum of J is strictly associated with the region of allowable values in the parameter space defined by G. However, if the parameters are constrained during the minimization, the algorithm must be written to cope with the case where the global minimum may lie on the boundary of the region. Alternatively, one can leave the minimization unconstrained and check the final estimates; this is the approach used in the applications in this dissertation. If the values returned as the optimal estimates by the minimization algorithm lie outside the allowable set, it is clearly indicative of either trouble with the algorithm or inadequacy of the chosen class of models to represent the system.

Let $\hat{\underline{a}}$ be the parameter vector calculated by the minimization algorithm, then for $\hat{\underline{a}}$ to be an optimal parameter vector, it must satisfy successively:

1) $\nabla J(\hat{\underline{a}}) = \underline{0}$ (stationary point)

2)′ $J(\hat{\underline{a}}) \leq J(\underline{a})$ for all $\underline{a}$ in some neighborhood of $\hat{\underline{a}}$ (local minimum)

3)′ $J(\hat{\underline{a}}) \leq J(\underline{a})$ for all $\underline{a}$ (global minimum)

Note that in general there is at least one point in parameter space at which the conditions 1), 2)′ and 3)′ are satisfied because $J(\underline{a})$ is a

continuous function of $\underline{a}$ and it is bounded below by zero. The exceptional case is where the global minimum occurs as $\|\underline{a}\|$ tends to infinity, which should not occur in practice.

To identify a system unambiguously, unique optimal estimates are necessary. This requires a refinement of 2)′ and 3)′ to give:

2) $J(\hat{\underline{a}}) < J(\underline{a})$ for all $\underline{a} \neq \hat{\underline{a}}$ in some neighborhood of $\hat{\underline{a}}$ (strict local minimum)

3) $J(\hat{\underline{a}}) < J(\underline{a})$ for all $\underline{a} \neq \hat{\underline{a}}$ (unique occurrence of global minimum)

The condition 2) excludes the possibility that $\hat{\underline{a}}$ is just one of a continuum of points giving the same minimum value of J, while 3) also excludes the possibility that the global minimum occurs at other local minima. Notice that the results of Appendix A imply that 2) and 3) would be guaranteed if the optimal model were locally identifiable and globally identifiable, respectively, and if the optimal output-error $\hat{\underline{v}}$ were zero. Unfortunately, measurement noise and model error make the latter a most unlikely event.

Ideally, the parameter vector $\hat{\underline{a}}$ calculated by the minimization algorithm should be required to satisfy conditions 1), 2) and 3), each successive condition being more restrictive than its predecessor. Each of these conditions is discussed in turn.

1) Stationary point:

Define the algorithm error by:

$$\underline{e}_J = \frac{1}{2} \nabla J(\hat{\underline{a}}) \qquad (2.4.9)$$

then $\underline{e}_J$ should ideally be zero but in practice it need only be small. Its effect on the accuracy of $\hat{\underline{a}}$ is shown in §2.4.4.

If ∇J is not used explicitly in the algorithm, the algorithm error can be calculated separately by using Eqs. (2.3.7) and (2.3.5), that is:

$$(\underline{e}_J)_k = -<\hat{\underline{v}}, \frac{\partial \hat{\underline{m}}}{\partial a_k}> + (\underline{i}_k, A(\hat{\underline{a}} - \hat{\underline{a}}_0)) \qquad (2.4.10)$$

where $\hat{\underline{v}} = \underline{y} - \underline{m}(\hat{\underline{a}}, \underline{z})$ and $\frac{\partial \hat{\underline{m}}}{\partial a_k} = \frac{\partial \underline{m}}{\partial a_k}\Big|_{\hat{\underline{a}}}$.

A small algorithm error can always be expected because of round-off errors when evaluating the quantities J, ∇J, etc, in the algorithm and because the minimum is always an approximate one. The latter situation arises either because only a finite number of iterations are performed (descent methods) or because of an approximation in the theory (filter methods).

2) Local minimum and local uniqueness

Assume $\underline{e}_J = \underline{0}$, then from a result in advanced calculus, $\hat{\underline{a}}$ gives a strict local minimum of J if the sensitivity matrix $\hat{S} = S(\hat{\underline{a}})$ is positive definite [see, for example, Eq. (2.4.32)]. This is a sufficient condition for a strict local minimum, but not a necessary one. However, it is necessary that $\hat{S}$ be positive semi-definite for a minimum at $\hat{\underline{a}}$. The sensitivity matrix can always be evaluated by substituting $\hat{\underline{a}}$ given by the algorithm into Eq. (2.3.8).

If the sensitivity coefficients $\frac{\partial \hat{\underline{m}}}{\partial a_j}$ are linearly independent and the output-error $\hat{\underline{v}}$ corresponding to $\hat{\underline{a}}$ is sufficiently small, then $\hat{S}$

will be positive definite. This follows from (2.3.8) and (2.3.9) because:

$$S(\underline{a}) = \tilde{S}(\underline{a}) + A - B(\underline{a}) \qquad (2.4.11)$$

where $$B_{jk}(\underline{a}) = \langle \underline{v}, \frac{\partial^2 \underline{m}}{\partial a_j \partial a_k} \rangle \qquad (2.4.12)$$

The weighting matrix A is positive semi-definite by definition and if the $\frac{\partial \underline{m}}{\partial a_j}\Big|_{\underline{a}=\hat{\underline{a}}}$ are linearly independent, $\tilde{S}(\hat{\underline{a}})$ is positive definite (Appendix A). Thus, if $B(\hat{\underline{a}})$ is sufficiently small, $\hat{S} = S(\hat{\underline{a}})$ will be positive definite.

Equation (2.4.11) also suggests that $\hat{S}$ could be made positive definite by a suitable choice of the weighting matrix A, at the risk of possibly biasing the estimates (see §2.4.3 and §2.4.4). Thus, prior knowledge could be used to force the parameter estimates to be locally unique.

3) Global minimum and global uniqueness:

Assuming that conditions 1) and 2) are satisfied, the remaining questions are whether the strict local minimum given by $\hat{\underline{a}}$ is also the global minimum of J and whether it is the only local minimum to give the global minimum. These questions are difficult to answer affirmatively, although if the model corresponding to $\hat{\underline{a}}$ is not globally identifiable, there is at least one other point in parameter space which gives the same minimum of J_0 as $\hat{\underline{a}}$.

The difficulties can be traced back to the nonlinear dependence

of the model output $\underline{m}(\underline{a},\underline{z})$ on the parameters $\underline{a}$ (§2.2.4). If the model output was a linear function of $\underline{a}$, then $B(\underline{a})$ would be identically zero in (2.4.11) and thus $S = \tilde{S} + A$ would be positive definite for all parameters $\underline{a}$, if the class of models were locally identifiable (see Appendix A). In this case, $J(\underline{a})$ would have a unique stationary point and this would give a minimum, so that once this point was located by the minimization algorithm, it would be guaranteed to give the global minimum and to be the only point to do so.

On the other hand, when the model output $\underline{m}(\underline{a},\underline{z})$ is a nonlinear function of $\underline{a}$, $B(\underline{a})$ is no longer identically zero. In general, $S(\underline{a})$ is not positive definite for all $\underline{a}$ and it is possible for J to have more than one local minimum. Thus, the global property of any calculated minimum cannot be ascertained, unless all of the local minima are found, or one of those which are found gives $J(\underline{a}) = 0$, both conditions being unlikely to be satisfied in practice. Similarly, the unique occurrence of the global minimum cannot be ascertained without determining all of the local minima, unless a minimum is found which gives $J(\underline{a}) = 0$ and it is known that the class of models is globally identifiable.

To illustrate the difficulties which can arise, consider an R-mode model of a linear structural system with N degrees of freedom (N>R). The measure-of-fit $J(\underline{a})$ will have a global minimum where the parameters of the R modes of the model are close to the parameters of the first R dominant modes of the linear system. However, $J(\underline{a})$ will also have other local minima where parameters

of the model are close to the parameters of any R modes of the linear system. Since most minimization algorithms will give a local minimum close to the initial estimates of the parameters, if the initial estimates of the modal parameters were not particularly good, the algorithm may converge only to a local minimum and thereby miss some of the important modes.

In conclusion, it is generally not possible to determine whether the global minimum of J has been found and whether it occurs at a unique point, unless an exhaustive search is made through the parameter space. The computation involved in such a search would, in many cases, be prohibitively expensive. It is common practice to be satisfied with finding a strict local minimum near the initial estimate $\hat{\underline{a}}_0$ of the parameter vector $\underline{a}$. This approach is also taken in the applications described later.

2.4.3. Measurement Noise and the Ideal Model

In this subsection, the effect of measurement noise on the accuracy of the optimal estimates of the parameters is considered. The term measurement noise is used to describe all those errors, both systematic and random, which lead to a difference between the history of the actual excitation or response at a point and the processed record of this used in the identification. Measurement noise therefore includes all those errors which arise in measuring, recording and digitizing which are not removed by subsequent data processing (see Fig. 2.1).

It is instructive to start with the hypothetical, ideal situation

where the records are noise-free and the chosen class of models, $\mathbb{m}$, is capable of modelling the behavior of the real system exactly, that is, the system has the same behavior as some model in $\mathbb{m}$ with the parameter values $\underline{a}^*$, say. In this case, the optimal estimates of the parameters, determined by minimizing $J_0(\underline{a})$ defined in Eq. (2.3.6), are equal to the true values of the parameters of the system, and $J_0(\hat{\underline{a}})$ is zero because a perfect output match is achieved. The equality of $\hat{\underline{a}}$ and $\underline{a}^*$ follows from a result given in Appendix A, under the assumption that the class $\mathbb{m}$ is identifiable. On the other hand, if the optimal estimates were determined by minimizing $J(\underline{a})$ instead of $J_0(\underline{a})$, they would generally be biased, because from (2.3.7):

$$\nabla J(\underline{a}^*) = \nabla J_0(\underline{a}^*) + 2A(\underline{a}^* - \hat{\underline{a}}_0) = 2A(\underline{a}^* - \hat{\underline{a}}_0) \tag{2.4.13}$$

whereas $$\nabla J(\hat{\underline{a}}) = \underline{0} \ .$$

A step towards the real situation can be taken by admitting that the theoretical model used will not represent the dynamics of the real system exactly. The expression "the true values of the parameters of the system" is therefore meaningless. However, an <u>ideal model</u> within the class $\mathbb{m}$ can be postulated which is the optimal model using J_0 and the true system input and output, $\underline{z}_0$ and $\underline{y}_s$, which are not affected by measurement noise (Fig. 2.3). The parameter values $\underline{a}^*$ corresponding to the ideal model are called the <u>ideal parameter values</u> and the difference $\underline{e}_M = \underline{y}_s - \underline{m}(\underline{a}^*, \underline{z}_0)$, between the true system output and the output of the ideal model, is called

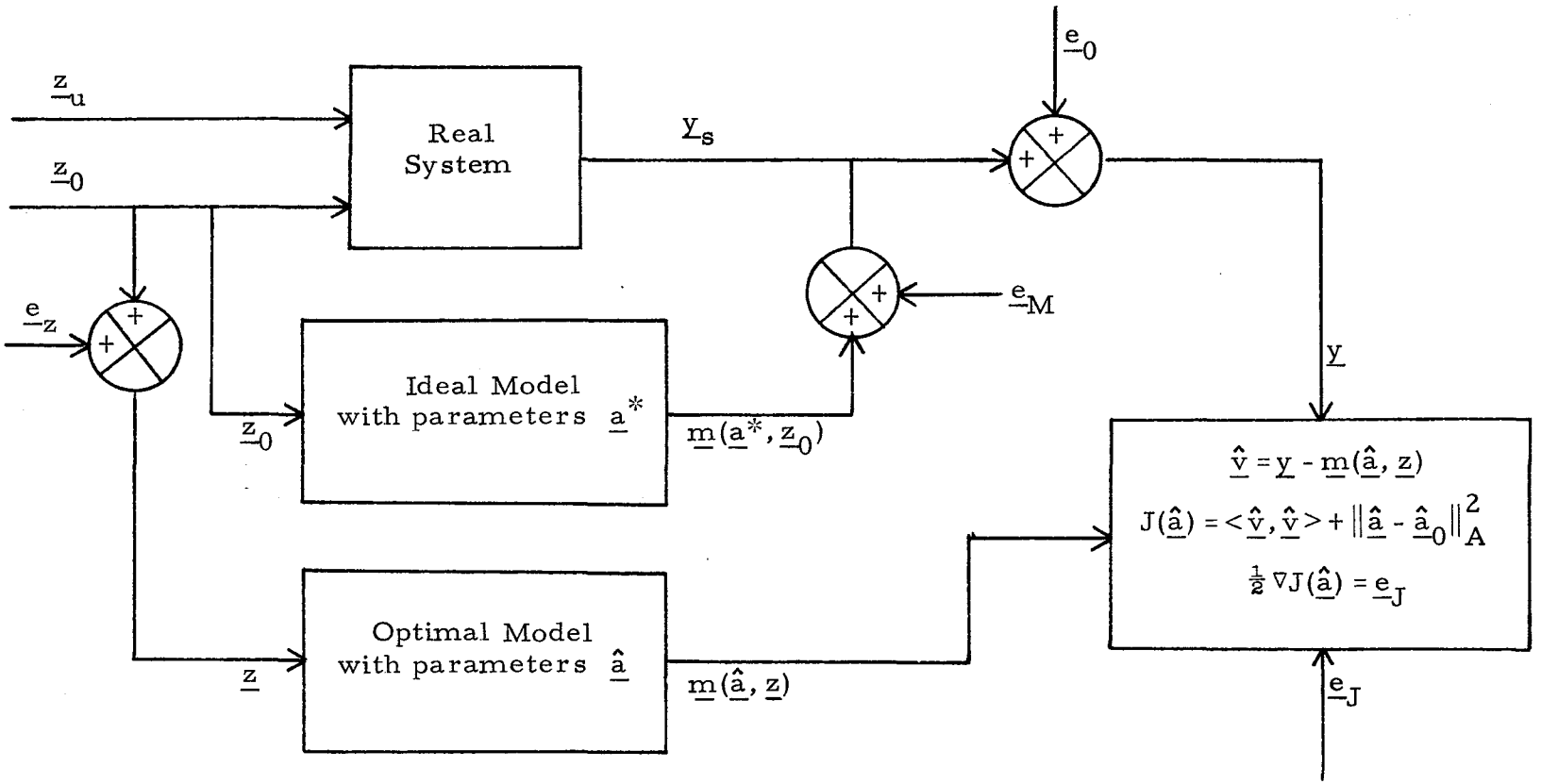


FIGURE 2.3. Block diagram illustrating the output-error approach at the stage when the minimization algorithm has been terminated. The symbols are defined in §2.4.4.

the ideal model-error.

To arrive at the real situation, input and output measurement-noise must also be considered (Fig. 2.3). In general, the optimal parameter values $\underline{\hat{a}}$, determined by minimizing $J_0(\underline{a})$, are now different from the ideal parameter values $\underline{a}^*$. The reason for this difference is that if any portion of the combination of measurement noise plus model error in the output-error of the ideal model can be treated as a possible model signal, the parameters will be changed from their ideal values to cancel this portion. The effect is to reduce the output-error and thus $J_0(\underline{\hat{a}})$ is less than $J_0(\underline{a}^*)$. These ideas are illustrated in more detail in the next section where an error analysis is performed.

In a sense, the ideal model given by $\underline{a}^*$ is the best model within the class of models because its definition is in terms of the true system input and output ($\underline{z}_0$ and $\underline{y}_s$), so it is not influenced by measurement noise. However, it is only a conceptual device because its determination would require complete knowledge of the measurement noise. In practice, one must be content with the optimal model, which is the best model within the given class for the records of the input and output ($\underline{z}$ and $\underline{y}$).

It is the ideal model which indicates how well the theoretical model can approximate the behavior of the real system over a given time segment $[T_i, T_f]$. To be able to judge the theoretical model from the quality of the optimal output-match, one must have confidence that measurement noise has not greatly affected the accuracy of the optimal estimates of the parameters. To gain this confidence, processing

of the data may be necessary to improve the signal-to-noise ratio. Of course, some prior knowledge of the characteristics of the noise and of the system are necessary to distinguish noise from signals in the records.

Considerable work has been done with strong-motion seismic records to identify the various sources of errors contributing to measurement noise, and to develop data-processing techniques to reduce their effects (Trifunac et al, 1970, 1971). It is believed that when these techniques are applied to the records obtained from standard accelerographs, the final processed records are of good quality over most of the frequency range of interest in structural engineering. For structural identification, the corrected data should be adequate over the frequency range from approximately 0.2 Hz to approximately 10 Hz. Difficulties can arise from measurement noise during identification of the fundamental modes of long-period structures (periods of the order of 5 seconds or more), and during identification of modes with short periods (periods of the order of 0.1 second or less). The difficulties in the latter case are partly due to the size of the typical sample interval, 0.02 second.

2.4.4. Deterministic Error Analysis

The following error analysis shows how measurement noise and the algorithm error (§2.4.2) affect the accuracy of the optimal parameter estimates. The ideal parameter values defined in the previous subsection are used to judge the accuracy of the optimal estimates.

The notation used is as follows (Fig. 2.3):

$\underline{a}$ = (vector of) parameter values

$\underline{a}^*$ = ideal parameter values

$\hat{\underline{a}}$ = optimal parameter estimates given by the minimization algorithm

$\hat{\underline{a}}_0$ = *a priori* estimate of $\underline{a}^*$ [Eq. (2.3.4)]

$\underline{e}_I$ = error in response of ideal model due to the transmission of $\underline{e}_z$ through the model

$\underline{e}_J$ = $\frac{1}{2}\nabla J\Big|_{\hat{\underline{a}}}$, the error of the minimization algorithm (§2.4.2)

$\underline{e}_M$ = ideal model error

$\underline{e}_0$ = output measurement-noise

$\underline{e}_z$ = input measurement-noise

$\underline{m}(\underline{a}, \underline{z})$ = output of a model with parameters $\underline{a}$ subjected to an input $\underline{z}$

$\underline{m}(\underline{a}^*, \underline{z}_0)$ = output of ideal model subjected to the true input

$\underline{m}(\hat{\underline{a}}, \underline{z})$ = output of optimal model subjected to the input record

$\underline{v}(\underline{a})$ = $\underline{y} - \underline{m}(\underline{a}, \underline{z})$, the output-error for a model with parameters $\underline{a}$

$\underline{v}^*$ = $\underline{v}(\underline{a}^*)$, the output-error for the ideal model

$\hat{\underline{v}}$ = $\underline{v}(\hat{\underline{a}})$, the output-error for the optimal model

$\underline{y}$ = output record corresponding to $\underline{y}_s$

$\underline{y}_s$ = true system output: response of system at the location of the output transducers

$\underline{z}$ = input record corresponding to $\underline{z}_0$

$\underline{z}_0$ = true system input: excitation of system at the location of the input transducers

$\underline{z}_u$ = excitation of system which affects its response but which is ignored in the models

Consider the output-error of a general model within a prescribed class, then:

$$\begin{aligned}\underline{v}(\underline{a}) &= \underline{y} - \underline{m}(\underline{a}, \underline{z}) \\ &= \underline{e}_0 + \underline{y}_s - \underline{m}(\underline{a}, \underline{z})\end{aligned} \tag{2.4.14}$$

This shows that $\underline{v}(\underline{a})$ consists of the output measurement-noise $\underline{e}_0$ plus the model error $[\underline{y}_s - \underline{m}(\underline{a}, \underline{z})]$ and that the model output $\underline{m}(\underline{a}, \underline{z})$ will attempt to match the combination of output measurement noise and true system output when $J_0 = \langle \underline{v}, \underline{v} \rangle$ is minimized. The output-error of the ideal model:

$$\underline{v}^* = \underline{v}(\underline{a}^*) = \underline{y} - \underline{m}(\underline{a}^*, \underline{z}) \tag{2.4.15}$$

can be expressed in the form:

$$\underline{v}^* = \underline{e}_M + \underline{e}_0 - \underline{e}_I \tag{2.4.16}$$

where $\underline{e}_M = \underline{y}_s - \underline{m}(\underline{a}^*, \underline{z}_0)$ is the ideal model error; $\underline{e}_0 = \underline{y} - \underline{y}_s$ is the output measurement noise; and $\underline{e}_I = \underline{m}(\underline{a}^*, \underline{z}) - \underline{m}(\underline{a}^*, \underline{z}_0)$ is the transmitted input measurement-noise, defined to be the difference between the outputs of the ideal model when it is subjected to the recorded input $\underline{z}$ and the true input $\underline{z}_0$ respectively. If a linear model is used, the output $\underline{m}$ is a linear function of the input, so that $\underline{e}_I = \underline{m}(\underline{a}^*, \underline{e}_z)$. In this case, $\underline{e}_I$ is therefore the response of the ideal

model to the input measurement noise $\underline{e}_z = \underline{z} - \underline{z}_0$.

Subtracting (2.4.14) and (2.4.15) leads to:

$$\underline{v}(\underline{a}) = \underline{v}^* + \underline{m}(\underline{a}^*, \underline{z}) - \underline{m}(\underline{a}, \underline{z}) \tag{2.4.17}$$

This equation indicates that the value of $J_0(\underline{a}) = <\underline{v}(\underline{a}), \underline{v}(\underline{a})>$ might be reduced from its value $J_0(\underline{a}^*) = <\underline{v}^*, \underline{v}^*>$ by changing the parameters from their ideal values so that the difference in the output cancels a portion of the combination of measurement noise and ideal model error. In the presence of measurement noise, $J_0(\hat{\underline{a}})$ would therefore be expected to be less than $J_0(\underline{a}^*)$ and the optimal estimates of the parameters would not be equal to the ideal values. The aim is to derive an "error equation" for the difference $\hat{\underline{a}} - \underline{a}^*$.

To arrive at the error equation, a truncated Taylor series of $\frac{1}{2}\nabla J$ about $\hat{\underline{a}}$ is made:

$$\frac{1}{2}\nabla J(\underline{a}^*) = \frac{1}{2}\nabla J(\hat{\underline{a}}) + \frac{1}{2}\nabla\nabla J(\hat{\underline{a}})(\underline{a}^* - \hat{\underline{a}}) + \underline{0}\left(\left\|\underline{a} - \hat{\underline{a}}\right\|^2\right) \tag{2.4.18}$$

where the last term accounts for the truncation error, with $\left\|\underline{a} - \hat{\underline{a}}\right\|^2 = (\underline{a} - \hat{\underline{a}}, \underline{a} - \hat{\underline{a}})$. If Eqs. (2.3.7), (2.3.8) and (2.4.9) are substituted into (2.4.18), the following equation can be derived:

$$\hat{S}(\hat{\underline{a}} - \underline{a}^*) = \underline{d}^* + \underline{0}\left(\left\|\hat{\underline{a}} - \underline{a}^*\right\|^2\right) \tag{2.4.19}$$

where $$\hat{S}_{kj} = S_{kj}(\hat{\underline{a}}) = <\frac{\partial \hat{\underline{m}}}{\partial a_k}, \frac{\partial \hat{\underline{m}}}{\partial a_j}> + A_{kj} - <\hat{\underline{v}}, \frac{\partial^2 \hat{\underline{m}}}{\partial a_k \partial a_j}> \tag{2.4.20}$$

$$d^*_k = (\underline{e}_j)_k + \langle \underline{v}^*, \frac{\partial \underline{m}^*}{\partial a_k} \rangle - (\underline{i}_k, A(\underline{a}^* - \hat{\underline{a}}_0)) \qquad (2.4.21)$$

and $$\frac{\partial \underline{m}^*}{\partial a_k} = \frac{\partial \underline{m}(\underline{a},\underline{z})}{\partial a_k}\bigg|_{\underline{a}^*}, \frac{\partial \hat{\underline{m}}}{\partial a_k} = \frac{\partial \underline{m}(\underline{a},\underline{z})}{\partial a_k}\bigg|_{\hat{\underline{a}}}, \frac{\partial^2 \hat{\underline{m}}}{\partial a_k \partial a_j} = \frac{\partial^2 \underline{m}(\underline{a},\underline{z})}{\partial a_k \partial a_j}\bigg|_{\hat{\underline{a}}} .$$

These equations show how the accuracy of the optimal estimates $\hat{\underline{a}}$, with respect to the ideal parameter values $\underline{a}^*$, depends on the sensitivity matrix $\hat{S}$, the algorithm error $\underline{e}_J$, the input and output measurement-noise $\underline{e}_I$ and $\underline{e}_0$, the ideal model error $\underline{e}_M$ and the weighting term in J involving the a priori estimates $\hat{\underline{a}}_0$ of the parameters.

If $\underline{e}_J$, $\underline{e}_I$, $\underline{e}_0$ and A are all set to zero, $\hat{\underline{a}} = \underline{a}^*$ should be a solution of (2.4.19) regardless of $\underline{e}_M$, because model error affects the accuracy only in the presence of noise. This is the case because d^*_k immediately reduces to $\langle \underline{e}_M, \frac{\partial \underline{m}^*}{\partial a_k} \rangle = \langle \underline{y}_s - \underline{m}(\underline{a}^*, \underline{z}_0), \frac{\partial \underline{m}(\underline{a}, \underline{z}_0)}{\partial a_k}\big|_{\underline{a}^*} \rangle$, which is zero because, by the definition of $\underline{a}^*$, it minimizes J_0 when $\underline{y}$ and $\underline{z}$ are replaced by $\underline{y}_s$ and $\underline{z}_0$.

The practical use of (2.4.19) is limited because $\underline{d}^*$ involves the unknown ideal parameter values and it is therefore difficult to bound. A bound which is useful to compare the accuracy of each estimate on a relative basis can be derived directly from (2.4.17), which implies:

$$<\hat{\underline{v}}, \frac{\partial \hat{\underline{m}}}{\partial a_k}> = <\underline{v}^*, \frac{\partial \hat{\underline{m}}}{\partial a_k}> + \sum_j <\frac{\partial \hat{\underline{m}}}{\partial a_k}, \frac{\partial \hat{\underline{m}}}{\partial a_j}> (a_j^* - \hat{a}_j)$$

$$+ 0\left(\|\underline{a}^* - \hat{\underline{a}}\|^2\right) \qquad (2.4.22)$$

This can be expressed in the matrix-vector form:

$$\overline{S}(\hat{\underline{a}} - \underline{a}^*) = \underline{d} + \underline{0}\left(\|\hat{\underline{a}} - \underline{a}^*\|^2\right) \qquad (2.4.23)$$

where

$$\overline{S}_{kj} = \widetilde{S}_{kj}(\hat{\underline{a}}) = <\frac{\partial \hat{\underline{m}}}{\partial a_k}, \frac{\partial \hat{\underline{m}}}{\partial a_j}> \qquad (2.4.24)$$

$$d_k = <\underline{v}^* - \hat{\underline{v}}, \frac{\partial \hat{\underline{m}}}{\partial a_k}> \qquad (2.4.25)$$

The matrix $\widetilde{S}(\underline{a})$ was defined in §2.3 and it is positive definite, and hence nonsingular, if the sensitivity coefficients are linearly independent (Appendix A). If it assumed that the errors are small enough for the second-order terms in Eq. (2.4.23) to be neglected, then this equation implies:

$$|\hat{a}_k - a_k^*| \leq \sum_j |(\overline{S}^{-1})_{kj} d_j| \qquad (2.4.26)$$

By the Schwarz inequality for scalar products:

$$|d_j| \leq \left[<\underline{v}^* - \hat{\underline{v}}, \underline{v}^* - \hat{\underline{v}}><\frac{\partial \hat{\underline{m}}}{\partial a_j}, \frac{\partial \hat{\underline{m}}}{\partial a_j}>\right]^{\frac{1}{2}} \qquad (2.4.27)$$

This gives the following bound:

$$|\hat{a}_k - a_k^*| \leq \sigma \sum_j |(\bar{S}^{-1})_{kj}| \, \bar{S}_{jj}^{\frac{1}{2}} \tag{2.4.28}$$

where
$$\sigma^2 = \langle \underline{v}^* - \hat{\underline{v}}, \, \underline{v}^* - \hat{\underline{v}} \rangle \tag{2.4.29}$$

The quantity σ is unknown because $\underline{v}^*$ is unknown and it does not appear likely that σ can be bounded by known quantities without making some assumptions about the level of measurement noise and model error. However, (2.4.28) does indicate relative bounds on the errors in the optimal estimates because the matrix $\bar{S}$ can always be calculated from (2.4.24).

If it should happen that the parameters are "orthogonal", in the sense that $\bar{S}_{kj} = 0$ if $k \neq j$, then (2.4.28) becomes:

$$|\hat{a}_k - a_k^*| \leq \sigma / \bar{S}_{kk}^{\frac{1}{2}} \tag{2.4.30}$$

This indicates that for good accuracy, the diagonal terms of $\bar{S}$ should be large. This point of view is taken up again in §2.4.5.

A statistical approach might also be taken to investigate the effect of measurement noise on the accuracy of the parameter estimates. In such an approach, one could imagine a series of hypothetical experiments where the same system input and output, $\underline{z}_0(t)$ and $\underline{y}_s(t)$, $t \in [T_i, T_f]$, were repeatedly measured, recorded and processed. For each such experiment at least a portion of the measurement noise would be different and hence the input and output records, $\underline{z}(t)$ and $\underline{y}(t)$, $t \in [T_i, T_f]$, would also be different. Over a large number of

experiments, a distribution of the optimal estimates of the parameters would be obtained.

The applicability of a statistical approach depends on the validity of the stochastic model used for the random portion of the measurement noise and whether the systematic portion can be identified properly. If the measuring and recording of data can be done only once for a given excitation of the system, it is difficult to test a hypothesized stochastic model for the measurement noise. This is the case when seismic records are used.

If a series of independent excitations were used to develop statistical information, the dependence of the ideal model on the excitation should also be taken into account (see §2.4.6). This would appear to require that the statistical approach be extended to include the excitation, the ideal model error and the ideal parameter values as random quantities. The difficult task of simultaneously identifying a dynamic model and stochastic models for the measurement noise and ideal model error would then be required.

In this dissertation, attention is focussed on the identification problem where the records used are from only one excitation of the system. No hypotheses are made about the detailed nature of the random or systematic portions of the measurement noise and ideal model error because, on the basis of one sample, it is unlikely that there is sufficient independent information to properly verify these assumptions.

2.4.5. Sensitivity Analysis

In the absence of assumptions about the level of measurement noise and ideal model error, one cannot expect to determine bounds on the error in each estimate of the parameters. Recognizing this fact, an alternative approach can be made which leads to partial but useful information about the accuracy of the estimates of the parameters. This approach is based on the sensitivity of the measure-of-fit J to variations in each parameter about its optimal estimate.

Suppose J is sensitive to a change in the value of a parameter a_k. The change in J from its value $J(\underline{a}^*)$ for the ideal model to its value $J(\hat{\underline{a}})$ for the optimal model, because of measurement noise, should then correspond to only a relatively small change in a_k from a_k^* to $\hat{a}_k$. Thus, if each error $(\hat{a}_k - a_k^*)$ is bounded by applying a sensitivity analysis for an assumed difference $(J(\underline{a}^*) - J(\hat{\underline{a}}))$, this will give a qualitative idea of the accuracy of the optimal estimates on a relative basis. Such a sensitivity analysis cannot be expected to give bounds on the actual error for each estimate since $J(\underline{a}^*)$ is an unknown quantity. From Eq. (2.3.4):

$$J(\underline{a}^*) = \langle \underline{v}^*, \underline{v}^* \rangle + (\underline{a}^* - \hat{\underline{a}}_0, A(\underline{a}^* - \hat{\underline{a}}_0)) \qquad (2.4.31)$$

and hence, with (2.4.16) in mind, $J(\underline{a}^*)$ depends on the measurement noise, the ideal model error and the a priori estimate $\hat{\underline{a}}_0$.

A Taylor series expansion of $J(\underline{a})$ about $\hat{\underline{a}}$ gives:

$$J(\underline{a}^*) = J(\hat{\underline{a}}) + (2\underline{e}_J, \underline{a}^* - \hat{\underline{a}}) + (\underline{a}^* - \hat{\underline{a}}, \hat{S}(\underline{a}^* - \hat{\underline{a}})) + 0\left(\|\underline{a}^* - \hat{\underline{a}}\|^3\right) \quad (2.4.32)$$

where $\hat{S} = S(\hat{\underline{a}})$ (Eq. (2.3.8)). If it is assumed that $\underline{e}_J$ is negligible, then (2.4.32) shows that the sensitivity of J to variations about the optimal estimates $\hat{\underline{a}}$ is governed by the sensitivity matrix $\hat{S}$.

The sensitivity matrix $\hat{S}$ can always be calculated after the optimal estimates have been obtained. It is useful to do so from several points of view. First, if it is positive definite, as assumed here, then $\hat{\underline{a}}$ is locally unique (§2.4.2). Furthermore, it is desirable that $\hat{S}$ be approximately diagonal and that the diagonal elements, which are necessarily positive, be large with respect to $J(\hat{\underline{a}})$.

The first property ensures that the parameters are "orthogonal", that is, that a large error in one parameter does not produce large errors in the other parameters. If two parameters a_k and a_j are not nearly "orthogonal", so that $\hat{S}_{kj}$ is comparable in magnitude with $\hat{S}_{kk}$ and $\hat{S}_{jj}$, then it is possible for both $\hat{a}_k$ and $\hat{a}_j$ to be in error by a considerable amount but for the combined effect of these errors on J to cancel. Thus, it would be expected that these parameters would be difficult to estimate accurately by minimizing J.

On the other hand, if $\hat{S}$ is approximately diagonal, then for a fixed $J(\underline{a}^*)$, the parameter errors will be governed directly by the size of the $\hat{S}_{kk}$, since from (2.4.32):

$$\sum_k \hat{S}_{kk}(\hat{a}_k - a_k^*)^2 \approx J(\underline{a}^*) - J(\hat{\underline{a}}) \qquad (2.4.33)$$

which implies that:

$$|\hat{a}_k - a_k^*| \leq [J(\underline{a}^*) - J(\hat{\underline{a}})]^{\frac{1}{2}}/\hat{S}_{kk}^{\frac{1}{2}} \qquad (2.4.34)$$

For good accuracy, the diagonal terms $\hat{S}_{kk}$ should therefore be large. This bound is analogous to the one in Eq. (2.4.30). In fact, the matrices $\overline{S}$ (Eq. (2.4.24)) and $\hat{S}$ (Eq. (2.4.20)) are approximately equal when $A = 0$. This is because the optimal output-error $\hat{\underline{v}}$ is normally relatively small so that in most cases the third term in (2.4.20) is negligible.

When $A \neq 0$, Eqs. (2.4.20) and (2.4.34) appear to imply that by taking A_{kk} large, the accuracy can be improved. However, this is not necessarily the case since $(J(\underline{a}^*) - J(\hat{\underline{a}}))$ also depends on A. In fact, in the limit as $A_{kk} \to \infty$, $\hat{a}_k \to \hat{a}_{0,k}$ while, from the definition of a_k^* in §2.4.3, the latter is independent of A. Thus, the error in the optimal parameter estimate $\hat{a}_k$ approaches the initial error $(\hat{a}_{0,k} - a_k^*)$.

Many of the above comments can be given a geometrical interpretation. If the sensitivity matrix $\hat{S}$ is positive definite, the contours of J in parameter space for constant $J(\underline{a}^*)$ close to $J(\hat{\underline{a}})$ are hyperellipses centered at the point $\hat{\underline{a}}$ and given by the quadratic form associated with $\hat{S}$. Although $\underline{a}^*$ will be unknown, it must lie on the hyperellipse given by $J(\underline{a}^*)$ and so the accuracy of the optimal parameter estimates will be governed by the shape and overall size of this

hyperellipse, which in turn is controlled by the properties of the sensitivity matrix.

If $\hat{S}$ is diagonal, the axes of the hyperellipse are parallel to the axes of the parameter space and it can be seen that the semi-axes of the hyperellipse give bounds on the errors $|\hat{a}_k - a_k^*|$. This leads directly to the bounds given in (2.4.34).

On the other hand, if $\hat{S}$ is not diagonal, the hyperellipse is oriented "obliquely", so that its axes are not parallel to the axes of the parameter space. However, the length of the axes of the hyperellipse continue to control the accuracy of the parameters. Since these lengths are inversely proportional to the square root of the eigenvalues of $\hat{S}$, the accuracy is ultimately controlled by these eigenvalues. In the "orthogonal" case, the eigenvalues of $\hat{S}$ are equal to the diagonal elements and the present interpretation reduces to the earlier one. It is clear from this geometrical interpretation that if any eigenvalue λ of $\hat{S}$ is almost zero, so that $\hat{S}$ is ill-conditioned, the corresponding axis of the hyperellipse is relatively large. Thus, all those parameters which have a significant component in the principal direction (eigenvector) associated with λ will be poorly estimated in general.

2.4.6. Effect of Model Limitations on Parameter Estimates

If the chosen theoretical model was capable of giving an exact description of the dynamic behavior of a system, the ideal model would be invariant with respect to the particular data used in the identification of the system, and the optimal model would change as different

data samples were used only to the extent of a change in the measurement noise. In practice, the theoretical model describes the behavior of the system only approximately, so that both the ideal and optimal models will change as different data from the system are used. For example, if a linear theoretical model is used to identify a nonlinear system, the estimated parameters can be expected to change as the level of the response changes.

In general, the optimal model can only be expected to predict the output of a system for an excitation with similar characteristics to that producing the data used to determine the model. Tests can be made to examine how well the optimal model determined from one sample of data is able to predict the output for other samples of data. Also, the optimal model can be determined from different samples to examine whether it is unduly sensitive to the particular data used. With each of these approaches, there is a fundamental difficulty in determining how much of the observed differences are due to limitations of the theoretical model in representing the behavior of the system and how much are due to measurement noise. The degree to which these effects can be separated depends on the amount of prior knowledge which is available about the characteristics of the noise and the system.

2.4.7. Final Remarks

One of the most difficult parts of system identification is assessing whether the parameter estimates are reliable. When an output-error approach is taken, it is suggested that this problem be

tackled in a number of steps, so that intermediate results may be examined. These steps may be summarized as follows:

1) The models used should be at least locally identifiable (§ 2.4.1).

2) The estimates returned by the output-error algorithm should be checked to determine whether they are the optimal estimates. A fundamental difficulty arises here in ensuring that a global minimum is found (§ 2.4.2).

3) The accuracy of the optimal estimates should be assessed. One difficulty is that there is no exact model to act as a basis for judging accuracy, so an ideal model is introduced as a substitute for an exact model (§2.4.3). Another difficulty is that often only limited data are available which makes it difficult to confirm assumptions about the character of the noise in the records and to estimate its level. However, the accuracy of each estimate may be compared on a relative basis by an error analysis (§ 2.4.4) or a sensitivity analysis (§ 2.4.5).

4) The final problem which should be considered is whether the optimal estimates are unduly sensitive to the particular data used to determine them because of the limitations of the model in describing the behavior of the system (§ 2.4.6).

In making an assessment of the parameter estimates as above, experience with the system or other similar systems is a great advantage. When using linear structural models, for example, there is a considerable amount of accumulated information which can be used to assess whether the estimates are reasonable. This is not the case for

the identification of nonlinear structural models, which is one of several reasons that make this a much more difficult problem.

It should also be mentioned that there are other ways of looking at the problems discussed above. One alternative to the ideal model is to assume that the system dynamics are described by the state equation of the theoretical model with an additive term called the equation-error (also called the plant noise or process noise in a stochastic setting. See, for example, Bowles and Straeter, 1972). A combined output-error/equation-error approach can then be used to estimate the parameters of the theoretical model, as in the filtering problem of state estimation theory. The ideal model error and the equation-error are two different treatments of the same problem, which is that any theoretical model will provide only an approximation to the dynamics of a real system, but the ideal model error is the appropriate concept to use in an output-error approach.

The approach to parameter estimation in this dissertation is primarily a deterministic one where noise is acknowledged but no assumptions are made about its character. There are a number of papers and books on system identification which provide a stochastic treatment. Eykhoff (1974) and Beck and Arnold (1977) are two examples which have been cited earlier.

When using seismic records, the sample base is so limited that there are inherent difficulties in verifying statistical assumptions and in judging the validity of the error estimates. However, for such cases, a deterministic framework can be used to derive the

equivalent of many of the results based on stochastic theory. This was pointed out, for example, in §2.2.3 with regard to the minimization problem [Eq. (2.2.4)] of the output-error approach. Again, the filters arising in the theory of state estimation, which may be specialized as output-error methods, can be derived on a deterministic basis (invariant-imbedding filter) or a stochastic basis (extended Kalman filter). A final example is that the sensitivity matrix (§2.4.5), which plays an important role in determining the accuracy of the parameter estimates within a deterministic setting, plays an equally important role in a stochastic error-analysis, where it is known as the (Fisher) information matrix.

REFERENCES

Bard, Y. (1970). Comparison of gradient methods for the solution of nonlinear parameter estimation problems. SIAM J. Numer. Anal. 7, 157-186.

Beck, J. V. and K. J. Arnold (1977). Parameter Estimation in Engineering and Science. Wiley and Sons, New York.

Bekey, G. A. (1970). System identification — an introduction and a survey. Simulation 15, 151-166.

Bellman, R. and K. J. Åström (1970). On structural identifiability. Math. Biosciences, 7, 329-339.

Bowles, R. L. and T. A. Straeter (1972). System identification computational considerations, in System Identification of Vibrating Structures, W. D. Pilkey and R. Cohen (eds). ASME, New York.

REFERENCES

Distefano, N. and A. Rath (1974). Modeling and Identification in Nonlinear Structural Dynamics. Report No. EERC 74-15, Univ. of Calif., Berkeley, California.

Eykhoff, P. (1974). System Identification. Wiley and Sons, New York.

Jackson, D. D. (1972). Interpretation of inaccurate, insufficient and inconsistent data. Geophys. J. R. Astr. Soc. 28, 97-109.

Jazwinski, A. H. (1970). Stochastic Processes and Filtering Theory. Academic Press, New York.

Matzen, V. C. and H. D. McNiven (1976). Investigation of the Inelastic Characteristics of a Single Story Steel Structure Using System Identification and Shaking Table Experiments. Report No. EERC 76-20, Univ. of Calif., Berkeley, California.

Murphy, L. M., ed. (1973). San Fernando, California, Earthquake of February 9, 1971, 1. U. S. Dept of Commerce, Washington, D. C.

Raggett, J. D. (1975). Estimating damping of real structures. J. Struc. Div., ASCE 101, ST9, 1823-1835.

Trifunac, M. D. (1970). Low Frequency Digitization Errors and a New Method for Zero Baseline Correction of Strong-motion Accelerograms. Report No. EERL 70-07, Calif. Inst. of Tech., Pasadena, California.

Trifunac, M. D., F. E. Udwadia and A. G. Brady (1971). High Frequency Errors and Instrument Corrections of Strong-motion Accelerograms. Report No. EERL 71-05, Calif. Inst. of Tech., Pasadena, California.

Udwadia, F. E. and D. K. Sharma (1978). Some uniqueness results related to building structural identification. SIAM J. Appl. Math., 34, 104-118.

Wood, J. H. (1972). Analysis of the Earthquake Response of a Nine-story Steel Frame Building during the San Fernando Earthquake. Report No. EERL 72-04, Calif. Inst. of Tech., Pasadena, California.

III. LINEAR STRUCTURAL MODELS

The identifiability of a class of linear structural models is considered in this chapter. It is shown that certain modal parameters are determined uniquely by the input and output of a model in this class, but that the stiffness and damping matrices are not determined uniquely in typical situations. A discussion is made of other difficulties arising in the application of these models to the identification of structures from seismic records. It is concluded that when linear models are used, they should be based on the dominant modes in the records of the response and not on the stiffness and damping matrices.

3.1. A Class of Linear Structural Models

Recall from §2.2.3 that the class of models used in an identification process is defined by a theoretical model, which determines the equation for the state of the model; an output equation, which relates the output of the model to the state; and a set of allowable values for the parameters of the models. Each aspect will be considered in turn for a class of linear structural models with N degrees of freedom, which is denoted by $\mathfrak{m}_N$ for convenience.

3.1.1. Theoretical Model

A discrete theoretical model is used which has the following equation of motion:

$$M\ddot{\underline{x}} + C\dot{\underline{x}} + K\underline{x} = -M\underline{b}\ddot{z}(t) \qquad (3.1.1)$$

This model has had a long history in analysis and design in structural dynamics. For a physical interpretation, it may be imagined that it represents a physical model consisting of a three-dimensional distribution of lumped masses linked by linear, massless springs and dashpots (viscous dampers), with the model sitting on a rigid base which moves in only one direction. The vector $\underline{x} = [x_1, x_2, \ldots x_N]^t$ then consists of the generalized displacement relative to the base of each degree of freedom of each lumped mass of the model, and $\ddot{z}$ is the acceleration of the base. To emphasize that each component of $\underline{x}$ has a specified direction associated with it, the x_i will be called the coordinates of the model. The components of the vector $\underline{b} = [b_1, b_2, \ldots, b_N]^t$ are the so-called pseudo-static influence coefficients which depend only on the geometry of the model (Ch. 27, Clough and Penzien, 1975). If z is a displacement of the base, $\underline{x} + \underline{b}z$ represents the corresponding total or absolute displacement of the masses. It will be assumed that the geometry of the model is prescribed so that $\underline{b}$ is known. The $N \times N$ matrices M, C and K are the mass, damping and stiffness matrices respectively and are parameters of the model. Equation (3.1.1) may be interpreted as expressing the balance between the inertia $M(\ddot{\underline{x}} + \underline{b}\ddot{z})$ of the physical model, its elastic restoring force $-K\underline{x}$ and its viscous damping force $-C\dot{\underline{x}}$, in accordance with Newton's Second Law.

The theoretical model given by Eq. (3.1.1) is often used as a planar model for buildings. In this case, the vector $\underline{x}$ is taken to represent the horizontal displacement at points in the structure, with

all displacements parallel to a fixed vertical plane, and $\ddot{z}$ is taken to be the horizontal component of base motion parallel to this plane. All the components of $\underline{b}$ are therefore unity. The theory in this chapter can be specialized to the planar interpretation by choosing these special values for the components of $\underline{b}$; indeed, this is done later with some illustrative examples.

The three-dimensional interpretation of $\underline{x}$ in (3. 1. 1) is emphasized here because ultimately a reduced form of this model based on the dominant modes of response is used. The planar interpretation is then unnecessarily restrictive because it precludes torsional response whereas the model based on modes does not. However, it is shown below that the simplified way in which the seismic excitation is represented by a single input in the model may cause some difficulties in identifying torsional modes.

There are several simplifications in the way that the seismic excitation is defined in the model. These include:

1) treating the base of the structure as rigid,

2) neglecting the three rotational components of motion of the base,

3) neglecting two of the translational components, one vertical and one horizontal, of the motion of the base.

The first two simplifications should lead to good approximations if there is no pronounced soil-structure interaction. Furthermore, the present data do not allow these features of the model to be improved because existing basement records consist of three orthogonal

translational components of motion at only one point of the base.

The effect of neglecting one horizontal component of the base motion is considered for the case in which the structural output consists of the horizontal motion in a fixed direction at certain points within the structure. Even though the output is "planar", the state $\underline{x}$ of the model can be treated as three-dimensional to include torsional response. The horizontal base motion in the given direction should represent the principal contribution of the seismic excitation to the output. This component of the base motion would therefore be used as the model input $\ddot{z}$. However, the horizontal base motion orthogonal to the given direction can also contribute to the output of the structure by being part of the excitation of either the torsional modes or any translational modes which have a pronounced three-dimensional character.

Some of the problems arising in the applications in Chapter 6 are attributed to inadequate modelling of the torsional contributions to the translational motion at the location of the accelerograph. To treat this feature, it will be necessary to extend the theoretical model to include both horizontal components of the base motion. This will introduce another participation factor for each mode, but the methods presented in this dissertation should remain applicable after some modifications.

3.1.2. Output Equation

The output $\underline{m}$ of each model in the class $\mathbb{m}_N$ is taken to be a vector consisting of the response at certain coordinates in the model. To determine identifiability, it is unnecessary to prescribe which

quantity x_i, $\dot{x}_i$ or $\ddot{x}_i$ is actually observed. It is therefore convenient to define the form of the output for each model by a set $\mathcal{J}$, a subset of the integers 1 to N, corresponding to those coordinates at which the response is measured. For example, if x_1, $\dot{x}_1$ and $\ddot{x}_3$ are measured, then $\mathcal{J} = \{1, 3\}$.

3.1.3. Allowable Values of the Parameters

The parameters of the theoretical model are the elements of the matrices M, K and C and the initial conditions. There are certain physical properties which the mathematical model should imitate and which are commonly used to place restrictions on M, K and C. These restrictions will be used in defining the allowable values of the parameters for the class $\mathfrak{M}_N$.

1) Mass Matrix: The x_i are assumed to correspond to the degrees of freedom of each lumped mass. The mass matrix M is therefore diagonal and positive definite, that is:

$$M = \begin{bmatrix} m_1 & & & 0 \\ & m_2 & & \\ & & \ddots & \\ 0 & & & m_N \end{bmatrix}, \quad m_i > 0 \qquad (3.1.2)$$

In this chapter, the mass matrix will be assumed known when trying to resolve the internal structure of the model from the input and output. The mass matrix for an assumed discretization of a structure is easier to determine a priori, using structural plans, than either the

stiffness or damping matrix. Furthermore, it is obvious from Eq. (3.1.1) that without some constraint on M, the stiffness and damping matrices cannot be determined from the input and output.

2) Stiffness Matrix: The stiffness matrix $K=[k_{ij}]$ is required to be symmetric and positive definite:

$$K = K^t \quad \text{or} \quad k_{ij} = k_{ji} \tag{3.1.3}$$

and

$$\underline{x}^t K \underline{x} > 0 \quad , \quad \forall \underline{x} \neq 0 \tag{3.1.4}$$

The symmetry follows from Betti's reciprocity law (Ch.11, Clough and Penzien, 1975). The necessity for (3.1.3) can also be shown by applying Newton's Third Law to each massless spring in the physical model of §3.1.1, noting that k_{ij} is the force in the direction of x_j given by unit displacement of x_i with all the other coordinates zero. The positive definiteness is imposed so that the equilibrium state of each model is stable.

3) Damping Matrix: The damping matrix $C=[c_{ij}]$ is required to be symmetric and positive semi-definite:

$$C = C^t \quad \text{or} \quad c_{ij} = c_{ji} \tag{3.1.5}$$

and

$$\dot{\underline{x}}^t C \dot{\underline{x}} \geq 0 \quad , \quad \forall \dot{\underline{x}} \tag{3.1.6}$$

The symmetry is imposed for a similar reason to that in 2). The positive semi-definiteness is imposed so that the rate of energy dissipation by the viscous damping forces is non-negative.

It is also assumed that the viscous damping is distributed throughout the model in such a way that "classical, oscillatory" modes exist. The "classical" part means that the modeshapes are the same in the damped and undamped cases, so they are the generalized eigenvectors of both K and C with respect to M. This property is equivalent to the following relation between M, K and C (Caughey and O'Kelly, 1965):

$$CM^{-1}K = KM^{-1}C \qquad (3.1.7)$$

The "oscillatory" part means that each mode is less than critically damped.

4) Initial Conditions: Recall from §2.2.1 that the initial conditions are also treated as parameters of the model. Unrestricted paremeters $\underline{x}_0$ and $\underline{v}_0$ are therefore required such that:

$$\underline{x}(0) = \underline{x}_0 \quad \text{and} \quad \dot{\underline{x}}(0) = \underline{v}_0 \qquad (3.1.8)$$

3.2. Modal Form of Theoretical Model

It will become apparent as the theory develops that a modal formulation plays an essential role in identification using linear theoretical models. In this section, the standard transformation of Eq. (3.1.1) to the uncoupled modal form is described, and a converse result is given. An equivalent formulation in terms of transfer functions is also presented.

3.2.1. Uncoupled Equations of Motion

Let $\Psi = [\underline{\psi}^{(1)}, \underline{\psi}^{(2)}, \ldots, \underline{\psi}^{(N)}]$ denote the modeshape matrix whose columns are the generalized eigenvectors of K, so:

$$K\underline{\psi}^{(r)} = \omega_r^2 M\underline{\psi}^{(r)} \quad , \quad r = 1, \ldots, N$$

or

$$K\Psi = M\Psi\Omega^2 \tag{3.2.1}$$

where

$$\Omega^2 \triangleq \begin{bmatrix} \omega_1^2 & & & 0 \\ & \omega_2^2 & & \\ & & \ddots & \\ 0 & & & \omega_N^2 \end{bmatrix}$$

and the $\omega_r > 0$ are the modal frequencies. The modes are labelled in order of increasing frequency. By an assumed property of C, each modeshape $\underline{\psi}^{(r)}$ also satisfies:

$$C\underline{\psi}^{(r)} = d_r M\underline{\psi}^{(r)} \quad , \quad r = 1, \ldots, N$$

or

$$C\Psi = M\Psi D \tag{3.2.2}$$

$$D \triangleq \begin{bmatrix} d_1 & & & 0 \\ & d_2 & & \\ & & \ddots & \\ 0 & & & d_N \end{bmatrix} = \begin{bmatrix} 2\zeta_1\omega_1 & & & 0 \\ & 2\zeta_2\omega_2 & & \\ & & \ddots & \\ 0 & & & 2\zeta_N\omega_N \end{bmatrix}$$

Here, the modal damping factors ζ_r have been introduced by defining:

$$\zeta_r = \frac{d_r}{2\omega_r} \tag{3.2.3}$$

Since C is positive semi-definite, each $d_r \geq 0$ and hence each $\zeta_r \geq 0$. Furthermore, the modes are assumed to be oscillatory so each $\zeta_r < 1$ [see Eq. (3.2.9)].

Since the $\underline{\psi}^{(r)}$ are linearly independent (or can be chosen to be such in the case of equal frequencies and damping factors), the matrix Ψ is nonsingular. A vector function $\underline{\xi}$ can therefore be defined by:

$$\underline{\xi}(t) = \Psi^{-1}\underline{x}(t)$$

so that:

$$\underline{x}(t) = \Psi\,\underline{\xi}(t) \tag{3.2.4}$$

or

$$x_i(t) = \sum_{r=1}^{N} \psi_i^{(r)} \xi_r(t) \ , \ i = 1, \ldots, N$$

Substituting (3.2.4) into (3.1.1) and pre-multiplying by Ψ^t:

$$(\Psi^t M \Psi)\ddot{\underline{\xi}} + (\Psi^t C \Psi)\dot{\underline{\xi}} + (\Psi^t K \Psi)\underline{\xi} = -\Psi^t M \underline{b}\ddot{z}(t) \tag{3.2.5}$$

Define the generalized mass matrix by:

$$\bar{M} = \Psi^t M \Psi \tag{3.2.6}$$

then substituting (3.2.1), (3.2.2) and (3.2.6) into (3.2.5):

$$\ddot{\underline{\xi}} + D\dot{\underline{\xi}} + \Omega^2 \underline{\xi} = -\underline{\alpha}\ddot{z}(t) \tag{3.2.7}$$

where

$$\underline{\alpha} \triangleq \bar{M}^{-1}\Psi^t M \underline{b} = \Psi^{-1}\underline{b} \tag{3.2.8}$$

is a vector of modal participation factors. In component form, (3.2.7) becomes:

$$\ddot{\xi}_r + 2\zeta_r \omega_r \dot{\xi}_r + \omega_r^2 \xi_r = -\alpha_r \ddot{z}(t) \ , \ r = 1, \ldots, N \tag{3.2.9}$$

The magnitude or norm of each $\underline{\psi}^{(r)}$ is so far arbitrary and so the scaling of each ξ_r is also arbitrary. It is useful for our purposes to express (3.2.4) and (3.2.9) in forms which are invariant with respect to the normalization introduced for each $\underline{\psi}^{(r)}$. This can be done by defining $x_i^{(r)}$, the contribution to x_i from the r^{th} mode, by:

$$x_i^{(r)}(t) = \psi_i^{(r)} \xi_r(t) \tag{3.2.10}$$

so that (3.2.4) may be written:

$$x_i(t) = \sum_{r=1}^{N} x_i^{(r)}(t) \tag{3.2.11}$$

and (3.2.9) leads to:

$$\ddot{x}_i^{(r)} + 2\zeta_r \omega_r \dot{x}_i^{(r)} + \omega_r^2 x_i^{(r)} = -\beta_i^{(r)} \ddot{z}(t) \tag{3.2.12}$$

where

$$\beta_i^{(r)} = \psi_i^{(r)} \alpha_r \tag{3.2.13}$$

The parameter $\beta_i^{(r)}$ will be called the <u>effective participation factor</u> for the r^{th} mode at the i^{th} coordinate.

Equations (3.2.11), (3.2.12) and (3.2.13) play an important role in both the theory and applications in this dissertation. Notice that the response x_i produced by $\ddot{z}$ depends on the parameters $\{\omega_r, \zeta_r, \beta_i^{(r)}, x_i^{(r)}(0), v_i^{(r)}(0): r = 1, \ldots, N\}$ where $v_i^{(r)} \equiv \dot{x}_i^{(r)}$.

The final points to be discussed relate to the generalized mass matrix $\bar{M}$ of Eq. (3.2.6). It is a standard result that $\bar{M}$ is diagonal

because the modeshapes are orthogonal with respect to the mass matrix M. Furthermore, the modeshapes can be normalized so that Eq. (3.2.6) may be rewritten as:

$$\Psi^t M \Psi = I \tag{3.2.14}$$

or

$$\sum_{i=1}^{N} m_i \psi_i^{(r)} \psi_i^{(s)} = \delta_{rs}$$

where I is the identity matrix of order N. This relation expressing orthogonality and normality is convenient for later use.

As a final remark, it can be shown that if there are no repeated modes (modes with a common frequency and damping factor), Ψ satisfying (3.2.1), (3.2.2) and (3.2.14) is unique to within a change of sign of each column in $\underline{\psi}^{(r)}$.

3.2.2 Construction of a Model from Modal Parameters

Given a model in the class $\mathbb{m}_N$, the modal parameters can be determined by solving an eigenvalue problem as in §3.2.1. Conversely, it is shown in this subsection that a unique model in $\mathbb{m}_N$ can be determined if modal quantities with the requisite properties are available.

Suppose the quantities ω_r, ζ_r and $\psi_i^{(r)}$ are known for $i, r = 1, \ldots, N$ and satisfy $\omega_r > 0$, $0 \leq \zeta_r < 1$ and Eq. (3.2.14). As suggested by (3.2.1), (3.2.2) and (3.2.14), take a linear model with the known mass matrix M and stiffness and damping matrices given by:

$$K = M\Psi\Omega^2\Psi^t M$$

and $$C = M\Psi D\Psi^t M \qquad (3.2.15)$$

In component form, these equations become:

$$k_{ij} = m_i m_j \sum_{r=1}^{N} \omega_r^2 \psi_i^{(r)} \psi_j^{(r)}$$

and $$c_{ij} = 2m_i m_j \sum_{r=1}^{N} \zeta_r \omega_r \psi_i^{(r)} \psi_j^{(r)} \qquad (3.2.16)$$

It is easy to show that this linear model is in $\mathbb{m}_N$ because K and C defined by (3.2.15) satisfy the conditions of §3.1.3. Furthermore, this is the only model in $\mathbb{m}_N$ which has the given modal parameters because if some other model with stiffness and damping matrices $\tilde{K}$ and $\tilde{C}$ has the same modal parameters, then from (3.2.1) and (3.2.14):

$$\tilde{K} = M\Psi\Omega^2\Psi^t M = K$$

and similarly, $\tilde{C} = C$.

3.2.3. Transfer Function Formulation:

It is convenient in proving the results of the next section to use the equivalent form of Eqs. (3.2.11) and (3.2.12) obtained by applying Laplace's transformation.

Let $X_i(s)$ and $\ddot{Z}(s)$ denote the Laplace transforms of the displacement x_i and base motion $\ddot{z}$, that is:

$$X_i(s) = \int_0^\infty x_i(t)e^{-st}dt$$

$$\ddot{Z}(s) = \int_0^\infty \ddot{z}(t)e^{-st}dt$$

where s may be complex. For later theory, it is necessary to determine a region in the complex plane where $X_i(s)$ and $\ddot{Z}(s)$ are analytic functions. If $f(t)$ is piecewise continuous on $(0, \infty)$, and of exponential order as $t \to \infty$, say $f(t) = 0(e^{ct})$, then the Laplace transform $F(s)$ of $f(t)$ exists and is an analytic function of s in $Re(s) > c$. It is assumed that $\ddot{z}(t)$ has finite duration, so the above result implies that $\ddot{Z}(s)$ is analytic on the whole s-plane. However, after the base motion has finished, the model will undergo free vibrations so $X_i(s)$ will have poles in the left-half plane. But from the result above it will be analytic on the right half-plane $\mathcal{D}^+ = \{s:Re(s)>0\}$, since $x_i(t)$ must remain bounded as $t \to \infty$. The proofs of the results in §3.3 rely on the fact that $X_i(s)$ (or $\ddot{Z}(s)$) on $\mathcal{D}^+$ are equivalent representations of $x_i(t)$ (or $\ddot{z}(t)$) on the time interval $(0, \infty)$.

From the transforms of Eqs. (3.2.11) and (3.2.12):

$$X_i(s) = \sum_{r=1}^{N} X_i^{(r)}(s) \tag{3.2.19}$$

$$s^2 X_i^{(r)}(s) - s x_i^{(r)}(0) - v_i^{(r)}(0) + 2\zeta_r \omega_r s X_i^{(r)}(s) - 2\zeta_r \omega_r x_i^{(r)}(0) + \omega_r^2 X_i^{(r)}(s) = -\beta_i^{(r)} \ddot{Z}(s) \tag{3.2.20}$$

Combining (3.2.19) and (3.2.20):

$$X_i(s) = G_i(s) + H_i(s)\ddot{Z}(s), \quad \forall s \in \mathcal{D}^+ \tag{3.2.21}$$

where $$G_i(s) \triangleq \sum_{r=1}^{N} [v_i^{(r)}(0) + 2\zeta_r \omega_r x_i^{(r)}(0) + s x_i^{(r)}(0)] H^{(r)}(s) \tag{3.2.22}$$

$$H_i(s) \triangleq -\sum_{r=1}^{N} \beta_i^{(r)} H^{(r)}(s) \tag{3.2.23}$$

$$H^{(r)}(s) \triangleq \frac{1}{s^2 + 2\zeta_r \omega_r s + \omega_r^2} \tag{3.2.24}$$

The function $H_i(s)$ is the transfer function between the base motion $\ddot{z}$ and the corresponding response x_i.

Each $H^{(r)}(s)$ has a pole at s_r and its complex conjugate $\bar{s}_r$ where:

$$s_r = -\zeta_r \omega_r + i(1 - \zeta_r^2)^{\frac{1}{2}} \omega_r \tag{3.2.25}$$

These poles are simple because $\zeta_r < 1$. The poles of $G_i(s)$ are those s_r and $\bar{s}_r$ for which $x_i^{(r)}(0) \neq 0$ or $v_i^{(r)}(0) \neq 0$, and the poles of $H_i(s)$ are those s_r and $\bar{s}_r$ for which $\beta_i^{(r)} \neq 0$. Assuming that all the s_r are distinct, the residue at s_r of $G_i(s)$ is:

$$\tfrac{1}{2} x_i^{(r)}(0) - \frac{1}{2(1-\zeta_r^2)^{\frac{1}{2}}\omega_r}[v_i^{(r)}(0) + \zeta_r \omega_r x_i^{(r)}(0)]i \qquad (3.2.26)$$

and the residue at s_r of $H_i(s)$ is:

$$\frac{\beta_i^{(r)}}{2(1-\zeta_r^2)^{\frac{1}{2}}\omega_r^2} i \qquad (3.2.27)$$

The residues at $\bar{s}_r$ are the complex conjugates of (3.2.26) and (3.2.27).

The points $\{s_r, \bar{s}_r : r = 1, \ldots, N\}$ are distinct unless there are repeated modes with the same frequency and damping factor. For an N degree-of-freedom model with R modes having a common frequency and damping factor, the R modes will appear as a single mode which has values of $x_i^{(r)}(0)$, $v_i^{(r)}(0)$ and $\beta_i^{(r)}$ equal to the sum of these quantities for the repeated modes. To be consistent with the theory to be developed, it is assumed that (R-1) modes are "missing" from the response. This is achieved by taking $x_i^{(r)}(0) = v_i^{(r)}(0) = \beta_i^{(r)} = 0$ for (R-1) values of r.

3.3. Uniqueness of Some Modal Parameters

In this section, some results are proved which form the basis of the approach suggested for identification using linear models. It is shown that certain modal parameters of any model in $\mathbb{m}_N$ are determined uniquely by the input and output of the model. The proofs do not give a practical way of actually determining the values of these parameters. This is left to later chapters.

It is first assumed that the models are initially at rest.

Proposition 1

Consider any model in the class $\mathbb{m}_N$ which is initially at rest but is then excited by a known base history $\ddot{z}$ of finite duration. If the output $\{x_i(t)$ or $\dot{x}_i(t)$ or $\ddot{x}_i(t)$: $t \geq 0$ and $i \in \mathcal{I}\}$ is known, then:

1) the $\beta_i^{(r)}$, $r = 1, \ldots, N$ and $i \in \mathcal{I}$, are determined uniquely,

2) ω_r and ζ_r are determined uniquely if the r^{th} mode makes a contribution to the output, that is, if $\beta_i^{(r)} \neq 0$ for some $i \in \mathcal{I}$.

The converse is also true.

Proof:

The proof is given for the case in which the output is the response at the single coordinate x_i. It can be generalized immediately to an arbitrary set of coordinates. Furthermore, only knowledge of

x_i is considered because the proof is almost identical for $\dot{x}_i$ and $\ddot{x}_i$ if the relations $\dot{X}_i(s) = sX_i(s)$ and $\ddot{X}_i(s) = s^2X_i(s)$ are used. For brevity in the development, some statements enclosed in brackets are included. These are standard results from the theory of complex variables, e.g. Churchill et al (1974).

Notice that under the hypothesis that the model is initially at rest, $G_i \equiv 0$ in Eq. (3.2.21). The proof of the converse is therefore immediate because if 1) and 2) of the proposition hold, the transfer function $H_i(s)$ in (3.2.21) is known. The converse was included to emphasize that the values of the parameters given in 1) and 2) give all the information about the model that is contained in the response x_i. Thus, the parameters listed are a complete set from this point of view.

The main result of the proposition is now proved. Suppose that $x_i(t)$, $t \geq 0$, is known and suppose that $\{\hat{\omega}_r, \hat{\zeta}_r, \hat{\beta}_i^{(r)}: r = 1, \ldots, N\}$ and $\{\tilde{\omega}_r, \tilde{\zeta}_r, \tilde{\beta}_i^{(r)}: r = 1, \ldots, N\}$ are both possible sets of values for the parameters of the model under study. This means that each set of values is consistent with measured input and output, $\ddot{z}$ and x_i. Let $\hat{H}_i(s)$ and $\tilde{H}_i(s)$ be the transfer functions corresponding to the two sets of values [Eqs. (3.2.23) and (3.2.24)]. The basic idea is to show that $\hat{H}_i \equiv \tilde{H}_i$ and to find the conditions under which this implies that $\hat{\omega}_r = \tilde{\omega}_r$, $\hat{\zeta}_r = \tilde{\zeta}_r$,

and $\hat{\beta}_i^{(r)} = \tilde{\beta}_i^{(r)}$.

By the hypotheses and Eq. (3.2.21), $\forall s \in \mathcal{D}^+$:

$$X_i(s) = \hat{H}_i(s)\ddot{Z}(s) \qquad (3.3.1)$$

and $$X_i(s) = \tilde{H}_i(s)\ddot{Z}(s)$$

Subtracting: $$0 = \Delta H_i(s)\ddot{Z}(s) \qquad (3.3.2)$$

where $$\Delta H_i \triangleq \hat{H}_i - \tilde{H}_i \qquad (3.3.3)$$

Since the zeros of an analytic function are isolated, there exists a domain $\mathcal{D}_0$ in $\mathcal{D}^+$ over which $\ddot{Z}(s) \neq 0$. This implies from (3.3.2) that $\Delta H_i = 0$ on $\mathcal{D}_0$, which in turn implies that $\Delta H_i = 0$ everywhere in the complex plane except possibly at the poles of $\hat{H}_i$ and $\tilde{H}_i$. [If F is analytic on a domain D and $F = 0$ on $D_0 \subset D$, then $F = 0$ on D]. It therefore follows that ΔH_i is analytic and zero everywhere. [If $F(s)$ is bounded and analytic throughout a domain $\{s: 0 < |s - s_0| < \delta\}$, then either F is analytic at s_0 or else s_0 is a removable singular point of F]. Each pole of $\hat{H}_i$ must therefore be cancelled by a pole of $\tilde{H}_i$ and vice versa. Recalling the results of §3.2.3, in particular (3.2.25) and (3.2.27), this can occur if and only if for each r such that $\hat{\beta}_i^{(r)} \neq 0$, the following equalities hold (relabelling if necessary):

$$\tilde{\omega}_r = \hat{\omega}_r, \tilde{\zeta}_r = \hat{\zeta}_r, \tilde{\beta}_i^{(r)} = \hat{\beta}_i^{(r)} \qquad (3.3.4)$$

and if $\hat{\beta}_i^{(r)} = 0$, then $\tilde{\beta}_i^{(r)} = 0$. This proves the proposition.

It should be noted that the condition that the input be of finite duration guarantees that it contains all but a countable number of frequencies. This is because a finite duration ensures that the Laplace transform of the input is analytic everywhere and so it has only isolated zeros. The Fourier transform, being the Laplace transform evaluated along the imaginary axis, therefore has the latter property also. Proposition 1 shows that, in theory, this property is sufficient to ensure unique determination of the modal parameters which control the output. In practice, the input would have to have a sufficiently strong signal over the bandwidth containing the modal frequencies so that the signal-to-noise ratio of the modes in the output would allow the modal parameters to be estimated reliably.

The theory can be extended to include the case of nonzero initial conditions. This situation is pertinent to the case where the initial portions of the base motion and response are not observed and so the initial values of the modal contributions $x_i^{(r)}$ are unknown. A complication in this case is that it is possible for the base motion to interact with the initial motion in such a way that two completely different models in $\mathbb{m}_N$ can have the same response for that base motion. This problem does not arise when the base motion $\ddot{z}$ of the models belongs to the class $\mathbb{C}_L$ of piecewise-linear time histories with

finite duration. This is the case in later applications where the input to the model is given by a linear variation between successive discrete data points.

Proposition 2

Consider any model in the class $\mathbb{m}_N$ which is excited by a known base motion history $\ddot{z}$ in the class of piecewise-linear functions $\mathcal{C}_L$. If the output $\{x_i(t)$ or $\dot{x}_i(t)$ or $\ddot{x}_i(t)$: $t \geq 0$ and $i \in \mathcal{I}\}$ is known, then:

1) the $\beta_i^{(r)}$, $x_i^{(r)}(0)$ and $v_i^{(r)}(0)$, $r=1,\ldots,N$ and $i \in \mathcal{I}$, are determined uniquely,

2) ω_r and ζ_r are determined uniquely if the r^{th} mode makes a contribution to the output, that is, if either $\beta_i^{(r)} \neq 0$ or $x_i^{(r)}(0) \neq 0$ or $v_i^{(r)}(0) \neq 0$ for some $i \in \mathcal{I}$.

The converse is also true.

Proof:

As before, the proof is given for one component $x_i(t)$, $t \geq 0$. It is similar for $\dot{x}_i$ and $\ddot{x}_i$ if $\dot{X}_i(s) = sX_i(s) - \sum_{r=1}^{N} x_i^{(r)}(0)$ and $\ddot{X}_i(s) = s^2X_i(s) - s\sum_{r=1}^{N} x_i^{(r)}(0) - \sum_{r=1}^{N} v_i^{(r)}(0)$ are used.

The proof is basically the same as that for Proposition 1. The main difference is that instead of Eq. (3.3.2), subtraction of the two

expressions for $X_i(s)$ leads to:

$$0 = \Delta G_i(s) + \Delta H_i(s)\ddot{Z}(s) \ , \ \forall s \in \mathcal{D}^+ \tag{3.3.5}$$

where $$\Delta G_i \triangleq \hat{G}_i - \tilde{G}_i \tag{3.3.6}$$

and $$\Delta H_i \triangleq \hat{H}_i - \tilde{H}_i \tag{3.3.7}$$

If it is assumed that ΔH_i is not identically zero on its whole domain of analyticity, its zeros are isolated and (3.3.5) leads to an expression for $\ddot{Z}(s)$ which is not consistent with its form for $\ddot{z}$ in $\mathcal{C}_L$. It can therefore be concluded that ΔH_i is zero everywhere except possibly at the poles of $\hat{H}_i$ and $\tilde{H}_i$. Repeating the arguments in the proof of Proposition 1 leads to the same results as in Eqs. (3.3.4).

Equation (3.3.5) therefore implies that ΔG_i is zero on $\mathcal{D}^+$. Repeating previous arguments, ΔG_i is zero everywhere and each pole of $\hat{G}_i$ is cancelled by a pole of $\tilde{G}_i$ and vice versa. Recalling the results of §3.2.3, in particular (3.2.25) and (3.2.26), this can occur if and only if for each r such that $\hat{x}_i^{(r)}(0) \neq 0$ or $\hat{v}_i^{(r)}(0) \neq 0$, the following equalities hold (relabelling if necessary):

$$\tilde{\omega}_r = \hat{\omega}_r, \tilde{\zeta}_r = \hat{\zeta}_r, \tilde{x}_i^{(r)}(0) = \hat{x}_i^{(r)}(0), \tilde{v}_i^{(r)}(0) = \hat{v}_i^{(r)}(0)$$

and if $\hat{x}_i^{(r)}(0) = 0 = \hat{v}_i^{(r)}(0)$, then $\tilde{x}_i^{(r)}(0) = 0 = \tilde{v}_i^{(r)}(0)$. (3.3.8)

The results due to ΔH_i and ΔG_i being identically zero are

combined to complete the proof of the proposition.

3.4. Identifiability of Models in $\mathbb{m}_N$

In this section, the identifiability of the class $\mathbb{m}_N$ of linear models is investigated. Recall that the identifiability of a model is determined by examining whether noise-free input and output of the model specify the parameters uniquely. Recall also that at least local identifiability is necessary to give unique optimal estimates when minimizing J_0 during the identification of a structure since, if the models are locally but not globally identifiable, it may be possible to use prior information about the parameters to choose the appropriate model from the finite number of solutions for an optimal model. This cannot be done if the models are not locally identifiable because there is then a continuum of solutions.

The main result of this section is that, in general, the models in $\mathbb{m}_N$ are neither globally nor locally identifiable unless the response is measured at half or more of the coordinates. Thus, the stiffness and damping matrices of a linear model of a structure typically cannot be determined from seismic records.

The approach taken in establishing these results is to find conditions under which the modal parameters of Proposition 1 determine the model, since these parameters give all the information about

the internal structure of the model that is contained in the input and output. It is assumed that the complete histories of the input and output are used and that the input has finite duration, because these were hypotheses of Proposition 1. Clearly, if a model is not identifiable when the complete histories are used, it is not identifiable when a portion of these histories are used. By applying Proposition 2, the results in this section can be shown to remain valid when the initial portions of the input and output are not available, provided the inputs to the models can be taken as piecewise-linear functions.

3.4.1. Identifiability, Controllability and Observability

It is assumed that a model in $\mathbb{M}_N$ is initially at rest and is then excited only by base motion. If $\beta_i^{(r)} = 0$, for each i in $\mathcal{I}$, the r^{th} mode will be missing from the output. Thus, the condition $\beta_i^{(r)} \neq 0$ at some measurement point is necessary if ω_r and ζ_r are to be determined from the input and output. This is also a sufficient condition according to Proposition 1. This condition, which can be written: $\forall\, r = 1, \ldots, N, \exists\, i \in \mathcal{I}$ such that $\beta_i^{(r)} \neq 0$, is equivalent to the three conditions:

(a) the model has no repeated modes;

(b) there are no modes with a zero participation factor;

(c) no mode has a node at each coordinate at which the

response is measured.

Conditions (a) and (b) are equivalent to the model being controllable and conditions (a) and (c) are equivalent to it being observable (Kalman, 1963). Thus, a necessary and sufficient condition for all of the modal frequencies and damping factors of a model to be determined from its input and output is that it be controllable and observable. Notice that if the input and output of a model show that it is not controllable and observable, then it is not possible to determine which of conditions (a), (b) and (c) are violated on the basis of these data alone.

It follows from the above that a necessary condition for both global and local identifiability of $\mathbb{m}_N$ is that each model be controllable and observable. Since there are obviously models in $\mathbb{m}_N$ which do not satisfy these conditions, the class of models is neither globally nor locally identifiable for any input. However, only the subclass of controllable and observable models is of interest for applications using data from actual structures. This is because the optimal model for the class $\mathbb{m}_N$ identified from structural data will always be constructed with N contributing modes. The optimal model is therefore automatically controllable and observable.

Recall that a model in $\mathbb{m}_N$ can be specified uniquely if all of its modal parameters are known (see §3.2.2). If a model is controllable and observable, its modal frequencies and damping factors can

be determined, but this still leaves the modeshapes. Proposition 1 indicates that the only other information in the output relating to the internal structure of the model is the $\beta_i^{(r)}$ for $r = 1, \ldots, N$ and for each i in $\mathcal{I}$. Thus, the output only directly specifies all the modeshape components $\psi_k^{(r)}$ when $\mathcal{I} = \{1, 2, \ldots, N\}$, that is, the output corresponds to the complete state $\underline{x}$ of the model. In all other cases, the modeshapes are not defined by the input and output alone. However, they must satisfy orthogonality, that is, satisfy the constraint (3.2.14) where the mass matrix M is known. It is therefore possible that this information, together with the values of the $\beta_i^{(r)}$ determined by the input and output, may be sufficient to determine the modeshapes. This is examined in the next two subsections, first for local identifiability and then for global identifiability.

3.4.2. Local Identifiability

The following result is proved in this subsection:

Proposition 3

Consider the subclass of controllable and observable models in $\mathbb{M}_N$ and let $N_0 \leq N$ be the number of coordinates at which the response is measured, which is equal to the number of integers in the output set $\mathcal{I}$, then:

The subclass is locally identifiable if and only if $N_0 \geq \frac{1}{2}N$, that is, the response is measured at no less than half the coordinates in a model.

Proof:

Suppose the input and output of a model in the subclass is measured, then from §3.4.1, all the modal frequencies and damping factors are determined. Consider the equations which are satisfied by the N^2 unknown modeshape components $\psi_k^{(r)}$. First, there are the equations given by the fact that $\beta_i^{(r)}$ is determined by the input and output for each of the N modes and for each i in $\mathcal{J}$. From (3.2.8) with $\bar{M}=I$, and from (3.2.13), this gives $N \times N_0$ quadratic equations:

$$\sum_{k=1}^{N} m_k b_k \psi_i^{(r)} \psi_k^{(r)} = \beta_i^{(r)} \tag{3.4.1}$$

However, only $N_0 \times (N-1)$ of these equations are independent since for each i:

$$\sum_{r=1}^{N} \beta_i^{(r)} = \sum_{r=1}^{N} \psi_i^{(r)} \alpha_r = (\Psi \underline{\alpha})_i = b_i \tag{3.4.2}$$

The constraint (3.2.14) gives $\frac{1}{2}N(N+1)$ independent quadratic

equations after symmetry is taken into account:

$$\sum_{k=1}^{N} m_k \psi_k^{(r)} \psi_k^{(s)} = \delta_{rs} \ , \ s \geq r \qquad (3.4.3)$$

There are therefore $N^2 + (N_0 - \frac{1}{2}N)(N-1)$ independent quadratic equations for the N^2 unknowns in the modeshape matrix Ψ.

Notice that for any solution for Ψ, there are $2^N - 1$ corresponding solutions in which the signs of the columns are changed, but all these solutions give the same model in $\mathbb{M}_N$. In particular, there are at least 2^N solutions of the equations, each corresponding to a modeshape matrix of the observed model.

If $N_0 < \frac{1}{2}N$, there are more unknowns than equations. Thus, there are free unknowns which can be arbitrarily assigned values and this leads, in general, to a whole continuum of real solutions for the modeshape matrix, and hence for K and C through Eqs. (3.2.15). Thus, if $N_0 < \frac{1}{2}N$, the subclass of controllable and observable models is not locally identifiable. However, there can be exceptional models which are locally identifiable. For these models, only a finite set of values for the free unknowns leads to real solutions for Ψ; all other values lead to complex solutions. This can be shown by using the theorem introduced in the next subsection.

If $N_0 = \frac{1}{2}N$, which can only occur if N is even, there are the

same number of unknowns as independent equations. From a standard result of algebraic geometry, this implies that there are only a finite number of real solutions for Ψ. Thus, the stiffness and damping matrices are locally unique since, for a finite number of solutions, each solution must be isolated in the parameter space. The subclass of controllable and observable models is therefore locally identifiable.

If $N_0 > \frac{1}{2}N$, there are fewer unknowns than equations, so the number of solutions must certainly remain finite and the subclass must again be locally identifiable. It might be thought that in this case there should be a unique solution for Ψ, to within an inconsequential change of sign of each column, and hence a unique solution for K and C. However, because of the nonlinearity of the equations, there are exceptional cases which prevent the whole subclass of models from being globally identifiable unless $N_0 = N$.

3.4.3. Global Identifiability

Because of the nonlinearity of the equations, conditions for global identifiability of the subclass of controllable and observable models in $\mathbb{M}_N$ cannot be determined by simply counting unknowns and equations. Instead, the equations must be solved to determine whether there is a unique real solution. The approach taken here is based on a theorem which is proved in Appendix B. The question posed is:

How many models in $\mathbb{m}_N$ would give rise to the same output as a known controllable and observable model when subjected to the same input? Notice that any model consistent with the input and output of the observed model must also be controllable and observable. The theorem reduces the above problem to solving a matrix problem which is easier to treat than the equations in §3.4.2.

Let $\Psi = [\underline{\psi}^{(1)}, \ldots, \underline{\psi}^{(N)}]$ be a modeshape matrix of a model in $\mathbb{m}_N$, then it is convenient to introduce the transformed modeshape matrix $\Phi = [\underline{\varphi}^{(1)}, \ldots, \underline{\varphi}^{(N)}]$ by defining:

$$\Phi = M^{\frac{1}{2}}\Psi \tag{3.4.4}$$

The transformed modeshapes $\underline{\varphi}^{(r)}$ therefore satisfy:

$$\begin{aligned} \underline{\varphi}^{(r)} &= M^{\frac{1}{2}}\underline{\psi}^{(r)} \\ \text{or} \qquad \varphi_i^{(r)} &= m_i^{\frac{1}{2}}\psi_i^{(r)} \end{aligned} \tag{3.4.5}$$

By Eq. (3.2.14):

$$\Phi^t\Phi = I \tag{3.4.6}$$

Since the left-hand inverse and right-hand inverse are always equal, this equation implies that $\Phi^{-1} = \Phi^t$. The transformed modeshape matrix Φ is therefore a real, unitary (or orthogonal) matrix. The roles of Φ and Ψ are equivalent since the mass matrix M is assumed known.

Theorem

Consider a controllable and observable model in $\mathbb{m}_N$ whose output for a known input is measured. Let $\mathcal{J}$ be the output set defining the coordinates at which the response is measured. Let Φ be a transformed modeshape matrix of the observed model. The number of models in $\mathbb{m}_N$ which are consistent with the observed data is equal to the number of solutions of the following matrix problem:

Find a nonsingular, real matrix B such that:

$$\text{(i)}\quad B^t \underline{e}_i = \underline{e}_i \quad , \quad \forall\, i \in \mathcal{J} \tag{3.4.7}$$

$$\text{(ii)}\quad B\underline{\rho} = \underline{\rho} \tag{3.4.8}$$

$$\text{(iii)}\quad (B\underline{\varphi}^{(r)})^t (B\underline{\varphi}^{(s)}) = 0 \quad , \quad r \neq s \tag{3.4.9}$$

where $\underline{e}_i$ is the unit vector given by $(\underline{e}_i)_k = \delta_{ik}$ and $\underline{\rho}$ is a known vector of dimension N with elements given by:

$$\rho_k = b_k m_k^{\frac{1}{2}} \tag{3.4.10}$$

Furthermore, for each solution B, the transformed modeshapes $\underline{\tilde{\varphi}}^{(r)}$ of the model in $\mathbb{m}_N$ which has the same output as the observed model are given by:

$$\underline{\tilde{\varphi}}^{(r)} = \frac{1}{\gamma_r} B\underline{\varphi}^{(r)} \tag{3.4.11}$$

where

$$\gamma_r^2 = (B\underline{\varphi}^{(r)})^t (B\underline{\varphi}^{(r)}) \tag{3.4.12}$$

The proof of this theorem is presented in Appendix B.

The theorem can be used to prove a number of results concerning the determination of models from their input and output. Notice that once the modeshape matrix $\tilde{\Psi}$ is determined from Eqs. (3.4.11), (3.4.12) and (3.4.4), the corresponding model can be determined from Eqs. (3.2.15) since it must have the same modal frequencies and damping factors as the observed model.

Proposition 4

Consider the subclass of controllable and observable models in $\mathbb{M}_N$ and let N_0 be as defined in Proposition 3, then:

The subclass is globally identifiable if and only if $N_0 = N$, that is, the response is observed at every coordinate of a model.

Proof:

If $N_0 = N$, then Eq. (3.4.7) in the theorem holds for all $i = 1, \ldots, N$ and so the only solution of the matrix problem is $B = I$, the identity matrix. The subclass is therefore globally identifiable. Notice that the hypothesis that the models are observable is redundant when $N_0 = N$ since a mode cannot have a node at every coordinate.

To show that $N_0 = N$ is a necessary condition for global identifiability, it is sufficient to show that if $N_0 = N - 1$, there are models which are not determined uniquely by their input and output.

Without loss of generality, the coordinates can be labelled so that the output set is $\mathcal{J} = \{1, 2, \ldots, N-1\}$. According to the theorem, the number of models consistent with the input and output of a given

model is equal to the number of nonsingular solutions to Eqs. (3.4.7), (3.4.8) and (3.4.9). From (3.4.7), the first (N - 1) rows of $B = [b_{ij}]$ are the same as the identity matrix and hence:

$$B\underline{\varphi}^{(r)} = [\varphi_1^{(r)}, \varphi_2^{(r)}, \ldots, \varphi_{N-1}^{(r)}, p_r]^t \tag{3.4.13}$$

where
$$p_r \triangleq (B\underline{\varphi}^{(r)})_N = \sum_{i=1}^{N} b_{Ni}\varphi_i^{(r)} \tag{3.4.14}$$

The first (N - 1) equations in (3.4.8) are satisfied identically and the last equation gives:

$$\sum_{i=1}^{N} b_{Ni}\rho_i = \rho_N \tag{3.4.15}$$

From the orthogonality condition (3.4.9):

$$p_r p_s = -\sum_{i=1}^{N-1} \varphi_i^{(r)}\varphi_i^{(s)} = \varphi_N^{(r)}\varphi_N^{(s)} \quad , \quad r \neq s \tag{3.4.16}$$

where the last equality follows from the orthogonality of the $\underline{\varphi}^{(r)}$.

To determine B, the b_{Ni}, $i = 1, \ldots, N$, could be determined from Eqs. (3.4.14), (3.4.15) and (3.4.16). However, it is sufficient to determine the solutions for the p_r, $r = 1, \ldots, N$, because there is a one-to-one correspondence between these unknowns and the b_{Ni}. This follows from (3.4.14), which can be written:

$$\underline{p} = \Phi^t \underline{b}_N \tag{3.4.17}$$

where $\underline{p} = [p_1, p_2, \ldots, p_N]^t$ and $\underline{b}_N = [b_{N1}, b_{N2}, \ldots, b_{NN}]^t$. Inverting this relation:

$$\underline{b}_N = \bar{\Phi}\underline{p} \tag{3.4.18}$$

Equation (3.4.15) may therefore be written:

$$\rho_N = \underline{\rho}^t \underline{b}_N = \underline{\rho}^t \bar{\Phi} \underline{p} \tag{3.4.19}$$

But from (3.2.8), (3.4.5) and (3.4.10):

$$\underline{\alpha} = \bar{\Phi}^t \underline{\rho} \tag{3.4.20}$$

and so: $$\rho_N = \underline{\alpha}^t \underline{p}$$

or $$\sum_{r=1}^{N} \alpha_r p_r = \rho_N \tag{3.4.21}$$

Equations (3.4.16) and (3.4.21) give the p_r in terms of the mode-shape components $\varphi_N^{(s)}$ and the participation factors α_s of the observed model, together with the constant $\rho_N = b_N m_N^{\frac{1}{2}}$ for the class of models.

The equations for the p_r can be uncoupled as follows. Multiply (3.4.21) by p_s and then use (3.4.16) to get the quadratic equation:

$$\alpha_s p_s^2 - \rho_N p_s + \sum_{\substack{r=1 \\ r \neq s}}^{N} \alpha_r \varphi_N^{(r)} \varphi_N^{(s)} = 0 \ , \ s = 1, \ldots, N \tag{3.4.22}$$

This can be simplified since from (3.4.20):

$$\Phi \underline{\alpha} = \underline{\rho}$$

or
$$\sum_{r=1}^{N} \alpha_r \varphi_i^{(r)} = \rho_i, \quad i = 1, \ldots, N \tag{3.4.23}$$

Substituting into (3.4.22):

$$\alpha_s p_s^2 - \rho_N p_s + \varphi_N^{(s)} (\rho_N - \alpha_s \varphi_N^{(s)}) = 0 \tag{3.4.24}$$

This gives two solutions for each p_s:

$$\text{(i)} \quad p_s = \varphi_N^{(s)} \quad , \quad \text{or} \quad \text{(ii)} \quad p_s = \frac{\rho_N - \alpha_s \varphi_N^{(s)}}{\alpha_s} \tag{3.4.25}$$

Note that $\alpha_s \neq 0$ because, by hypothesis, the observed model is controllable.

If solution (i) is taken for each $s = 1, \ldots, N$, it is easily verified using (3.4.23) that this gives a solution of (3.4.16) and (3.4.21). Furthermore, from (3.4.13), (3.4.11) and (3.4.12), this solution corresponds to the observed model. Uniqueness therefore depends on whether it is possible to have another solution for $\underline{p}$ where some of the p_s are given by (i) and some by (ii). This amounts to checking what combinations of (i) and (ii) can satisfy Eqs. (3.4.16) and (3.4.21). Because of the conditions on the $\beta_i^{(r)}$ for controllability and observability, it turns out that another solution is possible if and only if all but two modes have a node at coordinate N,

the only coordinate whose response is not measured. If q and r are such that:

$$\varphi_N^{(q)} \neq 0,\ \varphi_N^{(r)} \neq 0 \text{ but } \varphi_N^{(s)} = 0 \text{ for all other } s=1,\ldots,N \qquad (3.4.26)$$

then the other solution for $\underline{p}$ is given by:

$$p_q = \frac{\alpha_r}{\alpha_q}\varphi_N^{(r)},\ p_r = \frac{\alpha_q}{\alpha_r}\varphi_N^{(q)} \text{ and } p_s = 0 \text{ for all other } s=1,\ldots,N \qquad (3.4.27)$$

The corresponding transformed modeshapes $\tilde{\underline{\varphi}}^{(r)}$ can be found from (3.4.13), (3.4.11) and (3.4.12) and then the corresponding model can be constructed as in §3.2.2.

There are obviously models in the subclass of controllable and observable models which satisfy the conditions (3.4.26), so the subclass is not globally identifiable when $N_0 = N - 1$, and hence when $N_0 < N$. This completes the proof of Proposition 4.

It is unlikely that an optimal model determined during the identification of a structure would satisfy the requirements in (3.4.26) if N was large, although for small N, the requirement could be satisfied by "reasonable" models. As N_0 is reduced, it is expected that the conditions for nonuniqueness will become less stringent.

As an illustration of the results given by Eqs. (3.4.26) and (3.4.27), consider the case where $N=3$ and $N_0=2$. In particular, suppose that the observed model is an undamped chain model (Fig. 3.1). Damping consistent with $\mathbb{m}_3$ is not considered because its inclusion

is a trivial extension as far as uniqueness is concerned, since the damping matrix is constructed from the modeshapes in the same way as the stiffness matrix (§3.2.2). Taking $m_1 = 2m$, $m_2 = m_3 = m$ and $k_1 = k_2 = k_3 = k$, the model has a node at x_2 in its second mode. Thus, conditions (3.4.26) are satisfied if the response is measured only at coordinates x_1 and x_3. Recall that the pseudostatic influence coefficients are fixed by the prescribed geometry of the coordinates (§3.1.1), and in this case $b_1 = b_2 = b_3 = 1$. Starting with the substitution of the transformed modeshapes and the participation factors of the chain model into Eq. (3.4.27), the stiffness matrix of the other model with the same output can be calculated by following the steps given above. The calculations show that the stiffness matrices of the original model and its counterpart are given by:

$$K = k\begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix} \tag{3.4.28}$$

and

$$\tilde{K} = k\begin{bmatrix} 1.4545 & 0.09095 & -0.5455 \\ 0.09095 & 2.8183 & 0.09095 \\ -0.5455 & 0.09095 & 0.4546 \end{bmatrix}$$

Their common frequencies are: $\omega_1 = 0.4208\,\omega_0$, $\omega_2 = \omega_0$, $\omega_3 = 1.6802\,\omega_0$ where $\omega_0 = (k/m)^{\frac{1}{2}}$, and their common effective participation factors are: $\beta_1^{(1)} = 0.6483$, $\beta_1^{(2)} = 0.3333$, $\beta_1^{(3)} = 0.0184$, $\beta_3^{(1)} = 1.2966$, $\beta_3^{(2)} = -0.3333$, $\beta_3^{(3)} = 0.0367$. Notice that although the second model must lie in $\mathbb{M}_3$ and have the same geometry for its coordinates, it is not another

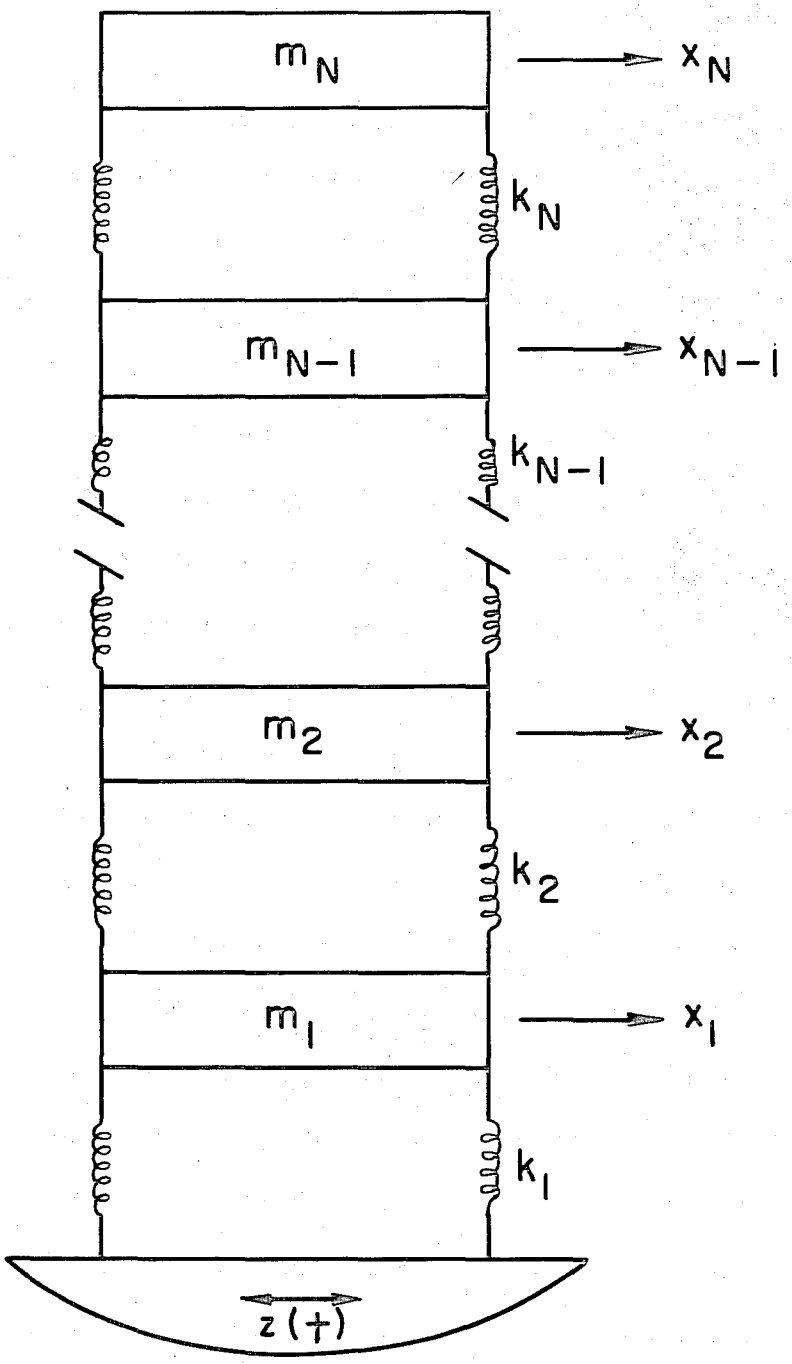


Figure 3.1. Undamped linear chain model with N degrees of freedom.

chain model because $\tilde{K}$ is not tridiagonal.

3.4.4. Identifiability of Linear Chain Models

The previous work demonstrates that the class of models $\mathbb{m}_N$ is too general to guarantee unique determination of a model from its input and output unless the complete state is measured. If the class is further restricted, the chances of it being identifiable can be enhanced, although this may also reduce the capabilities of adequately modelling an actual structure. One possibility is to restrict the models to the subclass of $\mathbb{m}_N$ given by the class of linear chain models (Fig. 3.1). In this case, the stiffness and damping matrices have the additional property of being tridiagonal. However, they are even further restricted because a general symmetric, tridiagonal matrix of order N has 2N-1 independent parameters but K and C for a chain model each have N independent parameters.

The question of uniqueness in the determination of linear chain models from their input and output has been studied by Udwadia and Sharma (1978) for models without damping, and Udwadia, Sharma and Shah (1978) for models with damping. Their damped models do not necessarily belong to $\mathbb{m}_N$ because they do not assume uncoupled modes; in fact, their work does not involve a modal approach. Udwadia and his colleagues consider the problem of determining the unknown stiffnesses k_i, i=1,...,N or unknown damping coefficients c_i, i=1,...,N, from knowledge of the base motion and the response at one coordinate or "floor". Their results show that the class of

linear models is globally identifiable if the output is the response of the first floor but it is only locally identifiable if the output is the response of any other floor.

The first result can be demonstrated by a simpler argument than the original proof by Udwadia and his colleagues. Consider the undamped case. Let w_i denote the absolute displacement of mass m_i which corresponds to the i^{th} "floor". In the notation of Fig. 3.1, $w_i = x_i + z$ and, for convenience, set $w_0 = z$. Consider the situation in the frequency domain as $\omega \to \infty$ and denote Fourier transforms by capital letters. The motion of each mass must be much smaller than that of the mass below it because the inertia of each mass restricts the transmission of the high-frequency motion up the model. In fact:

$$\frac{W_i(\omega)}{W_{i-1}(\omega)} \sim -\frac{k_i}{m_i}\frac{1}{\omega^2} \text{ as } \omega \to \infty \qquad (3.4.29)$$

because the inertia of each mass is balanced to the lowest order by the spring force set up by the motion of the mass below. Thus, at high frequencies, there is a progressive decrease of the order of $\frac{1}{\omega^2}$ in the motion at each successive mass higher up in the model.

From (3.4.29):

$$k_1 = -m_1 \lim_{\omega \to \infty} \frac{\omega^2 W_1(\omega)}{Z(\omega)} \qquad (3.4.30)$$

Since the base displacement z and the first floor response $w_1 = x_1 + z$ are known by hypothesis, k_1 can be determined. From the equation

of motion for mass m_1:

$$m_1\ddot{w}_1 = -k_1(w_1 - w_0) - k_2(w_1 - w_2) \qquad (3.4.31)$$

But $W_2(\omega)/W_1(\omega) = 0\left(\frac{1}{\omega^2}\right)$ as $\omega \to \infty$, and so as $\omega \to \infty$:

$$-m_1\omega^2 W_1(\omega) = -k_2[W_1(\omega) - W_0(\omega)] - k_2 W_1(\omega)$$

Thus, the stiffness of the second spring is given by:

$$k_2 = -k_1 + \lim_{\omega \to \infty} [m_1\omega^2 W_1(\omega) + k_1 Z(\omega)]/W_1(\omega) \qquad (3.4.32)$$

where all the quantities on the right-hand side are known. This allows w_2 to be determined from (3.4.31), that is:

$$w_2 = \frac{1}{k_2}[m_1\ddot{w}_1 + k_1(w_1 - z) + k_2 w_1] \qquad (3.4.33)$$

Since w_1 and w_2 are now known, the arguments can be repeated to determine k_3 and w_3, and so on.

This confirms that every undamped chain model is determined uniquely by knowledge of the base motion and the first floor response, and so in this case the class of chain models is identifiable for inputs of finite duration. However, it should be noted that the algorithm to construct the stiffnesses requires accurate knowledge of the motion at high frequencies where in earthquake records the signals are very small. (The limit process can be approximated by taking any ω much greater than ω_N, the highest modal frequency of the model.) In

almost all practical applications, high-frequency noise would prevent accurate determination of the stiffnesses. In fact, only the lower mode properties will be determined with reasonable accuracy (see §3.5). Consequently, the uniqueness result should not be used to govern the placement of a transducer in a building, as has been suggested, because the first floor will usually have the lowest signal-to-noise ratio for the response of each mode.

3.4.5. An Example: Two Degree of Freedom Models

To illustrate some aspects of nonuniqueness in the determination of models from their input and output, the class $\mathbb{m}_2$ of linear models having two degrees of freedom is considered in detail. Suppose that the geometry of the coordinates x_i is given by $b_1 = b_2 = 1$ and that the output is the response at x_1, then a general model in $\mathbb{m}_2$ can be depicted as a lumped-mass system connected by springs (Fig. 3.2). No special significance should be attached to the spatial arrangement of the masses in Fig. 3.2, although the coordinate directions must always be consistent with $b_1 = b_2 = 1$. Also, damping is not included because it does not provide additional insight into the nonuniqueness.

To revert temporarily to a more general situation, $\varkappa_{ij} (i \neq j)$ is used to denote the stiffness for relative motion between coordinates x_i and x_j when all other coordinates are fixed, and $\varkappa_{ii}$ is used to denote the stiffness for relative motion between coordinate x_i and the base, again when the other coordinates are fixed. For a class of

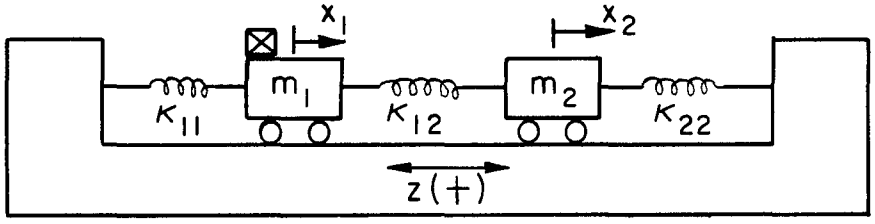


Figure 3.2. Undamped linear model with two degrees of freedom and pseudostatic influence coefficients $b_1 = b_2 = 1$. The symbol ⊠ indicates that the response is measured at that mass.

simple geometries, κ_{ij} will be the stiffness of the spring connecting x_i and x_j. If there is no spring between the coordinates, then κ_{ij} is set to zero. The κ_{ij} are to be distinguished from the elements k_{ij} of the stiffness matrix K. The following useful relation can be proved by deriving the equations of motion of the general mass-spring system using Lagrange's equations:

$$k_{ij} = -\kappa_{ij} \ , \ i \neq j$$
$$k_{ii} = \sum_{j=1}^{N} \kappa_{ij} \qquad (3.4.34)$$

The inverse relation, which will also be used in this subsection, has the same form:

$$\kappa_{ij} = -k_{ij} \ , \ i \neq j$$
$$\kappa_{ii} = \sum_{j=1}^{N} k_{ij} \qquad (3.4.35)$$

Equations (3.4.34) are useful for constructing the stiffness matrix for a general arrangement of point masses connected by linear, massless springs. The point masses can be replaced by finite, rigid lumped-masses provided they have no rotational degrees of freedom. There are also similar equations to (3.4.34) and (3.4.35) which give the relation between the dashpot coefficients and elements of the damping matrix.

Returning to the class $\mathbb{M}_2$, consider a model which is controllable, then ω_r and $\beta_1^{(r)} \neq 0$, $r = 1,2$, are determined by its input and output, the latter being the response at x_1. Two points to be noted are that this is a case with $N_0 = N - 1$ and that a coupled model with two degrees of freedom must be observable because it cannot have repeated modal frequencies and it cannot have a node at a coordinate. This last result also ensures that the conditions (3.4.26) for non-uniqueness are always satisfied because in the present case, the conditions reduce to $\varphi_2^{(1)} \neq 0$ and $\varphi_2^{(2)} \neq 0$. Let $\Phi = [\underline{\varphi}^{(1)}, \underline{\varphi}^{(2)}]$ be the transformed modeshape matrix of the observed model. From (3.4.27), (3.4.13) and (3.4.11), there is one other model consistent with the observed data, called the <u>companion model</u>, which is given by the transformed modeshape matrix $\tilde{\Phi} = [\underline{\tilde{\varphi}}^{(1)}, \underline{\tilde{\varphi}}^{(2)}]$ where:

$$\underline{\tilde{\varphi}}^{(1)} = \frac{1}{\gamma_1}\begin{bmatrix} \varphi_1^{(1)} \\ \dfrac{\alpha_2}{\alpha_1} \quad \varphi_2^{(2)} \end{bmatrix}, \quad \underline{\tilde{\varphi}}^{(2)} = \frac{1}{\gamma_2}\begin{bmatrix} \varphi_1^{(2)} \\ \dfrac{\alpha_1}{\alpha_2} \quad \varphi_2^{(1)} \end{bmatrix} \qquad (3.4.36)$$

and γ_1 and γ_2 are to be selected so that $\underline{\tilde{\varphi}}^{(1)}$ and $\underline{\tilde{\varphi}}^{(2)}$ have unit magnitude. Evaluating the γ_r and using the orthogonality of the $\underline{\tilde{\varphi}}^{(r)}$ and (3.4.20) to simplify (3.4.36), it can be shown that $\tilde{\Phi}$ can be expressed in the form:

$$\tilde{\Phi} = V\Phi \tag{3.4.37}$$

where $$V = \frac{1}{(\rho_1^2 + \rho_2^2)^{\frac{1}{2}}} \begin{bmatrix} \rho_1 & \rho_2 \\ \rho_2 & -\rho_1 \end{bmatrix}$$

Recall from (3.4.10) that $\rho_i = b_i m_i^{\frac{1}{2}}$, so that in the present case $\rho_1 = m_1^{\frac{1}{2}}$, $\rho_2 = m_2^{\frac{1}{2}}$ and $(\rho_1^2 + \rho_2^2)^{\frac{1}{2}} = (m_1 + m_2)^{\frac{1}{2}}$.

According to (3.2.15) and (3.4.4), the stiffness matrices of the observed and companion models are given by:

$$K = M^{\frac{1}{2}}\Phi\Omega^2\Phi^t M^{\frac{1}{2}} \tag{3.4.38}$$

and $$\tilde{K} = M^{\frac{1}{2}}\tilde{\Phi}\Omega^2\tilde{\Phi}^t M^{\frac{1}{2}}$$

Thus, from (3.4.37) and (3.4.38):

$$\tilde{K} = (M^{\frac{1}{2}} V M^{-\frac{1}{2}}) K (M^{\frac{1}{2}} V M^{-\frac{1}{2}})^t \tag{3.4.39}$$

Observe that:

$$M^{\frac{1}{2}} = \begin{bmatrix} \rho_1 & 0 \\ 0 & \rho_2 \end{bmatrix}$$

and so:

$$M^{\frac{1}{2}} V M^{-\frac{1}{2}} = \frac{1}{(1+r)^{\frac{1}{2}}} \begin{bmatrix} 1 & 1 \\ r & -1 \end{bmatrix} \tag{3.4.40}$$

where $r = m_2/m_1$. Substituting (3.4.40) into (3.4.39) gives the following relation between the stiffness elements of $\tilde{K}$ and K:

$$\begin{aligned}
\tilde{k}_{11} &= \frac{1}{1+r}\left[k_{11} + 2k_{12} + k_{22}\right] \\
\tilde{k}_{12} &= \tilde{k}_{21} = \frac{1}{1+r}\left[rk_{11} + (r-1)k_{12} - k_{22}\right] \\
\tilde{k}_{22} &= \frac{1}{1+r}\left[r^2 k_{11} - 2rk_{12} + k_{22}\right]
\end{aligned} \tag{3.4.41}$$

These relations can be interpreted in terms of spring stiffness in the two models by using (3.4.34) and (3.4.35):

$$\begin{aligned}
\tilde{\varkappa}_{11} &= \varkappa_{11} \\
\tilde{\varkappa}_{12} &= \frac{1}{1+r}\left(\varkappa_{22} - r\varkappa_{11}\right) \\
\tilde{\varkappa}_{22} &= r\varkappa_{11} + (1+r)\varkappa_{12}
\end{aligned} \tag{3.4.42}$$

Notice that the spring between mass m_1 and the base is the same in each model. This is consistent with a general result which can be proved using the theorem in Appendix B: if $b_k = 1$, $k = 1, \ldots, N$, and the response is measured at x_i, then the stiffness $\varkappa_{ii}$ for relative motion between x_i and the base is determined uniquely.

One interesting case is given by setting $\varkappa_{11} = 0$, so that the observed model is a chain model with the response measured at the "roof". In this case:

$$\tilde{\kappa}_{11} = 0$$

$$\tilde{\kappa}_{12} = \frac{1}{1+r}\,\kappa_{22} \qquad (3.4.43)$$

$$\tilde{\kappa}_{22} = (1+r)\,\kappa_{12}$$

The companion model is therefore a chain model as well. This is a more general case of the example given in § 2.4.1 which had $m_1 = m_2$, so $r = 1$.

Another interesting case is given by setting $\kappa_{22} = 0$, so that the observed model is now a chain model with the response measured at the first "floor". In this case:

$$\tilde{\kappa}_{11} = \kappa_{11}$$

$$\tilde{\kappa}_{12} = -\frac{r}{1+r}\,\kappa_{11} \qquad (3.4.44)$$

$$\tilde{\kappa}_{22} = (1+r)\;\kappa_{12}$$

Notice that the companion model is <u>not</u> a chain model. This is as it should be, because the result proved by Udwadia and his colleagues (§3.4.4) states that a chain model is determined uniquely by its first floor response.

This last example also illustrates another feature in the identification of general linear models; since $\kappa_{11} > 0$, (3.4.44) shows that $\kappa_{12} < 0$, so that the companion model has a spring with a negative stiffness. Indeed, the companion model is physically unreasonable if

one insists on interpreting its stiffness distribution in terms of physical springs between coordinates. However, this interpretation is too rigid because there are physically reasonable systems which require negative stiffnesses if their stiffness distribution is modelled by springs. The governing requirement is only that the stiffness matrix be positive definite.

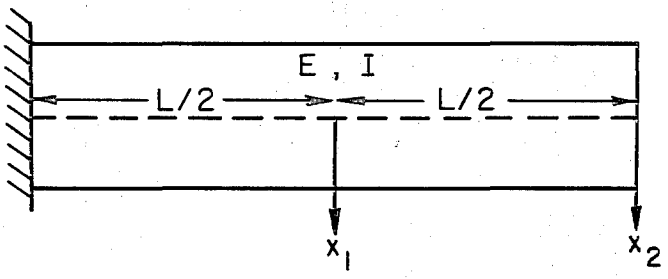
For example, consider a uniform cantilever beam undergoing lateral deflection. Let x_1 be the deflection half-way along the beam and x_2 the deflection at the free end (Fig. 3.3a). Using elementary beam theory, the corresponding stiffness matrix is:

$$K^* = k_0 \begin{bmatrix} 16 & -5 \\ -5 & 2 \end{bmatrix} \tag{3.4.45}$$

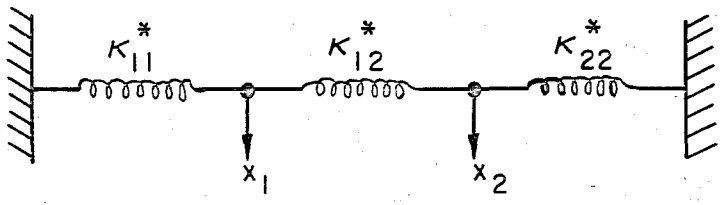
where $k_0 = \frac{48EI}{7L^3}$. The system of springs which has the same stiffness matrix (Fig. 3.3.b) is given by applying (3.4.35) to (3.4.45). This gives:

$$\begin{aligned} \kappa_{11}^* &= 11k_0 > 0 \\ \kappa_{12}^* &= 5k_0 > 0 \\ \kappa_{22}^* &= -3k_0 < 0 \end{aligned} \tag{3.4.46}$$

A spring with a negative stiffness is therefore required in order to model correctly the bending in the beam. Furthermore, three springs



(a)



(b)

Figure 3.3. (a) Uniform cantilever beam (b) System of springs with the same stiffness matrix for the coordinates x_1 and x_2.

are required; two springs can model shear but not bending.

As a final comment on this point, Eq. (3.4.35) can be used to prove the following result:

If K is a symmetric, positive definite matrix, then it corresponds to the stiffness matrix of a spring system with every spring stiffness positive if and only if all the off-diagonal elements of K are nonpositive and K is weakly diagonally dominant, that is,

$$k_{ii} \geq \sum_{\substack{j=1 \\ j \neq i}}^{N} |k_{ij}|, \quad i = 1, \ldots, N.$$

3.5. Determining Linear Models of Structures from Earthquake Records

The application of the class of linear models $\mathbb{m}_N$ to the identification of a structure from seismic records is considered in this section. It is shown that reliable estimates of the stiffness and damping matrices typically cannot be made from records of earthquake response because of basic limitations of the data. A practical strategy is then suggested for structural identification using linear models, which consists of two stages. In the first stage, parameters of the dominant modes in the records are estimated. In the second stage, these parameters are used to improve synthesized models, which are capable of giving more detailed estimates of the structural response.

3.5.1. Limitations of the Data and Models

The first limitation to be considered arises because the seismic response is usually measured at only a few points in a structure. In order for the optimal estimates of the stiffness and damping matrices, K and C, to be at least locally unique, the class $\mathbb{m}_N$ must be locally identifiable. From §3.4.2, the response must therefore be measured at $\frac{1}{2}N$ or more of the coordinates. This requirement will, in general, impose a severe restriction on the number of degrees of freedom allowable in the models. For example, if only one response record is available, which is the case for the two buildings examined in Chapter 6, the models cannot have more than two degrees of freedom. Thus, if the matrices K and C are taken as the parameters to be estimated uniquely, the number of degrees of freedom of the models in $\mathbb{m}_N$ will typically be so small that the mass, stiffness and damping distributions in the structure will be modelled very poorly.

On the other hand, the modal parameters for the structure, as given in Proposition 1, can theoretically be determined for models with any number of degrees of freedom. Furthermore, within the framework of a linear model, these parameters contain all the information about the structural properties that can be estimated directly

from the input and output records. It is therefore recommended that these modal parameters be estimated, rather than K and C.

Another important limitation is due to noise in the records which places an upper bound on the number of modes that can be estimated reliably. This relates back to the discussion in §2.4.1 of the compromise that must be made between resolution and variance of the parameter estimates. Experience with a few applications to multi-story buildings suggests that the bound will generally be of the order of ten modes or less. This is due to several factors which cause a deterioration in the signal-to-noise ratio of the higher modes. The fact that the distributed inertia forces induced by the earthquake motion all act in the same direction, but the modeshapes of the higher modes change sign up the building, limits the energy fed into these modes. This effect is seen in the linear models as a decrease in participation factor with mode number. In addition, the energy content of the ground motion falls off at high frequencies. These two factors result in a smaller signal for the higher modes. There is also a decrease in the signal-to-noise ratio due to an increase in the noise levels at higher frequencies because of limitations in the measuring, recording and data-manipulation processes. In terms of the modal approach recommended above, these factors imply that only the

dominant modes of the response can be estimated reliably from earthquake data.

These factors also imply that the higher-mode information in the stiffness and damping matrices will be unreliable if attempts are made to estimate these matrices from seismic records. Since Eqs. (3.2.16) show that it is the properties of the higher modes which dominate the values of K and C, the stiffness and damping matrices will be estimated poorly from structural data, even if in theory they can be determined uniquely by the input and output records. This is a general conclusion for structural identification using linear models, irrespective of whether the class $\mathbb{m}_N$ or some other class such as chain models is used.

Since the stiffness matrix cannot be estimated reliably from records of the earthquake response of a structure, the distribution of forces also cannot be estimated reliably. This is particularly unfortunate because the forces are of great interest to the earthquake engineer.

Two approaches are suggested to enable the forces to be determined or estimated. A purely empirical approach is to use appropriate transducers to measure directly the stress or strain in structural members. There are obviously practical limitations on the number of

such transducers which could be distributed throughout a structure, but they could be used to check the other approach to determine the forces. This is the one suggested in §2.1.2 in which the parameters of a synthesized model, such as a finite-element model, are adjusted so that the properties of its lower modes are equal to the same properties estimated from the structural data (see, for example, Collins et al, 1974). This will ensure that the output of the model is consistent with the recorded motion of the structure. The altered model may then be used to estimate the earthquake forces and also the structural motion at points where it was not recorded.

The final limitation to be mentioned is a consequence of the approximations inherent in the models in $\mathbb{m}_N$. Properties such as linearity, time-invariance and uncoupled modes are not features of real structures, which can exhibit amplitude nonlinearities, structural deterioration and other complications during an earthquake. The simplifications inherent in the linear modelling may be adequate at low levels of excitation, but for strong ground-motion they can be expected to lead to a pronounced dependence of the estimates of the parameters on the particular data sample chosen, as discussed in a general way in §2.4.6. However, there is still value in determining how well linear models can be made to fit the data because of their

dominant role in present structural design. They also serve as a useful first step towards the problems involved in using nonlinear models in structural identification.

The limitations discussed in this section are summarized in Table 3.1, together with suggestions to avoid these difficulties.

Limitation	Consequence	Suggested Approach
Limited number of records compared with desired number of coordinates in model	Limits resolution because modeshape information missing from data. Matrices K and C typically cannot be determined uniquely	Determine the modal parameters of Propositions 1 or 2, which contain all the information in the records
Noise in the records	Limits resolution because higher-mode information is strongly influenced by noise. Matrices K and C estimated poorly, even if they are identifiable	Must be content with estimating parameters of dominant modes in records of response
Model is only an approximation	Optimal model depends on data used to determine it and it may not predict response well for other excitations	Future research to identify more realistic models.

TABLE 3.1. Limitations when using linear models and earthquake records for structural identification.

3.5.2. Models Based on Dominant Modes

In the previous subsection, it was suggested that the practical strategy for structural identification with linear models is to take a modal approach wherein the parameters $\omega_r, \zeta_r, \beta_i^{(r)}, x_i^{(r)}(0)$ and $v_i^{(r)}(0)$ are estimated for the dominant modes in the seismic response records. This amounts to using a class of models defined by the theoretical model [Eq. (3.2.12)]:

$$\ddot{x}_i^{(r)} + 2\zeta_r \omega_r \dot{x}_i^{(r)} + \omega_r^2 x_i^{(r)} = -\beta_i^{(r)} \ddot{z}(t), \quad r = 1, \ldots, N \qquad (3.5.1)$$

with the initial displacement and velocity, $x_i^{(r)}(0)$ and $v_i^{(r)}(0)$, also treated as parameters, together with the output equation [Eq. (3.2.11)]:

$$x_i = \sum_{r=1}^{N} x_i^{(r)}, \quad \forall i \in \mathcal{J} \qquad (3.5.2)$$

where the output set $\mathcal{J}$ defines the coordinates (points and directions) at which the response $x_i, \dot{x}_i$ or $\ddot{x}_i$ is measured.

In this approach, the structural data are used to determine the number of modes, N, of the model as follows. A small number of modes is taken initially and the optimal estimates of the modal parameters are determined from the recorded input and output. Another mode is then added to the model and all of the modal parameters are

again estimated. This is continued until the change in the optimal measure-of-fit, J_{min}, with the addition of another mode, indicates that the output match is no longer significantly improved.

Clearly, some judgment is involved in determining the best value of N, so it is not a precisely defined quantity, but the above approach does give some indication of the resolution which can be achieved in the presence of the noise in the records. The parameters of the N modes included in the model by using the above criterion based on J_{min} are not necessarily estimated accurately because this depends on the nature of the noise. However, if additional modes were included, they would have such a small effect on the output that it is likely that the estimates of their parameters would be completely unreliable.

Some of the advantages of using models based on the dominant modes in structural identification are:

1) They deal directly with the parameters that control the structural output, as it is interpreted by a linear model.

2) The models are controllable, observable and identifiable, from the results of §3.3 and §3.4.1.

3) The order of the model is not arbitrarily defined, but is determined from the structural data.

4) The estimated parameters can be used to alter any synthesized linear model, whether it be discrete or continuous, to ensure that it is consistent with the structural data.

In the remaining chapters of this dissertation, attention will shift to structural models based on the dominant modes of the output. In Chapters 4 and 5, techniques are described which allow the modal parameters to be estimated from seismic records. The results obtained when these techniques were applied to seismic records from two buildings are discussed in Chapters 4 and 6.

REFERENCES

Caughey, T. K. and M. E. J. O'Kelly (1965). Classical normal modes in damped linear dynamic systems. J. Appl. Mech., ASME, 32, 583-588.

Churchill, R. V., J. W. Brown and R. F. Verhey (1974). Complex Variables and Applications, 3rd. Ed., McGraw-Hill, New York.

Clough, R. W. and J. Penzien (1975). Dynamics of Structures. McGraw-Hill, New York.

Collins, J. D., G. C. Hart, T. K. Hasselman and B. Kennedy (1974). Statistical identification of structures. Amer. Inst. of Aeron. Astron. Journ. 12, 185-190.

Kalman, R. E. (1963). Mathematical description of linear dynamical systems. SIAM J. Control, 1, 152-192.

REFERENCES

Udwadia, F. E. and D. K. Sharma (1978). Some uniqueness results related to building structural identification. SIAM J. Appl. Math., 34, 104-118.

Udwadia, F. E., D. K. Sharma and P. C. Shah (1978). Uniqueness of damping and stiffness distributions in the identification of soil and structural systems. J. Appl. Mech., ASME 45, 181-187.

IV. OPTIMAL FILTER METHOD

4.1. Introduction

The concept of an optimal filter plays an important role in modern control theory. Although a number of optimal filters can be defined, depending on the criterion for optimality, they all share the same basic property of estimating the state of a prescribed model for a system. At each instant of time, the filter gives the optimal estimate of the state based on the histories of the input and output of the system up until that time. If the values of parameters of the model are not known, the state and parameters can be estimated simultaneously by combining them into an augmented state and then determining the optimal filter for this new state. In this case, by a suitable choice of the optimality criterion, determination of the optimal filter is an output-error method for parameter estimation.

A deterministic, least-squares filter has been developed by Bellman et al (1966), and extended by Detchmendy and Sridhar (1966), by using the concept of invariant imbedding. Several investigations have recently been made to determine whether this filter is a useful technique to identify structures from seismic records (Distefano and Rath, 1974; Distefano and Pena-Pardo, 1976; Beck and Jennings, 1977). The method has several attractive features, including the ability to treat nonlinear models and to show how the estimates of the parameters change with time. The work with simulated response data, generated by calculating the response of a model to recorded ground motion, was

promising. However, applications to real seismic data reported later in this chapter show that there are weaknesses in the method when there is significant measurement noise or model error.

The contents of this chapter are as follows. A general formulation of the invariant-imbedding filter is given in §4.2 for the problem of estimating simultaneously the state and parameters. It is possible to introduce this as a special case of an output-error approach to state estimation, but the formulation given here emphasizes the similarities with the general output-error approach to parameter estimation which was given in §2.2. The filter equations are derived in §4.3 for a general type of model and are then specialized in §4.4 for a single degree-of-freedom linear model. The latter section also contains some results and conclusions from tests of the filter using simulated response records and seismic response records from a multi-story building.

4.2. Formulation

Recall from §2.2 that when the usual output-error approach to parameter estimation is applied, if the initial state is unknown, it is included along with the unknown model parameters in the vector $\underline{a}$ to be estimated from the input and output of the system. The approach is then strictly one to estimate simultaneously the state, $\underline{x}$, and the parameters, $\underline{\alpha}$, because once $\underline{a}$ is estimated, the complete history of the state can be determined from (2.2.1). The optimal filter method also estimates the state and parameters but it does so by combining the parameters and the final state into one vector $\underline{c}$. By estimating

$\underline{c}$ for a continuous succession of increasing subintervals of the complete data interval, sequential estimates of the parameters are obtained.

To be specific, the augmented state is defined by:

$$\underline{u}(t) = \begin{bmatrix} \underline{x}(t) \\ \\ \underline{\alpha} \end{bmatrix} \tag{4.2.1}$$

With this definition, the state Eq. (2.2.1) can be rewritten as:

$$\dot{\underline{u}}(t) = \underline{g}(\underline{u}, t) \triangleq \begin{bmatrix} \underline{f}(\underline{x}, \underline{z}, t; \underline{\alpha}) \\ \\ \underline{0} \end{bmatrix} \tag{4.2.2}$$

since the model parameters are constants. Also, to simplify the notation, the input $\underline{z}$ to the model has been omitted as an argument of $\underline{g}$. It is assumed that $\underline{f}$ allows the complete history of $\underline{x}$ to be determined from the value of $\underline{u}$ at any point.

Suppose τ lies in the interval $[T_i, T_f]$ corresponding to the portion of the data from the system which is to be used in the estimation. A vector $\underline{c}$ is defined by $\underline{c} = \underline{u}(\tau)$, then for the subinterval $[T_i, \tau]$, the measure-of-fit defined by (2.2.4) can be rewritten as:

$$J(\underline{c}, \tau) = \int_{T_i}^{\tau} \left\| \underline{y}(t) - \Gamma_1 \underline{u}(t) - \Gamma_2 \dot{\underline{u}}(t) \right\|_{V(t)}^2 dt + \left\| \underline{u}(T_i) - \hat{\underline{u}}_0 \right\|_A^2 \tag{4.2.3}$$

since, by hypothesis, $\underline{u}$ is completely determined by the final-value problem given by (4.2.2) and the end condition $\underline{u}(\tau) = \underline{c}$. Notice that $\underline{a}$ in the last term of (2.2.4) has been replaced by the equivalent vector $\underline{u}(T_i)$ and $\hat{\underline{a}}_0$ has been rewritten as $\hat{\underline{u}}_0$. Notice also that Γ_1 and Γ_2 are modified forms of the matrices appearing in Eq. (2.2.2) in which columns of zeros have been added so that the model output $\underline{m}$ is given by $\Gamma_1\underline{u} + \Gamma_2\dot{\underline{u}}$.

The optimal estimates of the final state and parameters for the subinterval $[T_i, \tau]$ of the data are given by the value of $\underline{c}$ which minimizes $J(\underline{c}, \tau)$. If this value is denoted by $\underline{e}(\tau)$, that is,

$$J(\underline{e}(\tau), \tau) = \min_{\underline{c}} J(\underline{c}, \tau) \qquad (4.2.4)$$

then $\underline{e}(\tau)$ is the optimal filter of $\underline{u}$ evaluated at time τ. The history of the optimal filter $\underline{e}(\tau)$, $\tau \in [T_i, T_f]$, gives sequential optimal estimates of the state and parameters based on increasing portions of the observed data.

Notice that initially, when τ is close to T_i, the term in (4.2.3) containing the initial estimates $\hat{\underline{u}}_0$ will dominate J and the optimal estimates will be biased towards the initial estimates. If the elements of A are small enough, or the interval $[T_i, T_f]$ is large enough, however, the integral term in J will eventually dominate as τ is increased. The optimal filter will therefore begin with the initial

estimates and "converge" to values which are controlled primarily by the data from the system, and the rate of "convergence" can be changed by altering the size of the elements of A. In the absence of model error and measurement noise, this behavior would represent an asymptotic approach to the true values of the parameters. In applications to real data the optimal filter need not converge to constant values of the parameters because the optimal estimates based on the data may change as more data are used, for the reasons discussed in §2.4.6.

Finally, it should be noted that the optimal estimates of the parameters in $\hat{\underline{a}}$ and $\underline{e}(T_f)$ are equal if the same measure-of-fit is used [equations (2.2.4) and (4.2.3)]. The only difference between the two output-error approaches is that one finds the minimum of J by defining the complete history of the state $\underline{x}$ in terms of an initial-value problem and the other uses a final-value problem.

4.3. Invariant-imbedding Filter Equations

The original derivation of the equations for the invariant-imbedding filter was given by Bellman et al (1966) and this was extended by Detchmendy and Sridhar (1966) to include an equation-error term in the state Eq. (4.2.2). Only the output-error approach is considered in this work and a similar derivation to that of Bellman et al is given.

The problem addressed is to solve the series of minimizations

of $J(\underline{c}, \tau)$ as τ ranges over the interval $[T_i, T_f]$. This is equivalent to determining the history of the optimal filter on this interval. The first step is to derive a partial differential equation for $J(\underline{c}, \tau)$. Since $\underline{u}$ is defined on the interval $[T_i, \tau]$ by the state Eq. (4.2.2) and the end condition $\underline{u}(\tau) = \underline{c}$, the notation $\underline{u}(t; \tau, \underline{c})$ is used. At time τ, $\underline{u}(t; \tau, \underline{c})$ has the value $\underline{c}$ and at time $\tau + \delta\tau$, it has the value $\underline{c} + \delta\underline{c}$ where:

$$\begin{aligned} \delta\underline{c} &= \dot{\underline{u}}(\tau; \tau, \underline{c})\, \delta\tau + 0(\delta\tau^2) \\ &= \underline{g}(\underline{c}, \tau)\, \delta\tau + 0(\delta\tau^2) \end{aligned} \tag{4.3.1}$$

But:

$$\underline{u}(t; \tau, \underline{c}) = \underline{u}(t; \tau + \delta\tau, \underline{c} + \delta\underline{c}) \tag{4.3.2}$$

on the interval $[T_i, \tau]$. Apply a Taylor series expansion to the right-hand side of (4.3.2) and substitute (4.3.1); then let $\delta\tau \to 0$. This leads to:

$$\underline{u}_\tau(t; \tau, \underline{c}) = -\underline{u}_{\underline{c}}(t; \tau, \underline{c})\, \underline{g}(\underline{c}, \tau) \tag{4.3.3}$$

where $\underline{u}_\tau$ is the vector with elements $\frac{\partial u_i}{\partial \tau}$ and $\underline{u}_{\underline{c}}$ is the matrix with elements $\frac{\partial u_i}{\partial c_j}$. By differentiating J and using (4.3.3), the following linear partial differential equation can be derived:

$$J_\tau(\underline{c}, \tau) + (J_{\underline{c}}(\underline{c}, \tau), \underline{g}(\underline{c}, \tau)) = \left\| \underline{y}(\tau) - \Gamma_1 \underline{c} - \Gamma_2 \underline{g}(\underline{c}, \tau) \right\|^2_{V(\tau)} \tag{4.3.4}$$

where $(\cdot, \cdot)$ is the Euclidean scalar product. The initial condition associated with (4. 3. 4) is given directly by (4. 2. 3):

$$J(\underline{c}, T_i) = \|\underline{c} - \underline{\hat{u}}_0\|_A^2 \tag{4. 3. 5}$$

A Taylor series expansion is used to solve Eqs. (4. 3. 4) and (4. 3. 5). Thus, with τ as a parameter, J and $\underline{g}$ may be expanded about $\underline{c} = \underline{e}(\tau)$ to get:

$$J(\underline{c}, \tau) = J[\underline{e}(\tau), \tau] + \|\underline{c} - \underline{e}(\tau)\|_{R(\tau)}^2 + \ldots \tag{4. 3. 6}$$

and

$$\underline{g}(\underline{c}, \tau) = \underline{g}[\underline{e}(\tau), \tau] + \underline{g}_{\underline{u}}[\underline{e}(\tau), \tau][\underline{c} - \underline{e}(\tau)] + \ldots \tag{4. 3. 7}$$

There is no first-order term in (4. 3. 6) because $\underline{e}(\tau)$ minimizes $J(\underline{c}, \tau)$ and this implies that:

$$J_{\underline{c}}[\underline{e}(\tau), \tau] = \underline{0} \tag{4. 3. 8}$$

The symmetric matrix $R(\tau)$ in (4. 3. 6) is defined by:

$$[R(\tau)]_{ij} = \frac{1}{2} \frac{\partial^2}{\partial c_i \partial c_j} J(\underline{c}, \tau) \Big|_{\underline{c} = \underline{e}(\tau)} \tag{4. 3. 9}$$

It is the sensitivity matrix of $J(\underline{c}, \tau)$ with respect to $\underline{c}$, the counterpart of $\hat{S} = S(\underline{\hat{a}})$ defined by (2. 3. 8). If $R(\tau)$ is positive definite, $\underline{e}(\tau)$ gives a strict local minimum of $J(\underline{c}, \tau)$.

If the infinite expansions (4. 3. 6) and (4. 3. 7) are substituted into (4. 3. 4), an infinite hierarchy of coupled ordinary differential

equations is obtained in which the n^{th} equation governs the behavior of an array of dimension $(n-1)$. The first three equations of this hierarchy are the scalar equation giving the measure-of-fit of the optimal estimate:

$$\min_{\underline{c}} J(\underline{c},\tau) \triangleq J[\underline{e}(\tau),\tau] = \int_{T_i}^{\tau} \left\| \underline{y}(t) - \Gamma_1 \underline{e}(t) - \Gamma_2 \underline{g}(\underline{e},t) \right\|_{V(t)}^2 dt \qquad (4.3.10)$$

the vector equation:

$$\dot{\underline{e}}(\tau) = \underline{g}(\underline{e},\tau) + Q(\tau) H^t(\underline{e},\tau) V(\tau) [\underline{y}(\tau) - \Gamma_1 \underline{e}(\tau) - \Gamma_2 \underline{g}(\underline{e},\tau)] \qquad (4.3.11)$$

and the matrix equation:

$$\begin{aligned} \dot{Q}(\tau) = {} & \underline{g}_{\underline{u}}(\underline{e},\tau) Q(\tau) + Q(\tau) \underline{g}_{\underline{u}}^t(\underline{e},\tau) + Q(\tau) K(\underline{e},\tau) Q(\tau) \\ & - Q(\tau) H^t(\underline{e},\tau) V(\tau) H(\underline{e},\tau) Q(\tau) - Q(\tau) E(\tau) Q(\tau) \end{aligned} \qquad (4.3.12)$$

Here $Q(\tau) = R^{-1}(\tau)$ is introduced to avoid numerical inversion of the matrix $R(\tau)$ and the matrices H, K and E are defined by:

$$H(\underline{e},\tau) = \Gamma_1 + \Gamma_2 \underline{g}_{\underline{u}}(\underline{e},\tau) \qquad (4.3.13)$$

$$K_{ij}(\underline{e},\tau) = (\Gamma_2 \underline{g}_{u_i u_j}(\underline{e},\tau),\ V(\tau)[\underline{y}(\tau) - \Gamma_1 \underline{e}(\tau) - \Gamma_2 \underline{g}(\underline{e},\tau)]) \qquad (4.3.14)$$

$$E_{ij}(\tau) = \frac{1}{2} \sum_k \frac{\partial^3 J(\underline{c},\tau)}{\partial c_i \partial c_j \partial c_k} \bigg|_{\underline{c} = \underline{e}(\tau)} (\dot{e}_k - g_k) \qquad (4.3.15)$$

An approximate solution for the optimal filter can be calculated

if the term $E(\tau)$ is dropped in (4.3.12). The first three equations of the hierarchy are then decoupled from the remaining equations. The approximate optimal filter, referred to from now on as the invariant-imbedding filter, is therefore defined by the initial-value problem given by Eqs. (4.3.11) and (4.3.12) with $E \equiv 0$, together with the initial conditions:

$$\underline{e}(T_i) = \underline{\hat{u}}_0 \quad ; \quad Q(T_i) = A^{-1} \tag{4.3.16}$$

which are derived by substituting (4.3.6) into (4.3.5). Notice that (4.3.16) implies that A must be nonsingular and, in particular, that it cannot be set to zero in (4.2.3). The initial estimates will therefore always have some influence during the minimization of $J(\underline{c}, \tau)$ when τ is close to T_i.

The optimal filter $\underline{e}(\tau)$ gives the optimal estimate of $\underline{u}(\tau)$ using the subinterval $[T_i, \tau]$ of the data. If the associated optimal estimate of the history of the augmented state is required, which is denoted by $\underline{\hat{u}}(t;\tau)$, $t \in [T_i, \tau]$, it must be determined by the state Eq. (4.2.2) and the end condition:

$$\underline{\hat{u}}(\tau;\tau) = \underline{e}(\tau) \tag{4.3.17}$$

The two time-parameters in the notation for $\underline{\hat{u}}$ are therefore interpreted as follows: t indicates the time at which the state $\underline{u}(t)$ is

estimated and τ indicates how much of the available data over the interval $[T_i, T_f]$ has been used in the estimation. In view of these comments, (4.3.10) is an interesting result because it states that the minimum of $J(\underline{c}, \tau)$, which, according to (4.2.3), is given by:

$$\min_{\underline{c}} J(\underline{c}, \tau) = \int_{T_i}^{\tau} \left\| \underline{y}(t) - \Gamma_1 \hat{\underline{u}}(t;\tau) - \Gamma_2 \underline{g}(\hat{\underline{u}}, t) \right\|_{V(t)}^2 dt$$

$$+ \left\| \hat{\underline{u}}(T_i;\tau) - \hat{\underline{u}}_0 \right\|_A^2 \qquad (4.3.18)$$

is also given by replacing $\hat{\underline{u}}(t;\tau)$ by $\underline{e}(t)$. However, it is clear from their respective definitions that these latter quantities are not equal. In fact, $\hat{\underline{u}}$ and $\underline{e}$ satisfy different differential equations, Eqs. (4.2.2) and (4.3.11) respectively.

In general, the invariant-imbedding filter is only an approximation to the optimal filter. However, the truncated equations are exact if the equation for the augmented state $\underline{u}$ is linear, that is,

$$\dot{\underline{u}}(t) = \underline{g}(\underline{u}, t) = B(t)\underline{u} + \underline{b}(t) \qquad (4.3.19)$$

where the matrix $B(t)$ and vector $\underline{b}(t)$ are known. In this case E is identically zero because $J(\underline{c}, \tau)$ is a quadratic function of $\underline{c}$. However, for parameter estimation, the equation for the augmented state is nonlinear for any nontrivial problem [see, for example, Eq. (4.4.3)].

In the above linear case, Eqs. (4.3.11) to (4.3.16) are formally equivalent to the equations for the stochastic filter of Kalman and Bucy (1961) if process noise is neglected. The Kalman-Bucy filter is in this case the minimum variance estimator of the state $\underline{u}$ of a linear system, whose parameters are known, in the presence of Gaussian white noise in the observations. For this filter, $Q(t)$ corresponds to the covariance matrix of $\underline{u}(t)$.

In the nonlinear case, the equations for the invariant-imbedding filter are formally equivalent to those for the extended Kalman filter if the term involving K in (4.3.12) is omitted. The present derivation of the invariant-imbedding equations does not include an equation-error term, corresponding to the process noise of the extended Kalman filter, but this term can be included (Detchmendy and Sridhar, 1966). The extended Kalman filter and other approximations for a stochastic nonlinear filter are discussed by Jazwinski (1970).

4.4. Single Degree-of-freedom Linear Model

The general filter equations of the last section are specialized in this section for the problem of estimating the state and parameters using a linear model with a single degree of freedom. This model is then used to examine the behavior of the invariant-imbedding filter for one of the simplest models useful in structural identification.

Consider the model given by the equations:

$$\dot{v} + a_2 v + a_1 x = -a_3 \ddot{z}(t)$$
$$v = \dot{x} \qquad (4.4.1)$$

where $a_1 = \omega^2$, $a_2 = 2\zeta\omega$, $a_3 = p$ and $\omega = 2\pi/T$; x and v are the displacement and velocity of the model; $\ddot{z}$ is the input; T is the undamped natural period of the model; ζ is its viscous damping factor; and the input multiplier p is called the participation factor of the model. Notice that the model Eq. (4.4.1) has the same form as the modal Eq. (3.5.1). In fact, the model is used later to estimate the corresponding parameters of the fundamental mode of a building.

From (4.2.1), the augmented state vector is given by:

$$\underline{u}(t) = [x(t), v(t), a_1, a_2, a_3]^t \qquad (4.4.2)$$

and the corresponding state equation is:

$$\dot{\underline{u}}(t) = \underline{g}(\underline{u}, t) = [v, -a_1 x - a_2 v - a_3 \ddot{z}(t), 0, 0, 0]^t \qquad (4.4.3)$$

This last equation is nonlinear in $\underline{u}$, even though the original model is linear in the response quantities and in the parameters.

It is assumed that the histories of the input $\ddot{z}$ and the relative displacement, velocity or acceleration, x_0, v_0 or a_0, are available over some time interval $[T_i, T_f]$. The output can therefore be described by:

$$\underline{y} = [k_1 x_0, k_2 v_0, k_3 a_0]^t \tag{4.4.4}$$

where $k_i = 1$ if the corresponding response quantity is to be used in the estimation, otherwise it is zero. The model output is:

$$\underline{m} = [k_1 x, k_2 v, k_3 \dot{v}]^t \tag{4.4.5}$$

or

$$\underline{m} = \Gamma_1 \underline{u} + \Gamma_2 \dot{\underline{u}} \tag{4.4.6}$$

where $\Gamma_1 = \begin{bmatrix} k_1 & 0 & 0 & 0 & 0 \\ 0 & k_3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$ and $\Gamma_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & k_3 & 0 & 0 & 0 \end{bmatrix}$ (4.4.7)

The measure-of-fit J which is to be minimized by the optimal filter becomes:

$$\begin{aligned} J(\underline{c}, \tau) = {} & k_1 V_{11} \int_{T_i}^{\tau} [x_0 - x]^2 dt + k_2 V_{22} \int_{T_i}^{\tau} [v_0 - v]^2 dt \\ & + k_3 V_{33} \int_{T_i}^{\tau} [a_0 - \dot{v}]^2 dt + A_{11} [x(T_i) - \hat{x}_0]^2 \\ & + A_{22} [v(T_i) - \hat{v}_0]^2 + A_{33} [c_3 - \hat{a}_{1,0}]^2 \\ & + A_{44} [c_4 - \hat{a}_{2,0}]^2 + A_{55} [c_5 - \hat{a}_{3,0}]^2 \end{aligned} \tag{4.4.8}$$

where the weighting matrices A and V(t) have been taken diagonal and constant, and x and v satisfy (4.4.1) with $x(\tau) = c_1$, $v(\tau) = c_2$,

$a_1 = c_3$, $a_2 = c_4$, $a_3 = c_5$. The V_{ii} are used to normalize each integral so that the effects of the different magnitudes of the response quantities are reduced. This is achieved by taking:

$$V_{11}^{-1} = (T_f - T_i)\left[\max |x_0(t)|^2\right] \qquad (4.4.9)$$

with similar expressions for V_{22} and V_{33}, and with the maximum taken over $[T_i, T_f]$. Alternatively, the integral-square response could have been used:

$$V_{11}^{-1} = \int_{T_i}^{T_f} x_0^2(t)\, dt \qquad (4.4.10)$$

with similar expressions for V_{22} and V_{33}. The square root of each integral term in (4.4.8) can be interpreted as the ratio of the root-mean-square (r.m.s.) response-error to the maximum response if (4.4.9) is used, and as the ratio of the r.m.s response-error to the r.m.s. response if (4.4.10) is used.

The optimal filter is the value of $\underline{c}$ which minimizes $J(\underline{c}, \tau)$ and it leads to the sequential optimal estimates $\underline{e}(\tau)$:

$$\underline{e}(\tau) = [\hat{x}(\tau;\tau), \hat{v}(\tau;\tau), \hat{a}_1(\tau), \hat{a}_2(\tau), \hat{a}_3(\tau)]^t \qquad (4.4.11)$$

as τ ranges over the interval $[T_i, T_f]$. Equations (4.3.11) and (4.3.12) for the invariant-imbedding filter become:

$$\dot{e}_j(\tau) = g_j(\underline{e}, \tau) + k_1 V_{11} Q_{j1}(\tau)[x_0(\tau) - e_1(\tau)]$$

$$+ k_2 V_{22} Q_{j2}(\tau)[v_0(\tau) - e_2(\tau)] + k_3 V_{33} P_{j2}(\underline{e}, \tau)[a_0(\tau) - g_2(\underline{e}, \tau)] \tag{4.4.12}$$

and $$\dot{Q}_{jk}(\tau) = P_{jk}(\underline{e}, \tau) + P_{kj}(\underline{e}, \tau) - k_1 V_{11} Q_{j1}(\tau) Q_{k1}(\tau) - k_2 V_{22} Q_{j2}(\tau) Q_{k2}(\tau)$$

$$- k_3 V_{33} P_{j2}(\underline{e}, \tau) P_{k2}(\underline{e}, \tau) - k_3 V_{33} N_{jk}(\tau)[a_0(\tau) - g_2(\underline{e}, \tau)] \tag{4.4.13}$$

where $$P_{jk} = Q_{j2} \quad , \quad \text{if } k = 1$$

$$= -(e_3 Q_{j1} + e_4 Q_{j2} + e_1 Q_{j3} + e_2 Q_{j4} + \ddot{z} Q_{j5}), \quad \text{if } k = 2$$

$$= 0, \quad \text{otherwise} \tag{4.4.14}$$

and $$N_{jk} = Q_{j1} Q_{k3} + Q_{j3} Q_{k1} + Q_{j2} Q_{k4} + Q_{j4} Q_{k2} \tag{4.4.15}$$

The initial conditions are given by Eqs. (4.3.16), which become:

$$\underline{e}(T_i) = [\hat{x}_0, \hat{v}_0, \hat{a}_{1,0}, \hat{a}_{2,0}, \hat{a}_{3,0}]^t \tag{4.4.16}$$

where the vector on the right-hand side contains the prescribed initial estimates, together with:

$$Q_{jk}(T_i) = \frac{1}{A_{jj}} \delta_{jk} \tag{4.4.17}$$

The last term in (4.4.13) arises from the matrix K defined by (4.3.14) and it makes a contribution only if $k_3 = 1$, that is, acceleration matching is included in J. The term is shown here for completeness but it was not included in the results reported below. With the term

omitted, the equations are identical to those for the extended Kalman filter.

4.4.1. Tests: Simulated Data

Numerous tests of the filter method were made using response data generated by numerically solving the equations of motion (4.4.1). The oscillator was taken to be initially at rest and $\ddot{z}$ was taken to be the first 10 seconds of the N-S component of the 1940 El Centro earthquake record (Figure 4.1).

A normalized version of the initial value problem for the invariant-imbedding filter given by Eqs. (4.4.12) to (4.4.17) was solved using a standard subroutine for systems of first-order differential equations. It is only necessary to solve the Ricatti equation for Q for each Q_{ij} with $j \geq i$ because Q is symmetric. The subroutine used is based on the Adams-Moulton predictor-corrector algorithm with a variable time-step capability. The time-step is selected so that the local truncation error in the numerical solution of the equations is nominally less than a prescribed amount, which was taken as 1% of the current calculated solution. The subroutine therefore selects small time-steps when the matrix $Q(t)$ is initially changing rapidly and then automatically increases the time-step as the rate of change of $Q(t)$ decreases.

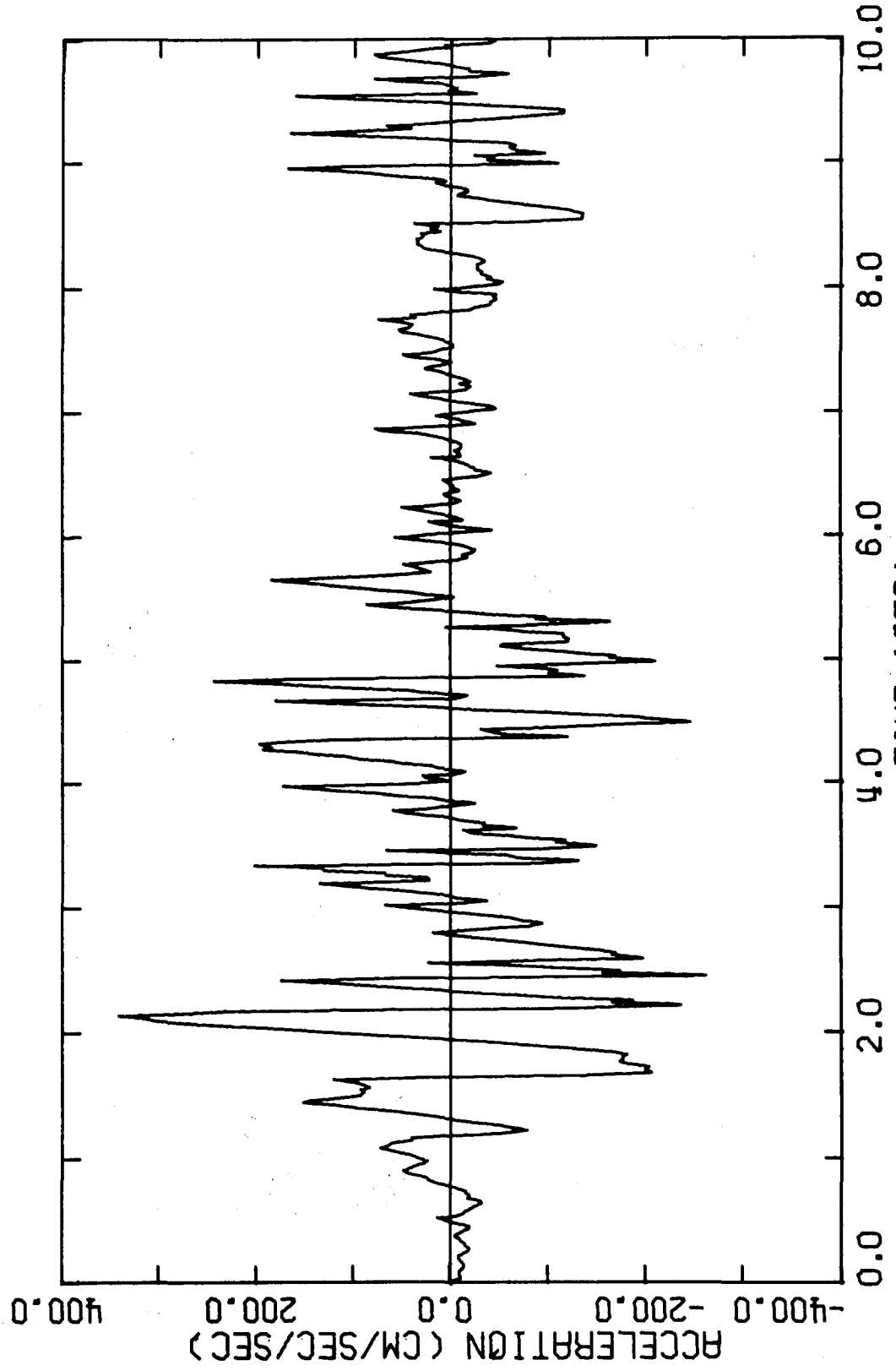


Figure 4.1. North-south component of the record of the El Centro earthquake, May 1940.

The filter equations were solved in a normalized, nondimensional form given by scaling $\underline{e}(t)$ and $Q(t)$ by the initial estimates of the parameters and by the maximum of the recorded response. This helps when selecting appropriate values for $Q(0)$ or A. In addition, the relative sizes of the diagonal terms of the inverse sensitivity matrix $Q(t)$ then give an indication of the accuracy of the estimates relative to one another.

The first illustration of the results obtained uses the relative displacement of a linear oscillator with $T = 1.0$ sec., $\zeta = 5\%$ and $p = 1.0$. To investigate the performance of the filter when both the parameters and the initial conditions are unknown, the estimates of the initial displacement and velocity were taken to be 10% of x_m and $\hat{\omega}_0 x_m$ respectively, where x_m is the "recorded" peak displacement. The values of the $Q_{jj}(0)$ in the normalized version of (4.4.17) were chosen to be:

$$Q_{jj}(0) = A_{jj}^{-1} = \begin{cases} 10^2 , & j = 1, 2 \\ 10^4 , & j = 3, 4, 5 \end{cases} \qquad (4.4.18)$$

The initial estimates of the model parameters and the estimates given by the filter after the first 5 seconds of the data have been used are shown in Table 4.1. The corresponding estimates of the period, damping factor and participation factor are also shown. It took

	Parameters in Model			Physical Parameters		
	a_1	a_2	a_3	T	ζ(%)	p
True values	39.4784	0.6283	1.0000	1.0	5.0	1.0
Initial estimates Relative error	49.3 +25%	0.314 -50%	1.50 +50%	0.894 -11%	2.24 -55%	1.50 +50%
Final estimates Relative error	39.52 +0.1%	0.625 -0.5%	0.999 -0.1%	0.9995 -0.05%	4.97 -0.6%	0.999 -0.1%

TABLE 4.1. Estimates of parameters of linear oscillator using 5 seconds of displacement

approximately 15 seconds of CPU time on an IBM 370/158 to compute the first 5 seconds of the filter, which are typical figures for all the computer runs with a single-degree-of-freedom model. The behavior of the sequential estimates of the parameters, which are given by the filter components $e_3(t)$, $e_4(t)$ and $e_5(t)$, are shown in Fig. 4.2. The estimates have essentially converged to the true values in about $3\frac{1}{2}$ cycles. The slower rate of convergence for the damping coefficient in Fig. 4.2 is typical of all the results obtained with lightly-damped oscillators and it is presumably due to the fact that the response is less sensitive to the damping. The estimates of the response converge much more rapidly. Figure 4.3 shows that the filter component $e_1(t)$ converges to the actual displacement in a quarter of a cycle. The velocity component, $e_2(t)$, took one cycle to converge.

It was found that using any of the relative displacement, velocity and acceleration of the single-degree-of-freedom oscillator leads to almost the same rate of convergence and accuracy of the filter, provided the normalized filter equations are used. This is illustrated in Table 4.2 which shows the estimates of the physical parameters after using the first 6.5 seconds of the response of a linear oscillator with $T = 1.0$ sec., $\zeta = 5\%$ and $p = 1.0$. The initial estimates of the model parameters a_1, a_2 and a_3 were taken to be 50% greater than the true values, and the estimates of the initial displacement and velocity

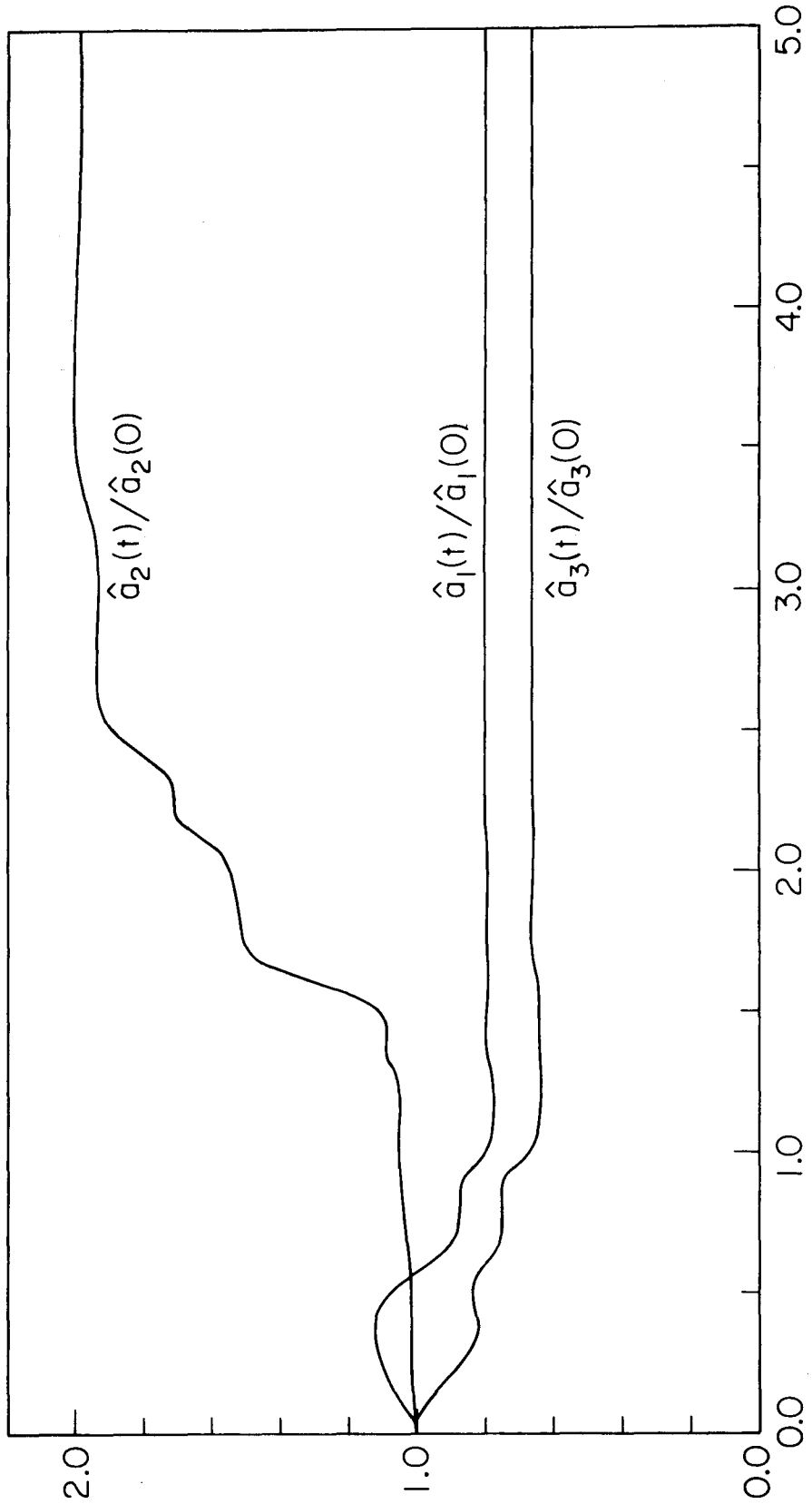


Figure 4.2. Sequential estimates of the model parameters for the linear oscillator of Table 4.1, using its relative displacement.

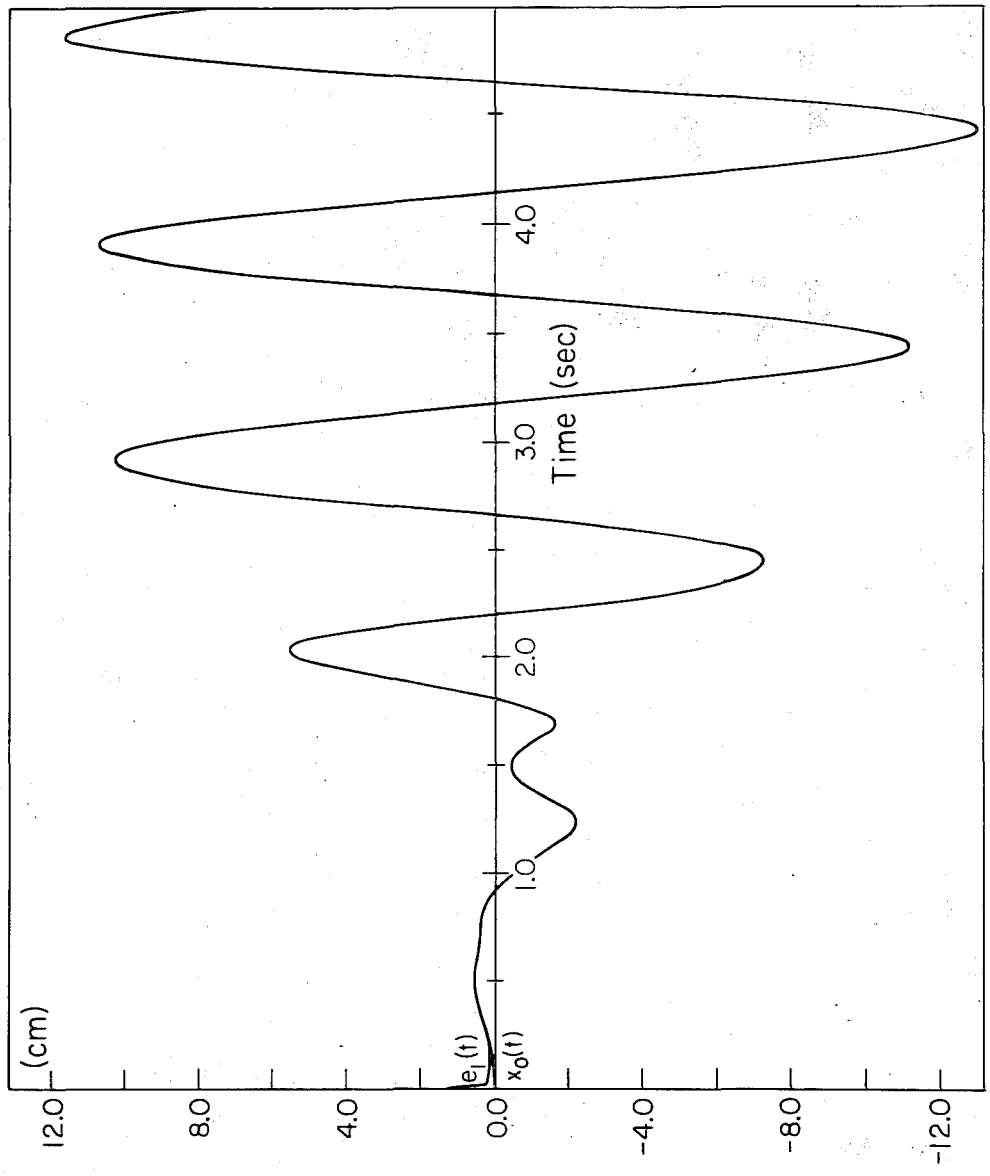


Figure 4.3. Displacement component $e_1(t)$ of the filter and the actual displacement $x_0(t)$ for the same problem as in Fig. 4.2. The curves are distinguishable only for small times.

Record used	λ	$\hat{T}$(sec)	$\hat{\zeta}$(%)	$\hat{p}$	Q_{33}	Q_{44}	Q_{55}	CT(sec)
Displacement	1	0.9998	6.3	1.06	0.0033	0.33	0.11	>6.5
	10	1.0003	5.1	1.03	0.0026	0.44	0.13	6.1
	10^2	1.0004	4.994	1.009	0.0021	0.41	0.14	5.8
	10^4	0.9999	4.990	1.001	0.0019	0.41	0.12	3.2
	10^6	0.9996	4.994	1.0006	0.0019	0.41	0.12	2.2
Velocity	10^2	0.9998	5.010	1.006	0.0031	0.62	0.19	5.6
	10^4	0.9996	4.988	0.9996	0.0028	0.62	0.17	3.2
Acceleration	10^2	1.0000	5.005	1.007	0.0049	0.70	0.18	5.3
	10^4	0.9996	4.988	0.9985	0.0045	0.69	0.17	2.4

TABLE 4.2. Final estimates of the parameters of the linear oscillator using 6.5 seconds of response. The Q_{jj} are their final values, λ is defined by Eq. (4.4.19), and CT is the time to reach a "steady-state", that is, within 1% of the final estimate.

were set to zero. The values of the $Q_{jj}(0)$ were taken to be:

$$Q_{jj}(0) = A_{jj}^{-1} = \begin{cases} 10^{-6} & , \quad j = 1,2 \\ \lambda & , \quad j = 3,4,5 \end{cases} \qquad (4.4.19)$$

where λ is given in Table 4.2. The case where the relative displacement is used illustrates that when the $Q_{jj}(0)$ are relatively small, the optimal estimates can be influenced by the initial estimates via their terms in J, whereas once the $Q_{jj}(0)$ are sufficiently large, the optimal estimates are controlled by the response data via the integral terms in J.

It was also found that the filter technique was capable of successfully estimating the parameters of a linear oscillator over the range of damping from $\zeta = 0\%$ to $\zeta = 100\%$. Figure 4.4 (a) corresponds to an oscillator with $T = 1.0$ sec., $\zeta = 5\%$ and $p = 1.0$ and Fig. 4.4(b) corresponds to an oscillator with the same parameters except that the damping is increased to $\zeta = 50\%$. In both cases, the relative acceleration was used as the output to be matched by the model and the initial estimates of the parameters a_1, a_2 and a_3 were taken to be 50% greater than the true values. For the same initial values of the $Q_{jj}(0)$ [Eq. (4.4.19) with $\lambda = 10^4$], the estimates of the parameters took about $2\frac{1}{2}$ cycles ($\zeta = 5\%$) and $\frac{1}{2}$ cycle ($\zeta = 50\%$) to converge. The estimates of the model parameters after 3 seconds of response data were accurate

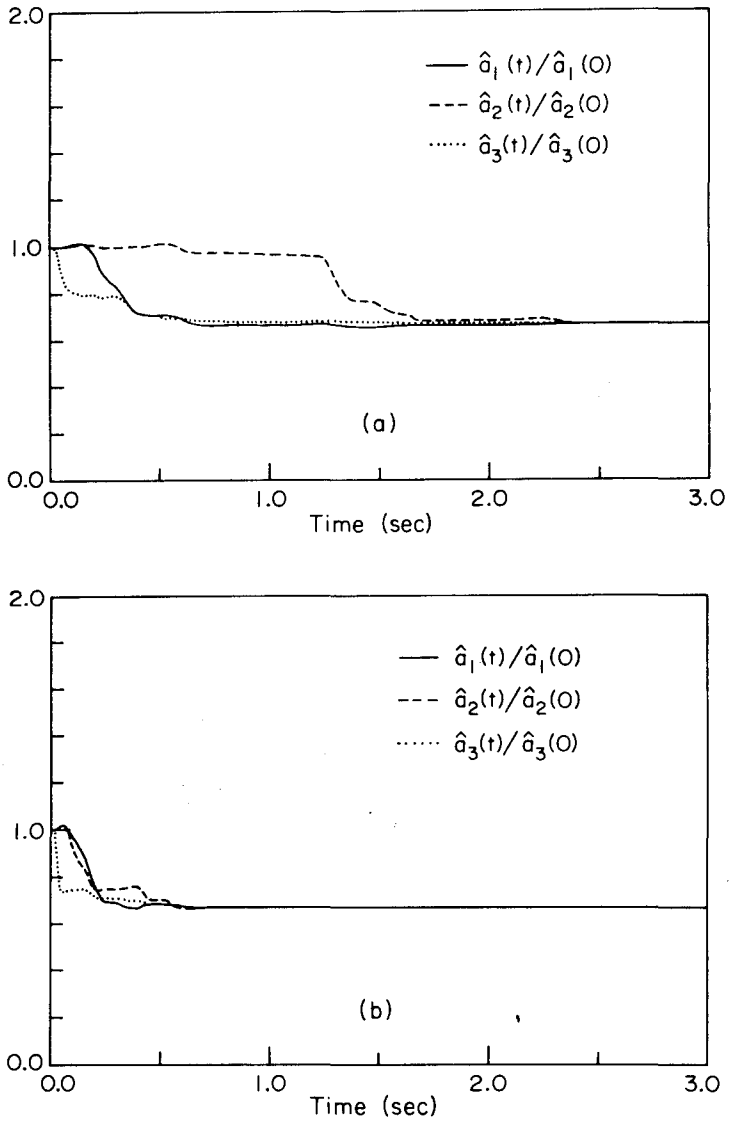


Figure 4.4. Sequential estimates of the model parameters for two linear oscillators, (a) $T = 1.0$, $\zeta = 5\%$, $p = 1.0$ and (b) $T = 1.0$, $\zeta = 50\%$, $p = 1.0$, using the relative acceleration of each oscillator.

to $\frac{1}{2}\%$ or less in each case.

It was observed that in every case where the initial matrix $Q(0)$ was taken diagonal, the sequential estimates of the parameters approached constant values. If the filter correctly calculated the optimal estimates, it would converge to the true values of the parameters because there is no model error and no significant measurement noise. The reasons for this were discussed in §4.2. However, it was found that when the $Q_{jj}(0)$ were relatively small, the steady-state values calculated by the filter were not always close to the true values of the parameters. It appears that $Q(\tau)$ approaches the singular solution $Q = 0$ of the Ricatti equation (4.4.13), regardless of whether the filter values are optimal, so that eventually the data are not used to update $\underline{e}(\tau)$ in (4.4.12). Notice from the examples in Table 4.2 that the approach to $Q = 0$ is eventually independent of the initial conditions on Q.

A similar behavior of $Q(t)$ occurs when the Kalman-Bucy filter is applied to a linear model of the form of (4.3.19), which can lead to "divergence" wherein the filter output gradually departs from the recorded output (Jazwinski, 1970). Although the same effect occurs in the present case, it has a different cause. In the Kalman-Bucy filter, it is due to model error. The filter values are always optimal because the equations for the filter give the exact solution for a linear

model of the form of (4.3.19). In the present application, there is no model error and the "divergence" of the filter output, or the "pseudo-convergence" of the estimates of the parameters, is due to the sub-optimal nature of the nonlinear filter.

4.4.2 Tests: Real Data

The computer program employed above was next used to estimate the parameters of the fundamental longitudinal mode of the Union Bank Building, Los Angeles, from records obtained during the 1971 San Fernando earthquake. The response of this building is studied in more detail in the next chapter using a different technique. The purpose of the present section is to show that the estimates of the parameters calculated by the filter can be in considerable error compared with the optimal estimates when there is significant model error or measurement noise. This was originally discovered during an investigation of the sensitivity of J with respect to the model parameters, but a different approach is taken here in order to demonstrate the magnitude of the errors in the estimates given by the filter in specific cases.

The results presented in Table 4.3 include the final estimates given by the filter for time segments from 5 to 15 seconds and from 15 to 25 seconds of the longitudinal component of the relative displacement at the 19^{th} floor. The longitudinal component of the absolute

acceleration in the sub-basement was taken as the input to the models. The two time-segments represent respectively the worst and best performance of the filter in finding the optimal estimates for a number of time segments. The optimal estimates in Table. 4.3 were obtained by minimizing J, Eq. (4.4.8), using the technique to be described in the next chapter. The value of J_0 in the Table is the contribution to J from the first integral term in (4.4.8) since here $k_1 = 1$ and $k_2 = k_3 = 0$. The Table also shows that increasing $Q_{jj}(0)$ does not improve the performance of the filter, in contrast to the behavior of the filter with simulated data.

To determine the values of J or ∇J for the estimates given by the filter, the complete displacement history for the model is required. Thus, Eqs. (4.4.1) must be solved backwards in time because $\underline{e}(T_f)$ gives estimates of the final values of the displacement and velocity and not the initial conditions. Since this is equivalent to computing a forward solution with negative damping, the solution can become unstable. The unstable nature of the backwards problem can also be understood by noting that the solution of (4.4.1) for a forward problem in time is eventually independent of the initial conditions. The seriousness of the instability depends on the size of the quantity $\zeta\omega T_f$. Tests showed that sufficient accuracy was obtained in the present problem because the damping is small and the duration is just over two cycles. For

Time Segment	Value	$\hat{T}_1$	$\hat{\zeta}_1$(%)	$\hat{p}_1$	$\hat{x}^{(1)}(0)$	$\hat{v}^{(1)}(0)$	Q_{33}	Q_{44}	Q_{55}	$J_0 \times 10^2$	$J \times 10^2$
5-15 [smaller Q(0)]	Initial	4.5	3.0	0.7	0.70	-16.0	10^2	10^2	10^2		
	Final	4.46	5.2	0.65	-1.6	3.0	0.20	80	2.1	2.9	3.4
	Optimal	4.36	3.5	0.76	1.7	-8.9				0.54	0.60
5-15 [larger Q(0)]	Initial	4.5	3.0	0.7	0.70	-16.0	10^4	10^4	10^4		
	Final	4.25	8.4	0.85	12.1	-22.9	0.71	1050	11	6.6	6.9
	Optimal	4.39	5.8	0.81	1.9	-9.4				0.52	0.53
15-25	Initial	4.5	3.7	0.76	21.7	18.9	10^2	10^2	10^2		
	Final	4.621	4.5	0.513	19.2	20.7	0.09	15	40	0.309	0.335
	Optimal	4.627	4.2	0.533	19.2	20.5				0.307	0.332

TABLE 4.3. Final estimates given by the filter compared with the optimal estimates for portions of the relative displacement record, longitudinal direction, at the 19th floor of the Union Bank building.

larger damping or longer durations, the instability might cause difficulties in evaluating the performance of the filter.

The estimates given by the filter for the time segment of the displacement record from 5 to 15 seconds were particularly poor. It is shown in the next chapter that there is significant long-period noise in this portion of the record, in addition to the model error due to ignored higher modes and nonlinearities in the response of the building. These effects are thought to explain why the filter had difficulty in estimating the damping coefficient, as suggested by the large final values of Q_{44} in Table 4.3. The smaller value of J_0 shows that these effects are not so pronounced in the time segment from 15 to 25 seconds. The optimal estimates of the displacement of the fundamental mode, given by minimizing J_0 for each time segment using the method in the next chapter, are given in Figs. 4.5 and 4.6 to indicate the magnitude of the model error plus measurement noise.

Figure 4.7 is included to show the behavior of the sequential estimates of the model parameters in a case where the filter is presumably giving optimal, or near optimal, estimates. The plots were generated during the run of the filter program which gave the results in Table 4.3 for the portion of the relative displacement record from 15 to 25 seconds. The variation of the estimates during the latter part of the time segment show that there is interaction between the estimates

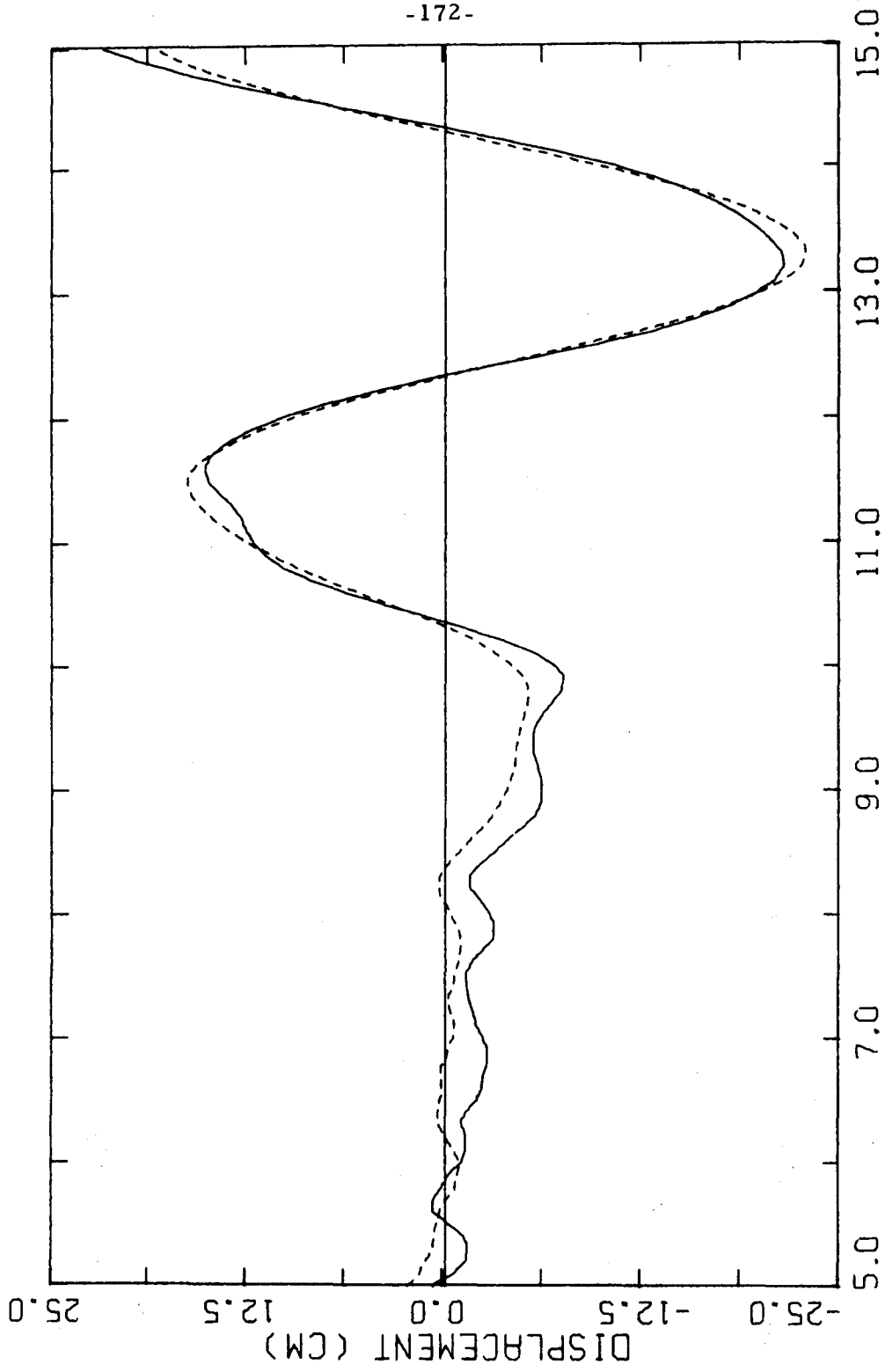


Figure 4.5. Optimal estimate (----) of the displacement of the fundamental longitudinal mode of the Union Bank building compared with the recorded relative displacement (——) at the 19th floor.

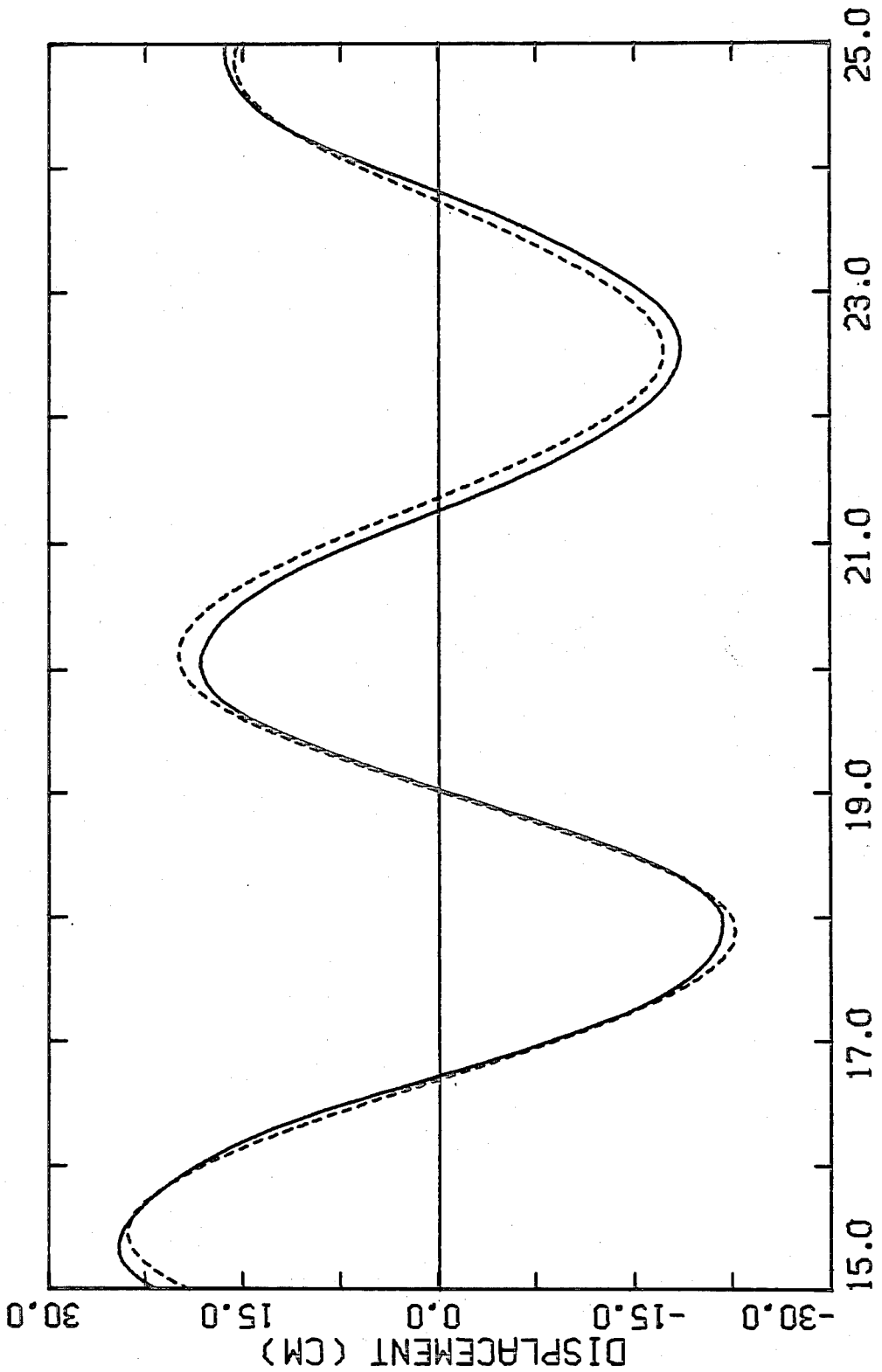


Figure 4.6. Optimal estimate (----) of the displacement of the fundamental longitudinal mode of the Union Bank building compared with the recorded relative displacement (——) at the 19th floor.

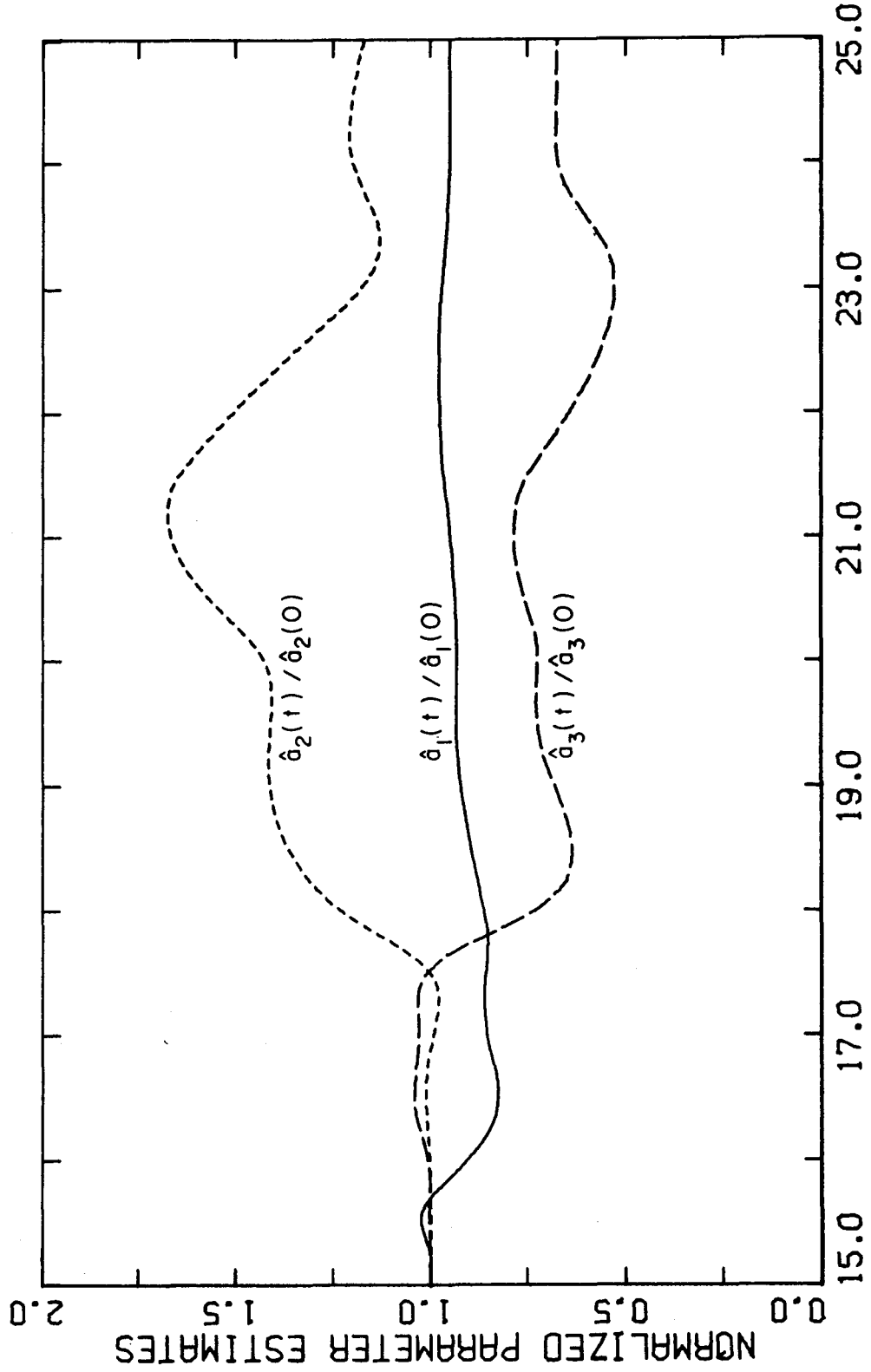


Figure 4.7. Sequential estimates of the model parameters for the fundamental longitudinal mode of the Union Bank building. The relative displacement record at the 19th floor was used to obtain the estimates.

of the damping coefficient a_2 and the participation factor a_3, these estimates tending to increase or to decrease together. Plots such as Fig. 4.7, together with the information given by the inverse sensitivity matrix $Q(t)$, are useful features of the filter method when it can be relied upon to produce optimal estimates.

It was found that rerunning the filter using the final estimates of the preceding run as the initial estimates of the new run led to some improvement in the estimates. This is illustrated in Table 4.4 for two additional runs of the filter. If the filter was truly optimal, repeating this process should produce convergence to the optimal estimates corresponding to J_0, since the initial estimates should eventually equal the final estimates and there should be no contribution to the minimum of J from the terms weighting the initial estimates. The results in Table 4.4 show that the approximate filter need not behave in this manner.

The observed behavior of the filter during the tests with simulated and real data is consistent with the expected behavior of the error due to neglecting the term containing E in (4.3.12). From (4.3.5) and (4.3.15), this error is initially zero. However, the terms $\left.\frac{\partial^3 J(\underline{c}, \tau)}{\partial c_i \partial c_j \partial c_k}\right|_{\underline{c} = \underline{e}(\tau)}$ can be expected to become nonzero because the model response is a nonlinear function of the parameters. Thus, $E(\tau)$ will also become nonzero and the approximation will deteriorate with

Value	$\hat{T}_1$	$\hat{\zeta}_1(\%)$	$\hat{p}_1$	$\hat{x}^{(1)}(0)$	$\hat{v}^{(1)}(0)$	Q_{33}	Q_{44}	Q_{55}	$J_0 \times 10^2$	$J \times 10^2$
Initial	4.5	3.0	0.7	0.70	-16.0	10^2	10^2	10^2		
Final	4.46	5.2	0.65	-1.6	3.0	0.20	80	2.1	2.9	3.4
Initial	4.46	5.2	0.65	-1.6	3.0	10^2	10^2	10^2		
Final	4.39	6.2	0.80	1.8	-7.9	0.25	63	3.7	0.56	0.73
Initial	4.39	6.2	0.80	1.8	-7.9	10^2	10^2	10^2		
Final	4.41	8.0	0.80	-0.58	-6.0	0.25	52	2.9	0.88	0.91
Optimal	4.40	6.5	0.84	2.1	-9.8				0.52	-

TABLE 4.4. Final estimates given by the filter for successive runs using the portion from 5 to 15 seconds of the relative displacement at the 19th floor of Union Bank building. The optimal estimates correspond to J_0.

time, unless the output of the exact filter converges rapidly to the recorded output. In this case, Eq. (4.3.11) rapidly approaches $\dot{\underline{e}} = \underline{g}$ and hence E quickly tends to zero. The term involving E in (4.3.12) should therefore never get the chance to grow relatively large and the approximate filter should have almost the same behavior as the exact filter. The conditions for good performance of the approximate filter therefore appear to be that the model output is capable of giving a good fit to the recorded output, and the weighting matrix A is small enough that convergence of the model output to the recorded output is not slowed too much by the influence of the initial estimates.

The following conclusion is suggested by the above discussion and by the tests of the filter: The method should produce optimal, or near optimal, estimates if the model is capable of matching the recorded data well, but it can be unreliable if the optimum output-error is large, either because of measurement noise or model error. Unfortunately, significant model error is a possibility when linear models are applied to identify structures using strong-motion records. Thus, following the tests of the filter method, it was felt that a more reliable technique was required for the desired applications. A technique was therefore developed which involves no approximations in the theory. It is both more reliable and more efficient numerically than the filter method, although the useful feature of obtaining sequential estimates of

the parameters is lost. The technique is introduced in the next chapter.

REFERENCES

Beck, J. L. and P. C. Jennings (1977). An optimal filter approach to identification in structural dynamics. Proc. Symp. on Applications of Computer Methods in Engineering, 1, 251-260, Univ. of Southern Calif., Los Angeles, California.

Bellman, R. E., H. H. Kagiwada, R. E. Kalaba and R. Sridhar (1966). Invariant imbedding and nonlinear filtering theory. J. Astronaut. Sci. 13, 110-115.

Detchmendy, D. M. and R. Sridhar (1966). Sequential estimation of states and parameters in noisy nonlinear dynamical systems. J. Basic Engng, Trans. ASME, 88, 362-368.

Distefano, N. and A. Rath (1974). Modeling and Identification in Nonlinear Structural Dynamics. Report No. EERC 74-15, Univ. of Calif., Berkeley, California.

Distefano, N. and B. Pena-Pardo (1976). System identification of frames under seismic loads. J. Engng Mech. Div., ASCE, 102, 313-330.

Jazwinski, A. H. (1970). Stochastic Processes and Filtering Theory. Academic Press, New York.

Kalman, R. E. and R. S. Bucy (1961). New results in linear filtering and prediction theory. J. Basic Engng, Trans. ASME 83, 95-108.

V. MODAL MINIMIZATION METHOD

5.1. Introduction

Difficulties encountered in the application of the optimal filter method to the earthquake response of the Union Bank building emphasized the need for a reliable technique which would be guaranteed to find a minimum of the measure-of-fit J while remaining numerically efficient. This led to the development of the modal minimization method for multi-degree-of-freedom linear models.

The method is an extension of a well-known iterative approach to minimizing a function $J(a_1, \ldots, a_n)$, whereby a series of one-dimensional minimizations are performed with respect to each a_i (Bekey, p. 157, 1970). In the basic approach, one first solves the problem $\min_{a_1} J(a_1, a_2^{(0)}, \ldots, a_n^{(0)})$ and then, with the minimizing argument denoted by $a_1^{(1)}$, $\min_{a_2} J(a_1^{(1)}, a_2, a_3^{(0)}, \ldots, a_n^{(0)})$ is found and so on. After one sweep through the parameters, giving new estimates $a_1^{(1)}$, $a_2^{(1)}$, $a_3^{(1)}$, $\ldots$, $a_n^{(1)}$ for the minimizing point of J, successive sweeps can be performed until convergence is achieved.

Bekey (1970) points out that this method is slow to converge if the axes of the contours of J near the minimum are not aligned closely with the axes of the parameters. However, for the present problem of estimating the modal parameters of linear structural models, the

properties of the models can be exploited to reduce this effect and to make other modifications of the basic method which lead to better numerical efficiency.

In the next section, the output-error approach is specialized for the class of problems treated in the applications. The modal minimization method is then described. Finally, the results of applying the method to simulated data are given.

5.2. Formulation of Problem

The modal minimization method is an output-error method for estimating the modal parameters of a linear model. The case considered here is that in which the input and output consist of one component of the base excitation and the parallel component of the response at some point in the structure respectively. The applications given later are of this kind, but the method is easily extended to include multiple inputs and multiple outputs.

The parameters to be estimated are the modal parameters $\underline{a}^{(r)} = [a_1^{(r)}, a_2^{(r)}, a_3^{(r)}, a_4^{(r)}, a_5^{(r)}]^t$, $r = 1, 2, \ldots, N$, where:

$$\ddot{x}^{(r)} + a_2^{(r)} \dot{x}^{(r)} + a_1^{(r)} x^{(r)} = -a_3^{(r)} \ddot{z}(t) \qquad (5.2.1)$$

and

$$x^{(r)}(T_i) = a_4^{(r)} \quad , \quad \dot{x}^{(r)}(T_i) = a_5^{(r)} \qquad (5.2.2)$$

This is the model given in §3.5.2 with:

$$a_1^{(r)} = \omega_r^2 \quad , \quad a_2^{(r)} = 2\zeta_r \omega_r \quad , \quad a_3^{(r)} = \beta_i^{(r)} \qquad (5.2.3)$$

Also, the subscript i on each $x_i^{(r)}$ has been omitted because the response of only one coordinate is used.

The recorded output is assumed to consist of the history of

$$\underline{y} = [k_1 x_0, k_2 v_0, k_3 a_0]^t \tag{5.2.4}$$

over some interval $[T_i, T_f]$. Thus, any combination of the displacement, velocity or acceleration records of one component of the structural response at a point can be used by choosing each k_i as either 1 or 0, as in §4.4. The corresponding model output is:

$$\underline{m} = [k_1 x, k_2 \dot{x}, k_3 \ddot{x}]^t \tag{5.2.5}$$

where, from (3.5.2):

$$x(t;\underline{a}^{(1)}, \ldots, \underline{a}^{(N)}) = \sum_{r=1}^{N} x^{(r)}(t;\underline{a}^{(r)}) \tag{5.2.6}$$

The measure-of-fit to be used for the output matching is given by Eqs. (2.2.3), (2.2.4), (5.2.4) and (5.2.5):

$$\begin{aligned} J(\underline{a}^{(1)}, \underline{a}^{(2)}, \ldots, \underline{a}^{(N)}) = {} & k_1 V_{11} \int_{T_i}^{T_f} (x_0 - x)^2 dt \\ & + k_2 V_{22} \int_{T_i}^{T_f} (v_0 - \dot{x})^2 dt \\ & + k_3 V_{33} \int_{T_i}^{T_f} (a_0 - \ddot{x})^2 dt \end{aligned} \tag{5.2.7}$$

This has to be minimized subject to Eqs. (5.2.1), (5.2.2) and (5.2.6). Notice that the weighting matrix A in (2.2.4) has been set to zero and V(t) has been taken diagonal and constant, although a more general case could also be treated if desired. The V_{ii} were chosen to normalize each integral as in (4.4.9). This allows a comparison to be made between the optimal values of J for different time segments and for different response quantities.

Some results of interest can be derived by using the notation of §2.3, so that:

$$J(\underline{a}^{(1)}, \underline{a}^{(2)}, \ldots, \underline{a}^{(N)}) = \langle \underline{v}, \underline{v} \rangle \qquad (5.2.8)$$

where $$\underline{v} = \underline{y} - \underline{m}(\underline{a}^{(1)}, \ldots, \underline{a}^{(N)}; z) \qquad (5.2.9)$$

and $\underline{y}$ and $\underline{m}$ are given by Eqs. (5.2.4) and (5.2.5). At the global minimum $\underline{a}^{(r)} = \hat{\underline{a}}^{(r)}$, $r = 1, \ldots, N$, of J, $\nabla J = 0$ and hence from (2.3.7):

$$\langle \hat{\underline{v}}, \frac{\partial \hat{\underline{m}}}{\partial a_k^{(r)}} \rangle = 0 \ , \ k = 1, 2, \ldots, 5 \text{ and } r = 1, 2, \ldots, N \qquad (5.2.10)$$

The sensitivity coefficients are therefore orthogonal to the output-error at the optimal estimates of the parameters. This is true for any model, linear or nonlinear, but an additional result can be derived for linear models. From Eqs. (5.2.5) and (5.2.6), and from Eqs. (5.3.7) and (5.3.8) given in the next section:

$$\underline{m} = \sum_{r=1}^{N} \sum_{k=3}^{5} a_k^{(r)} \frac{\partial \underline{m}}{\partial a_k^{(r)}} \qquad (5.2.11)$$

so that (5.2.10) implies:

$$<\hat{\underline{v}}, \hat{\underline{m}}> = 0 \qquad (5.2.12)$$

The output of the optimal model and the optimal output-error are therefore orthogonal under the scalar product $<\cdot,\cdot>$. For example, if only displacement matching is considered ($k_1=1$, $k_2=0$, $k_3=0$), then:

$$\int_{T_i}^{T_f} (x_0 - \hat{x})\hat{x}\, dt = 0$$

5.3. Minimization Method

The method used to minimize $J(\underline{a}^{(1)},\ldots,\underline{a}^{(N)})$ is described in four parts: modal sweeps, single-mode minimization, one-dimensional (1-D) minimization, and numerical evaluation of J.

1) Modal sweeps

Initial estimates are made for the $\underline{a}^{(r)}$, $r=1,\ldots,N$, then J is minimized with respect to $\underline{a}^{(1)}$, the parameters of the first mode, while the parameters of the other modes are held fixed. Using the new estimate of $\underline{a}^{(1)}$ and the initial estimates of the other $\underline{a}^{(r)}$, J is then minimized with respect to $\underline{a}^{(2)}$. By estimating one mode at a time in this manner, new estimates of the modal parameters are available after one sweep of the N modes. Successive sweeps are

performed until J is no longer reduced significantly. From three to five modal sweeps were usually sufficient to give adequate convergence to the minimum of J.

The method may be summarized as consisting of repeated applications of the following sequence of minimizations:

$$J(\hat{\underline{a}}^{(1)}, \underline{a}^{(2)}, \ldots, \underline{a}^{(N)}) = \min_{\underline{a}^{(1)}} J(\underline{a}^{(1)}, \underline{a}^{(2)}, \ldots, \underline{a}^{(N)})$$

$$J(\hat{\underline{a}}^{(1)}, \hat{\underline{a}}^{(2)}, \ldots, \underline{a}^{(N)}) = \min_{\underline{a}^{(2)}} J(\hat{\underline{a}}^{(1)}, \underline{a}^{(2)}, \ldots, \underline{a}^{(N)})$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$J(\hat{\underline{a}}^{(1)}, \hat{\underline{a}}^{(2)}, \ldots, \hat{\underline{a}}^{(N)}) = \min_{\underline{a}^{(N)}} J(\hat{\underline{a}}^{(1)}, \hat{\underline{a}}^{(2)}, \ldots, \underline{a}^{(N)})$$

Notice that this method of minimizing J can be considered as an application of the basic approach outlined in §5.1 except that it is at the modal level instead of at the level of each parameter. The procedure used to minimize J with respect to a given mode is described in the next part.

2) Single-mode minimization

Consider the stage in a modal sweep where J is to be minimized with respect to the parameters of the r^{th} mode. From (5.2.6) and (5.2.7), it follows that the latest estimates of the parameters of the other modes are used to subtract all but the r^{th} mode from the recorded response. The remaining portion of the response is then used to determine the new estimate of $\underline{a}^{(r)}$. Thus, the minimization of J with respect to $\underline{a}^{(r)}$ is equivalent to minimizing the function:

$$J_r(\underline{a}^{(r)}) = k_1 V_{11} \int_{T_i}^{T_f} \left[x_0^{(r)}(t) - x^{(r)}(t;\underline{a}^{(r)}) \right]^2 dt$$

$$+ k_2 V_{22} \int_{T_i}^{T_f} \left[v_0^{(r)}(t) - \dot{x}^{(r)}(t;\underline{a}^{(r)}) \right]^2 dt \qquad (5.3.1)$$

$$+ k_3 V_{33} \int_{T_i}^{T_f} \left[a_0^{(r)}(t) - \ddot{x}^{(r)}(t;\underline{a}^{(r)}) \right]^2 dt$$

subject to the constraints of (5.2.1) and (5.2.2). Here,

$$x_0^{(r)} = x_0 - \sum_{\substack{s=1 \\ s \neq r}}^{N} x^{(s)}$$

$$v_0^{(r)} = v_0 - \sum_{\substack{s=1 \\ s \neq r}}^{N} \dot{x}^{(s)} \qquad (5.3.2)$$

$$a_0^{(r)} = a_0 - \sum_{\substack{s=1 \\ s \neq r}}^{N} \ddot{x}^{(s)}$$

are all known quantities, because the modal contributions are given by the latest estimates of the $\underline{a}^{(s)}$, $s \neq r$.

The linearity of the model can be exploited in the minimization of $J_r(\underline{a}^{(r)})$ to enable the parameters $a_3^{(r)}$, $a_4^{(r)}$ and $a_5^{(r)}$ to be determined explicitly in terms of $a_1^{(r)}$ and $a_2^{(r)}$. Define a linear operator by:

$$L^{(r)} \equiv \frac{d^2}{dt^2} + a_2^{(r)} \frac{d}{dt} + a_1^{(r)} \qquad (5.3.3)$$

and define the functions $s_k^{(r)}(t;a_1^{(r)},a_2^{(r)})$, $k=3,4,5$, by:

$$L^{(r)}s_3^{(r)} = -\ddot{z}(t)\ ,\ s_3^{(r)}(T_i)=0\ ,\ \dot{s}_3^{(r)}(T_i)=0 \qquad (5.3.4)$$

$$L^{(r)}s_4^{(r)} = 0\ ,\ s_4^{(r)}(T_i)=1\ ,\ \dot{s}_4^{(r)}(T_i)=0 \qquad (5.3.5)$$

$$L^{(r)}s_5^{(r)} = 0\ ,\ s_3^{(r)}(T_i)=0\ ,\ \dot{s}_5^{(r)}(T_i)=1 \qquad (5.3.6)$$

From these definitions, and the linearity of Eqs. (5.2.1) and (5.2.2):

$$x^{(r)}(t;\underline{a}^{(r)}) = \sum_{k=3}^{5} a_k^{(r)} s_k^{(r)}(t;a_1^{(r)},a_2^{(r)}) \qquad (5.3.7)$$

and

$$s_k^{(r)} = \frac{\partial x^{(r)}}{\partial a_k^{(r)}}\ ,\ k=3,4,5 \qquad (5.3.8)$$

The term $a_3^{(r)}s_3^{(r)}$ is the forced vibration component of the modal response $x^{(r)}$ and $a_4^{(r)}s_4^{(r)}+a_5^{(r)}s_5^{(r)}$ is the free vibration component due to the initial conditions. For <u>fixed</u> $a_1^{(r)}$ and $a_2^{(r)}$, it follows from (5.3.1), (5.3.7) and (5.3.8) that the global minimum of $J_r(\underline{a}^{(r)})$ is given by the solution for $a_3^{(r)}$, $a_4^{(r)}$ and $a_5^{(r)}$ of the linear system of equations:

$$\sum_{k=3}^{5} b_{jk}^{(r)} a_k^{(r)} = c_j^{(r)}\ ,\ j=3,4,5 \qquad (5.3.9)$$

where $$b_{jk}^{(r)} = k_1 V_{11}\int_{T_i}^{T_f} s_j^{(r)} s_k^{(r)}\,dt + k_2 V_{22}\int_{T_i}^{T_f} \dot{s}_j^{(r)} \dot{s}_k^{(r)}\,dt$$

$$+ k_3 V_{33}\int_{T_i}^{T_f} \ddot{s}_j^{(r)} \ddot{s}_k^{(r)}\,dt \qquad (5.3.10)$$

and $$c_j^{(r)} = k_1 V_{11} \int_{T_i}^{T_f} s_j^{(r)} x_0^{(r)}\, dt + k_2 V_{22} \int_{T_i}^{T_f} \dot{s}_j^{(r)} v_0^{(r)}\, dt$$

$$+ k_3 V_{33} \int_{T_i}^{T_f} \ddot{s}_j^{(r)} a_0^{(r)}\, dt \qquad (5.3.11)$$

There is a unique solution to the equations given by (5.3.9) since the 3x3 matrix $[b_{jk}^{(r)}]$ is non-singular. This follows from the fact that this matrix corresponds to the matrix $\tilde{S}$ defined by (2.3.9), which, according to Appendix A, is positive definite because the sensitivity coefficients $s_3^{(r)}$, $s_4^{(r)}$ and $s_5^{(r)}$ are clearly linearly independent in view of (5.3.7), (5.2.1) and (5.2.2).

By using the solution to (5.3.9) for any given $a_1^{(r)}$ and $a_2^{(r)}$, a function f_r can be defined by:

$$f_r(a_1^{(r)}, a_2^{(r)}) = \min_{(a_3^{(r)}, a_4^{(r)}, a_5^{(r)})} J_r(\underline{a}^{(r)}) \qquad (5.3.12)$$

The original problem of solving $\min_{\underline{a}^{(r)}} J_r(\underline{a}^{(r)})$ therefore reduces to finding the minimum of f_r. This is achieved by applying the basic iterative approach outlined in §5.1. Thus, a series of 1-D minimizations are performed by minimizing f_r alternately with respect to $a_1^{(r)}$ and with respect to $a_2^{(r)}$. This process, which is indicated schematically in Fig. 5.1, is continued until a consecutive pair of 1-D minimizations results in a fractional decrease in f_r of less than ϵ, where ϵ is specified. The procedure used to carry out each 1-D minimization is described in the next part.

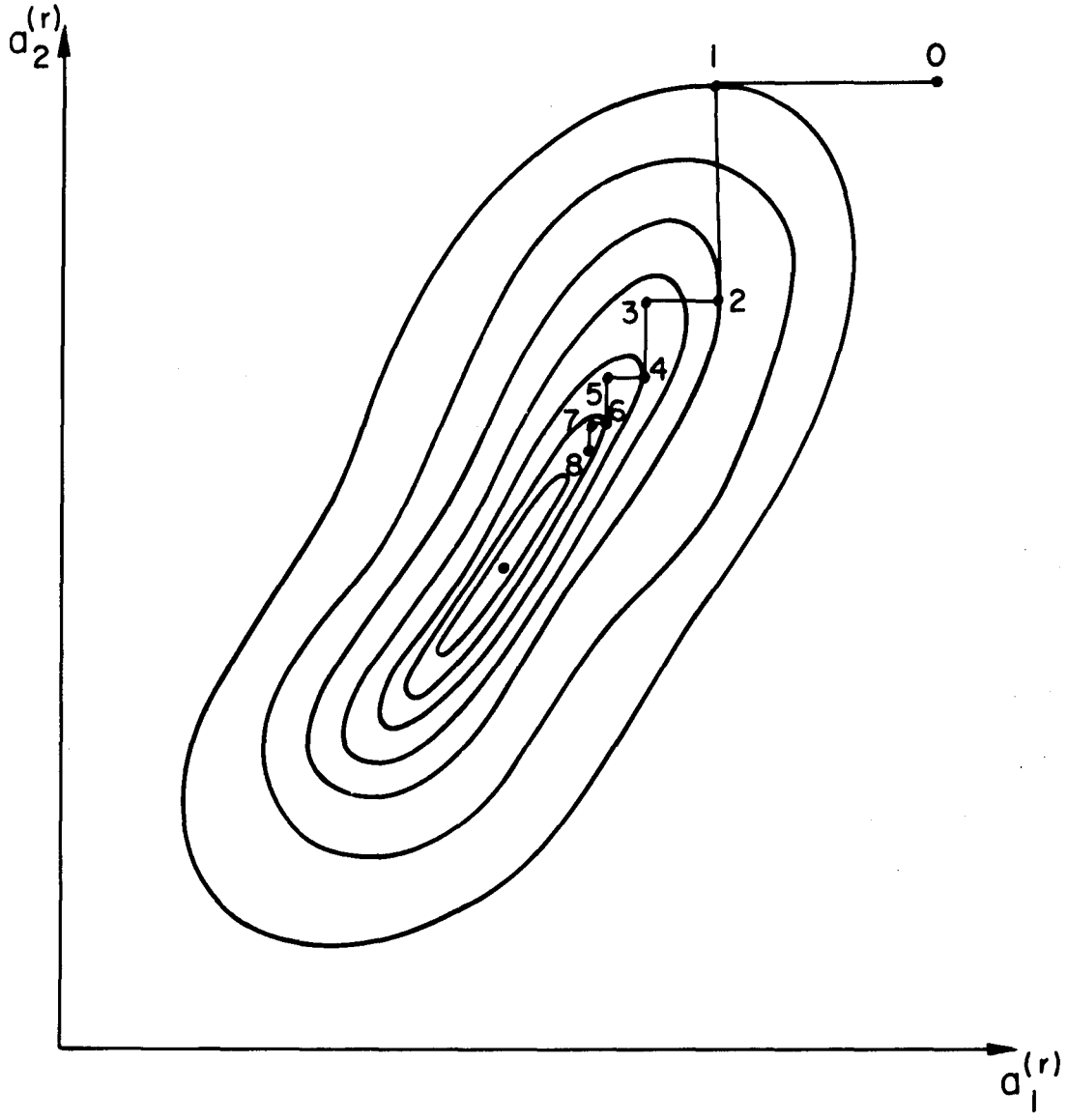


Figure 5.1. Schematic diagram of contours of $f_r(a_1^{(r)}, a_2^{(r)})$ [Eq. (5.3.12)] showing a path of convergence.

The criterion for convergence in terms of the relative change in f_r was chosen instead of the change in the estimates of the parameters $a_1^{(r)}$ and $a_2^{(r)}$ because the latter can cause difficulties with the higher modes. The response may be so insensitive to these parameters that the changes in f_r reach the roundoff level of the computer before the parameters have converged to the specified accuracy, and this may prevent the algorithm from "converging". The criterion in terms of f_r, on the other hand, automatically takes into account the fact that the resolution differs from one parameter to another. In the results reported later, ϵ was set equal to 10^{-4} in the criterion for convergence.

The parameters $a_1^{(r)}$ and $a_2^{(r)}$ can be replaced by the modal period, T_r, and damping factor, ζ_r, in the algorithm to minimize f_r so that the 1-D minimizations are with respect to these latter parameters. This was done in the applications of the method because of the convenience of working directly with the parameters of interest.

3) <u>One-dimensional minimization</u>

A method is given for the minimization of a function $f(\alpha)$ of a single parameter α. The minimization of $f_r(a_1^{(r)}, a_2^{(r)})$ is then achieved by applying the method alternately to $f(\alpha) \equiv f_r(\alpha, a_2^{(r)})$ and to $f(\alpha) \equiv f_r(a_1^{(r)}, \alpha)$. Notice that the gradient of $f_r(a_1^{(r)}, a_2^{(r)})$ is not required to determine the direction of the search for the minimum of f_r because this direction is always parallel to one of the axes of the parameters. A method to minimize $f(\alpha)$ was therefore selected which only involved evaluating the function.

The method starts with an initial estimate, α_0, of the minimum of $f(\alpha)$, and a step-size $\delta > 0$. If $f(\alpha_0 - \delta) \geq f(\alpha_0)$, then α is incremented continually by δ, that is, $\alpha = \alpha_0 + \delta$, $\alpha_0 + 2\delta, \ldots$, but if $f(\alpha_0 - \delta) < f(\alpha_0)$, then α is decremented continually by δ. The value of f is calculated at each step and the stepping is continued until $f(\alpha)$ no longer decreases, showing the minimum has been passed. The minimizing argument for a parabola through the last three values of f is then calculated and this is taken as the minimizing argument of f.

A key point is the choice of δ. If it is too large, the estimate of the minimum of $f(\alpha)$ will be poor and an excessive amount of time may be spent iterating on $a_1^{(r)}$ and $a_2^{(r)}$. On the other hand, it if is too small, too much time may be spent in the stepping required to find the minimum of $f(\alpha)$. A choice which was found to work well is the following. If the 1-D minimization is with respect to $a_1^{(r)}$, δ is taken to be half the change that occurred in the estimates of $a_1^{(r)}$ given by the last two 1-D minimizations in this direction. A similar choice is made for the value of δ for the 1-D minimizations with respect to $a_2^{(r)}$. To illustrate this choice of δ, for the minimization determining $a_1^{(r)}$ to locate point 5 in Fig. 5.1, δ would be half the difference between the values of $a_1^{(r)}$ at points 1 and 3. To start the process during each modal sweep, δ is given an initial value $\delta_1^{(r)}$ while determining point 1 and for point 3, δ is set equal to $\frac{1}{2}\delta_1^{(r)}$. Similarly, an initial value $\delta_2^{(r)}$ is used to start the 1-D minimizations with respect to $a_2^{(r)}$. For the first modal sweep, $\delta_1^{(r)}$ and $\delta_2^{(r)}$ are prescribed for each mode

and they are then halved for each successive modal sweep.

4) Evaluation of J

The only calculations required for minimizing J are those involved in evaluating $f_r(a_1^{(r)}, a_2^{(r)})$. The algorithm used is discussed briefly here.

At the beginning of the minimization of J with respect to the parameters of the r^{th} mode, the current contribution of this mode to the response is added to the current output-error to form whichever of the quantities in Eqs. (5.3.2) are required. These quantities are then kept fixed during the estimation of $\underline{a}^{(r)}$.

To evaluate $f_r(a_1^{(r)}, a_2^{(r)})$ for a given $a_1^{(r)}$ and $a_2^{(r)}$, Eqs. (5.3.4) to (5.3.6) are first solved. The method used to solve all three linear equations is the efficient transition-matrix approach introduced by Nigam and Jennings (1969). This turns out to give the exact solutions at each time step, apart from the roundoff in the arithmetic calculations, because of the way the continuous record $\ddot{z}(t)$ is defined as a linear interpolation of the digitized data. Equations (5.3.9) are then solved for $a_3^{(r)}, a_4^{(r)}$ and $a_5^{(r)}$ using Gaussian elimination, and the new estimate of the contribution of the r^{th} mode is calculated from (5.3.7). The value of $f_r(a_1^{(r)}, a_2^{(r)})$ is then calculated from (5.3.1) using Simpson's rule for numerical integration. This procedure for evaluating $f_r(a_1^{(r)}, a_2^{(r)})$ is repeated for each pair $(a_1^{(r)}, a_2^{(r)})$ arising in the

minimization of f_r by the method described in parts 2) and 3).

5.3.1 Comments on Method

The modal minimization method has several advantages in comparison with the optimal filter method of Chapter 4. The most important advantage is its reliability. Convergence to a local minimum occurred in every application of the method. Another advantage is that its convergence is easy to examine by assessing J and the current estimates of the parameters of the r^{th} mode every time J, or equivalently, $f_r(a_1^{(r)}, a_2^{(r)})$, is evaluated. In contrast, it is difficult to determine whether the invariant-imbedding filter has given the optimal estimates of the parameters because the effects of the approximation in the theory are not easy to assess in a particular case.

Part 2) of the method is similar to the method used by Raggett (1974). He examines one mode at a time by filtering the response using a narrow band-pass filter centered on the estimated modal frequency. A single-degree-of-freedom model is then used to estimate the modal parameters from the filtered response. The advantage of the present method in comparison to this approach is that the properties of a number of modes are estimated simultaneously by the minimization of J, so that modal interference can be reduced to an acceptable level. Raggett's results using simulated data for a linear system with three

degrees of freedom show that modal interference has a significant effect on the accuracy of the estimates when using his approach. Using a three-degree-of-freedom model, the present approach is capable of giving the exact values of the parameters from the simulated data used by Raggett.

An interesting feature of the method for finding the minimum of the function $f_r(a_1^{(r)}, a_2^{(r)})$ is that after the first 1-D minimization has been performed, it is equivalent to the method of steepest descent. In the latter approach, the gradient of f_r would be evaluated to determine the direction of steepest descent and then a 1-D minimization would be performed in this direction. Referring to Fig. 5.1, for the method of part 2) the line from point 0 to point 1 must be tangential to the contour of f_r at point 1. The next 1-D minimization in the direction of $a_2^{(r)}$ is therefore in a direction normal to the contour of f_r at point 1, which is the direction of steepest descent. All subsequent directions of search for the minimum also behave in this manner, proving the original assertion.

It is well known that the rate of convergence of the method of steepest descent can be very slow if there is significant interaction between the parameters near the minimum. An illustration of this problem can be seen in Fig. 5.1, where the interaction, or lack of "orthogonality" (see §2.4.5), between $a_1^{(r)}$ and $a_2^{(r)}$ has been

exaggerated. In the applications of the method, the interaction between T_r and ζ_r was not pronounced. The few times that convergence with respect to T_r and ζ_r was slow were always cases where the model was having difficulty interpreting the data for the higher modes, which was reflected in unreasonable values returned for some of the modal parameters.

The two parameters $a_2^{(r)}$ and $a_3^{(r)}$ would cause slow convergence if $J_r(\underline{a}^{(r)})$ was minimized by applying the method of steepest descent to all five parameters in $\underline{a}^{(r)}$, because there is strong interaction between these two parameters. The interaction arises because the major effect of both $a_2^{(r)}$ and $a_3^{(r)}$ is to alter the amplitude of the contribution of the r^{th} mode to the response. The effect of this interaction is reduced by using the explicit method to determine the minimizing value of $a_3^{(r)}$.

Another source of interaction is that between the modes, which could cause slow convergence during the modal sweeps. However, this would not be expected to be a problem if the modes have widely-spaced modal frequencies. The applications support this conclusion.

5.4. Tests with Simulated Data

5.4.1 Single Degree-of-freedom Linear Oscillator

The modal minimization method was first tested using the same

model and data as in §4.4.1. The linear oscillator used to generate the data numerically had values for its parameters of $T = 1.0$ sec, $\zeta = 5\%$ and $p = 1.0$. It was found that the method gave nearly exact values for the parameters from two cycles of data, regardless of whether the displacement, velocity or acceleration of the oscillator were used in the measure-of-fit J. One cycle and a half-cycle of the displacement were also used and were found to give nearly exact results. The final estimates of the parameters are shown in Table 5.1., together with the initial estimates. Because of the way the method works, initial estimates of the participation factor, initial displacement and initial velocity were not required. The portions of the excitation and response which were used can be seen by examining Figs. 4.1 and 4.3. Of course, the results are for a special case in which the only source of error is roundoff in the computations.

Time segment (seconds)	Estimates of Parameters		
	$\hat{T}$	$\hat{\zeta}(\%)$	$\hat{p}$
2.0 - 4.0	1.00000	4.9999	1.0000
3.0 - 4.0	1.00000	4.9994	0.9999
3.0 - 3.5	1.00000	4.9989	0.9998
Initial estimates	1.12	3.82	-
True values	1.0	5.0	1.0

TABLE 5.1. Optimal estimates of the parameters using different portions of the displacement of the linear oscillator.

5.4.2. Ten Degree-of-freedom Linear Chain System

The next tests of the modal minimization method used the "roof" response computed for a ten-degree-of-freedom linear chain system (Fig. 3.1) which was initially at rest and then subjected to a base acceleration given by the first 10 seconds of the 1940 El Centro earthquake record (Fig. 4.1). The modal properties of the uniform chain system are given in Table 5.2. The modal participation factor p_r is the quantity $a_3^{(r)}$ [Eq. (5.2.1)] or $\beta_{10}^{(r)}$ [Eq. (3.2.13)] and it is independent of the normalization of the modeshapes. If the modeshapes are normalized to unity at the roof, p_r is equal to the conventional participation factor. Also shown in Table 5.2 are some of the peak modal contributions to the relative displacement, velocity and

Mode	T_r	ζ_r(%)	p_r	$\lvert x^{(r)} \rvert_{max}$ (cm)	$\lvert \dot{x}^{(r)} \rvert_{max}$ (cm/sec)	$\lvert \ddot{x}^{(r)} \rvert_{max}$ (g)
1	1.0000	5	1.2673	16.0 (100%)	117 (100%)	0.66 (100%)
2	0.3358	5	-0.4068	0.76 (5%)	13.5(12%)	0.26 (39%)
3	0.2045	5	0.2259	0.15(1.0%)	4.1(3.5%)	0.14 (21%)
4	0.1495	5	-0.1429	0.05(0.3%)	1.8(1.5%)	0.08 (12%)
5	0.1199	5	0.0934		0.89(0.8%)	0.06 (9%)
6	0.1019	5	-0.0601			0.03 (5%)
7	0.0904	5	0.0366			0.02 (3%)
8	0.0829	5	-0.0199			0.01 (1.5%)
9	0.0782	5	0.0087			
10	0.0756	5	-0.0021			

TABLE 5.2. Modal properties of the ten-degree-of-freedom uniform chain system used in the tests.

acceleration at the roof.

Some of the results of applying the modal minimization method to various portions of the simulated response records are shown in Table 5.3, 5.4 and 5.5. The parameters and initial conditions were estimated for each mode included in a model but only the results for the modal parameters are given. The measure-of-fit J for each model is also given. Recall that because of the normalization of J by the V_{ii} in Eq. (5.2.7), $J^{\frac{1}{2}}$ represents the ratio of the r.m.s. output-error to the maximum response.

The errors in the parameters are primarily due to the model error created by neglecting the higher modes of the chain system in each model. For a given number of modes and for a given time segment, this model error is greatest when the acceleration is used and hence the errors in the estimates tend to be the largest in this case. There is also "measurement noise" because the equations of motion for the uniform chain system were solved only to within an accuracy of 1% of the exact response. This noise may therefore affect the accuracy of those modes whose signals are relatively small.

Observe that the modal periods are always estimated very accurately, at least for the six modes investigated, and the damping factor and participation factor are estimated quite accurately for each mode in a model except for the highest mode, which is most affected by

Modal parameter		Record used to estimate parameters						
		Displacement		Velocity		Acceleration		
		0-10 sec	2-4 sec	0-10 sec	2-4 sec	0-10 sec	2-6 sec	2-4 sec
(a)	$\hat{T}_1$	-0.02	-0.08	0.000	-0.3	-0.09	-0.4	-1.3
	$\hat{\zeta}_1$	0.06	-0.7	0.7	-10	-3	-5	-50
	$\hat{p}_1$	0.4	-0.5	1	-4	-3	-8	-30
$J \times 10^4$		1.8	5.9	10.4	51.6	96.4	187	337
(b)	$\hat{T}_1$	0.001	-0.01	0.000	-0.04	0.01	0.1	0.2
	$\hat{\zeta}_1$	-0.02	0.1	-0.1	-2	0.4	1	0.2
	$\hat{p}_1$	-0.06	0.1	-0.2	-1	0.4	2	2
	$\hat{T}_2$	-0.1	-0.3	-0.09	-0.2	-0.2	-0.6	-0.4
	$\hat{\zeta}_2$	-7	-2	-6	-2	-11	-12	-4
	$\hat{p}_2$	-4	0.2	-2	-0.01	-6	-10	-6
$J \times 10^6$		7.5	23	119	532	2650	5330	8750

TABLE 5.3. Relative errors (%) in the optimal parameter estimates using (a) one-mode model, and (b) two-mode model, of the uniform chain system of Table 5.2.

Modal parameter	Record used to estimate parameters				
	Velocity		Acceleration		
	0-10 sec	2-4 sec	0-10 sec	2-6 sec	2-4 sec
$\hat{T}_1$	-0.001	0.000	-0.004	0.001	0.03
$\hat{\zeta}_1$	-0.04	-0.1	-0.2	-0.2	2
$\hat{p}_1$	-0.05	-0.1	-0.2	-0.2	0.9
$\hat{T}_2$	0.006	-0.07	0.02	-0.04	-0.1
$\hat{\zeta}_2$	-0.04	0.3	0.2	0.02	0.2
$\hat{p}_2$	0.2	1	0.5	-0.07	0.4
$\hat{T}_3$	0.01	-0.01	0.01	0.04	-0.01
$\hat{\zeta}_3$	-1	-3	-2	-4	-4
$\hat{p}_3$	-1	-6	-2	-5	-7
$\hat{T}_4$	0.2	0.02	0.2	0.1	-0.1
$\hat{\zeta}_4$	-13	-18	-15	-13	-26
$\hat{p}_4$	-10	-9	-11	-10	-14
$J \times 10^6$	2.5	10	140	260	420

TABLE 5.4. Relative errors (%) in the optimal parameter estimates using a four-mode model of the uniform chain system of Table 5.2.

Mode	Five-mode model			Six-mode model		
	$\hat{T}_r$	$\hat{\zeta}_r$	$\hat{p}_r$	$\hat{T}_r$	$\hat{\zeta}_r$	$\hat{p}_r$
1	1.0000	5.001	1.2677	1.0000	5.0001	1.2674
2	0.3358	4.994	-0.4065	0.3358	4.9952	-0.4066
3	0.2046	5.009	0.2267	0.2045	4.9996	0.2260
4	0.1495	4.84	-0.138	0.1495	4.969	-0.1420
5	0.1204	3.74	0.072	0.1200	4.83	0.089
6				0.1024	3.18	-0.037

TABLE 5.5. Optimal estimates of the modal parameters for a five-mode model and a six-mode model of the uniform chain system of Table 5.2. The first ten seconds of the acceleration were matched.

the model error. Observe also that there is some interaction between the estimates of the damping factor and participation factor of each mode. With few exceptions, they are consistently both too large or both too small in magnitude, and often by roughly the same percentage. This might be expected since the height of the resonant peaks in the amplitude of the transfer function between the base and the roof are controlled by the ratio $p_r/2\zeta_r$. The difference between the effects of p_r and ζ_r is that the former scales the forced vibration component of the modal response uniformly in time, whereas the latter has an accumulative effect with time on both the forced and free vibration components.

The initial estimates for the results given in the Tables for the model with (R+1) modes were taken to be the optimal estimates for the model with R modes. For the new mode, initial estimates are required only for the period and damping factor if the modal sweeps are started with this mode instead of with the first mode. The initial estimate for the period of the first mode was 1.12 sec and the initial estimate for the damping factor of all the modes was 3.82%. These values were chosen to give unrounded numbers for initial errors. For the higher modes, the period ratios for a uniform shear beam were used to give the initial estimates of the periods.

With these choices of the initial estimates, it was found that the number of modal sweeps required to give convergence of J to within

0.01% was normally three to five sweeps for displacement matching and velocity matching and three to seven sweeps for acceleration matching. The number of modal sweeps required tended to increase as the length of the record decreased. The computer time per sweep for ten seconds of data ranged from 4 seconds for a two-mode model to 14 seconds for a six-mode model, using an IBM 370/158. These times are also representative for the applications reported in Chapter 6 which used earthquake records from buildings.

Figure 5.2 shows that the plots of the displacement at the roof of the uniform chain system and the displacement of an optimal model with two modes are indistinguishable. The optimal model is given by the parameters in the first column of Table 5.3. Figures 5.3 and 5.4 show a similar match of the velocity and acceleration is achieved by an optimal model with four modes (see Table 5.4, column 1).

An illustration is given in Fig. 5.5 of a profile of the measure-of-fit J computed for a single-mode model by varying the modal period T while keeping the damping factor constant at $\zeta = 5\%$. The acceleration record from 2.0 to 4.0 seconds was used in J and the plot is strictly one of $f_1(T, \zeta)$ [Eq. (5.3.12)]. Local minima corresponding to the first four modes can be observed. The local minimum at a period of about $T = 0.6$ sec is a spurious one due to the interaction of the first and second modes, that is, a single-mode model with a period

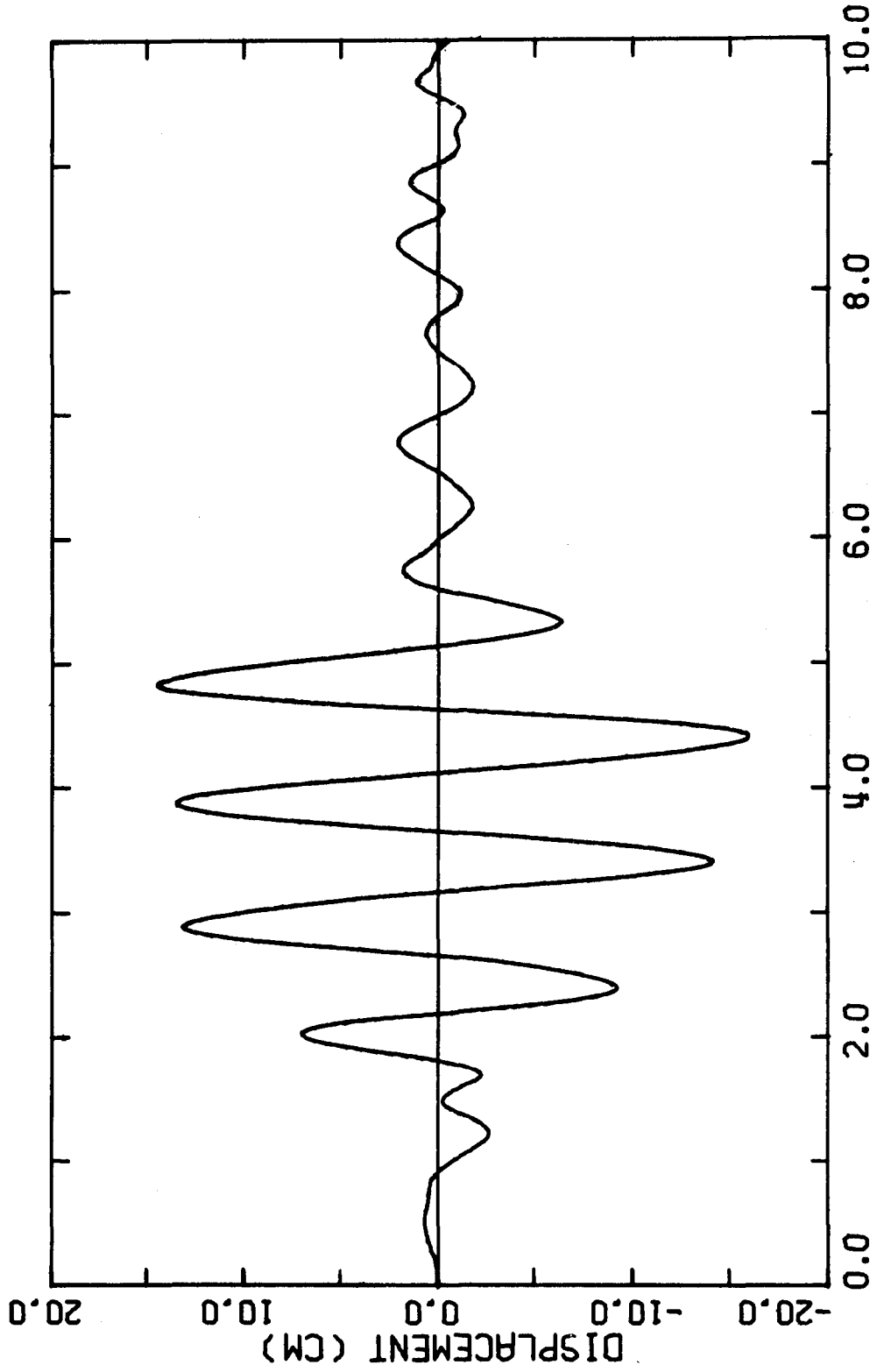


Figure 5.2. Relative displacement (——) at the roof of the uniform chain system and relative displacement (----) of the optimal model with two modes which was determined by matching displacements. The two curves are indistinguishable.

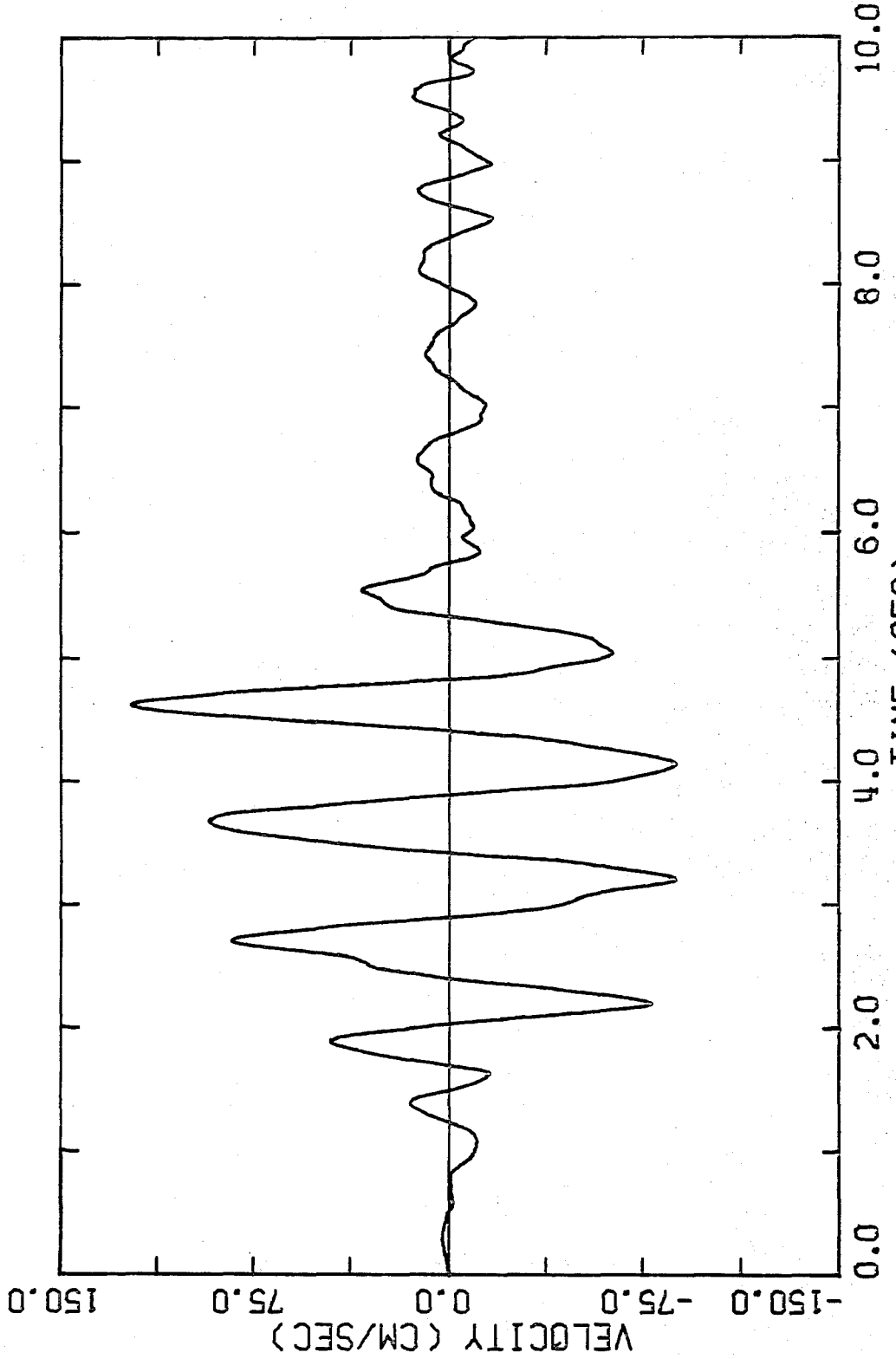


Figure 5.3. Relative velocity (——) at the roof of the uniform chain system and relative velocity (----) of the optimal model with four modes which was determined by matching velocities. The two curves are indistinguishable.

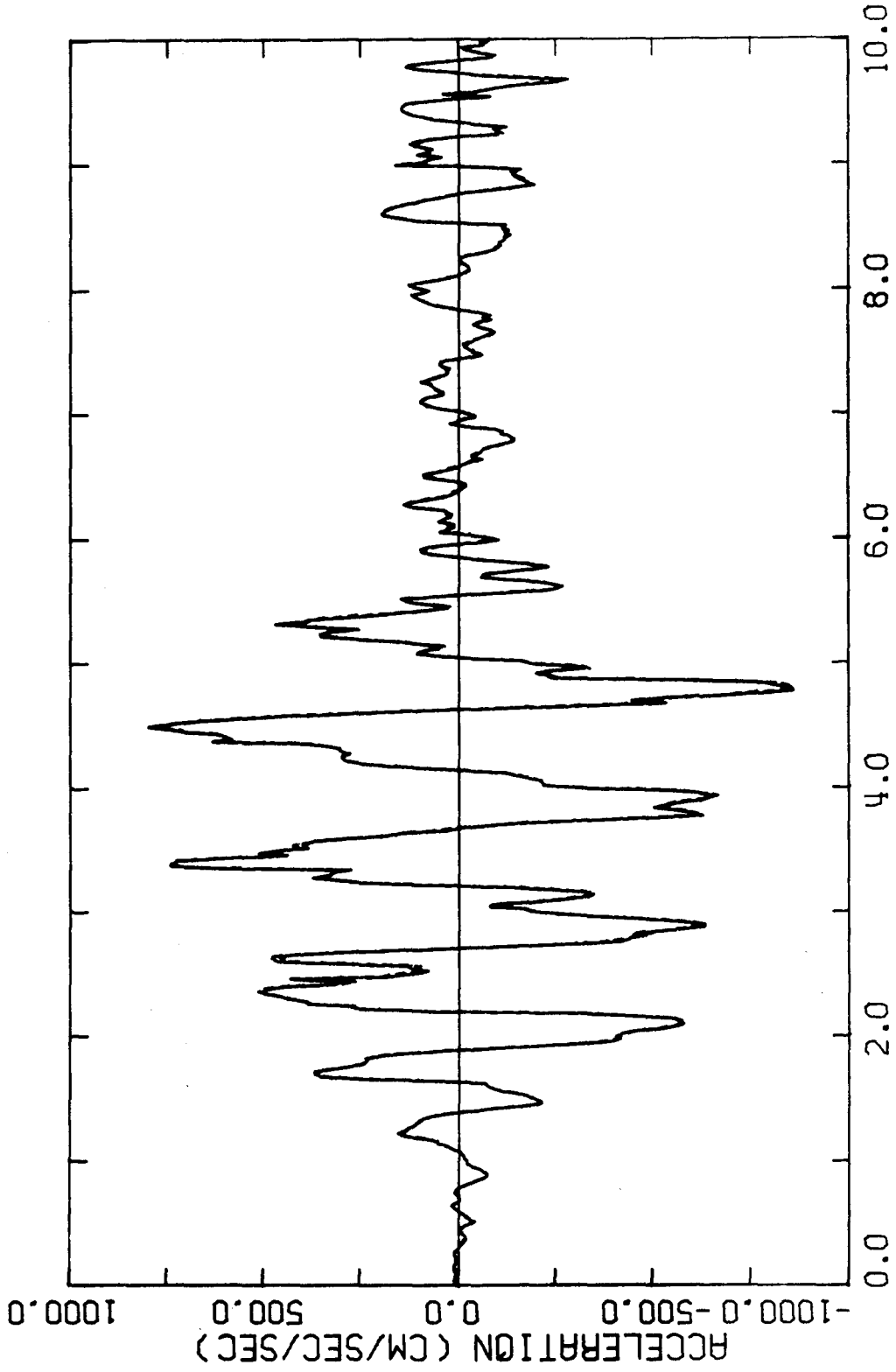


Figure 5.4. Relative acceleration (——) at the roof of the uniform chain system and relative acceleration (----) of the same model as in Fig. 5.3. The two curves are barely distinguishable.

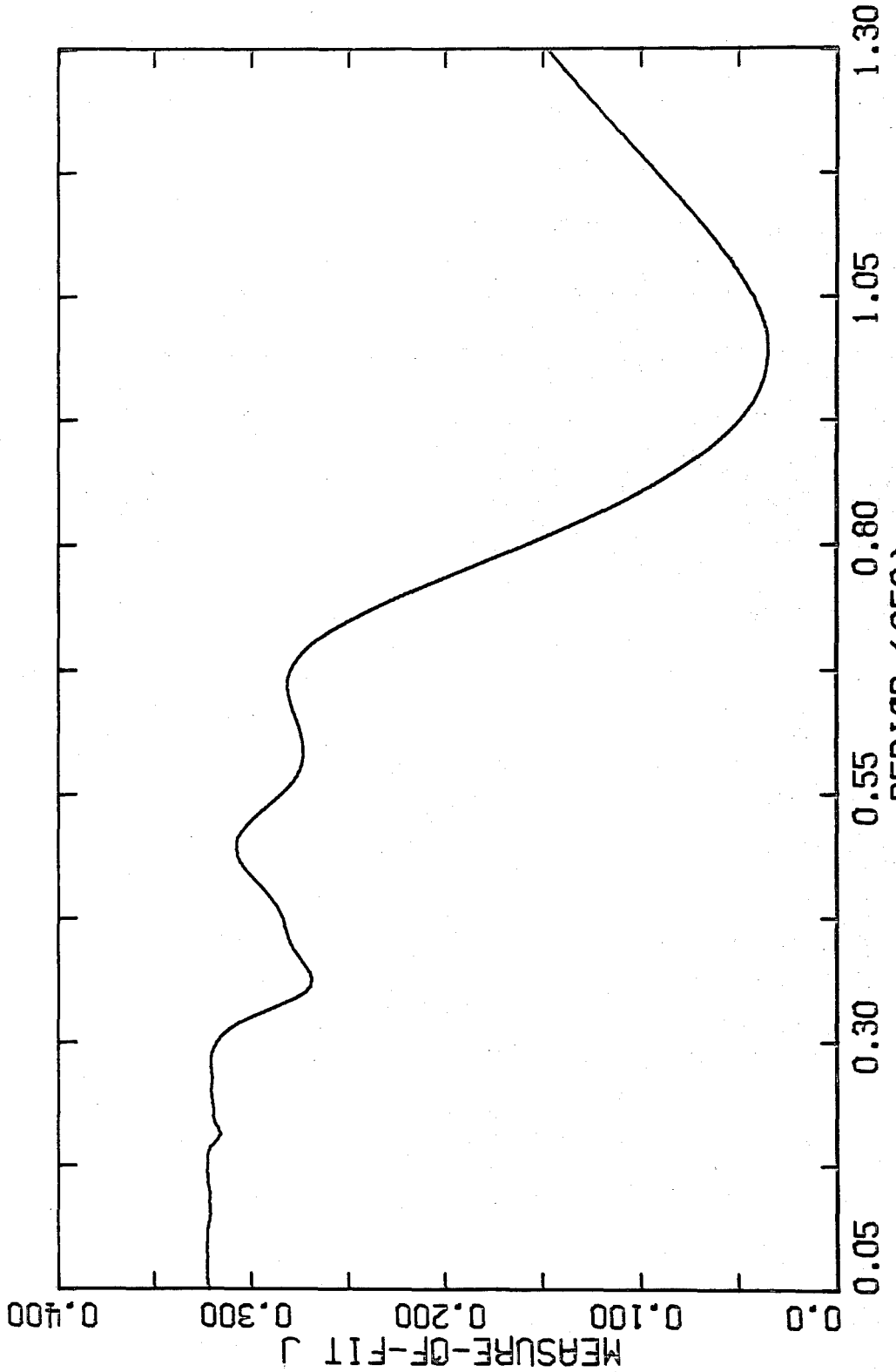


Figure 5.5. Profile of the measure-of-fit J for a single-mode model with a constant damping factor of 5%. The portion from 2 to 4 seconds of the roof acceleration of the uniform chain system was used in calculating J.

of about 0.6 sec and a damping factor of 5% is able to cancel enough of both the first and second mode to produce a local minimum in J. The spurious minimum disappears when the first mode contribution is subtracted from the acceleration response, as it would be during the identification of the second mode by the modal minimization method. Furthermore, when the first mode is subtracted, the local minima become sharper and another one appears in the plot of the profile of J which corresponds to the fifth mode.

Encouraged by the results of applying the modal minimization method to simulated data, the method was applied to earthquake records from some multi-story buildings. The results are reported in the next chapter.

REFERENCES

Bekey, G. A. (1970). System identification — an introduction and a survey. Simulation 15, 151-166.

Nigam, N. C. and P. C. Jennings (1969). Calculation of response spectra from strong-motion earthquake records. Bull. Seism. Soc. Am. 59, 909-922.

Raggett, J. D. (1974). Time domain analysis of structural motions. Meeting Preprint 2209, ASCE National Structural Engineering Meeting, Cincinnati, Ohio.

VI. APPLICATIONS TO BUILDINGS

The modal minimization method is used in this chapter to identify the modal properties of linear models for two multi-story buildings. The records used were obtained during the 1971 San Fernando earthquake, California.

6.1. Union Bank Building, Los Angeles

The Union Bank building is a 42-story steel-frame structure in downtown Los Angeles which experienced peak accelerations at mid-height of 20% g (transverse direction) and 13% g (longitudinal direction) during the 1971 San Fernando earthquake ($M_L = 6.3$). Only minor nonstructural damage occurred. Features of the building and its earthquake response are discussed by A. C. Martin and Associates (1973) and by Foutch et al (1975).

At the time of the San Fernando earthquake, strong-motion accelerographs with synchronized timing were installed in the sub-basement, on the 19th floor and on the 39th floor, but the instrument on the 39th floor failed to record. The S38^{o}W components of the digitized relative acceleration, velocity and displacement at the 19th floor were used as the response data in the analysis . These components correspond to the longitudinal direction of the building (Fig. 6.1). The sub-basement absolute acceleration, S38^{o}W component,

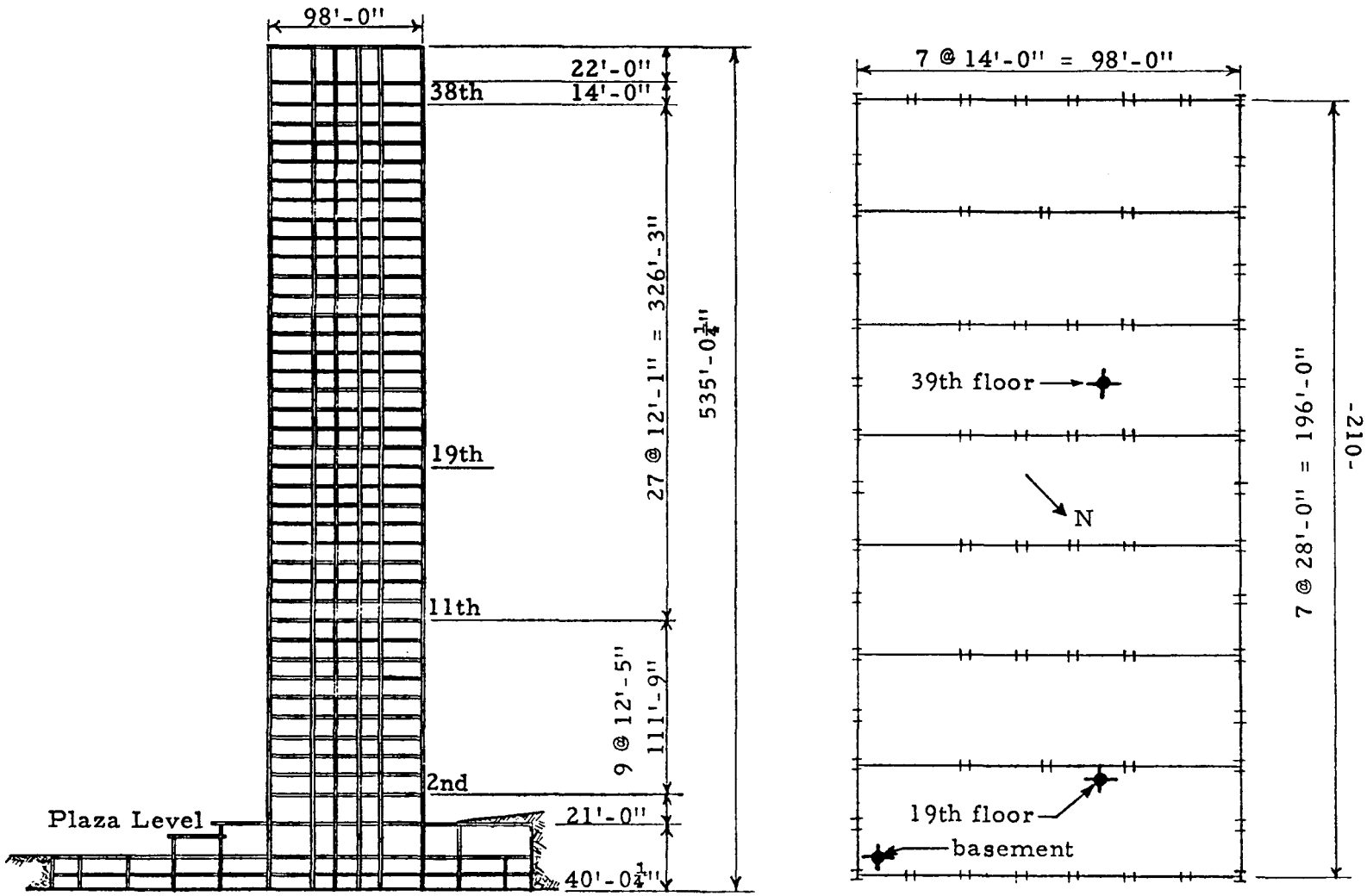


Figure 6.1. Transverse section and typical floor plan of Union Bank building. ✦ — Location of accelerograph. (from A.C. Martin and Associates, 1973)

was used as the input to the models (Fig. 6. 2).

The Fourier amplitude spectrum for 40. 96 seconds of the absolute acceleration at the 19th floor is shown in Fig. 6. 3 and the amplitude of the estimated transfer function is shown in Fig. 6. 4. The first five dominant peaks of each plot are given in Tables 6. 1 and 6. 2. Only the frequency range 0. 2-2. 5 Hz was used in constructing the Tables and the spectral amplitude ratios in Table 6. 1 for the absolute velocity and displacement were deduced by using the simple expression for transforms of derivatives. The erratic behavior of the unsmoothed transfer function in Fig. 6. 4 is typical for those estimated from seismic records.

The interpretation of the peaks in Tables 6. 1 and 6. 2 is based primarily on the period ratios for a uniform shear beam, since past work with ambient and forced vibration tests has shown that these ratios serve as a rough guide to identification of the resonant peaks of the lower modes of tall framed structures. The absence of the third longitudinal mode in Table 6. 2 might be expected since for this mode the 19th floor should be close to a node. It appears in Table 6. 1 because there is a relatively large peak in the Fourier amplitude spectrum of the sub-basement motion at a frequency of about 1 Hz. The tentative identification of the torsional mode is based on a simple

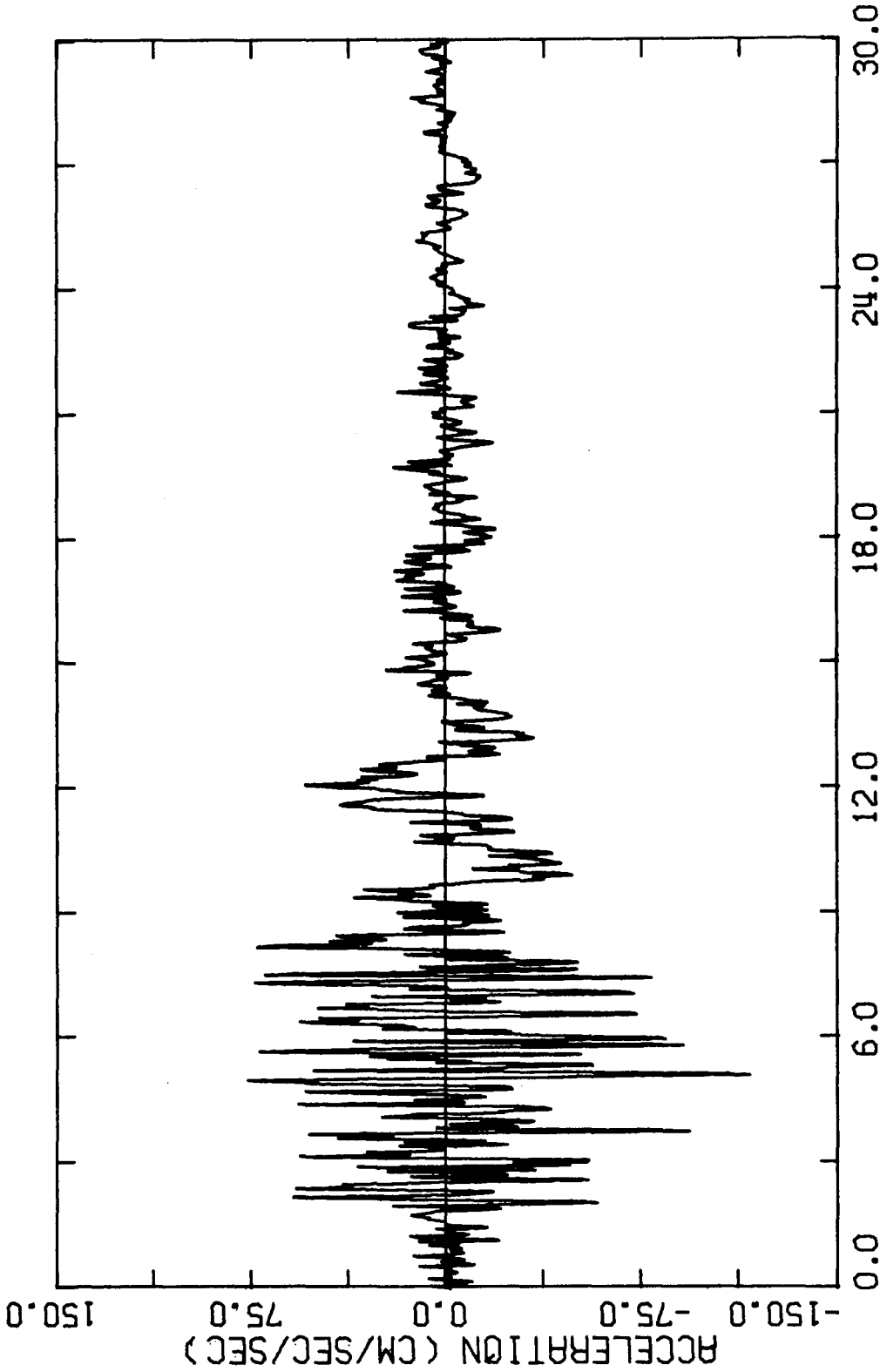


Figure 6.2. S38°W component of the sub-basement absolute acceleration in the Union Bank building (1971 San Fernando earthquake).

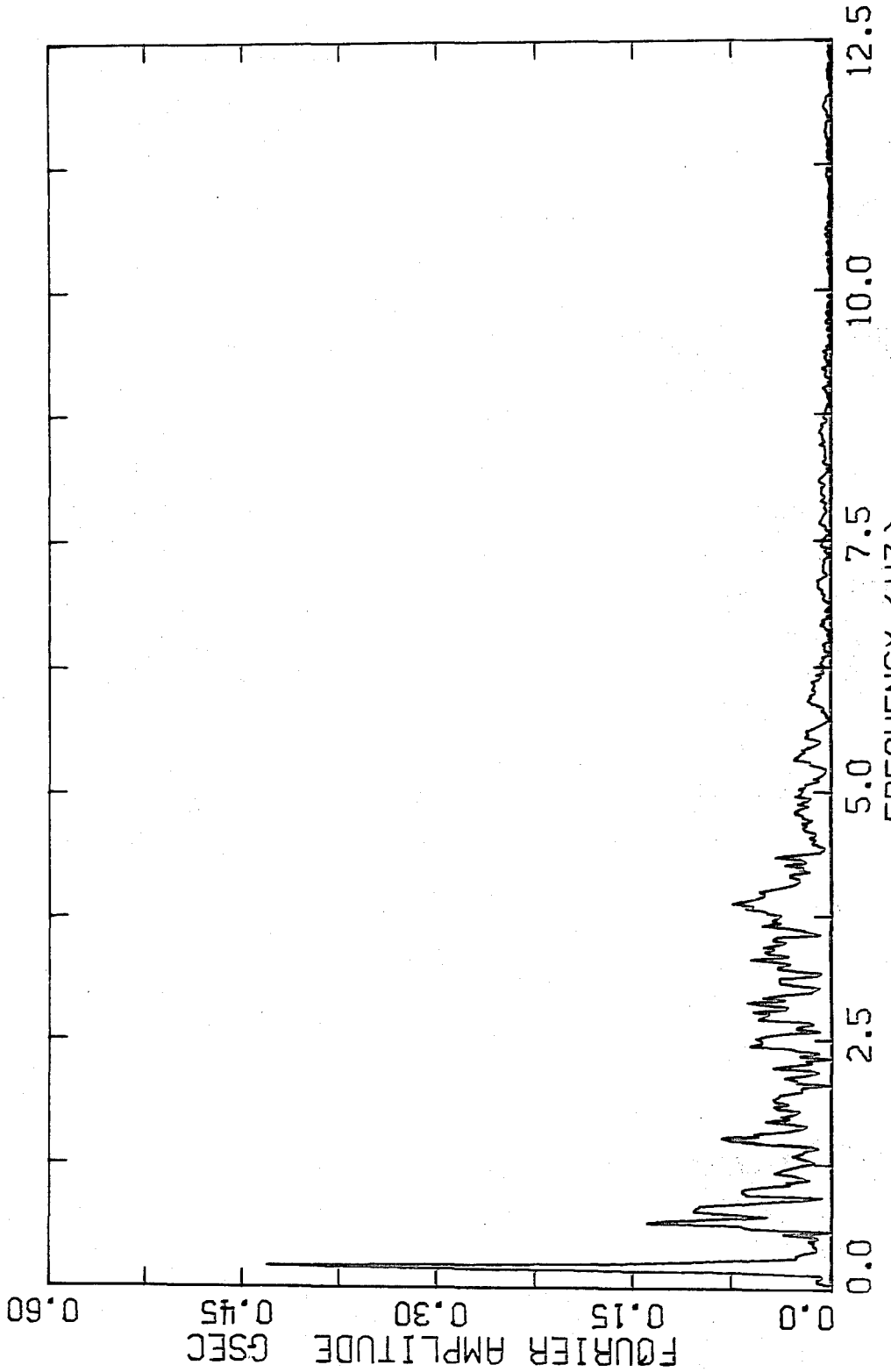


Figure 6.3. Fourier amplitude spectrum of the absolute acceleration, S38°W component, at the 19th floor of the Union Bank building (1971 San Fernando earthquake).

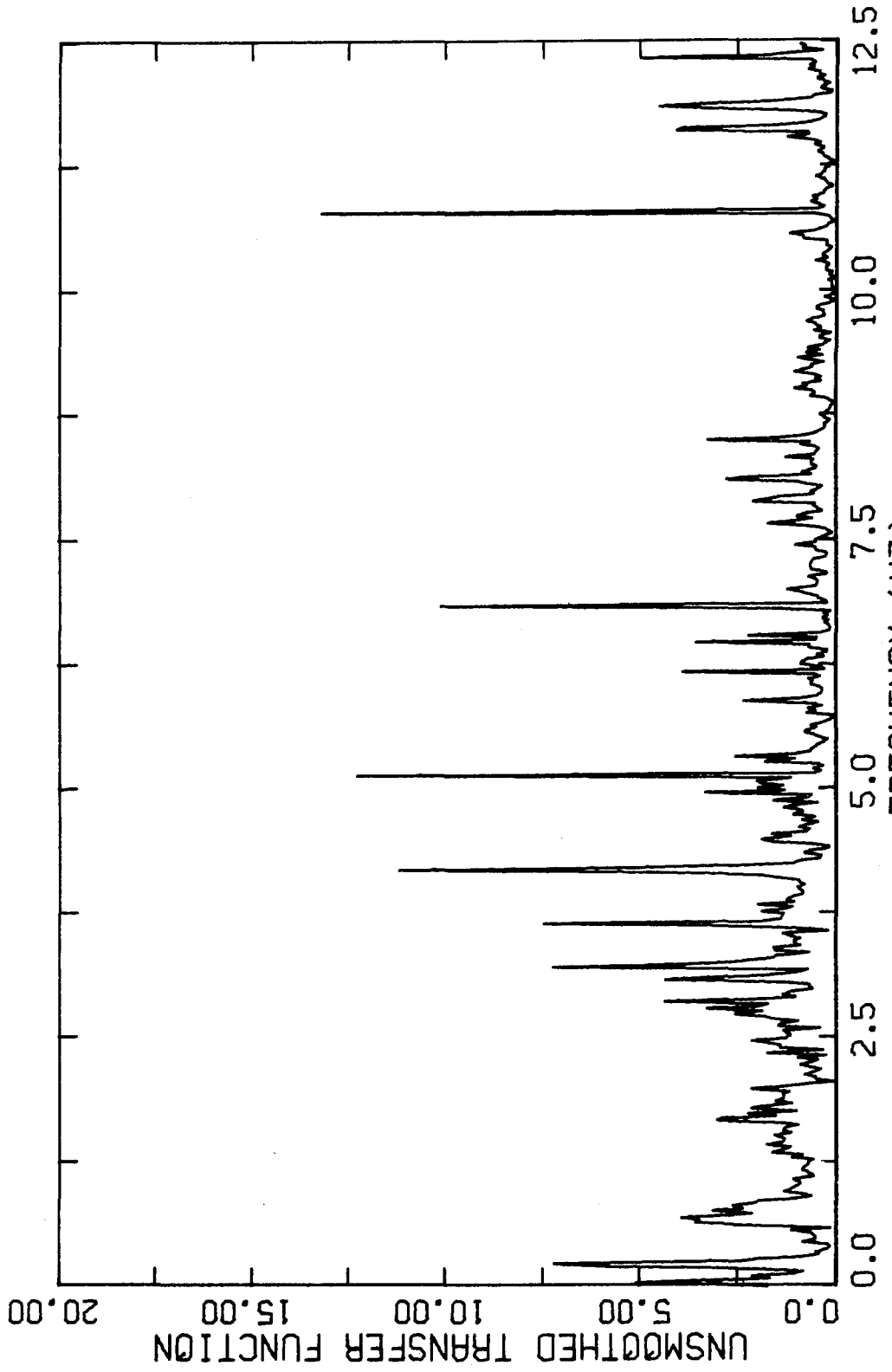


Figure 6.4. Amplitude of the unsmoothed transfer function between the S38°W components of the absolute acceleration in the sub-basement and at the 19th floor, Union Bank building.

Period (sec)	Period ratio	Ratio of spectral values(%)			Interpretation of mode
		Acceleration	Velocity	Displacement	
4.8	1.0	100	100	100	1st longitudinal
1.5	3.2	29	9	2.8	2nd longitudinal
1.25	3.8	21	5.5	1.5	2nd torsional
0.66	7.3	17	2.3	0.3	4th longitudinal
1.0	4.8	15	3.1	0.7	3rd longitudinal

TABLE 6.1. First five dominant peaks of the Fourier amplitude spectrum of the S38°W component of the absolute acceleration at the 19th floor of Union Bank building.

Period (sec)	Period ratio	Amplitude ratio (%)	Interpretation of mode
4.8	1.0	100	1st longitudinal
1.5	3.2	54	2nd longitudinal
1.3	3.7	43	2nd torsional
0.6	8.0	42	4th longitudinal
0.4	12.0	30	?

TABLE 6.2. First five dominant peaks of the amplitude of the transfer function between the S38°W components of the absolute acceleration in the sub-basement and at the 19th floor.

model of the structure wherein the interstory stiffness of each moment-resisting frame is replaced by a single spring and the mass and stiffness properties are assumed to be uniform with height. If it is assumed that the motion of longitudinal modes is purely in the longitudinal direction and the motion of the torsional modes is purely rotational, and if the experimental observation of equal fundamental translational periods is employed, the period of the r^{th} torsional mode can be shown to be about 80% of that of the r^{th} longitudinal mode.

6.1.1. Time-invariant Models

The parameters were first estimated for the major segment of the records, from 5 to 30 seconds. Following the general procedure described in §3.5, a succession of models was taken in which the number of modes was increased one at a time and the optimal estimates from one model were used as the initial estimates for the next model. As discussed in §5.4.2, when this is done, initial estimates are required only for the period and damping factor of the new mode. The initial estimates of the periods were those in Table 6.1 and the initial estimates of the damping factors were 4%.

The intention was to add the modes to the models in the order of their dominance in Table 6.1, although some difficulties were encountered. The results for the optimal models determined by

displacement matching and by velocity matching are given in Table 6.3, while those for acceleration matching are given in Table 6.4. The p_r are the effective participation factors at the 19^{th} floor [Eq. (3.2.13)]. The measure-of-fit J for each model is also given. Recall that $J^{\frac{1}{2}}$ represents the ratio of the r.m.s. output-error to the maximum response, and the values in the Tables give this ratio as a percentage. For example, Table 6.3 shows that for a single-mode model determined by velocity matching, the r.m.s. velocity-error is 9% of the peak velocity.

Only a one-mode model was determined by matching the recorded and model displacements because the signal of the second mode was so small. The quality of the match is shown in Fig. 6.5. The initial displacement is not equal to the "recorded" value because the initial conditions for each modal contribution are estimated along with the other modal parameters and these initial conditions are used in the calculations of the response of an optimal model. The calculated displacement and velocity for the two-mode model determined by velocity matching are compared with the actual displacement and velocity in Figs. 6.6a and 6.6b. Figure 6.6.a shows that a good displacement match is obtained even when the model is determined by matching velocities.

Modal parameter	One-mode model (Displacement Match)	One-mode model (Velocity Match)	Two-mode model (Velocity Match)	Three-mode model (Velocity Match)
$\hat{T}_1$	4.60	4.62	4.62	4.62
$\hat{\zeta}_1$	3.2	3.9	3.9	3.8
$\hat{p}_1$	0.78	0.81	0.83	0.82
$\hat{T}_2$			1.49	1.50
$\hat{\zeta}_2$			5.1	5.9
$\hat{p}_2$			0.40	0.50
$\hat{T}_3$				1.0
$\hat{\zeta}_3$				35
$\hat{p}_3$				-0.5
$J^{\frac{1}{2}}$ (%)	8.3	9.0	6.5	5.7

TABLE 6.3. Optimal estimates of the parameters of the longitudinal modes using the portion of the Union Bank records from 5 to 30 seconds.

Modal parameter	Two-mode model	Three-mode model	Four-mode model	Five-mode model
$\hat{T}_1$	4.59	4.62	4.62	4.61
$\hat{\zeta}_1$	3.5	4.4	4.4	4.2
$\hat{p}_1$	0.74	0.86	0.87	0.84
$\hat{T}_2$	1.50	1.49	1.49	1.49
$\hat{\zeta}_2$	4.2	4.6	5.7	5.8
$\hat{p}_2$	0.31	0.39	0.48	0.46
$\hat{T}_2^t$				1.2
$\hat{\zeta}_2^t$				7.7
$\hat{p}_2^t$				-0.12
$\hat{T}_3$		0.9	1.1	0.95
$\hat{\zeta}_3$		27	19	13
$\hat{p}_3$		-0.4	-0.29	-0.13
$\hat{T}_4$			0.66	0.66
$\hat{\zeta}_4$			7.2	6.6
$\hat{p}_4$			-0.17	-0.15
$J^{\frac{1}{2}}$(%)	10.2	8.8	8.3	8.2

TABLE 6.4. Optimal estimates of the parameters using the portion from 5 to 30 seconds of the Union Bank acceleration record, longitudinal direction. The torsional mode is distinguished by the superscript t.

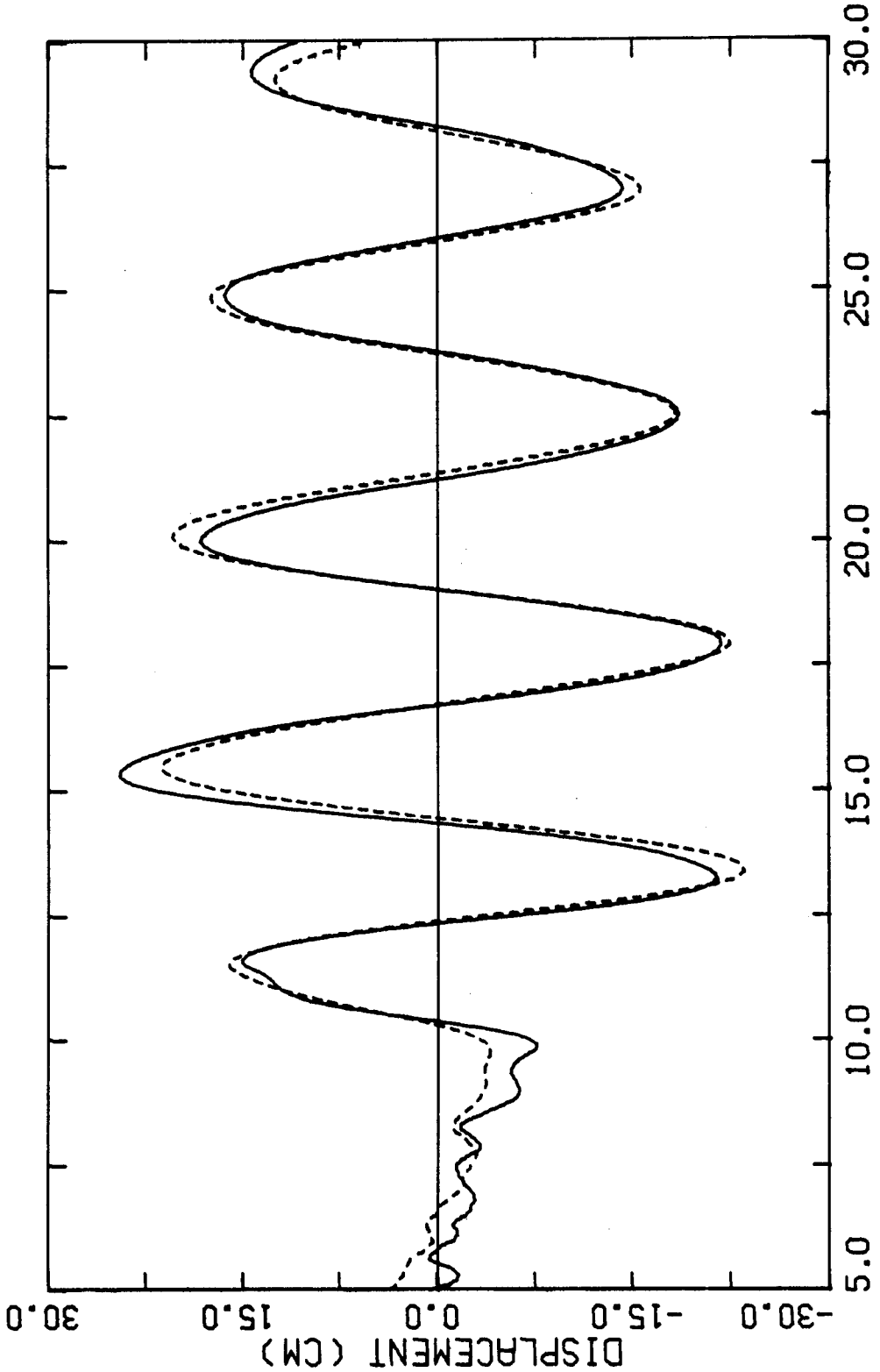


Figure 6.5. Relative displacement (——) at the 19th floor of Union Bank building and calculated displacement (----) of the optimal model with one mode determined by matching displacements over the interval.

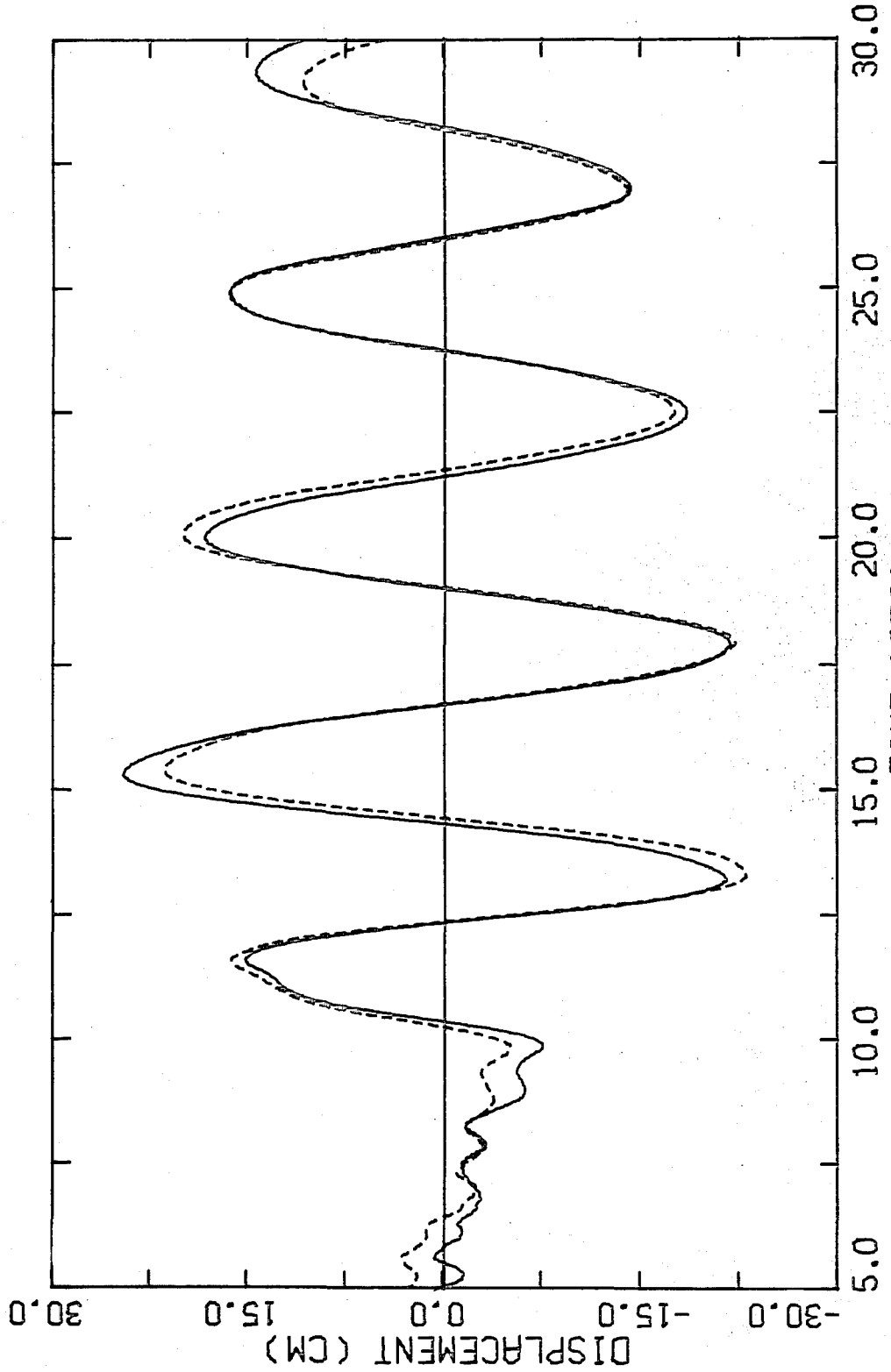


Figure 6.6.a. Relative displacement (——) at 19th floor of Union Bank building and calculated displacement (----) of the optimal model with two modes determined by matching velocities over the interval.

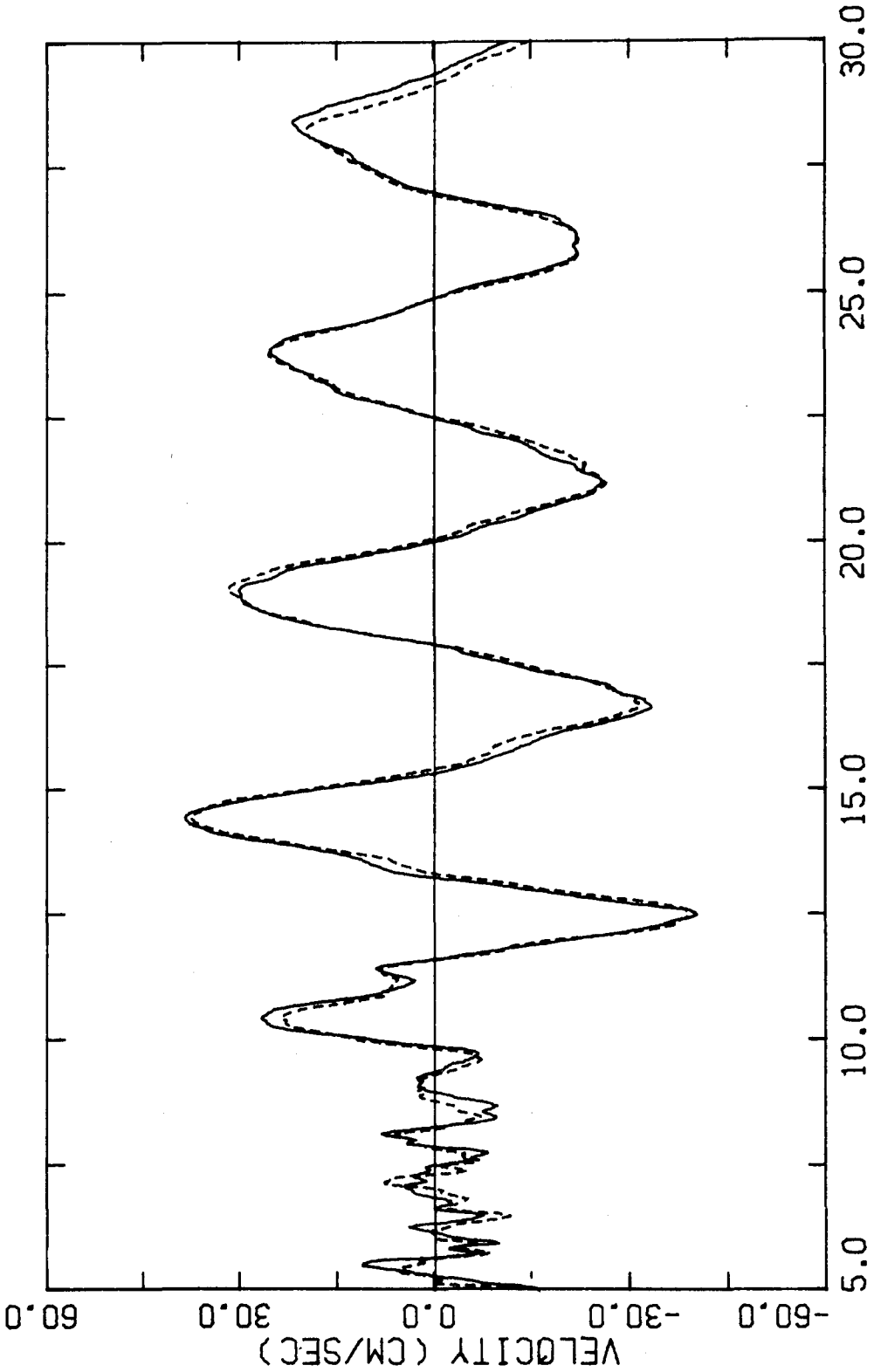


Figure 6.6.b. Relative velocity (——) at 19th floor of Union Bank building and calculated velocity (----) of the same model as in Fig. 6.6.a.

An attempt was made to identify the second torsional mode by using a three-mode model and velocity matching. However, for the third dominant mode, the modal minimization method converged to a period of 1.0 second and a very high damping factor of 35% (Table 6.3). A possible explanation for this behavior is that because the signals in the velocity record corresponding to the higher modes are small (Table 6.1), the method chooses a local minimum of J which arises from partial compensation by the model of up to three of the higher modes in the velocity record with a period around 1 sec. The high damping of 35% would allow the identified "mode" to do this because it produces a resonant peak in the frequency domain with a broad bandwidth.

The difficulties in identifying the torsional mode also occurred when the models were determined from the acceleration record (Table 6.4). Even though the torsional period was included as an initial estimate for the new mode for the models with three and four modes, the modal minimization method converged to one of the less dominant modes of Table 6.1. It was only when a five-mode model was taken that the second torsional mode appeared. Furthermore, it produced only a small change in J. This suggests that the chosen class of models is not capable of producing a torsional signal of comparable strength to that appearing in the actual records, possibly

because the excitation of the torsional mode by the transverse base motion is not considered in the modelling (§3.1.1).

The calculated velocity and acceleration for the four-mode model determined by matching accelerations are compared with the actual records in Figs. 6.7.a and 6.7.b. The displacement of the four-mode model was almost identical to the displacement of the two-mode model determined by velocity matching (Fig. 6.6.a). Also, a comparison of Figs. 6.6.b and 6.7.a shows that the velocities of the four-mode model determined by matching accelerations and the two-mode model determined by matching velocities produce nearly the same agreement with the recorded velocity.

It is concluded from the results that a time-invariant linear model with a small number of modes can reproduce the strong-motion records at the 19^{th} floor surprisingly well. The number of modes required to give a very good approximation of the relative displacement, velocity and acceleration are one, two and four modes respectively. The respective optimal models give calculated response which have an r.m.s. error of about 8% or less of the peak response. The quality of the match of recorded and model responses was not expected prior to the identification; the match given by one of the two-dimensional dynamic models used in the design of the building was not nearly as good, as seen in Fig. 6.8 [from A. C. Martin and Associates (1973)].

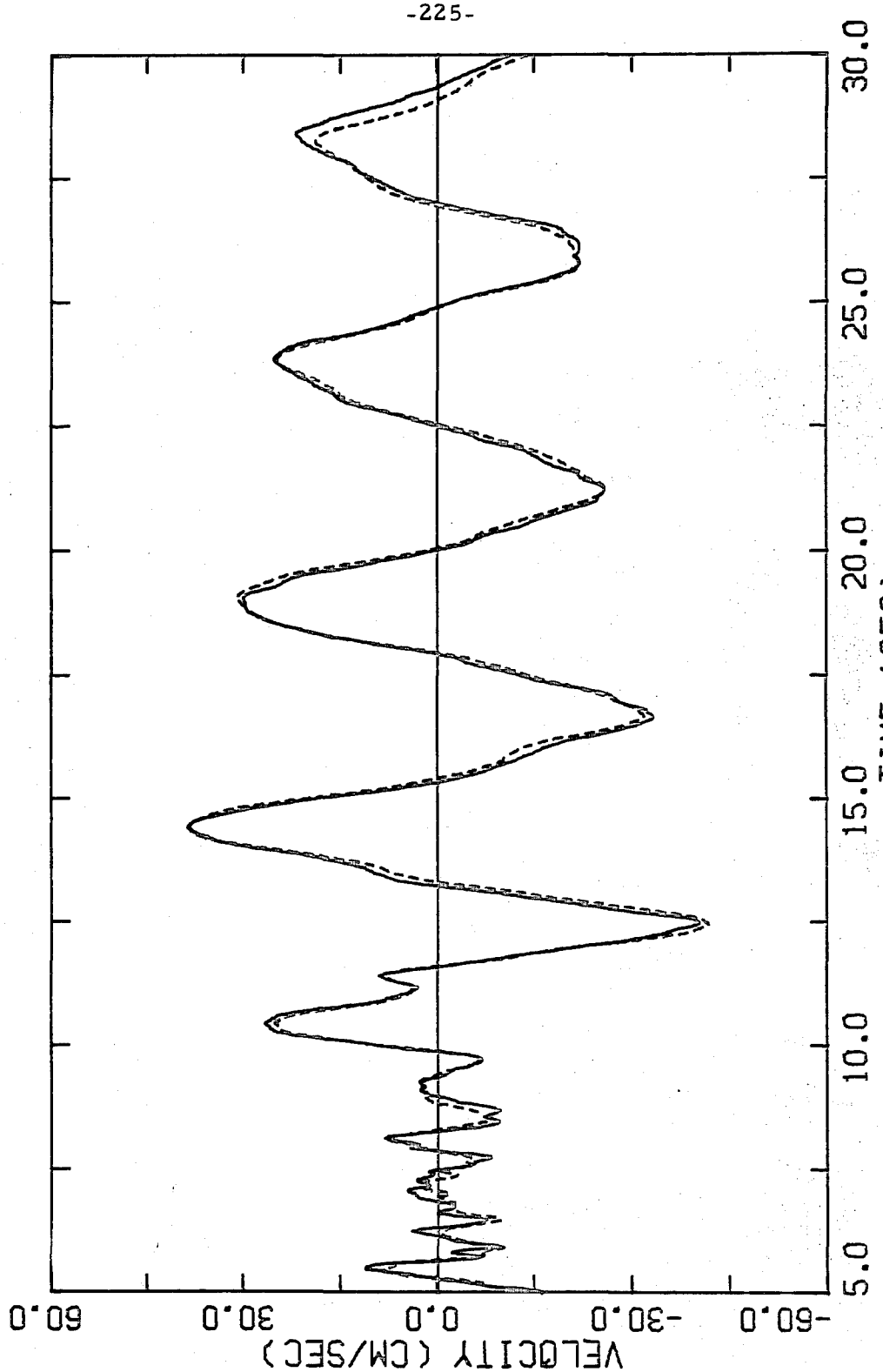


Figure 6.7.a. Relative velocity (——) at the 19th floor of Union Bank building and calculated velocity (----) of the optimal model with four modes determined by matching accelerations over the interval.

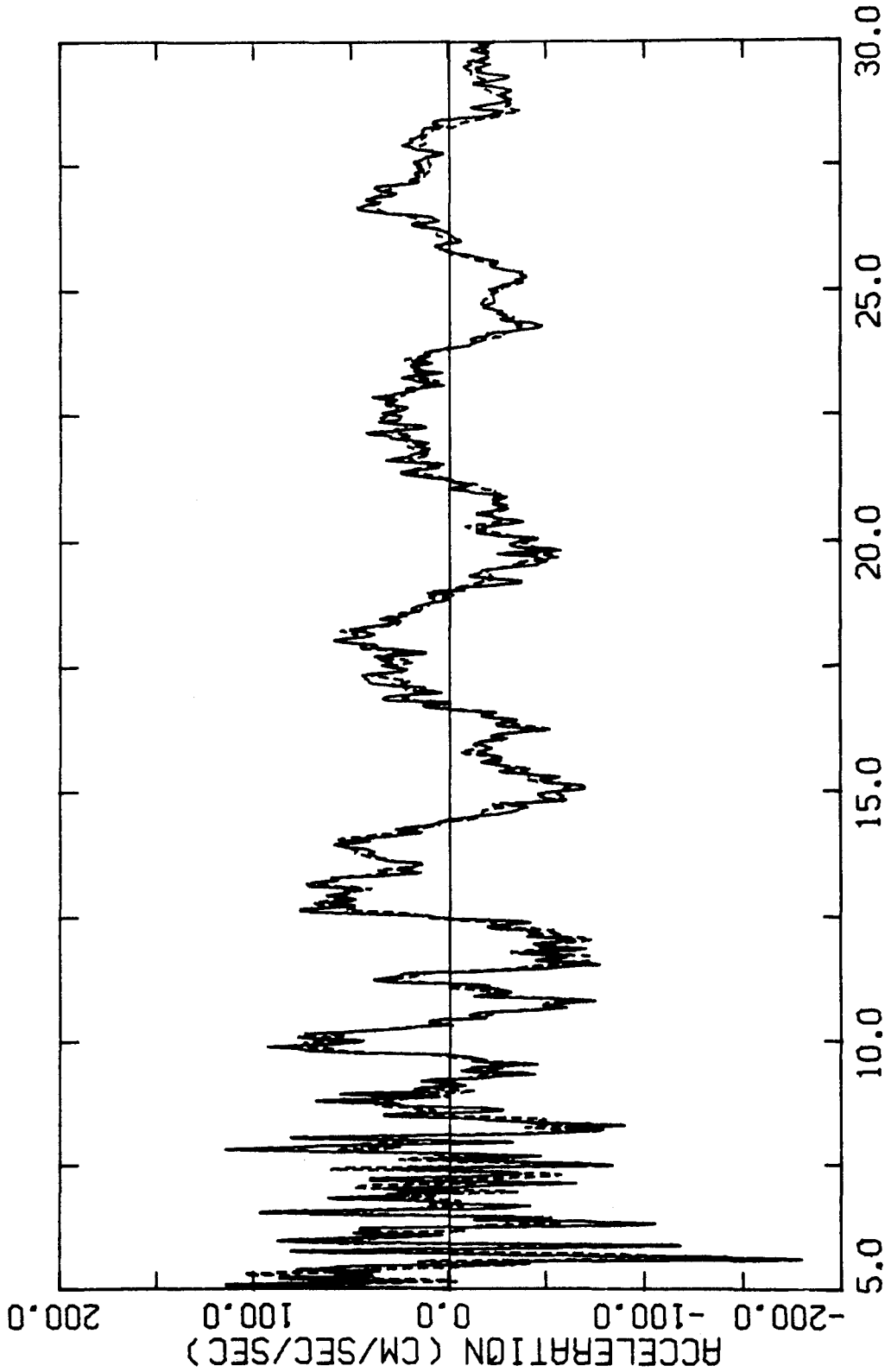


Figure 6.7.b. Relative acceleration (——) at the 19th floor of Union Bank building and calculated acceleration (----) of the same model as in Fig. 6.7.a.

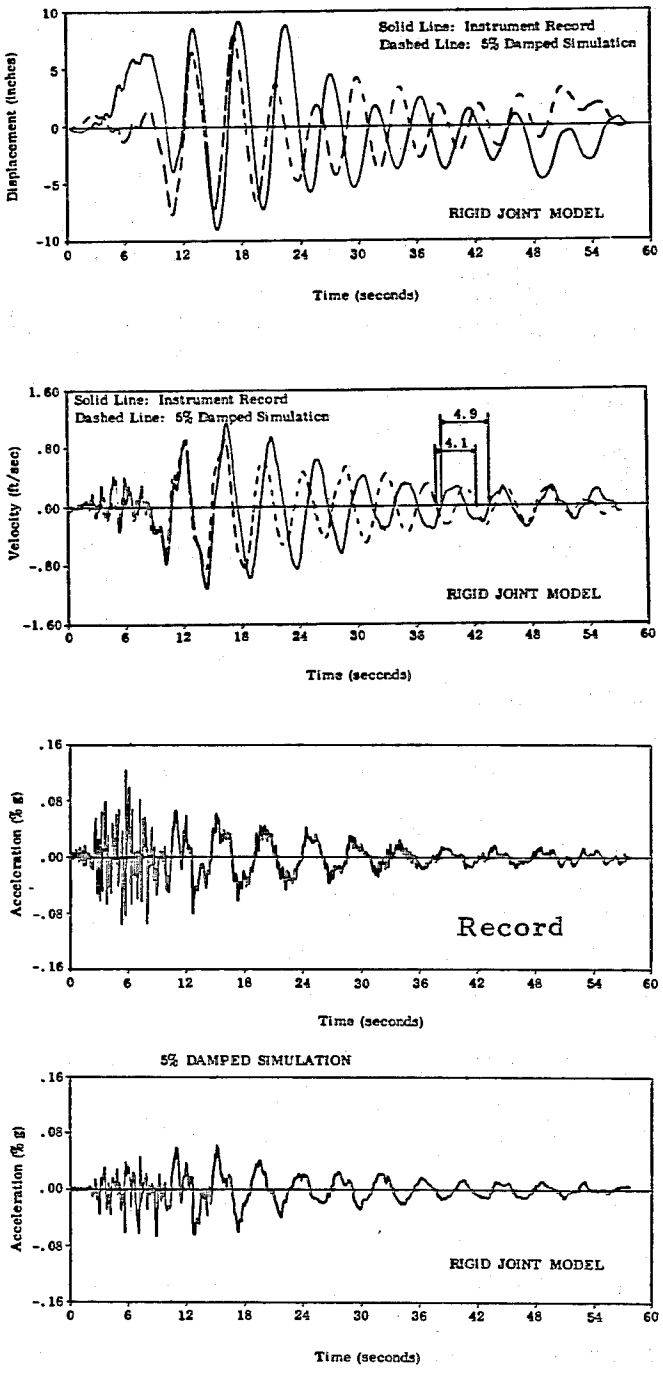


Figure 6.8. Absolute response at the 19th floor of Union Bank building during the 1971 San Fernando earthquake compared with the calculated response of a structural model used in the design of the building (from A. C. Martin and Associates, 1973).

Of course, a better fit of the response histories should be expected for a model determined from the earthquake records than for a model synthesized from structural plans.

The periodic nature of the difference between the displacement record and the displacement of the optimal model in Fig. 6.5 or 6.6.a suggests that the discrepancy between these two histories is primarily due to long period errors of about 8 to 10 seconds in the record. This conclusion is also consistent with the large peak in the estimated transfer function in Fig. 6.4 which is at a lower frequency than the fundamental frequency of the building.

An early study of long-period errors in accelerograms suggested that the errors in the derived displacement records should be relatively small up to a period of about 16 seconds (Trifunac, 1970b), but a later study, using the large amount of data from the 1971 San Fernando earthquake, showed that there were significant errors in some records at periods less than 16 seconds (Hanks, 1973). Thus, early records processed at the California Institute of Technology, which include the Union Bank records, were high-pass filtered with a cut-off frequency corresponding to a period of about 14 seconds, whereas for most of the later records, the cut-off frequency corresponded to a period of 8 seconds. To remove the long-period components, the Union Bank records were filtered with a roll-off

termination frequency of 0.125 Hz and a roll-off bandwidth of 0.05 Hz and this led to a considerable improvement in the optimal displacement match for a one-mode model, as can be seen by comparing Figs. 6.5 and 6.9. However, the estimates of the parameters using the filtered displacement and a one-mode model changed only slightly from their values in column 1 of Table 6.3 to the values: $\hat{T}_1 = 4.61$ sec, $\hat{\zeta}_1 = 3.4\%$, $\hat{p}_1 = 0.79$. These results suggest that if sufficiently long segments of the records are taken, the method is not sensitive to measurement noise which is at frequencies significantly different from the modal frequencies, as might have been anticipated from the least-squares nature of the approach.

The optimal estimates of the parameters of the four longitudinal modes determined by matching the acceleration record with a five-mode model are compared in Table 6.5 with other available values for these parameters. The participation factors for the synthesized structural model and for the ambient vibration tests were calculated from the known mass distribution of the building and the published modeshapes. There was clearly a reduction in the stiffness of the structure during the San Fernando earthquake which was only partially recovered after the earthquake. The degradation in stiffness is thought to be due to changes in the nonstructural elements such as partitions. Notice that the period ratios are roughly the same during the

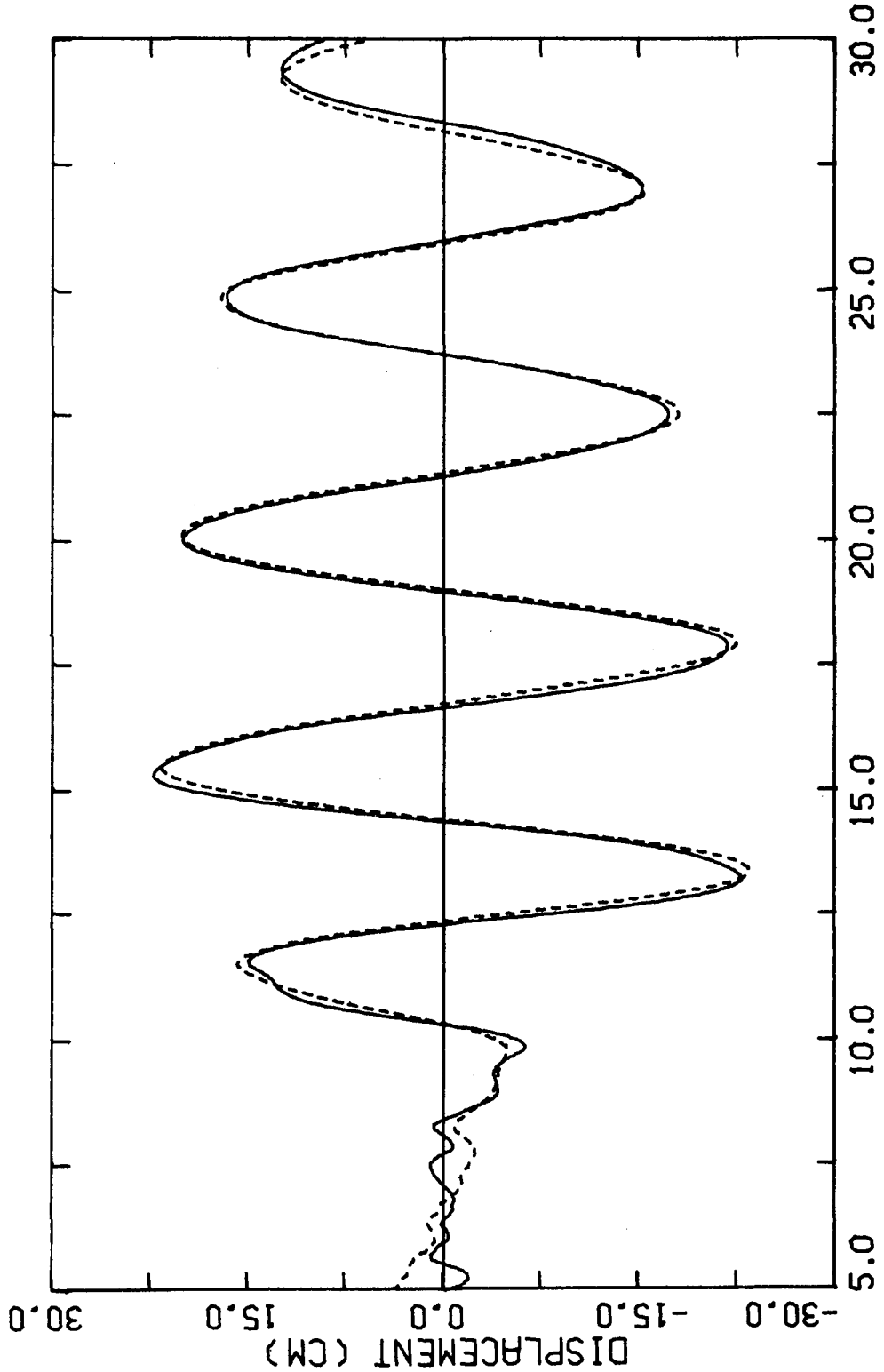


Figure 6.9. Relative displacement (——) at the 19th floor of Union Bank building and calculated displacement (- - -) of the optimal model with one mode determined by displacement matching. The filtered records were used.

Source of values	Period (sec)				Damping factor (%)				Participation factor			
	T_1	T_2	T_3	T_4	ζ_1	ζ_2	ζ_3	ζ_4	P_1	P_2	P_3	P_4
Optimal estimates from acceleration records, 1971 San Fernando earthquake	4.61	1.49	1.0	0.66	4.2	5.8	13(?)	6.6	0.84	0.46	-0.13(?)	-0.15
	T_1/T_r=	3.1	4.6	7.0								
Uniform shear beam	4.61	1.54	0.92	0.66					0.90	0.30	-0.18	-0.13
	T_1/T_r=	3.0	5.0	7.0								
Synthesized structural model assuming rigid joints (A. C. Martin and Associates, 1973)	4.13	1.45	0.81	0.63	5% in all modes				0.66			
	T_1/T_r=	2.9	4.8	6.6								
Ambient vibration tests:												
(a) Before 1971 San Fernando earthquake (Trifunac, 1970a)	3.1	1.1	0.6	0.4	1.7	1.5	1.8	2.0	0.72	0.45	0.01	-0.2
	T_1/T_r=	2.9	5.1	7.2					(calculated from measured modeshapes)			
(b) After 1971 San Fernando earthquake (Udwadia and Trifunac, 1973)	3.8	1.3	0.8	0.5								
	T_1/T_r=	2.9	5.0	6.9								

TABLE 6.5. Comparison of the optimal estimates of the parameters of the first four longitudinal modes (from Table 6.4) with other available data for Union Bank building.

earthquake as before and after the event, suggesting that the reduction in stiffness was approximately uniform over the structure. Furthermore, these period ratios are close to those for a uniform shear beam.

The estimates of the damping factors from the seismic records are much greater than those for the ambient vibration tests. This is consistent with experience for other buildings and might be expected in view of the much larger amplitudes of the structural motion during the earthquake. The damping factors also increase slightly with the higher modes for this building.

It is often assumed that the modeshapes, and hence the participation factors, will not change appreciably as the amplitude of the motion increases. This was the case with the participation factor of the second mode, but the participation factor of the fundamental longitudinal mode during the earthquake was quite different from its value in the pre-earthquake ambient vibration tests. The corresponding modeshape in the ambient tests was almost a straight line. The first two values for p_1 in Table 6.5 suggest that the fundamental modeshape may have been more like that for a uniform shear beam during the earthquake.

On the basis of all the results, it is believed that the optimal estimates in Table 6.5 of the parameters for the first, second and

fourth longitudinal modes are reliable values for the strong-motion behavior of the building, as interpreted by a time-invariant linear model. It is more difficult to assess the estimates for the parameters of the second torsional mode and third longitudinal mode. The periods appear reasonable but the large damping estimate for the third mode is questionable. Furthermore, because the estimate of the damping factor is most probably too large, the magnitude of the participation factor may also be too large.

6.1.2. Time-varying Models

The results presented in the previous subsection show that there was a degradation in the stiffness of the Union Bank building during the 1971 San Fernando earthquake. To further investigate this effect, optimal linear models were determined for four successive, overlapping subintervals in the time interval from 5 to 30 seconds. The variation of the optimal estimates with each time segment then shows how the equivalent linear parameters changed during the earthquake due to nonlinearities in the structural response.

Time windows of ten seconds were used, moving the window by five seconds each time. This gives time segments of just over two cycles of the fundamental mode. This choice, suggested by the results using simulated data, is a compromise between the desire to take a

small interval so that the instantaneous structural properties can be approximated, and the necessity to take a sufficiently large interval for the parameters to be estimated reliably.

The optimal models were determined by matching displacements (one mode) and by matching velocities (one and two modes). The records used were those discussed above which were high-pass filtered with a roll-off termination frequency of 0.125 Hz, corresponding to a period of 8 seconds. The optimal estimates of the modal parameters and the corresponding measure-of-fit are presented in Table 6.6. The estimates of the parameters for a given time segment are in good agreement for the three different models, except for the low damping in the first row of Table 6.6. The reason for the latter discrepancy is not clear, although it may be due partly to the small signal over half of the interval (see Fig. 6.10.a). The greater variation in the estimates of p_1 for the last time segment in Table 6.6 is to be expected since the determination of p_1 becomes ill-conditioned for later portions of the records. This is because the basement acceleration is small for these time intervals and the structural motion is dominated by the free-vibration components which do not depend on the p_r.

The results in Table 6.6 and plots of the recorded and model responses suggest that the lengthening in the period of the fundamental longitudinal mode during the interval from 5 to 30 seconds was

Time interval	Record used	$\hat{T}_1$	$\hat{\zeta}_1$	$\hat{p}_1$	$\hat{T}_2$	$\hat{\zeta}_2$	$\hat{p}_2$	$J^{\frac{1}{2}}$(%)
5-15	Displacement	4.41	2.2	0.73				5.1
	Velocity	4.46	4.2	0.79				11.0
	Velocity	4.42	4.1	0.80	1.50	4.6	0.42	6.1
10-20	Displacement	4.58	4.0	0.83				3.8
	Velocity	4.55	3.9	0.70				8.2
	Velocity	4.57	4.3	0.83	1.47	4.8	0.41	2.8
15-25	Displacement	4.64	3.7	0.72				1.9
	Velocity	4.65	3.7	0.72				5.0
	Velocity	4.64	4.0	0.80	1.49	5.0	0.29	2.5
20-30	Displacement	4.72	2.8	0.58				1.6
	Velocity	4.74	2.9	0.68				4.9
	Velocity	4.75	2.8	0.72	1.60	2.9	0.33	2.1

TABLE 6.6. Optimal estimates of the modal parameters for different segments of the records from Union Bank building, San Fernando earthquake. One-mode and two-mode models were used.

progressive rather than abrupt. The change in the fundamental period from the first to the last time segment was $7\frac{1}{2}\%$, and the period of the second mode changed by 7%. For the first three time intervals, the damping factors of the fundamental and second modes were approximately 4% and 5% respectively, dropping to about 3% over the last time interval.

A comparison is given in Figs. 6.10.a, b and c of the recorded response and the calculated response of the optimal model with two modes determined by matching velocities over the interval from 5 to 15 seconds. A similar comparison is given in Figs. 6.11.a, b and c for the optimal model with two modes determined by again matching velocities, but over the interval from 20 to 30 seconds. The two cases presented in Figs. 6.10 and 6.11 represent respectively the worst and best match of velocities for a two-mode model. Ignored higher modes are evident in the velocity and acceleration comparisons in the Figures.

6.1.3. Sensitivity Analyses, Union Bank Building

The full sensitivity matrix $\hat{S}$ [the Hessian matrix, Eq. (2.3.8)] and the reduced sensitivity matrix $\tilde{S}$ [the partial Hessian matrix, Eq. (2.3.9)], involving derivatives of J [Eq. (5.2.7)] with respect to the model coefficients $a_i^{(r)}$, were evaluated in several cases. It was found that corresponding elements of $\hat{S}$ and $\tilde{S}$ were similar because the last term of Eq. (2.3.8) was small. Conclusions regarding the sensitivity of J with respect to the $a_i^{(r)}$ were therefore unchanged

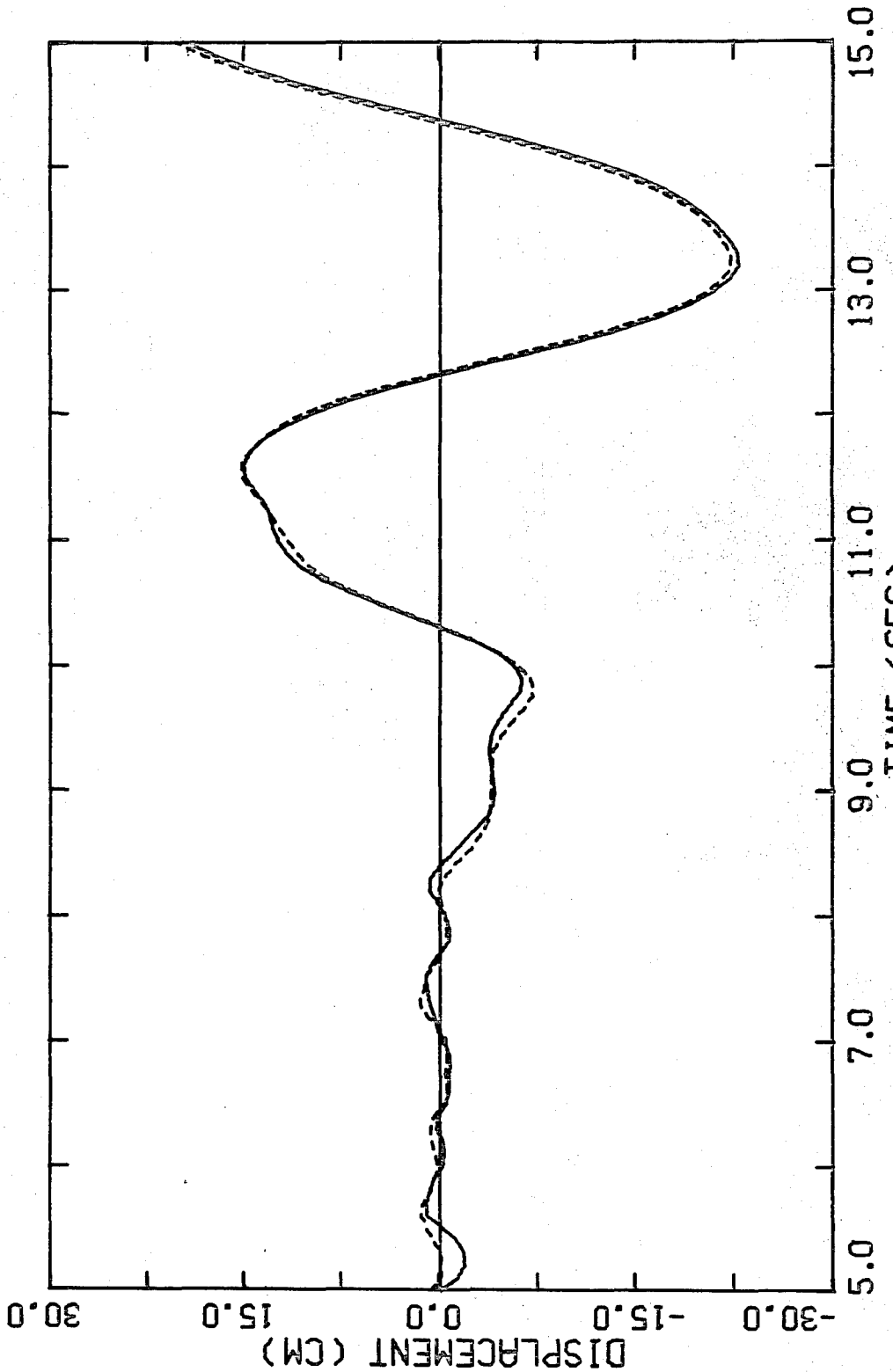


Figure 6.10.a. Relative displacement (——) at the 19th floor of Union Bank building and calculated displacement (----) of the optimal model with two modes determined by matching velocities over the interval.

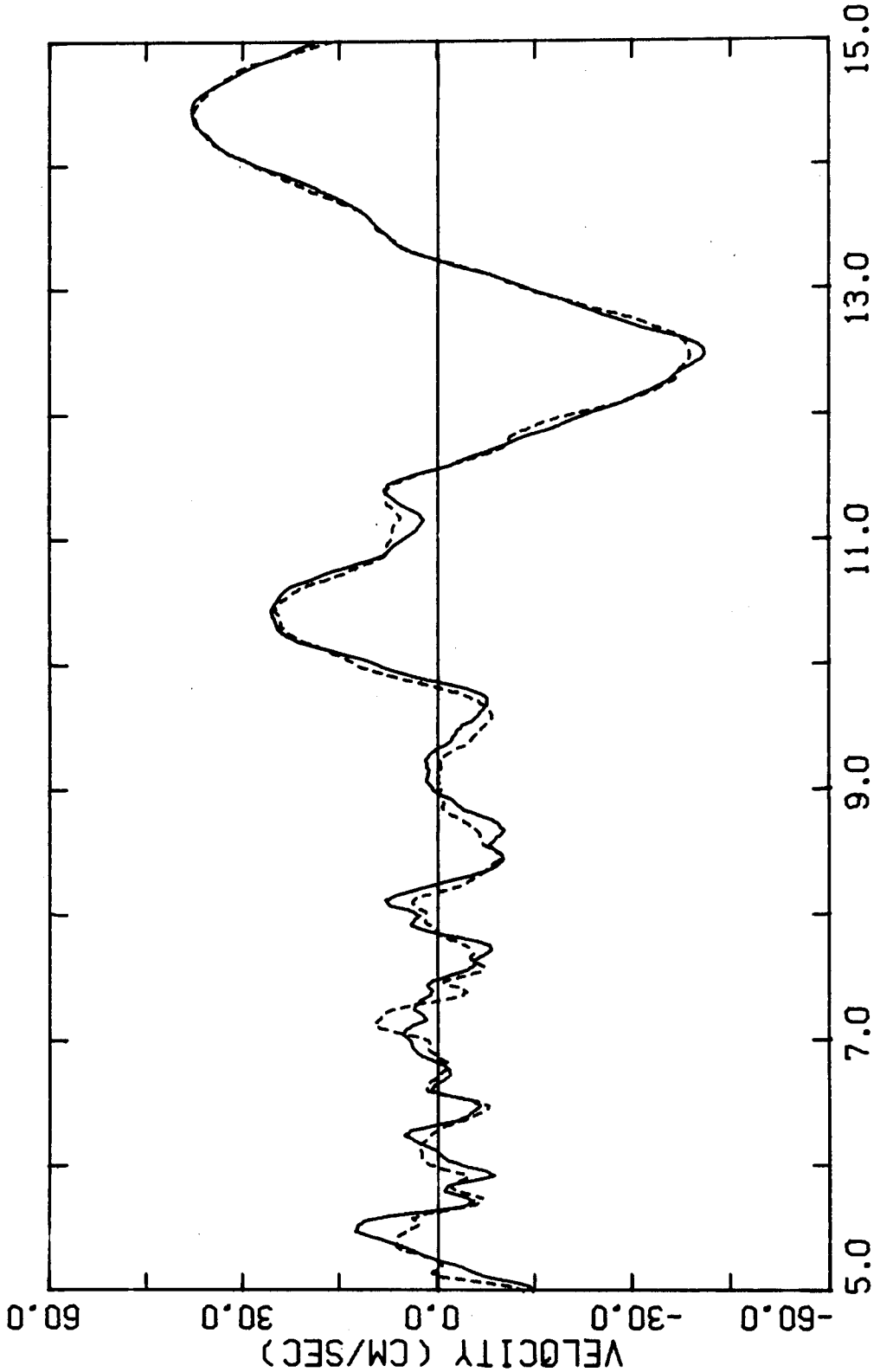


Figure 6.10.b. Relative velocity (——) at the 19th floor of Union Bank building and calculated velocity (----) of the same model as in Fig. 6.10.a.

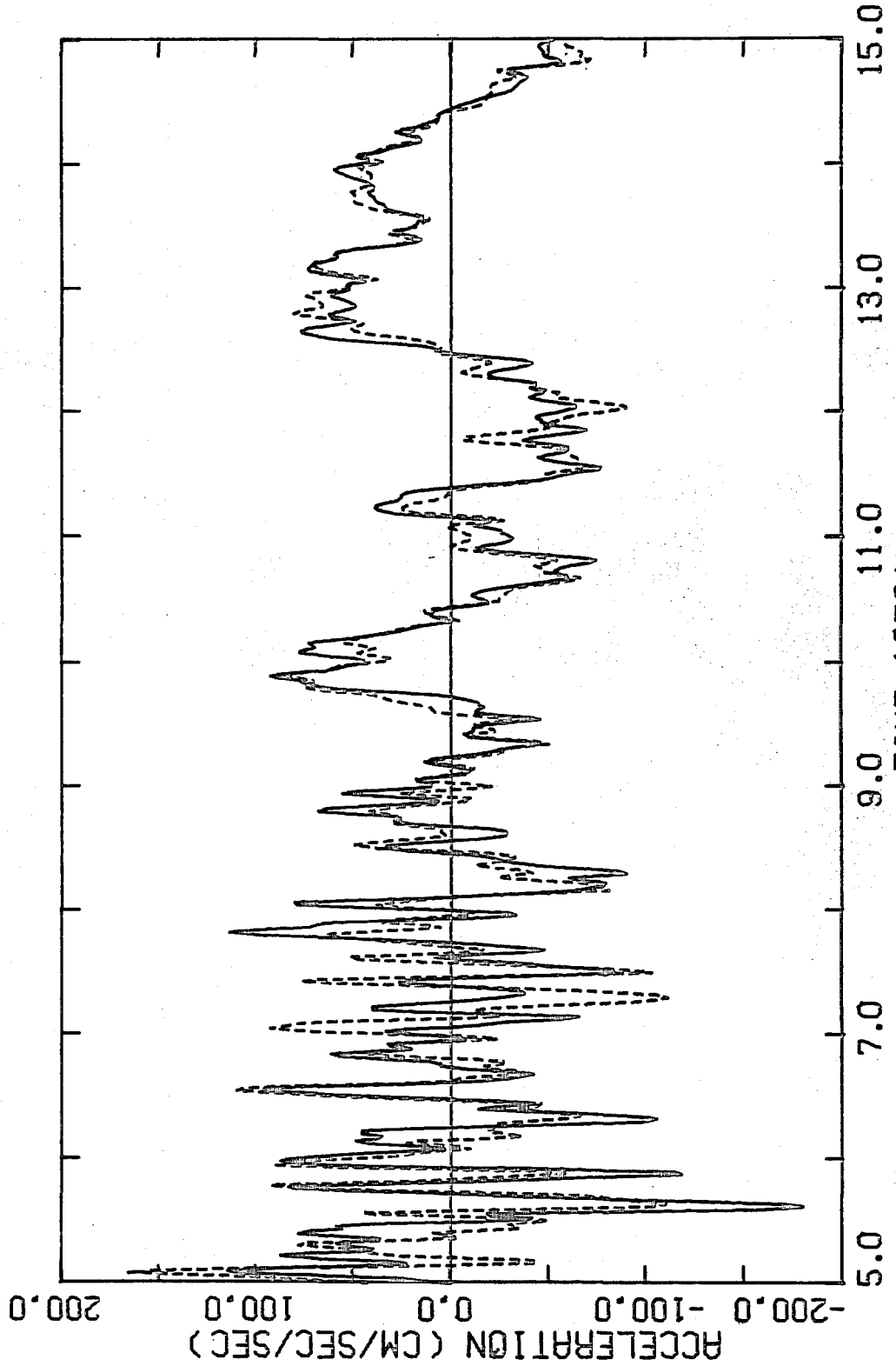


Figure 6.10.c. Relative acceleration (——) at the 19th floor of Union Bank building and calculated acceleration (----) of the same model as in Fig. 6.10.a.

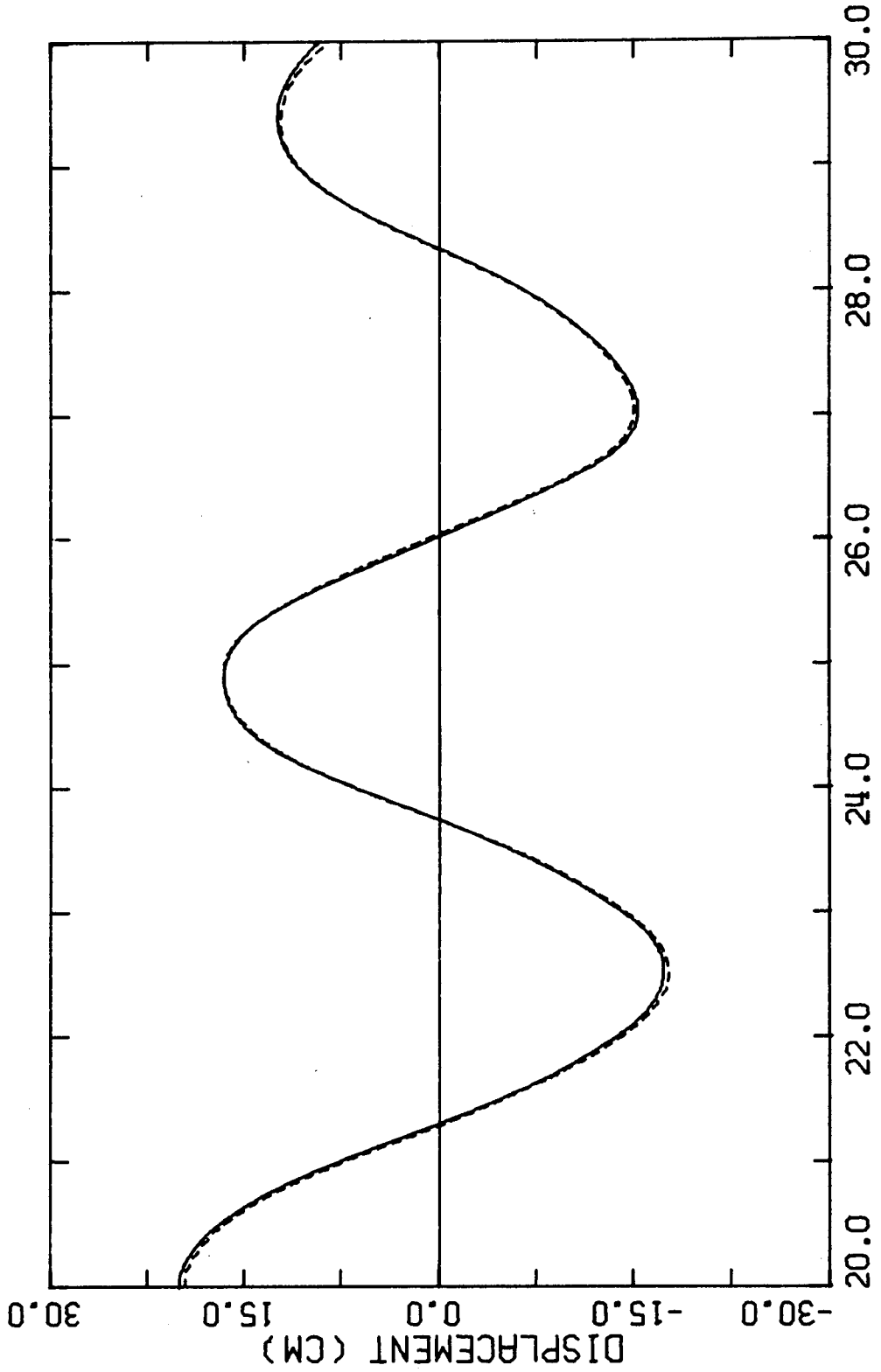


Figure 6.11.a. Relative displacement (——) at the 19th floor of Union Bank building and calculated displacement (----) of the optimal model with two modes determined by matching velocities over the interval.

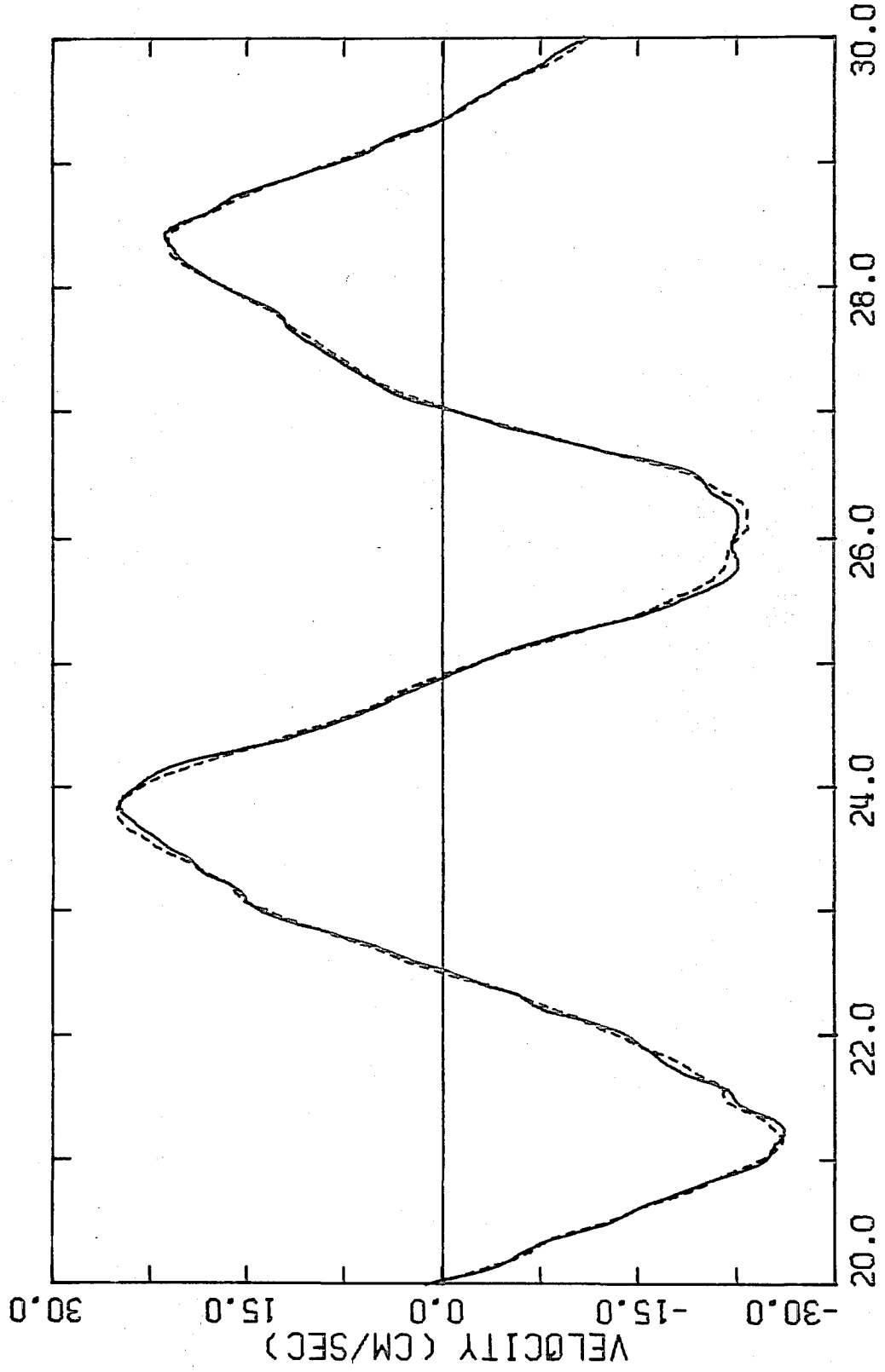


Figure 6.11.b. Relative velocity (——) at the 19th floor of Union Bank building and calculated velocity (----) of the same model as in Fig. 6.11.a.

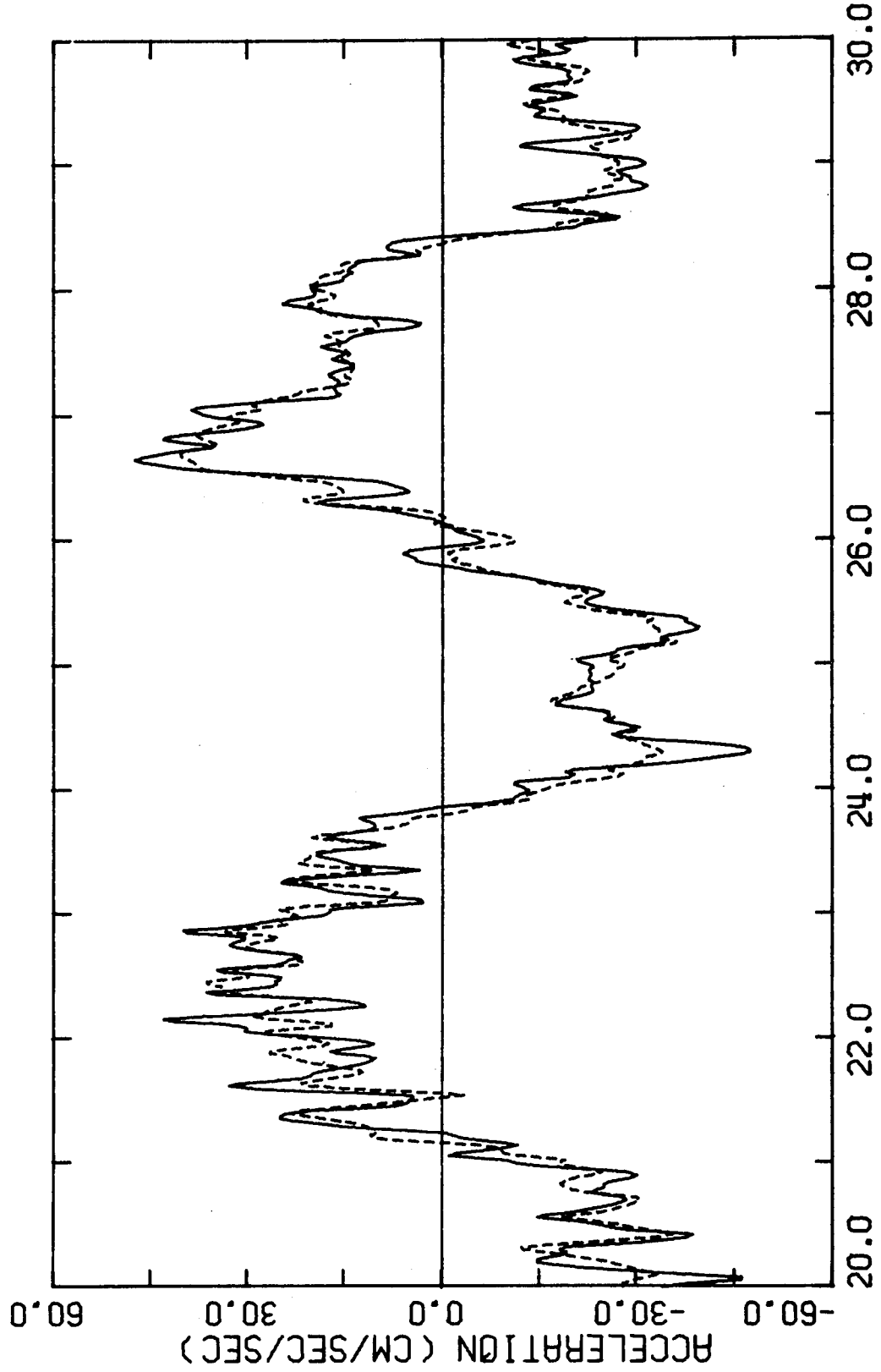


Figure 6.11.c. Relative acceleration (——) at the 19th floor of Union Bank building and calculated acceleration (----) of the same model as in Fig. 6.11.a.

when $\tilde{S}$ was used instead of $\hat{S}$. The reduced sensitivity matrix was also calculated in several cases by replacing derivatives with respect to the model coefficients $a_1^{(r)}$ and $a_2^{(r)}$ [Eq. (5.2.3)] with derivatives with respect to T_r and ζ_r, since it is the latter parameters which are of major interest. To calculate the sensitivity matrices, the derivatives of the model response with respect to the parameters are required. These are the so-called sensitivity coefficients and they were determined by computing the response of a number of single-degree-of-freedom linear systems, which include those given in Eqs. (5.3.4) to (5.3.6). Finally, the first derivatives of J were also evaluated at the optimal estimates as a check, and in each case they were found to be suitably small.

The sensitivity matrices gave quantitative confirmation of points which have already been noted and which can be inferred using other arguments. These points are summarized below.

1) For a given mode, J is much more sensitive to the period than to the other modal parameters.

2) The interaction between T_r and ζ_r, and between T_r and p_r, is generally small, but the interaction between ζ_r and p_r is quite pronounced.

3) The interaction between the parameters of different modes is generally small.

4) J gets progressively less sensitive to the modal parameters as the mode number increases. The rate of decrease in the sensitivity with mode number is greatest for displacement matching and least for

acceleration matching.

5) For the fundamental mode, J is most sensitive to the modal parameters when displacement matching is used. For the second mode, the sensitivity of J is least for displacement matching and about the same for velocity and acceleration matching. For the third and higher modes, the sensitivity of J is greatest for acceleration matching.

These conclusions regarding the sensitivity of J can be interpreted directly in terms of the expected accuracy of the estimates of the parameters, as discussed in §2.4.5. For example, points 1 and 2 indicate that the modal periods will be estimated much more accurately than the damping and participation factors.

The diagonal elements of the reduced sensitivity matrix are presented in (a) of Table 6.7 for the optimal estimates of the parameters given by a three-mode model determined by matching the velocity record over the interval 5 to 30 seconds (column 5, Table 6.3). These values have been normalized by multiplying by the optimal estimates so that the sensitivities can be directly compared, without regard to the magnitude of the parameters. For example, Eq. (2.4.32) and the values in Table 6.7 (a) show that an $\epsilon\%$ change in T_1 from its optimal value will produce a change of $22.9 \times (\epsilon/100)^2$ in J, whereas an $\epsilon\%$ change in ζ_1 will produce a change of only $0.03 \times (\epsilon/100)^2$, or almost $\frac{1}{1000}$ of the previous change. In view of Eqs. (2.4.30) and (2.4.34), the square root of the diagonal elements are more indicative of the accuracy of the estimates. For example, it can be deduced

	r	Sensitivity w. r. t. T_r	Sensitivity w. r. t. ζ_r	Sensitivity w. r. t. p_r	Interaction T_r and ζ_r	Interaction T_r and p_r	Interaction ζ_r and p_r
(a)	1	22.9	0.03	0.24	0.08	0.03	0.80
	2	0.60	0.002	0.004	0.04	-0.08	0.57
	3	0.006	0.0008	0.0008	0.27	-0.15	0.76
(b)	1	1.44	0.004	0.22	0.22	0.31	0.77
	2	1.06	0.002	0.005	0.06	-0.19	0.56
(c)	1	1.92	0.01	0.008	-0.08	-0.15	0.09
	2	0.36	0.0003	0.001	-0.04	-0.22	0.82

TABLE 6.7. Interaction coefficients and diagonal elements of the reduced sensitivity matrix for the optimal estimates given by matching the Union Bank velocity record over the intervals (a) 5 to 30 seconds with a 3-mode model, (b) 5 to 15 seconds with a 2-mode model, and (c) 20 to 30 seconds with a 2-mode model.

from the above that the bound on the relative error in $\hat{T}_1$ is about 4% and 10%, respectively, of the bounds on the relative error in $\hat{\zeta}_1$ and $\hat{p}_1$.

Also shown in Table 6.7 are the interaction coefficients, which are defined for two parameters a_i and a_j by the ratio $-\tilde{S}_{ij}/(\tilde{S}_{ii}\tilde{S}_{jj})^{\frac{1}{2}}$. These coefficients are introduced to indicate the extent of the interaction between the two parameters and they are the analog of correlation coefficients in statistical theory. It can be shown that the magnitude of the interaction coefficients cannot be greater than unity, and the larger the magnitude, the greater the interaction between the corresponding parameters. If an interaction coefficient has unit magnitude, there is a straight line in parameter space along which the two parameters can be varied without changing the value of J, all other parameters remaining fixed. The values in Table 6.7 support point 2 given above.

To illustrate how the sensitivities are affected when the parameters are estimated by matching smaller segments of the record, results for matching velocities over the time intervals from 5 to 15 secs and 20 to 30 secs are also presented in Table 6.7 in (b) and (c) respectively. As was to be expected, the sensitivities, and hence the accuracy of the estimates, are decreased by taking smaller intervals of data. Observe from the sensitivities that the parameters T_2 and ζ_2 of the second mode should be estimated more accurately using the first ten-second time segment rather than the last, but the opposite is true for the fundamental mode. This is because the

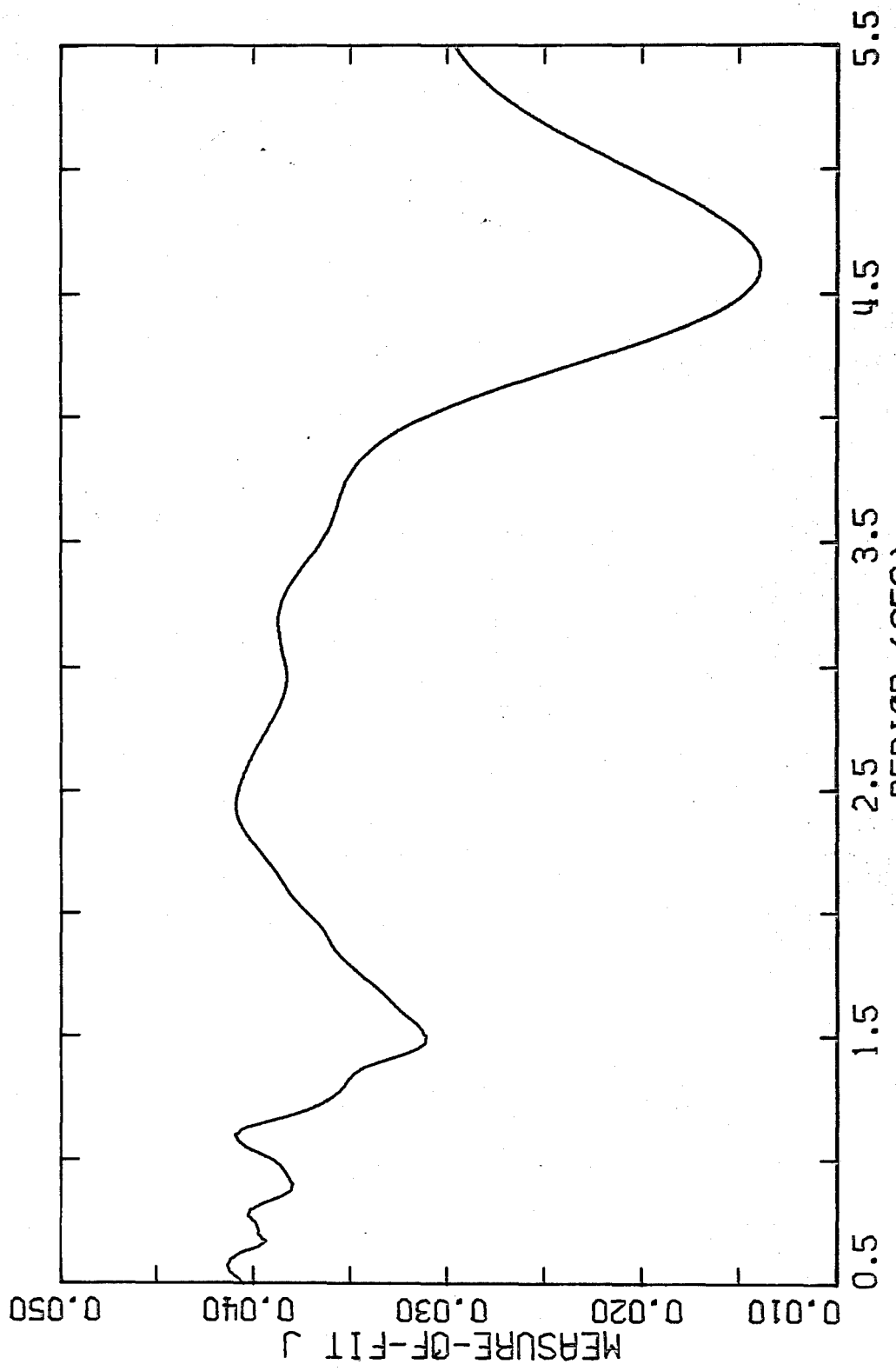


Figure 6.12. Profile of the measure-of-fit J for matching of the acceleration record of Union Bank over the interval 5 to 30 seconds. The period of a single-mode model was varied, with the damping factor held constant at 4%.

higher-mode signals decrease in time compared with the fundamental mode. The participation factors are estimated poorly in the last time segment because their determination is ill-conditioned (§6.1.2).

Finally, a profile of the measure-of-fit J, or more strictly $f_1(T,\zeta)$ [Eq. (5.3.12)], is presented in Fig. 6.12 for matching of the acceleration record over 5 to 30 seconds. The damping factor was held constant at $\zeta = 4\%$ and the period was incremented in steps of 0.025 sec. Observe that the profile is very smooth and that local minima corresponding to the first four longitudinal modes are apparent. The reason for the spurious minimum at $T = 3$ sec is similar to that given at the end of Chapter 5. The absence of a minimum corresponding to the second torsional mode is consistent with the discussion in §6.1.1. A small minimum did appear at 1.2 seconds when the profile of J was replotted after subtracting the contributions of the first and second mode from the acceleration record.

6.2. Building 180, Jet Propulsion Laboratory, Pasadena

Building 180 is a 9-story steel-frame structure on the campus of the Jet Propulsion Laboratory, Pasadena, California, which is located approximately 15 miles from the epicenter of the 1971 San Fernando earthquake. The amplitude of the acceleration response of the building during the earthquake was about twice that of the Union Bank building, but damage was limited to minor nonstructural cracking. Features of the design are discussed by Wood (1972) who also developed two-dimensional models of the building. These analytical

models suggest that the peak stresses in the structural frame approached, but did not exceed, the yield point during the San Fernando earthquake. Building 180 was investigated in this work because Wood's study was available for comparison and, in addition, the peak acceleration was among the largest recorded in a building during the San Fernando earthquake.

Strong-motion accelerographs with synchronized timing were installed in the basement and on the roof (Fig. 6.13). The S82°E components of the digitized relative acceleration, velocity and displacement at the roof were used as the response data in the analysis. These components correspond to the longitudinal direction of the building. The peak acceleration in this direction during the San Fernando earthquake was about 40%g compared with a peak in the transverse direction of about 20%g. The basement absolute acceleration in the longitudinal direction was used as the input to the models (Fig. 6.14).

The Fourier amplitude spectra for 40.96 seconds of the absolute acceleration in the basement and at the roof are shown in Figs. 6.15 and 6.16 respectively, and their ratio is shown in Fig. 6.17. The first six dominant peaks in Fig. 6.16 are presented in Table 6.8 and the first four dominant peaks over the frequency range 0-4Hz in Fig. 6.17 are presented in Table 6.9. There are several points to note in regard to these Tables. First, it will be explained later why both peaks with a period of about 0.4 sec are labelled as the second longitudinal mode. Second, the fourth and fifth peaks listed in Table 6.8 are not resonant peaks because they essentially vanish in the

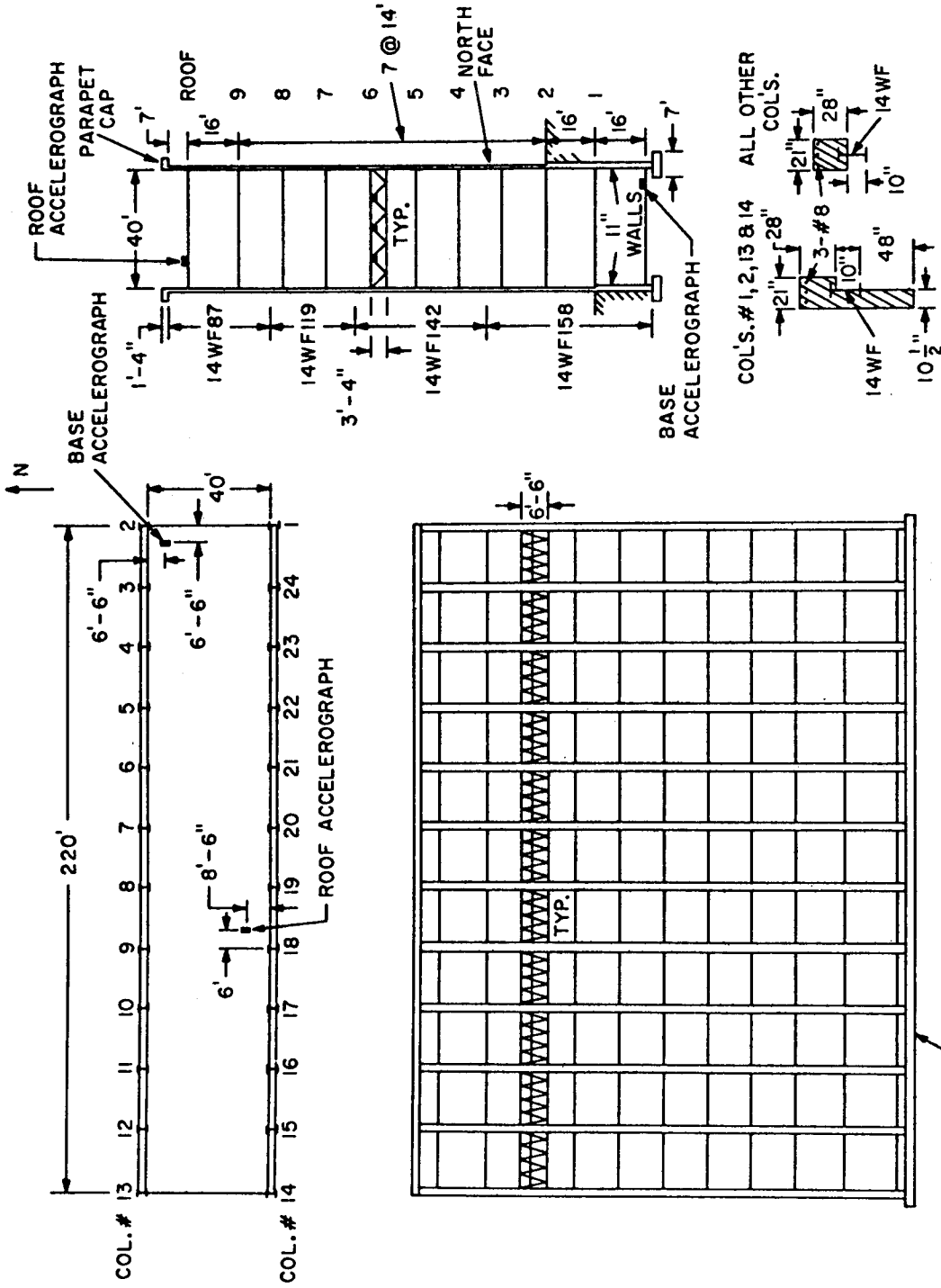


Figure 6.13. Typical floor plan and longitudinal and transverse sections of JPL Building 180 (from Wood, 1972).

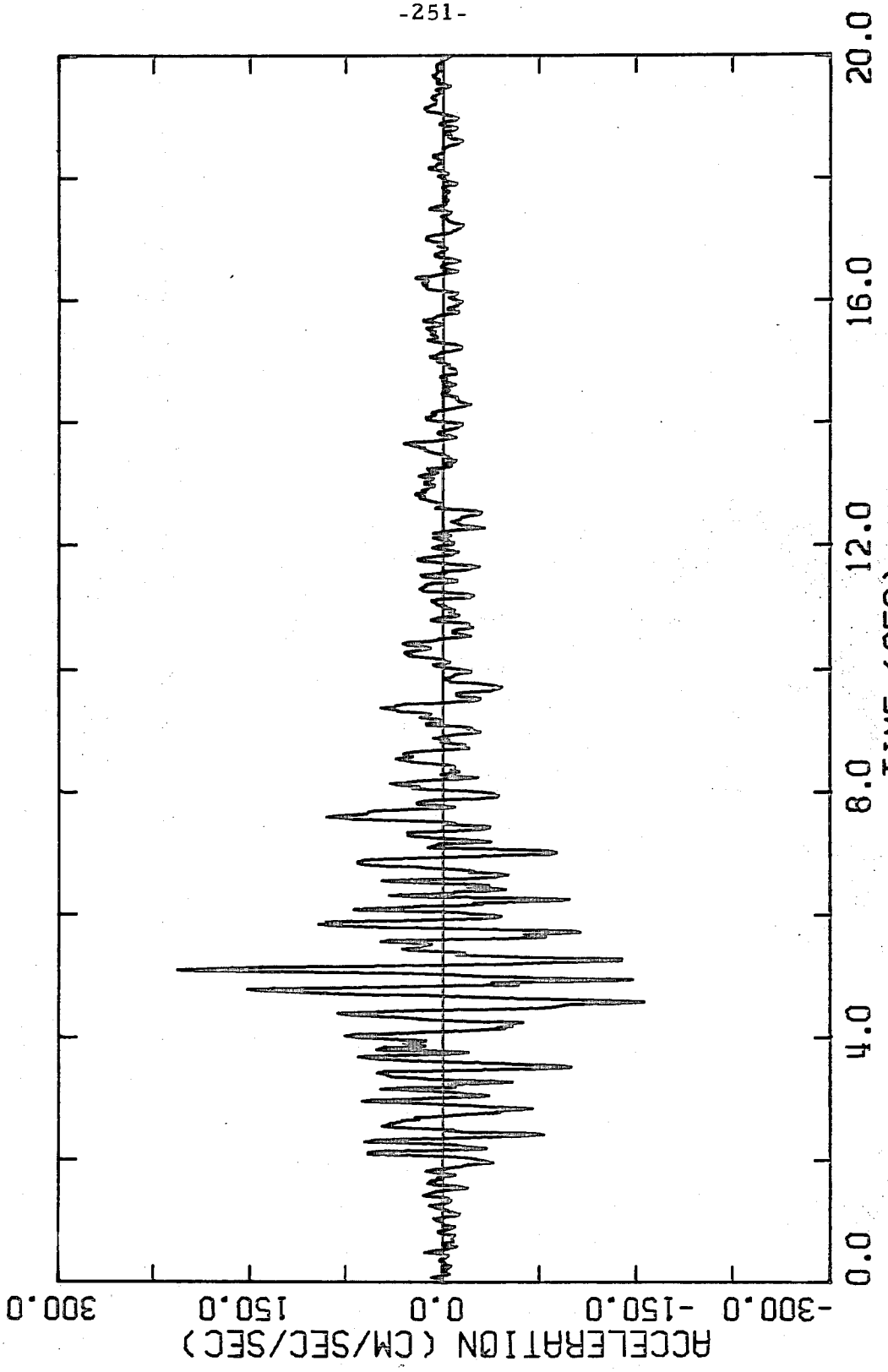


Figure 6.14. S82°E component of the basement absolute acceleration in JPL Building 180 (1971 San Fernando earthquake).

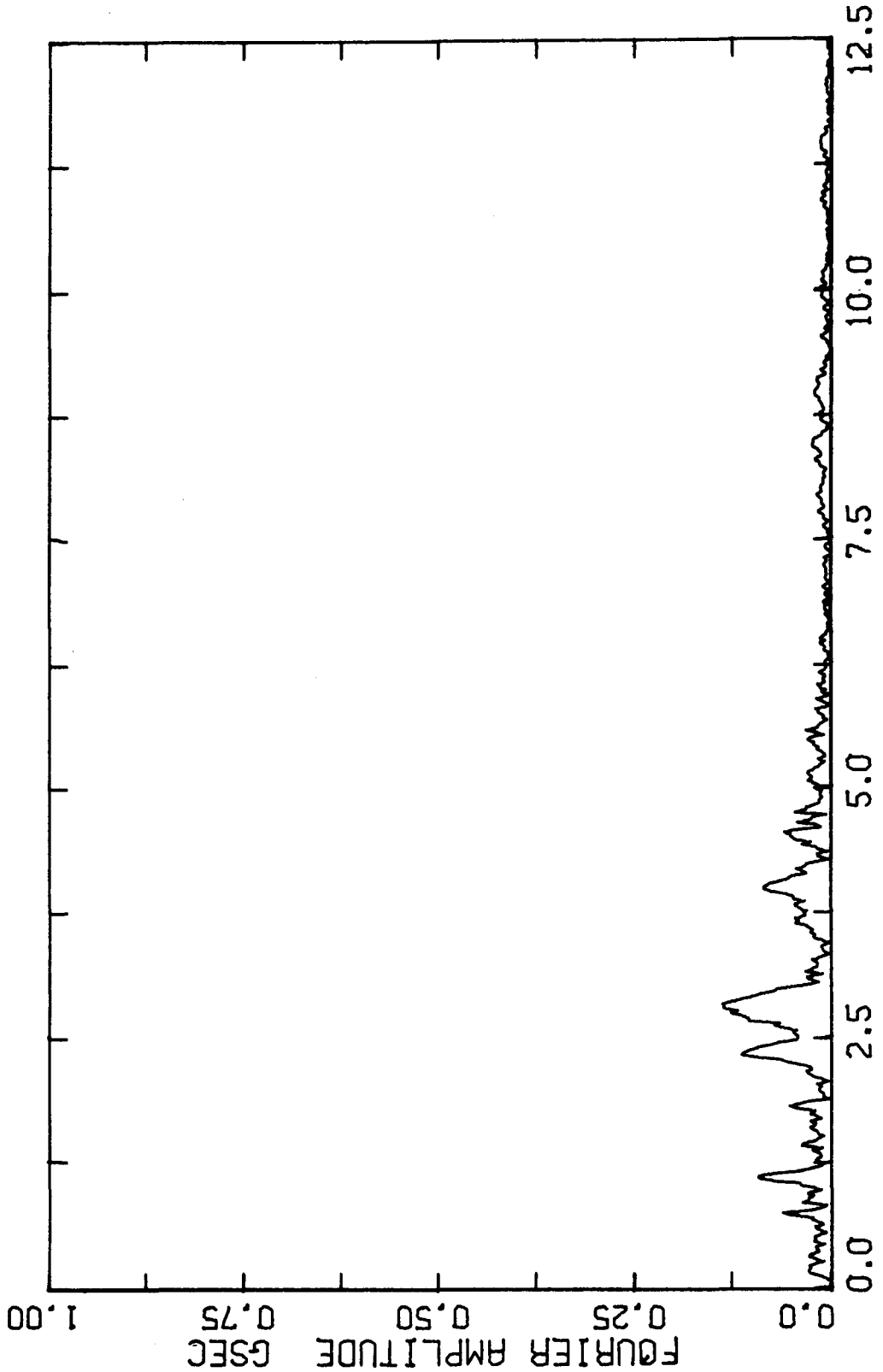


Figure 6.15. Fourier amplitude spectrum of the absolute acceleration, S82°E component, in the basement of JPL Building 180 (1971 San Fernando earthquake).

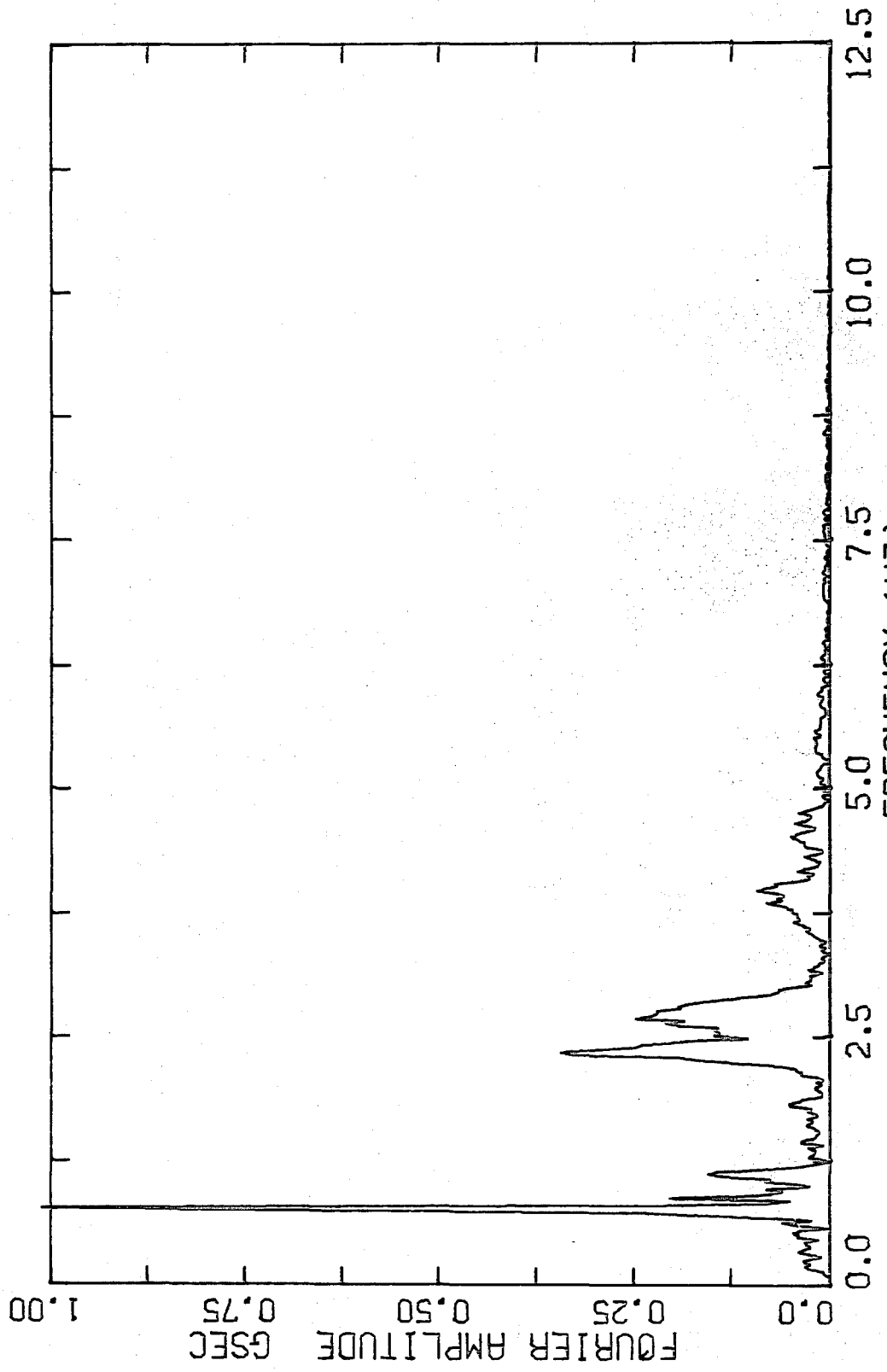


Figure 6.16. Fourier amplitude spectrum of the absolute acceleration, S82°E component, on the roof of JPL Building 180 (1971 San Fernando earthquake).

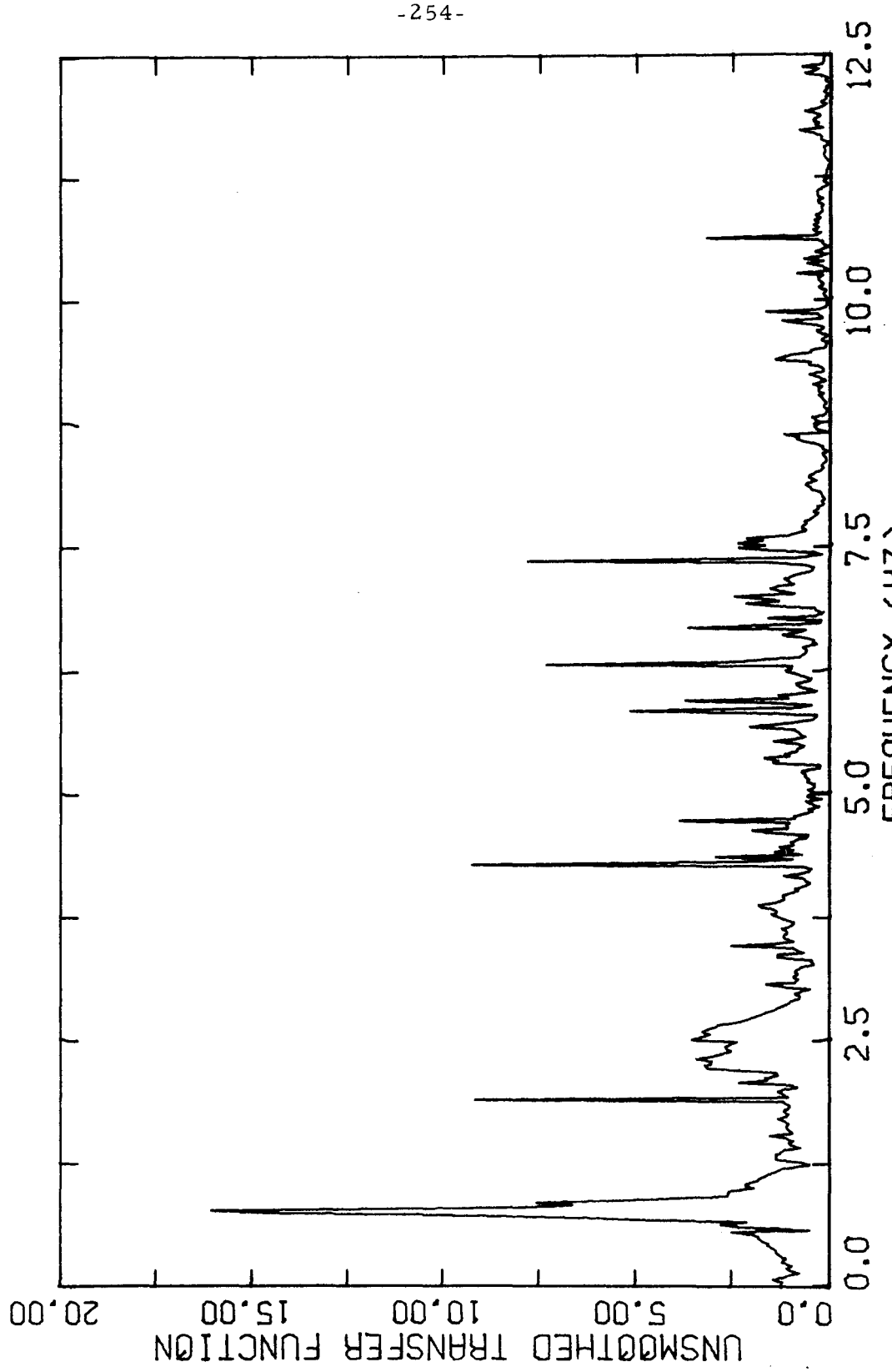


Figure 6.17. Amplitude of the unsmoothed transfer function between the S82°E components of the absolute acceleration in the basement and on the roof, JPL Building 180.

Period (sec)	Period ratio	Ratio of spectral values (%)			Interpretation
		Acceleration	Velocity	Displacement	
1.28	1.0	100	100	100	1st longitudinal
0.42	3.0	34	11	4	2nd longitudinal
0.38	3.4	25	7	2	2nd longitudinal
1.13	1.1	20	18	17	------
0.90	1.4	16	11	8	------
0.26	4.9	9	2	0.4	3rd longitudinal

TABLE 6.8. First six dominant peaks of the Fourier amplitude spectrum of the S82°E component of the absolute acceleration at the roof of Building 180.

Period (sec)	Period ratio	Amplitude ratio (%)	Interpretation of mode
1.28	1.0	100	1st longitudinal
0.43	3.0	21	2nd longitudinal
0.39	3.3	22	2nd longitudinal
0.26	4.9	12	3rd longitudinal

TABLE 6.9. First four dominant peaks of the amplitude of the transfer function between the S82°E components of the absolute acceleration in the basement and at the roof.

transfer function. They apparently come from amplification by the first mode of peaks in the Fourier amplitude spectrum of the base motion. Finally, sharp peaks in the unsmoothed transfer function associated with near zeros in the basement spectrum were ignored because the calculation of the ordinate of the amplitude of the transfer function is ill-conditioned in this case.

6.2.1. Time-invariant Models

The parameters were first estimated for the time segment of the records from 0 to 20 seconds. The general procedure of successively adding modes to the models was followed. The initial estimates of the periods were taken from Table 6.8 (with $\hat{T}_2 = 0.42$ sec) and the initial estimates of the damping factors were 4%.

The results for the optimal models determined by matching the relative displacement, velocity and acceleration records are presented in Table 6.10. The p_r are the effective participation factors at the roof [Eq. (3.2.13)], which are equal to the conventional participation factors α_r for modeshapes normalized to unity at the roof. The values in parentheses are explained later. In marked contrast to the estimates for the Union Bank building (Tables 6.3 and 6.4), the estimates of the damping factors and participation factors determined from different response quantities are not in good agreement. Furthermore, for the optimal models determined by matching displacement and velocity, the measure-of-fit J is significantly greater for a specified number of modes than its counterpart for the Union Bank

Modal parameter	1-mode model (Displacement match)	1-mode model (Velocity match)	1-mode model (Acceleration match)	2-mode model (Velocity match)	2-mode model (Acceleration match)	3-mode model (Acceleration match)
$\hat{T}_1$	1.27	1.26	1.26	1.26	1.25	1.25
$\hat{\zeta}_1$	2.0(2.9)	2.6(3.1)	3.5(3.5)	2.5(3.3)	4.2(3.6)	4.2(3.6)
$\hat{p}_1$	0.9(1.3)	1.1(1.3)	1.3(1.3)	1.0(1.3)	1.5(1.3)	1.5(1.3)
$\hat{T}_2$				0.35	0.37	0.38
$\hat{\zeta}_2$				14(12.4)	13(12.7)	5.3(10.2)
$\hat{p}_2$				-0.52(-0.46)	-0.47(-0.46)	-0.24(-0.46)
$\hat{T}_3$						0.3
$\hat{\zeta}_3$						12
$\hat{p}_3$						-0.4
$J^{\frac{1}{2}}(\%)$	15.1	14.2	12.6	12.2	8.2	6.9

TABLE 6.10. Optimal estimates of the parameters of the longitudinal modes using the first 20 seconds of the S82°E records of JPL Building 180. The values in parentheses are explained in the text.

building (Table 6.3). Part of the difference between the recorded and model displacements appears to be a long-period error in the record of period 6 to 7 seconds, which is apparent over the interval from 10 to 20 seconds in the comparison of displacements. Another part is due to a significant change in the fundamental period with time which shows up most clearly in the displacement comparisons (see, for example, Fig. 6.18.a). A surprising result is that for a given number of modes, the acceleration record can be matched better than the displacement and velocity records (see $J^{\frac{1}{2}}$ in Table 6.10), which was not the case with the Union Bank building.

Other problems with the estimates in Table 6.10 are apparent. For example, the values $\hat{p}_1 = 0.9$ and $\hat{p}_1 = 1.0$ are suspect if judged on the basis of their values for linear models. An examination of the expression for α_1 [Eq. (3.28)] shows that p_1 is greater than unity for a linear model if the modeshape of the fundamental mode has its greatest value at the roof. In fact, for a uniform shear beam, $p_1 = 1.27$. Furthermore, the estimate in Table 6.10 for the participation factor of the third dominant mode has the opposite sign to what would be expected for a third translational mode (for a uniform shear beam, $p_3 = 0.25$). A curious result is that the period estimate for the third dominant mode corresponds to a trough in Figs. 6.16 and 6.17, although its large damping would give the corresponding modal peak a broad bandwidth which could account for contributions in the response at nearby frequencies. It is thought that the difficulty in identifying the third longitudinal mode is partially due to its relatively small

signal in the records (Table 6.8).

The results in Table 6.10 suggest strongly that there is a pronounced interaction between $\hat{p}_r$ and $\hat{\zeta}_r$ in the calculations since the small damping estimates are associated with small participation factors and the large estimates are similarly correlated. This interaction could be suppressed by fixing the p_r on the basis of prior information during the minimization of the measure-of-fit J. Alternatively, since there is reason to believe that the interaction is almost linear, $\hat{\zeta}_r$ and $\hat{p}_r$ can be scaled by the same factor so that $\hat{p}_r$ is equal to some prior value. For example, if the $\hat{\zeta}_r$ and $\hat{p}_r$ in Table 6.10 are scaled to give the participation factors from one of Wood's synthesized structural models, the damping estimates are in much better agreement, as shown by the values in parentheses in Table 6.10. The high damping factor for the second mode may be a consequence of the large change in its period (§6.2.2). In effect, the method may attempt to include the two broad peaks at 2.5 Hz in Fig. 6.16.

A question of considerable interest is why the interaction of the damping and participation factors is so pronounced compared with the Union Bank results. A sensitivity analysis similar to §6.1.3 showed, somewhat unexpectedly, that the sensitivities were almost the same for the two buildings, regardless of whether displacement, velocity or acceleration matching were used. The interaction coefficents involving T_r, and the interaction coefficient for ζ_2 and p_2,

however, were larger for Building 180. The results are illustrated in Table 6.11 which can be compared with their counterparts in (a) of Table 6.7.

r	Sensitivity w. r. t. T_r	Sensitivity w. r. t. ζ_r	Sensitivity w. r. t. p_r
1	67.4	0.04	0.14
2	0.23	0.004	0.005
r	Interaction T_r and ζ_r	Interaction T_r and p_r	Interaction ζ_r and p_r
1	0.04	-0.42	0.78
2	0.14	-0.22	0.90

TABLE 6.11. Results corresponding to Table 6.7 but for the Building 180 velocity record over the interval from 0 to 20 seconds. The corresponding optimal estimates are given in column 5, Table 6.10.

It is thought that the above difficulties may be associated with the much stronger response of Building 180. It is shown in §6.2.2 by using time-varying models that the structure exhibited a marked nonlinear or time-varying behavior during the earthquake. The temporal change of the equivalent linear parameters may possibly allow more interaction to occur between the damping and participation factors during the matching of the responses.

Despite the difficulties with the damping and participation factors, the model responses were in good agreement with the records.

The calculated response of the optimal model with two modes determined by matching velocities is compared in Figs. 6.18.a,b and c with the recorded response. The optimal estimates of the initial displacement and velocity are nonzero because in the presence of model error (ignored higher modes and nonlinearities in the structural response), these values give the smallest value of J, although this is achieved at the expense of a poor match over the first two seconds. A point of interest is that the acceleration of this model was very similar to that of the optimal model with two modes determined by matching accelerations, but the velocities of these two models were significantly different. The parameters for the two models are given in columns 5 and 6 of Table 6.10.

The optimal estimates of the parameters of the three longitudinal modes, which were determined from the relative acceleration record, are compared in Table 6.12 with values from other sources. The parameters of the first three modes of Wood's refined model were determined by Wood (1972) as follows: the participation factors were determined from a synthesized model; the periods were estimated from the transfer function calculated from the earthquake records; and the damping factors were estimated by attempting to match through trial and error the resonant peaks of the Fourier amplitude spectra of the recorded and model accelerations. Difficulties were encountered in the latter approach because the resonant peaks from the records were much broader than those from the model response; presumably because of the change in time of the modal frequencies. Wood

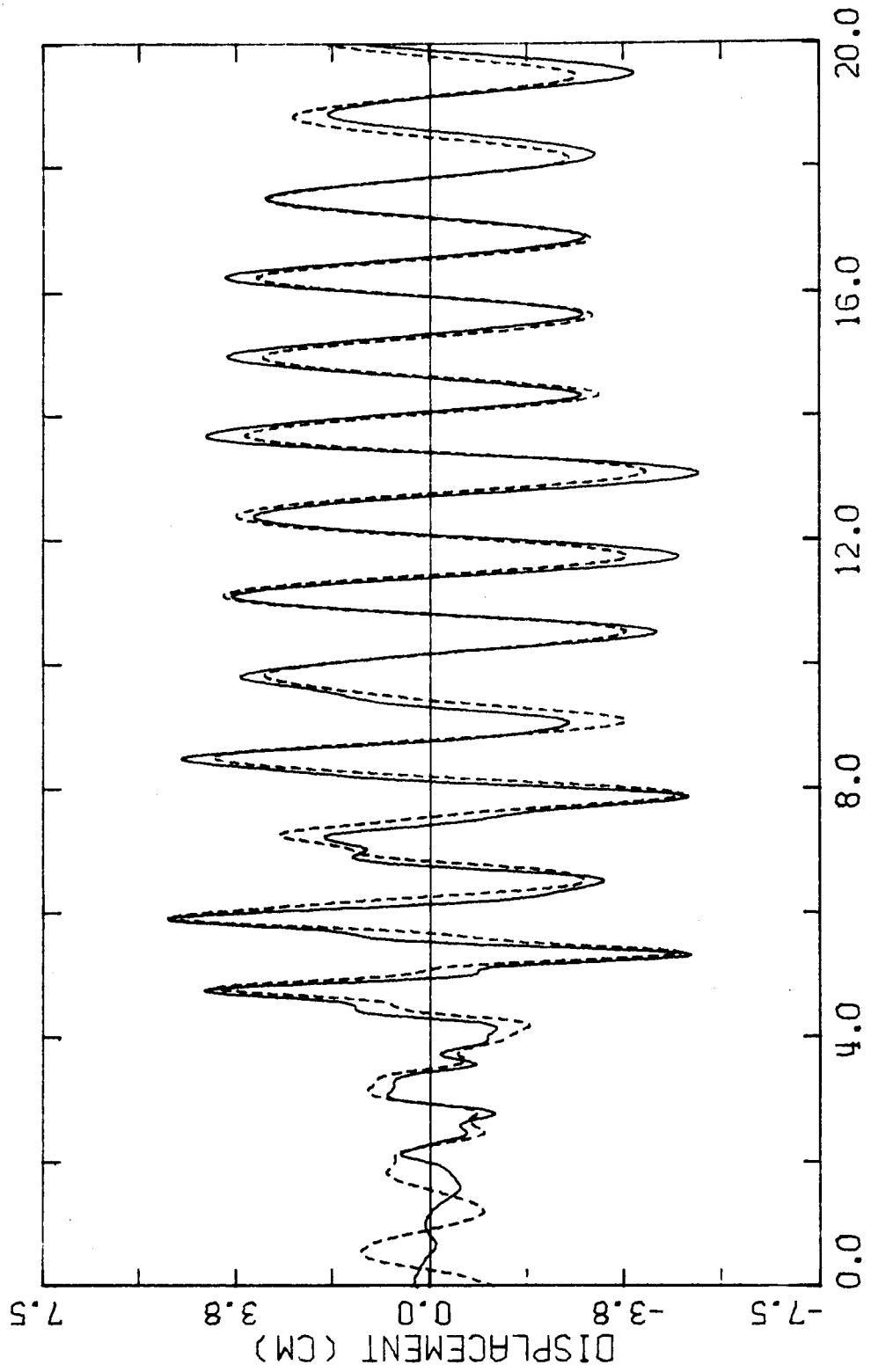


Figure 6.18.a. Relative displacement (——) at the roof of JPL Building 180 and calculated displacement (----) of the optimal model with two modes determined by matching velocities over the interval.

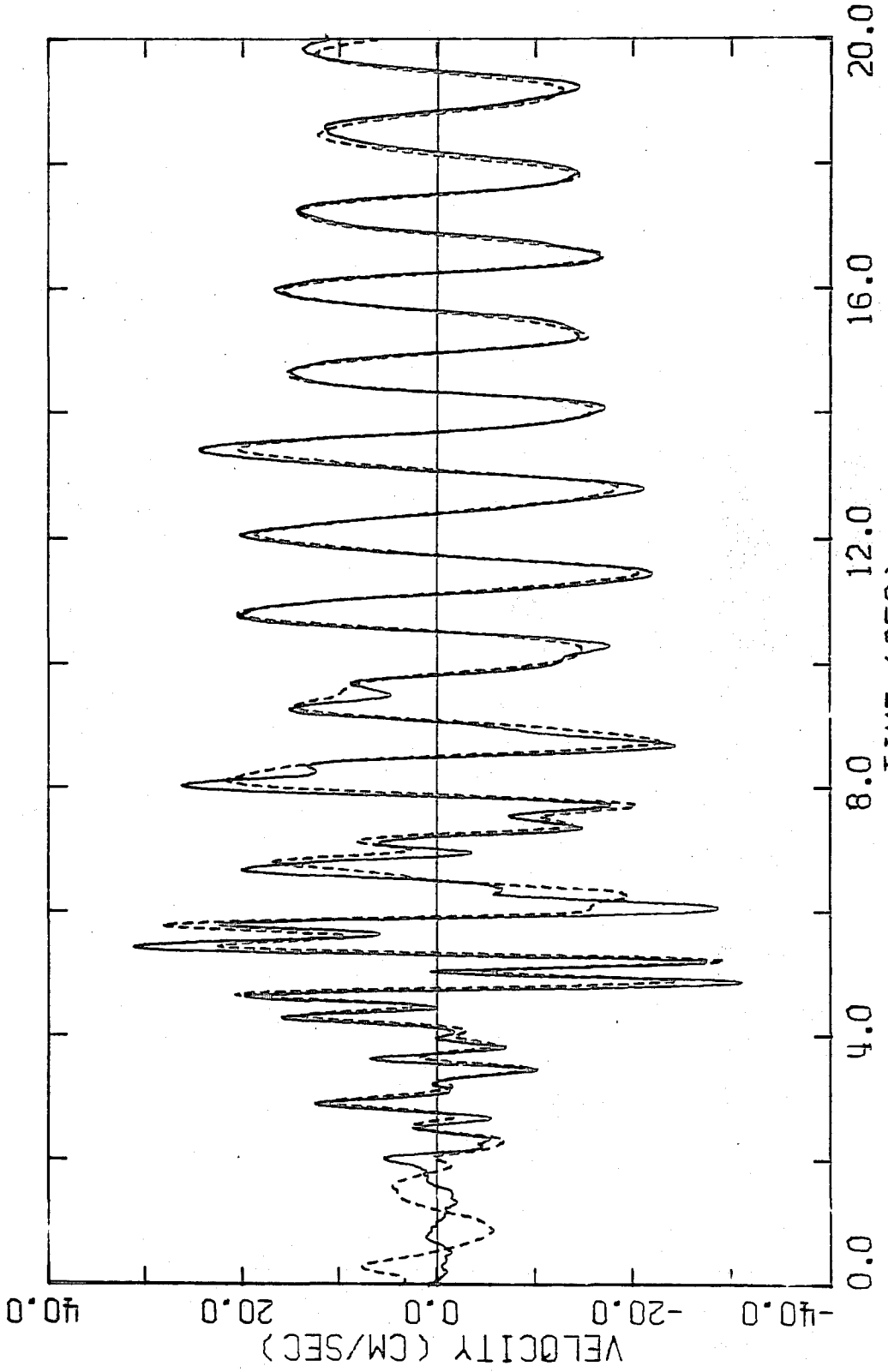


Figure 6.18.b. Relative velocity (——) at the roof of JPL Building 180 and calculated velocity (----) of the same model as in Fig.6.18.a.

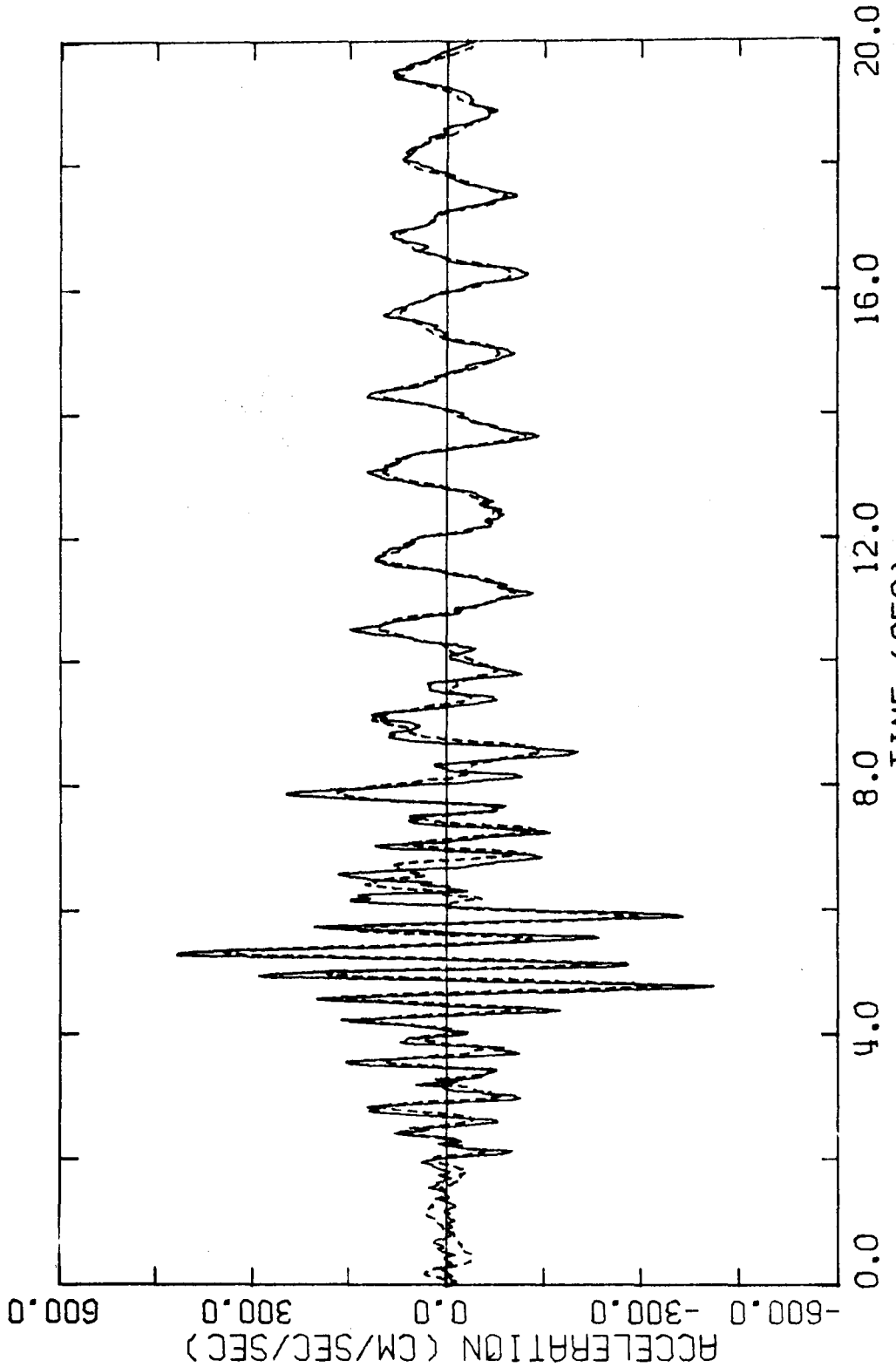


Figure 6.18.c. Relative acceleration (——) at the roof of JPL Building 180 and calculated acceleration (----) of the same model as in Fig. 6.18.a.

Source of values	Period (sec)			Damping factor (%)			Participation factor		
	T_1	T_2	T_3	ζ_1	ζ_2	ζ_3	p_1	p_2	p_3
Optimal estimates from acceleration records, 1971 San Fernando earthquake	1.25	0.38	0.3(?)	4	5	12(?)	1.5	-0.2	-0.4(?)
	T_1/T_r= 3.3		4.2						
Uniform shear beam	1.25	0.42	0.25				1.27	-0.42	0.26
	T_1/T_r= 3.0		5.0						
Wood's refined model based on records from 1971 San Fernando earthquake (Wood, 1972)	1.29	0.42	0.26	4	6	6	1.30	-0.46	0.24
	T_1/T_r= 3.1		5.0						
Forced vibration test (Nielsen, 1964)	0.99	0.33	0.20	0.5	1	2			
	T_1/T_r= 3.0		5.0						

TABLE 6.12. Comparison of the optimal estimates of the parameters of the first three longitudinal modes (from Table 6.10) with other available data for JPL Building 180.

(i) Lengthening in periods & larger damping than for small-amplitude forced vibration tests.

chose to keep the participation factors constant and to vary the damping until the spectral peaks of the second and third mode were about 30% higher than the corresponding peaks from the records. His approach suppresses the interaction effects discussed above.

It was mentioned earlier that there are twin peaks in the Fourier amplitude spectrum of the acceleration record at a period of about 0.4 sec (Fig. 6.16). Wood chose the spectral peak at 0.42 sec as the period of the second longitudinal mode in his refined model, whereas the modal minimization method chose a second-mode period corresponding to the peak at 0.38 sec. It is shown in the next section that the second-mode period changed considerably during the earthquake response and that the value $\hat{T}_2 = 0.38$ sec corresponds to the initial strong-motion portion while the value $\hat{T}_2 = 0.42$ sec corresponds to the later portion of the response, which is almost free vibrations.

The match of the recorded acceleration given by the optimal model with three modes (Tables 6.10 and 6.12) and the corresponding match for the three lowest modes of Wood's refined model are presented in Figs. 6.19.a and b respectively. To allow the calculated responses to be compared on the same basis, both models were started from rest, although the estimates of the initial conditions for the contribution of each mode in the optimal model were nonzero. The calculated velocities for the same models are presented in Figs. 6.20.a and b. The optimal model clearly gives a much better fit to the recorded response than Wood's model. This is primarily because the model response is very sensitive to the modal periods and the

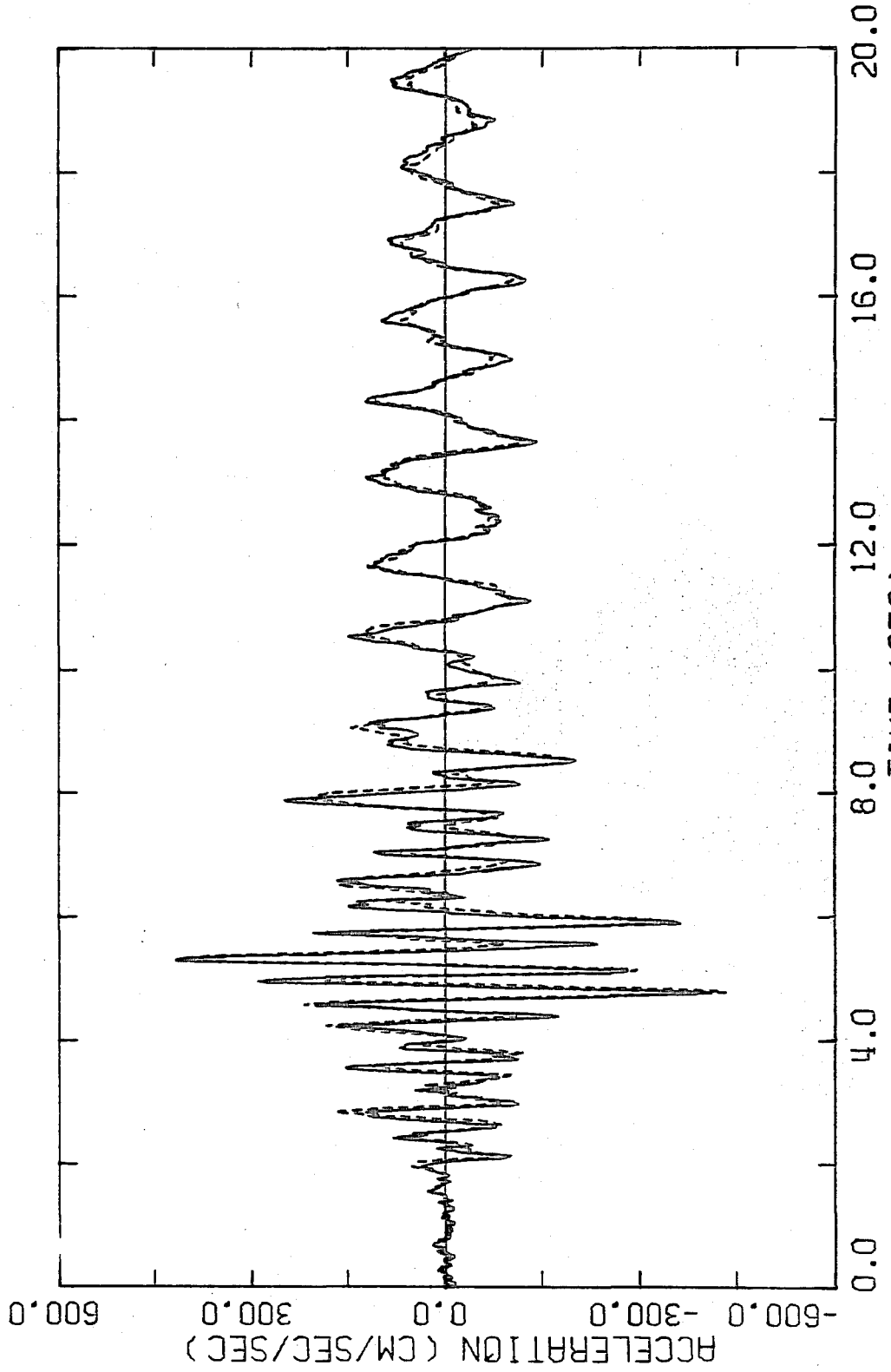


Figure 6.19.a. Relative acceleration (——) at the roof of JPL Building 180 and calculated acceleration (----) of the optimal model with three modes determined by matching accelerations over the interval. The model started from rest. ($J^{\frac{1}{2}} = 7.3\%$).

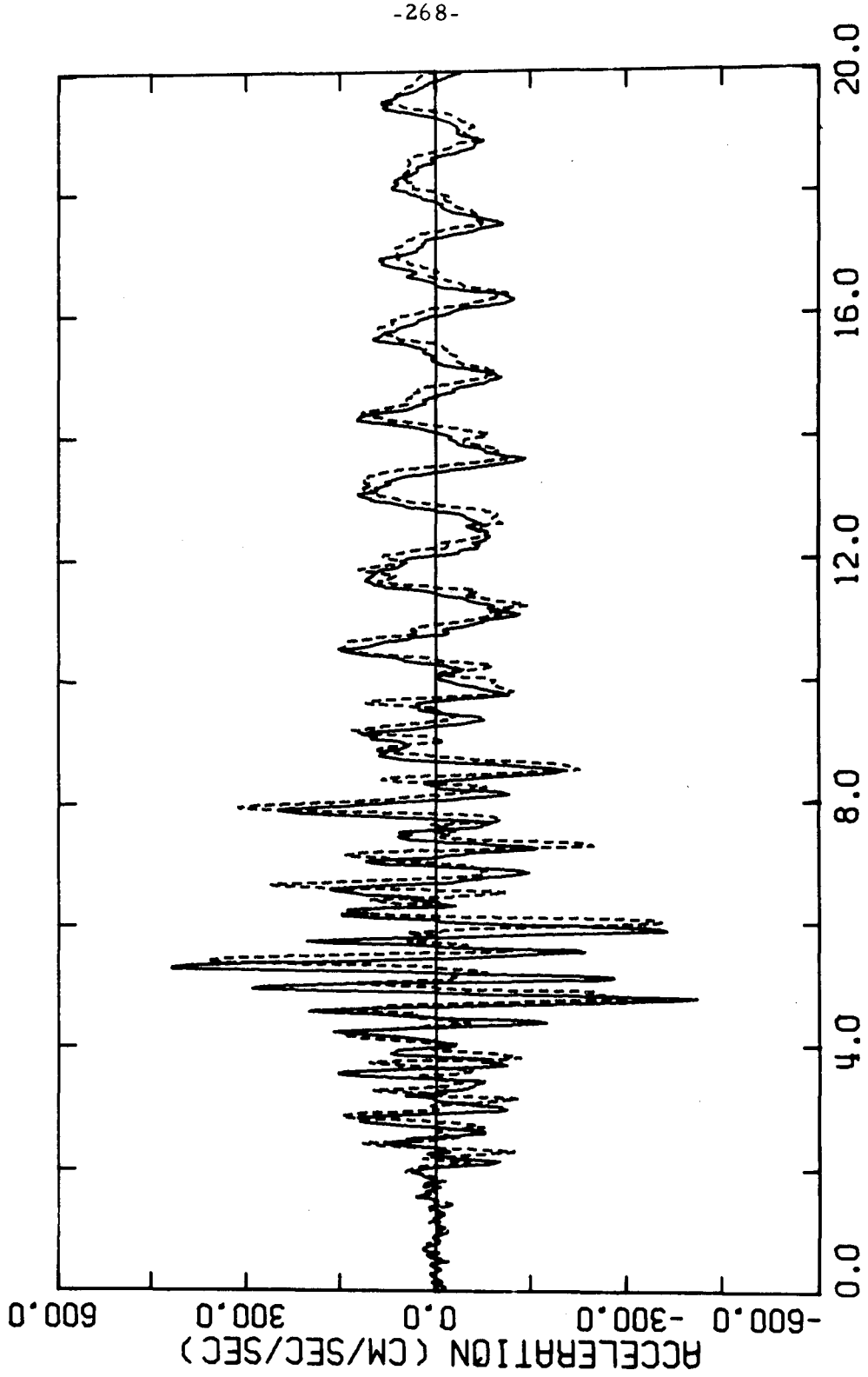


Figure 6.19.b. Relative acceleration (——) at the roof of JPL Building 180 and calculated acceleration (- - - -) due to the three lowest modes of Wood's refined model. The model started from rest. (J^2= 19.8%).

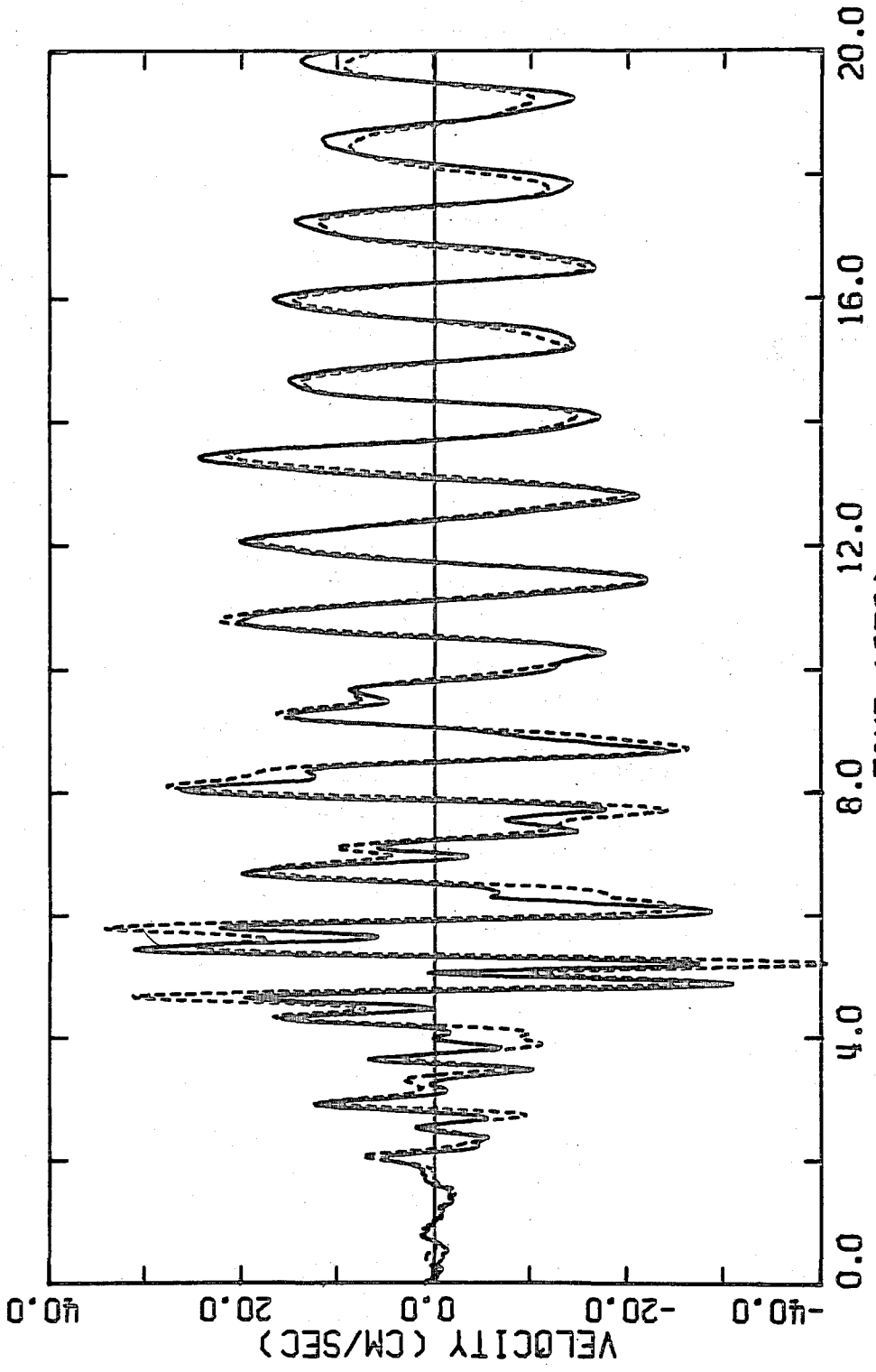


Figure 6.20.a. Relative velocity (——) at the roof of JPL Building 180 and calculated velocity (----) of the same model as in Fig. 6.19.a. ($J^{\frac{1}{2}} = 14.3\%$).

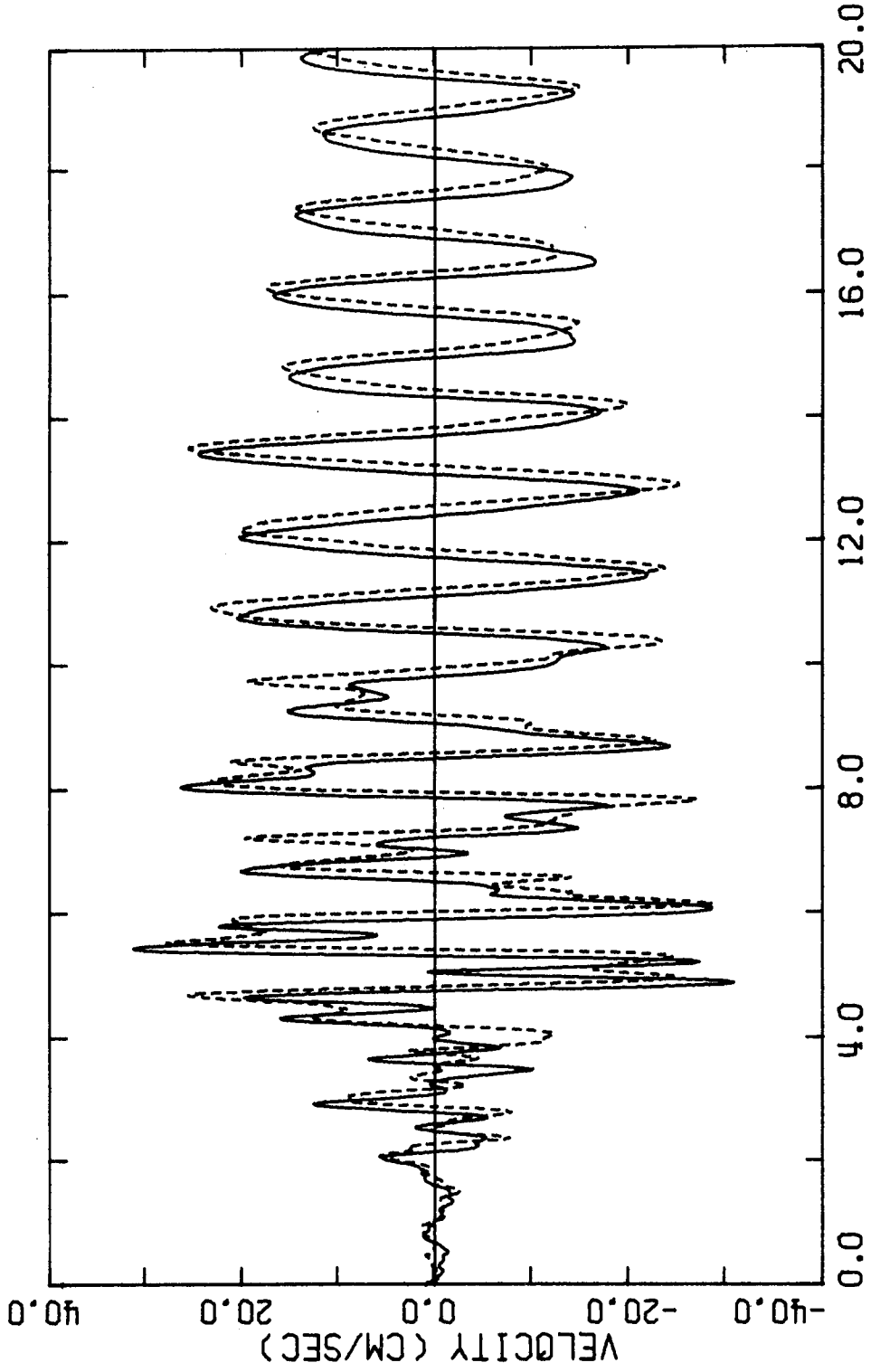


Figure 6.20.b. Relative velocity (——) at the roof of JPL Building 180 and calculated velocity (----) of the same model as in Fig. 6.19.b. ($J_{\text{本}}$=28.4%).

modal minimization method finds the values for the periods which optimize the response match (acceleration in this case). It is of interest that the velocity of this optimal model is also in good agreement with the record.

In conclusion, the time-invariant models for Building 180 did not perform as well as they did for the Union Bank building. There are at least two possible reasons for this. First, the earthquake response of Building 180 was considerably stronger than that of the Union Bank and so effects not included in the linear models may be more pronounced. The second reason relates to the number of cycles of the record which were used to determine the time-invariant models. For the Union Bank building, the 25-second portion used is only about $5\frac{1}{2}$ cycles of the fundamental longitudinal mode, whereas the 20-second segment used for Building 180 corresponds to about 16 cycles of the fundamental mode. Thus, even if the average relative change in the modal parameters per fundamental cycle was approximately the same in the two buildings, the overall change would be greater for Building 180. The change in the modal parameters with time is investigated in the next subsection by determining the optimal estimates for a succession of short time segments of the records.

6.2.2. Time-varying Models

In the initial investigation into how the equivalent linear parameters changed during the earthquake, time-windows of five seconds were used, which corresponds to about 4 cycles of the fundamental

mode. Over the first 10 seconds of response, the optimal estimates of the damping and participation factors exhibited even greater interaction than for the time-invariant models and corresponding estimates for velocity matching and for acceleration matching were not consistent. However, the modal periods were consistent and gave interesting results, so these are presented in Table 6.13. It can be seen that over the first 10 seconds or so of response, the fundamental period changed from 1.1 sec to 1.3 sec, or by about 20%, and the period of the second mode changed from 0.33 sec to 0.42 sec, or by about 30%. The value 0.33 sec is the same as that determined by Nielsen (1964) from forced vibration tests performed on Building 180 before the architectural work was completed (Table 6. 12).

It was found that time-windows of ten seconds produced more consistency between the various estimates of the damping and participation factors. The results for these segments of the records are presented in Table 6. 14. The damping of the fundamental mode apparently decreased from about 5% in the first 10 seconds to about 3% in the interval from 10 to 20 seconds. The corresponding decrease in the damping of the second mode was from about 12% to about 4%. The high damping factor of the second mode in the first time segment may be a spurious effect due to the considerable change in the period of the second mode over this interval (Table 6. 13)

The calculated velocity and acceleration of the optimal model with two modes determined by matching the first ten seconds of the velocity record are compared with the corresponding recorded responses in Figs. 6. 21. a and b. A similar comparison is given in

Time interval	Record used	$\hat{T}_1$	$\hat{T}_2$	$\hat{T}_1/\hat{T}_2$
0-5	Velocity	1.09		
	Acceleration	1.10		
	Velocity	1.08	0.33	3.3
	Acceleration	1.08	0.32	3.4
2-7	Velocity	1.19		
	Acceleration	1.19		
	Velocity	1.20	0.37	3.2
	Acceleration	1.21	0.37	3.3
4-9	Velocity	1.24		
	Acceleration	1.25		
	Velocity	1.25	0.38	3.3
	Acceleration	1.25	0.38	3.3
6-11	Velocity	1.27		
	Acceleration	1.28		
	Velocity	1.29	0.42	3.1
	Acceleration	1.29	0.42	3.1
10-15	Velocity	1.26		
	Acceleration	1.26		
	Velocity	1.26	0.42	3.0
	Acceleration	1.26	0.42	3.0
15-20	Velocity	1.27		
	Acceleration	1.27		
	Velocity	1.27	0.42	3.0
	Acceleration	1.27	0.41	3.1

TABLE 6.13. Optimal estimates of the modal periods for different segments of the records from JPL Building 180, San Fernando earthquake. One-mode and two-mode models were used.

Time interval	Record used	$\hat{T}_1$	$\hat{\zeta}_1$	$\hat{p}_1$	$\hat{T}_2$	$\hat{\zeta}_2$	$\hat{p}_2$	$J^{\frac{1}{2}}$ (%)
0-10	Velocity	1.23	4.6	1.22				16.5
	Velocity	1.23	4.6	1.21	0.36	11.6	-0.43	12.9
	Acceleration	1.23	5.2	1.33				17.1
	Acceleration	1.23	8.0	1.60	0.37	12.4	-0.48	10.1
10-20	Velocity	1.27	2.9	1.24				5.0
	Velocity	1.27	2.8	1.16	0.41	3.6	-0.33	4.4
	Acceleration	1.27	3.0	1.25				8.7
	Acceleration	1.27	3.0	1.20	0.41	4.0	-0.34	6.2

TABLE 6.14. Optimal estimates of the modal parameters for different segments of the records from JPL Building 180, San Fernando earthquake. One-mode and two-mode models were used.

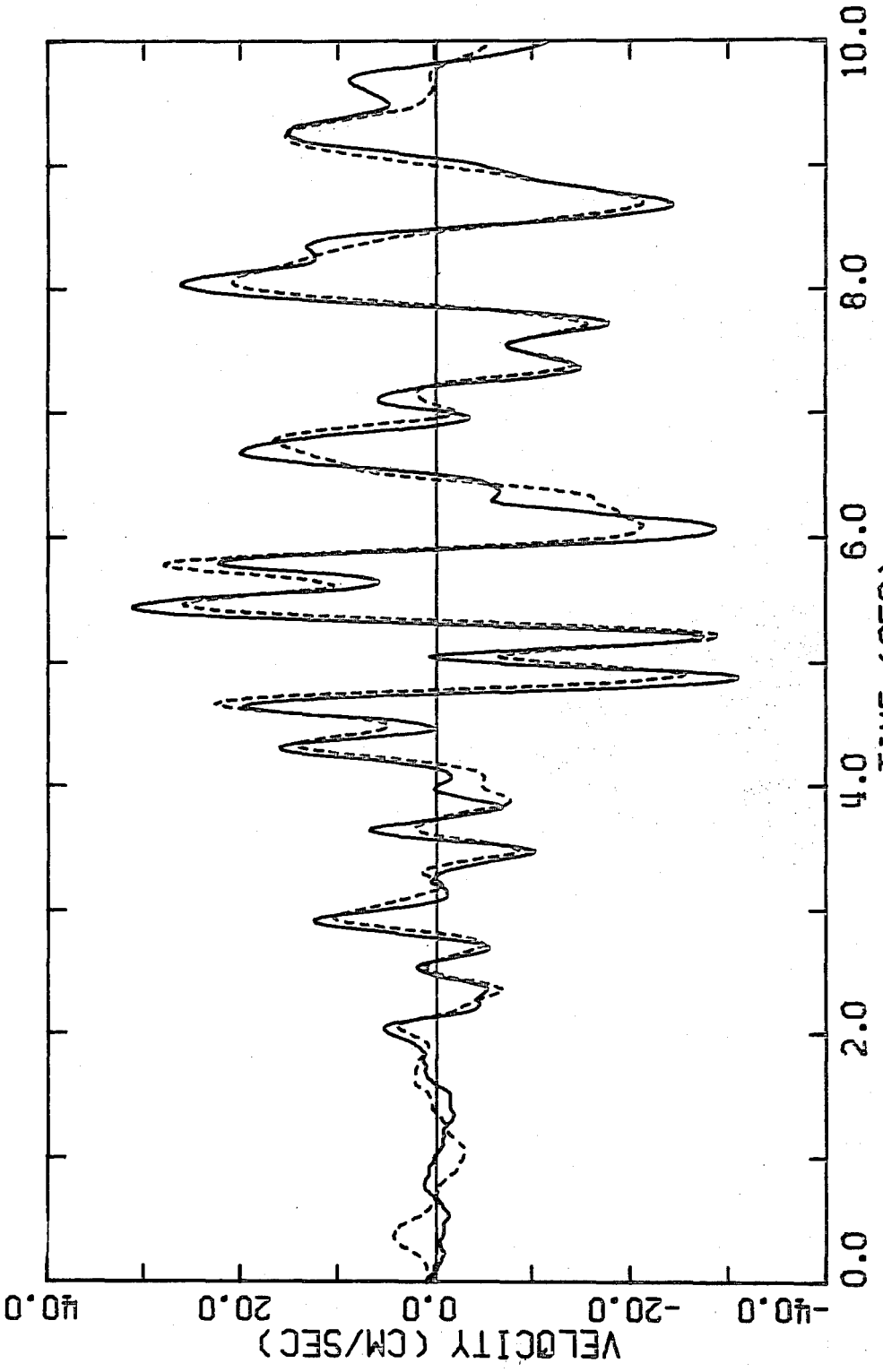


Figure 6.21.a. Relative velocity (——) at the roof of JPL Building 180 and calculated velocity (----) of the optimal model with two modes determined by matching velocities over the interval.

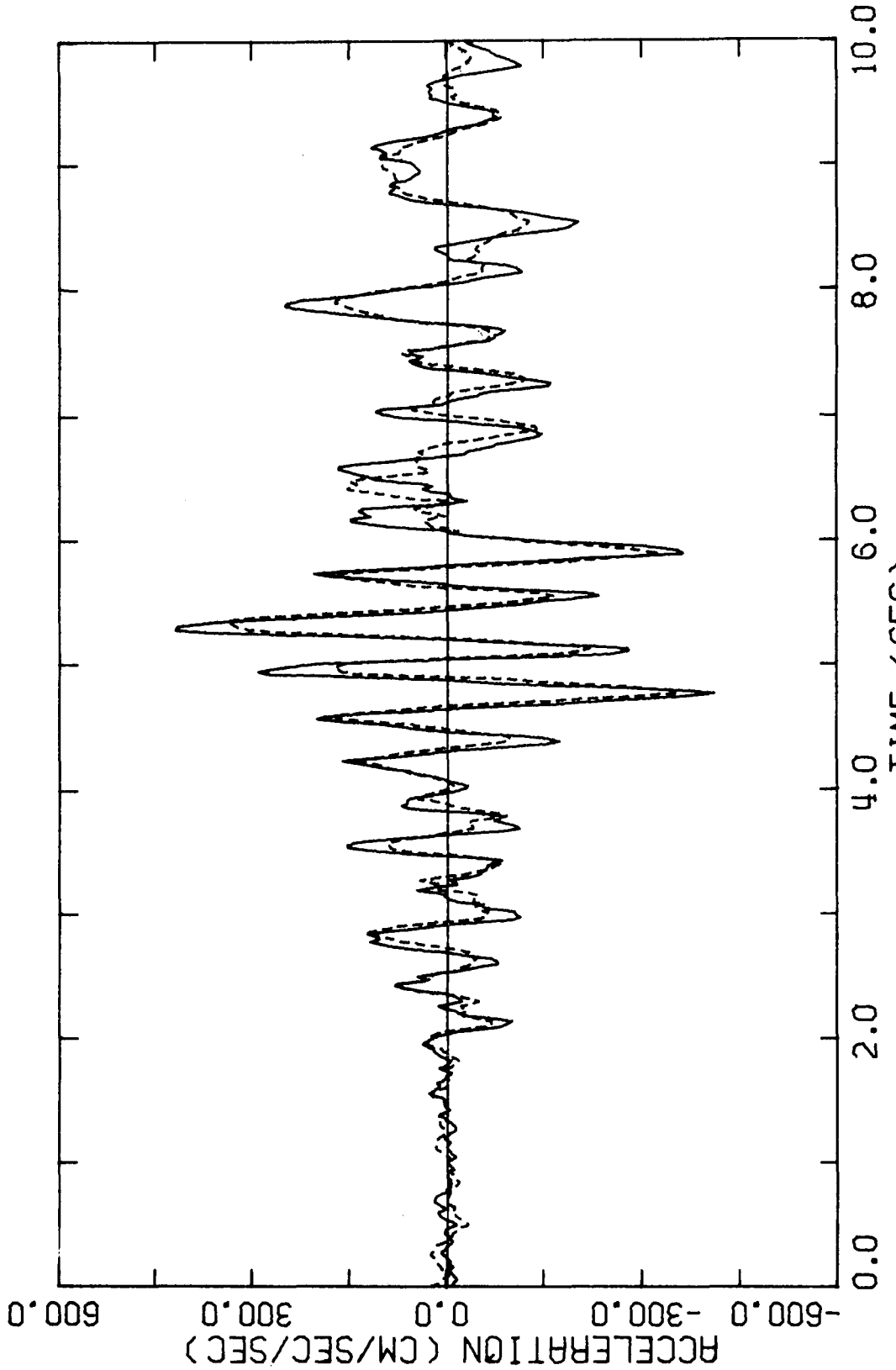


Figure 6.21.b. Relative acceleration (——) at the roof of JPL Building 180 and calculated acceleration (----) of the same model as in Fig. 6.21.a.

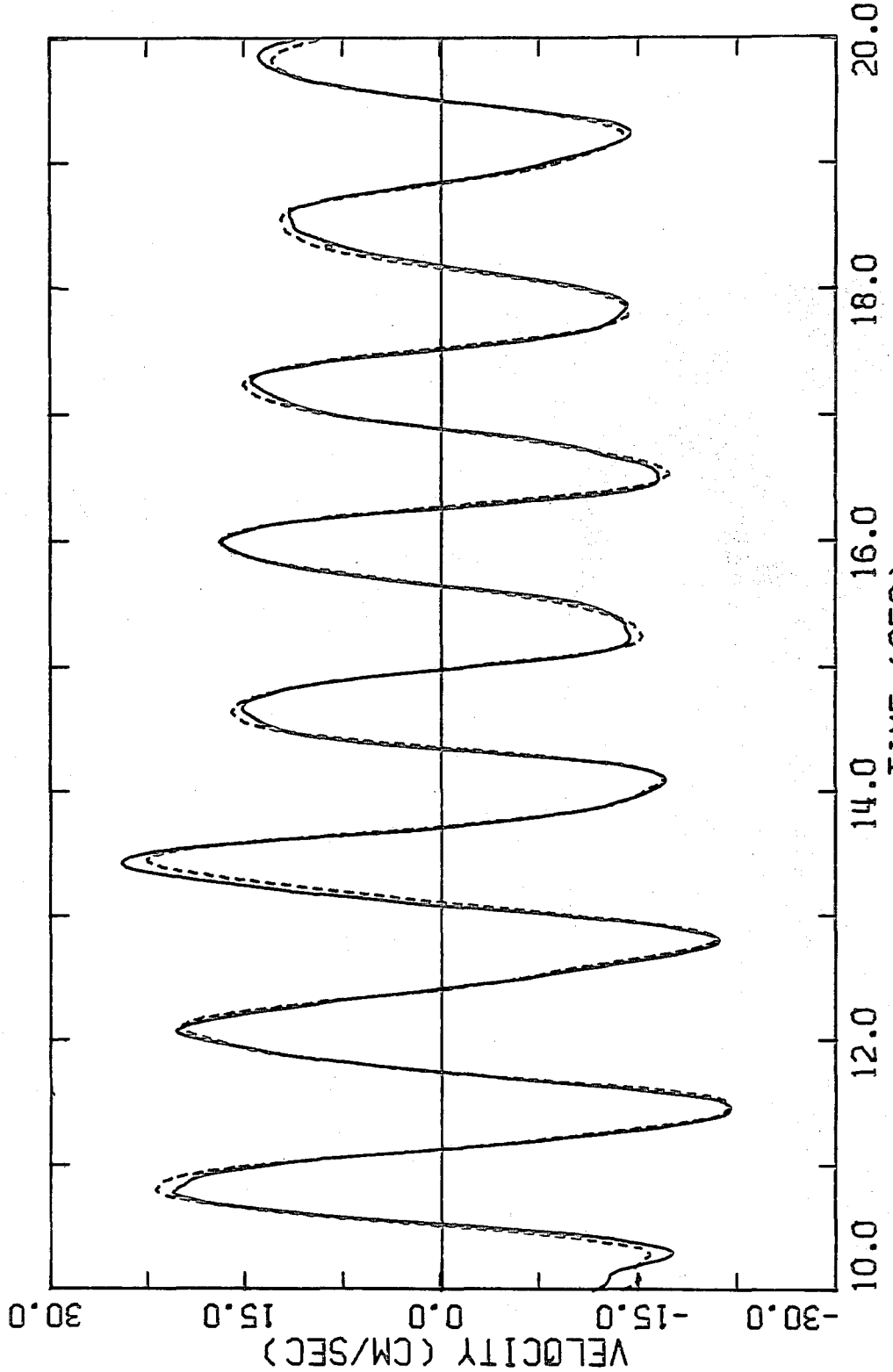


Figure 6.22, a. Relative velocity (——) at the roof of JPL Building 180 and calculated velocity (----) of the optimal model with two modes determined by matching velocities over the interval.

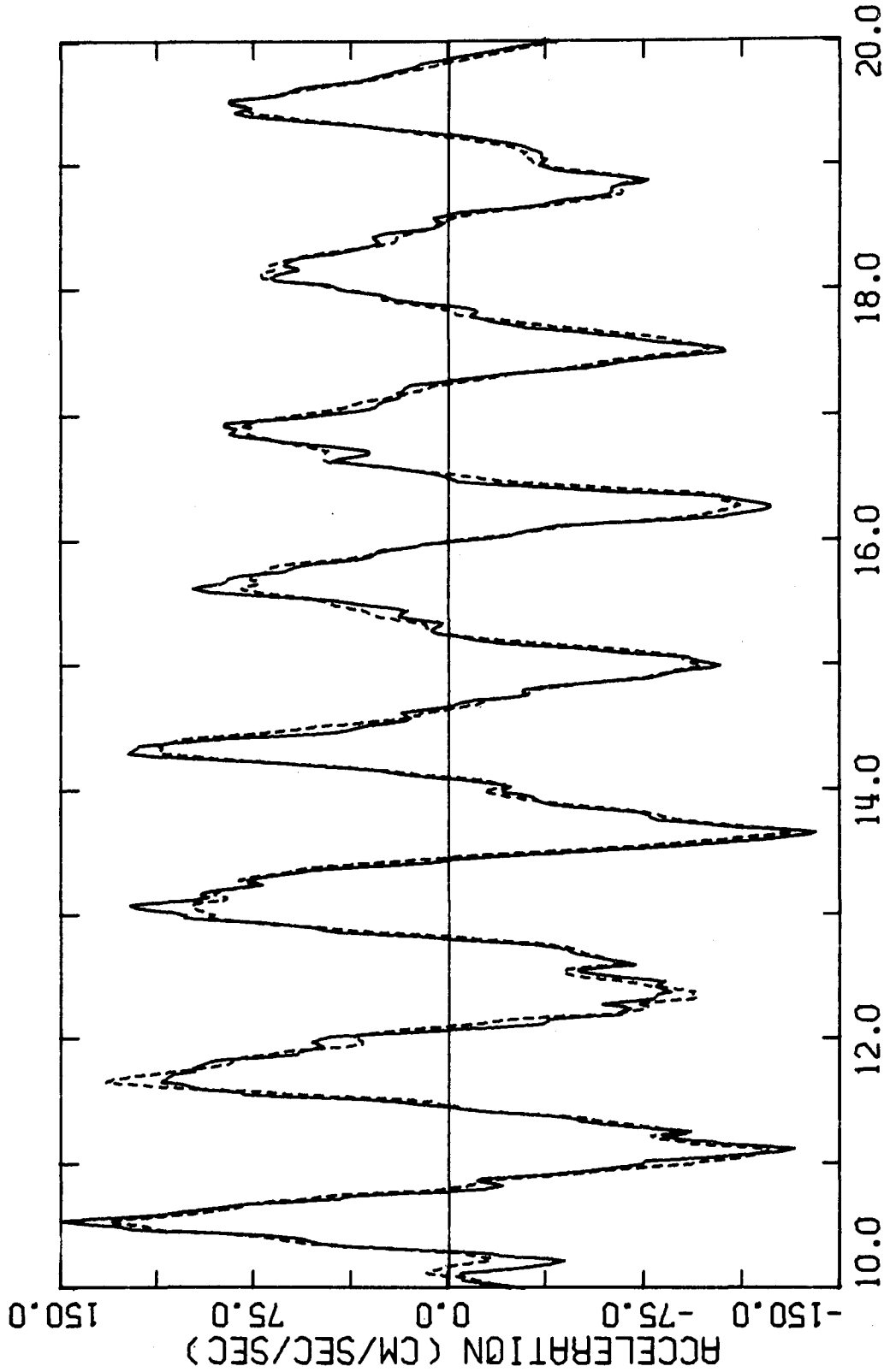


Figure 6.22.b. Relative acceleration (——) at the roof of JPL Building 180 and calculated acceleration (----) of the same model as in Fig. 6.22.a.

Figs. 6.22.a and b for the optimal model with two modes determined by matching the segment of the velocity record from 10 to 20 seconds.

On the basis of all the results, it is thought that the optimal estimates in Table 6.12 of the periods of the fundamental and second modes are reliable values, although it is to be remembered that these are the best periods for a time-invariant model over most of the strong-motion portion of the response and that the periods actually changed considerably during the earthquake. The damping factor of 4% in Table 6.12 for the fundamental mode is also considered to be representative for the strong-motion response of Building 180 during the San Fernando earthquake. There are difficulties in estimating the damping of the second mode because of the marked change in its period, which is also thought to have produced the twin peaks in the Fourier amplitude spectrum of the roof acceleration record. The estimates in Table 6.12 for the parameters of the third mode are considered to be unreliable. This mode has little effect on the response, and for this building, it is concluded that only two modes are required to give a good approximation of the relative velocity and acceleration records, and one mode is sufficient for the displacement.

REFERENCES

Foutch, D.A., G.W. Housner and P.C. Jennings (1975). Dynamic Responses of Six Multistory Buildings during the San Fernando Earthquake. Report No. EERL 75-02, Calif. Inst. of Tech., Pasadena, California.

REFERENCES

Hanks, T. C. (1973). Current assessment of long period errors, in Strong Motion Earthquake Accelerograms, Part G., II. Report No. 73-52, Calif. Inst. of Tech., Pasadena, California.

Martin, A. C. and Associates (1973). Union Bank Square, in San Fernando, California, Earthquake of February 9, 1971, L. M. Murphy (ed), 1, Part B, 575-596, U. S. Dept. of Commerce, Washington, D. C.

Nielsen, N. N. (1964). Dynamic Response of Multi-story Buildings. Earthq. Engng Res. Lab. Report, Calif. Inst. of Tech., Pasadena, California.

Trifunac, M. D. (1970a). Ambient Vibration Test of a Thirty-Nine Story Steel Frame Building. Report No. EERL 70-02, Calif. Inst. of Tech., Pasadena, California.

Trifunac, M. D. (1970b). Low Frequency Digitization Errors and a New Method for Zero Baseline Correction of Strong-motion Accellerograms. Report No. EERL 70-07, Calif. Inst. of Tech., Pasadena, California.

Udwadia, F. E. and M. D. Trifunac (1973). Ambient vibration tests of full-scale structures. Proc. 5th Wld. Conf. Earthq. Engng, 1430-1439, Rome.

Wood, J. H. (1972). Analysis of the Earthquake Response of a Nine-story Steel Frame Building during the San Fernando Earthquake. Report No. EERL 72-04, Calif. Inst. of Tech., Pasadena, California.

VII. CONCLUSION

In this dissertation, a practical strategy has been devised and implemented for systematically determining best estimates of parameters of linear structural models from records of base motion and response during an earthquake. The investigation was set within the framework of a general output-error approach (Chapter 2). In this approach, the model parameters are estimated by using a suitable computer algorithm which systematically varies the parameters until some measure-of-fit between the structural output and model output, such as the integral-squared difference, is minimized. The parameter values so calculated are called the optimal estimates for the class of models employed.

The question of whether this approach allows the parameters to be determined uniquely and reliably was studied for a general class of linear structural models for which the mass matrix was assumed known (Chapter 3). It was shown that reliable estimates of the stiffness and damping matrices for these models usually cannot be made from records of the earthquake response of a structure, because of several basic limitations of the data. It was also shown that the modal periods, modal damping factors and the effective participation factors at the points of measurement give all the information about the stiffness and

damping distributions that is contained in the structural records. These modal parameters, rather than the stiffness and damping matrices, should be estimated when identifying a structure from earthquake records. Furthermore, in order to obtain reasonable accuracy, only the parameters of the dominant modes in the records should be estimated, and this should be done by performing a series of identifications in which modes are successively added to the models until the measure-of-fit is no longer significantly decreased. Otherwise, the higher modes of the model may be primarily matching noise in the records. It is also desirable to estimate the parameters by separately matching the displacement, velocity and acceleration records, so that a check can be made on whether the linear models produce consistent results.

It was concluded that the distribution of earthquake forces throughout the structure generally cannot be estimated reliably from a few records of the structural motion (§3.5). If estimates of the forces experienced by a structure during a particular earthquake are required, and records of the structural motion are available, the forces can be estimated by using a synthesized model which has been modified so that the parameters of its lower modes are equal to the corresponding values determined from the records.

Two output-error techniques for determining the modal

parameters from seismic response records were investigated. The first technique studied was the optimal filter method (Chapter 4). Advantages of this method are that it provides sequential estimates of the parameters, thereby showing how the estimates depend on the length of record used, and it can also be applied to nonlinear models. It was found from tests of the method that because of approximations which must be made in the theory in order to produce a feasible algorithm, the estimates calculated by the filter can be quite different from the optimal estimates when there is significant measurement noise or model error. It was concluded after the tests that a more reliable technique is required when identifying linear models of structures from earthquake records.

A new output-error technique is introduced for estimating the modal parameters, which was called the modal minimization method (Chapter 5). This method can be relied upon to find the optimal estimates of the modal parameters and it is more efficient numerically than the optimal filter method, but it is limited to linear structural models.

The modal minimization method was employed to study the strong-motion response of two multi-story buildings during the 1971 San Fernando earthquake (Chapter 6). Within the framework of linear models, new information was obtained from the records concerning

the properties of the higher modes and also the time-varying character of the equivalent linear parameters.

The response in the longitudinal direction of a 40-story steel-frame building was first studied. The building experienced a peak acceleration at midheight of about 20% g but suffered no structural damage. It was found that a time-invariant linear model with a small number of modes could reproduce the strong-motion records remarkably well over the 25-second interval investigated. The matching of the recorded and model responses was considerably better than anticipated on the basis of other studies where less systematic techniques were applied to attempt to achieve a good match with linear models. The number of modes required to give a good approximation of the relative displacement, velocity and acceleration records was one, two and four modes respectively.

The modal minimization method also gave new information about the higher-mode damping and participation factors for the building, as well as more reliable information about the modal periods. The estimates of the modal damping factors from 25 seconds of the acceleration record ranged from 4% for the fundamental longitudinal mode to 7% for the fourth longitudinal mode. Although there was a reduction in the stiffness of the structure during the earthquake, the period ratios among the first four longitudinal modes were roughly the

same as during ambient vibration tests before and after the seismic shaking. This suggests that the degradation in stiffness was approximately uniform over the structure. Furthermore, the period ratios during the earthquake remained close to those for a uniform shear beam.

The other building studied was a 9-story steel-frame structure which experienced a peak acceleration of about 40% g, one of the largest recorded in a building during the San Fernando earthquake. Damage was limited to minor nonstructural cracking. A time-invariant linear model was again able to match the response well over the twenty-second interval of the longitudinal records which was studied. The match was considerably better than that given by a previously reported model which was essentially based on the traditional frequency-domain approach for determining the modal periods and damping factors. However, the optimal estimates of the modal damping and participation factors from records of different response quantities were less consistent than in the first building investigated. It is thought that this may be due to the stronger nonlinear, or time-varying, dynamic behavior of the 9-story structure.

The time variation of the structural properties during an earthquake can be studied by determining the optimal estimates of the modal parameters for short time segments of the records. By using

successive time windows, the results can show how the equivalent linear parameters change because of nonlinear effects.

This approach was applied to both buildings described above. It was found that the average increase of the periods per fundamental cycle for the two lowest modes was about 1% for the taller building and about 3% for the smaller building. For the latter structure, almost all of the change occurred in the first 10 or so seconds, during which the fundamental period increased from 1.1 sec to 1.3 sec, or by about 20%, while the period of the second mode increased from 0.33 sec to 0.42 sec, or by about 30%. The latter change was thought to have produced the twin peaks at 0.38 sec and 0.42 sec in the Fourier amplitude spectrum of the roof acceleration.

On the basis of the results for the two buildings, it is tentatively concluded that when the optimal estimates of the modal parameters are used, a time-invariant linear model based on a small number of modes can adequately reproduce a building's strong-motion response if structural damage does not occur. This is useful information for structural design employing linear models, even though the best values for the modal parameters are difficult to determine prior to measuring the response to the earthquake. From the results of Chapter 6, it is suggested that for design calculations, this difficulty might best be treated by using typical values for the damping factors obtained from

studies of the earthquake response of similar structures, and by using participation factors from a synthesized structural model, since the model response is not particularly sensitive to these parameters. However, the response is very sensitive to changes in the modal periods, which are difficult to synthesize accurately. It is therefore suggested that the overall stiffness of the structural model should be varied so that the periods of the lowest modes, particularly the fundamental, cover a representative range of values.

In continuing research in this area, it is considered that it would be most fruitful to concentrate on improving features of the linear models and in applying the approach to records from a wide variety of buildings. In particular, the models could be modified to include both horizontal components of the translational motion of the base because this might allow better modelling of any torsional response shown during an earthquake. There may also be other applications in earthquake engineering for which the method could be fruitfully employed. For example, the method may be applied to the recorded response of soil layers to investigate local site effects, provided earthquake records, such as those from bore holes, are also available to serve as input to the linear models. The method may also be a useful approach to determine the dynamic properties of large earth dams.

It is also desirable to investigate the identification of nonlinear models so that structural properties such as strength and ductility during an earthquake can be studied. Some other output-error technique, such as the Gauss-Newton method, is then required to determine the optimal estimates of the parameters, but some of the difficulties discussed in this work for linear models, such as those arising from lack of identifiability and from limited resolution, will have their counterparts in structural identification with nonlinear models. Nonlinear modelling is a particularly challenging area of research because it is difficult to formulate models which include such observed phenomena as amplitude-dependent stiffness, hysteresis and structural deterioration. In addition, the lack of experience in this area will make it difficult to assess whether the estimates of the parameters are reliable. Research is also inhibited by the scarcity of response records of structural motions well into the inelastic range or approaching failure, and such data from tests employing large-scale shaking tables can make a valuable contribution.

In conclusion, it is felt that the modal minimization method has proven to be a useful technique to investigate the dynamic properties of buildings from their strong-motion records. In particular, it has shown that by using the optimal estimates of the modal parameters, time-invariant linear models of the two buildings studied can reproduce their strong-motion records surprisingly well.

APPENDIX A: IDENTIFIABILITY

Consider a class of models $\mathbb{M}$ with its associated set of allowable parameter values denoted by $\mathbb{A}$, and let $\mathbb{C}$ be a class of inputs to the models. The class $\mathbb{M}$ is defined by the function $\underline{m}(\underline{a},\underline{z})$ relating the output $\underline{m}$ of a model with parameters $\underline{a}$ to the input $\underline{z}$ in $\mathbb{C}$. In the following work, $\underline{z}$ and $\underline{m}$ will denote the histories of the input and output over a specified time interval.

Let M^* be a model in $\mathbb{M}$ given by the parameters $\underline{a}^*$ in $\mathbb{A}$, then the definitions in §2.4.1 may be written as:

(i) M^* is globally identifiable for $\mathbb{C}$

$\Leftrightarrow \forall \underline{z} \in \mathbb{C}$,

$\underline{m}(\underline{a},\underline{z}) \neq \underline{m}(\underline{a}^*,\underline{z})$, $\forall \underline{a} \in \mathbb{A}$ with $\underline{a} \neq \underline{a}^*$

(ii) M^* is locally identifiable for $\mathbb{C}$

$\Leftrightarrow \forall \underline{z} \in \mathbb{C}$, $\exists$ a neighborhood $\mathcal{N}(\underline{a}^*) \subset \mathbb{A}$ such that:

$\underline{m}(\underline{a},\underline{z}) \neq \underline{m}(\underline{a}^*,\underline{z})$, $\forall \underline{a} \in \mathcal{N}(\underline{a}^*)$ with $\underline{a} \neq \underline{a}^*$

In Fig. A.1, the models corresponding to $\underline{a}_1, \underline{a}_2$ and $\underline{a}_3$ are respectively globally identifiable, locally identifiable, and neither globally nor locally identifiable. In order for the whole class of models to be globally or locally identifiable, the appropriate definition above has to hold for each $\underline{a}^*$ in $\mathbb{A}$.

The following results have been used in this dissertation. Recall the definition of $J_0(\underline{a})$ given by (2.3.5) and (2.3.6):

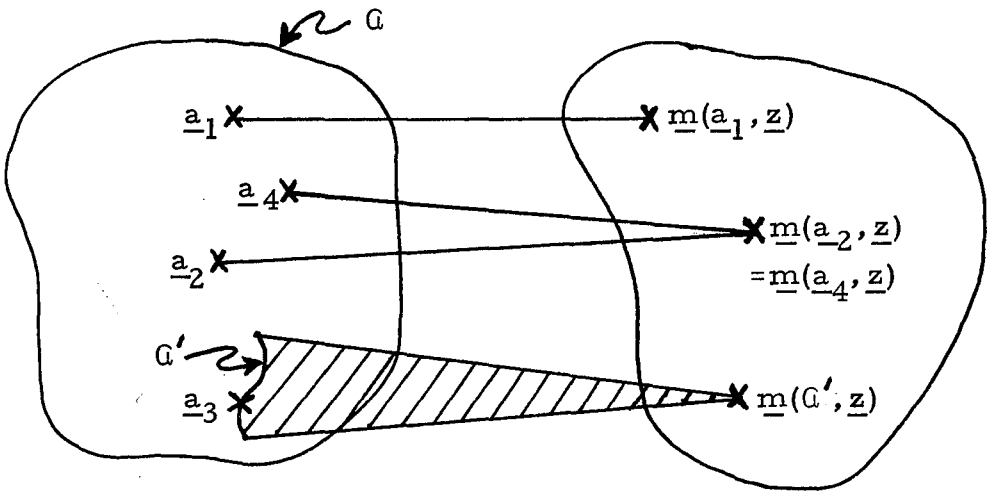


Figure A. 1. Schematic diagram illustrating identifiability in terms of the mapping $\underline{m}(\underline{a}, \underline{z})$ between the set of allowable parameter values and the set of outputs.

$$J_0(\underline{a}) = \langle \underline{y} - \underline{m}(\underline{a}, \underline{z}) , \underline{y} - \underline{m}(\underline{a}, \underline{z}) \rangle \qquad \text{(A. 1)}$$

where $\langle \cdot, \cdot \rangle$ is a scalar product and $\underline{y}$ is the output to be matched by minimizing J_0. The identifiability of a model can be expressed in terms of the minimization of J_0 using the model output for $\underline{y}$ as follows:

Theorem A. 1

(i) M^* is globally identifiable for $\mathcal{C}$

$\Leftrightarrow \forall \underline{z} \in \mathcal{C}$, the global minimum of:

$$J_0(\underline{a}) = \langle \underline{m}(\underline{a}^*, \underline{z}) - \underline{m}(\underline{a}, \underline{z}), \underline{m}(\underline{a}^*, \underline{z}) - \underline{m}(\underline{a}, \underline{z}) \rangle \qquad \text{(A. 2)}$$

occurs only at $\underline{a} = \underline{a}^*$, that is,

$$J_0(\underline{a}) > J_0(\underline{a}^*), \ \forall \underline{a} \in \mathcal{G} \quad \text{with} \quad \underline{a} \neq \underline{a}^*$$

(ii) M^* is locally identifiable for $\mathcal{C}$

$\Leftrightarrow \forall \underline{z} \in \mathcal{C}$, $J_0(\underline{a})$ has a strict local minimum at $\underline{a} = \underline{a}^*$, that is, $\exists$ a neighborhood $\mathcal{N}(\underline{a}^*)$ such that:

$$J_0(\underline{a}) > J_0(\underline{a}^*), \ \forall \underline{a} \in \mathcal{N}(\underline{a}^*) \quad \text{with} \quad \underline{a} \neq \underline{a}^*$$

Proof

The proof is given for (ii). It is easily modified for (i).

(a) Necessity: By hypothesis, M^* is locally identifiable, so there is a neighborhood $\mathcal{N}(\underline{a}^*)$ of $\underline{a}^*$ such that $\underline{m}(\underline{a}, \underline{z}) \neq \underline{m}(\underline{a}^*, \underline{z})$ for each $\underline{a} \neq \underline{a}^*$ in $\mathcal{N}(\underline{a}^*)$. Notice that $J_0(\underline{a}) \geq J_0(\underline{a}^*) = 0$. Suppose there exists $\underline{a}' \neq \underline{a}^*$ in $\mathcal{N}(\underline{a}^*)$ such that $J_0(\underline{a}') = J_0(\underline{a}^*)$, then $J_0(\underline{a}') = 0$ and hence, by the positive-definite property of $\langle \cdot, \cdot \rangle$, $\underline{m}(\underline{a}', \underline{z}) = \underline{m}(\underline{a}^*, \underline{z})$, which is

a contradiction. Thus, $J_0(\underline{a}) > J_0(\underline{a}^*)$, $\forall \underline{a} \in \eta(\underline{a}^*)$ such that $\underline{a} \neq \underline{a}^*$.

(b) Sufficiency: By hypothesis,

$$J_0(\underline{a}) > J_0(\underline{a}^*) = 0, \quad \forall \underline{a} \in \eta(\underline{a}^*) \quad \text{with} \quad \underline{a} \neq \underline{a}^*$$

Suppose there exists $\underline{a} \neq \underline{a}^*$ in $\eta(\underline{a}^*)$ such that $\underline{m}(\underline{a}, \underline{z}) = \underline{m}(\underline{a}^*, \underline{z})$, then $J_0(\underline{a}) = 0$, which is a contradiction. Thus, M^* is locally identifiable.

The results proved in Theorem A. 1 are used as definitions by Bellman and Åström (1970) except that they omitted to state in their version of (i) that the global minimum must occur at only one point. The advantage of the definitions given in §2. 4. 1 are that they make it clear that identifiability is a property of the model and it is independent of the particular scalar product chosen for J_0. For example, with continuous data, $\langle \cdot, \cdot \rangle$ in Theorem A. 1 could be defined in terms of the quantities in the time domain or in terms of their transforms in the frequency domain.

Another result related to identifiability which has been stated in this dissertation is the following:

Theorem A. 2

The sensitivity coefficients $\frac{\partial \underline{m}(\underline{a}, \underline{z})}{\partial a_k}$ are linearly independent over the data interval if and only if the reduced sensitivity matrix $\tilde{\underline{S}}(\underline{a}, \underline{z})$ is positive definite.

Proof

From (2. 3. 9), for an arbitrary vector $\underline{\lambda}$ of appropriate

dimension:

$$\underline{\lambda}^t \tilde{S}(\underline{a}, \underline{z})\underline{\lambda} = \sum_{j,k} \lambda_j \lambda_k \left\langle \frac{\partial \underline{m}}{\partial a_j}, \frac{\partial \underline{m}}{\partial a_k} \right\rangle$$

$$= \left\langle \sum_j \lambda_j \frac{\partial \underline{m}}{\partial a_j}, \sum_k \lambda_k \frac{\partial \underline{m}}{\partial a_k} \right\rangle \qquad \text{(A. 3)}$$

$$\geq 0$$

Thus, $\tilde{S}$ is always positive semi-definite. Furthermore, using the positive-definite property of a scalar product:

$\underline{\lambda}^t \tilde{S}(\underline{a}, \underline{z})\underline{\lambda} = 0 \Leftrightarrow \sum_j \lambda_j \frac{\partial \underline{m}}{\partial a_j} = 0$ over the date interval, from which the desired result follows.

The first part of the statement in Theorem A. 2 is used as a definition of identifiability by Beck and Arnold (1977). It is a stronger statement than local identifiability, as the following result shows:

Theorem A. 3

The sensitivity coefficients $\left.\frac{\partial \underline{m}}{\partial a_k}\right|_{\underline{a}^*}$ are linearly independent

$\Rightarrow$ M^* is locally identifiable for $\underline{z}$

Proof

First, from Eq. (A. 2):

$$J_0(\underline{a}^*) = 0 \quad \text{and} \quad J_0(\underline{a}) \geq 0 \, .$$

Thus, J_0 has a local minimum at $\underline{a}^*$. From (2. 3. 8) and (2. 3. 9),

$$\tfrac{1}{2} \nabla\nabla J_0(\underline{a}^*) = \tilde{S}(\underline{a}^*, \underline{z})$$

since here $\underline{v}(\underline{a}^*) = \underline{m}(\underline{a}^*, \underline{z}) - \underline{m}(\underline{a}^*, \underline{z}) = 0$. Thus, by hypothesis and

and Theorem A. 2, $\nabla\nabla J_0(\underline{a}^*)$ is positive definite. From a result in advanced calculus, this ensures that J_0 has a strict local minimum at $\underline{a}^*$ and hence the desired result follows from Theorem A. 1, part (ii).

The converse of the theorem is not true in general. However, it can be shown that the two statements in Theorem A. 3 are equivalent in the special case where $\underline{m}(\underline{a}, \underline{z})$ is a linear function of the parameters $\underline{a}$. In this case, each statement is a necessary and sufficient condition for uniqueness of the optimal estimates of the model parameters.

REFERENCES

Beck, J. V. and K. J. Arnold (1977). Parameter Estimation in Engineering and Science. Wiley and Sons, New York.

Bellman, R. and K. J. Åström (1970). On structural identifiability. Math. Biosciences, 7, 329-339.

APPENDIX B: PROOF OF THEOREM IN §3.4.3

Theorem

Consider a controllable and observable model in $\mathbb{M}_N$ whose output for a known input is measured. Let $\mathcal{J}$ be the output set defining the coordinates at which the response is measured. Let $\underline{\Phi}$ be a transformed modeshape matrix of the observed model. The number of models in $\mathbb{M}_N$ which are consistent with the observed data is equal to the number of solutions of the following matrix problem:

Find a nonsingular, real matrix B such that:

(i) $B^t \underline{e}_i = \underline{e}_i \quad , \quad \forall i \in \mathcal{J}$ (3.4.7)

(ii) $B\underline{\rho} = \underline{\rho}$ (3.4.8)

(iii) $(B\underline{\varphi}^{(r)})^t (B\underline{\varphi}^{(s)}) = 0 \quad , \quad r \neq s$ (3.4.9)

where $\underline{e}_i$ is the unit vector given by $(\underline{e}_i)_k = \delta_{ik}$ and $\underline{\rho}$ is a known vector of dimension N with elements given by:

$$\rho_k = b_k m_k^{\frac{1}{2}} \tag{3.4.10}$$

Furthermore, for each solution B, the transformed modeshapes $\underline{\tilde{\varphi}}^{(r)}$ of the model in $\mathbb{M}_N$ which has the same output as the observed model are given by:

$$\underline{\tilde{\varphi}}^{(r)} = \frac{1}{\gamma_r} B\underline{\varphi}^{(r)} \tag{3.4.11}$$

where

$$\gamma_r^2 = (B\underline{\varphi}^{(r)})^t (B\underline{\varphi}^{(r)}) \tag{3.4.12}$$

Proof

Let $\mathcal{L}_N \subset \mathcal{M}_N$ denote the set of all models which are consistent with the observed data and let $\mathcal{B}_N$ denote the set of all matrices B which satisfy the matrix problem given by Eqs. (3.4.7), (3.4.8) and (3.4.9). To prove the theorem, it is shown that a one-to-one onto mapping f can be constructed between $\mathcal{B}_N$ and $\mathcal{L}_N$. This also handles the case where the number of solutions is infinite. The proof is in three steps: (a) f is defined, (b) f is shown to be onto, (c) f is shown to be one-to-one.

(a) There are two preliminaries. First, the $\tilde{\underline{\varphi}}^{(r)}$ defined by (3.4.11) and (3.4.12) are possible transformed modeshapes since:

$$(\tilde{\underline{\varphi}}^{(r)})^t \, \tilde{\underline{\varphi}}^{(s)} = \frac{1}{\gamma_r \gamma_s} (B\underline{\varphi}^{(r)})^t \, (B\underline{\varphi}^{(s)}) = \delta_{rs} \quad ,$$

using (3.4.9) in addition to (3.4.11) and (3.4.12). Notice that $\gamma_r \neq 0$ since if $\gamma_r = 0$, then $B\underline{\varphi}^{(r)} = 0$; since B is nonsingular, this implies $\underline{\varphi}^{(r)} = \underline{0}$, which is a contradiction. Second, by Proposition 1, or Proposition 2 with $\mathcal{C}_L$, a model in $\mathcal{M}_N$ is consistent with the measured input and output if and only if it has the same values of ω_r, ζ_r and $\beta_i^{(r)}$ as the observed model, $\forall\, r = 1, \ldots, N$ and $\forall i \in \mathcal{I}$.

Define a mapping f from $\mathcal{B}_N$ to $\mathcal{L}_N$ as follows. Let $B \in \mathcal{B}_N$, then define f(B) as that model in $\mathcal{M}_N$ which has the transformed modeshapes $\tilde{\underline{\varphi}}^{(r)}$ given by (3.4.11) and (3.4.12), and the modal frequencies and damping factors ω_r and ζ_r of the observed model. From (3.4.5) and §3.2.2, the model is defined uniquely. Furthermore,

the fact that the sign of γ_r is arbitrary is consistent with the result stated in §3.2.1 that the modeshapes of a model are only unique to within a change of sign.

To show $f(B) \in \mathcal{S}_N$, it remains to prove that it has the correct values of $\beta_i^{(r)}$, $\forall\, r = 1, \ldots, N$ and $\forall i \in \mathcal{J}$.

From (3.2.13) and (3.4.5)

$$\tilde{\beta}_i^{(r)} = m_i^{-\frac{1}{2}} \tilde{\varphi}_i^{(r)} \tilde{\alpha}_r \tag{B.1}$$

Consider first $\tilde{\varphi}_i^{(r)}$ where i is in $\mathcal{J}$. From (3.4.7):

$$(\underline{\varphi}^{(r)})^t \, B^t \, \underline{e}_i = (\underline{\varphi}^{(r)})^t \, \underline{e}_i$$

and thus from (3.4.11):

$$(\tilde{\underline{\varphi}}^{(r)})^t \underline{e}_i = \frac{1}{\gamma_r} (\underline{\varphi}^{(r)})^t \underline{e}_i$$

or

$$\tilde{\varphi}_i^{(r)} = \frac{1}{\gamma_r} \varphi_i^{(r)} \tag{B.2}$$

Consider next $\tilde{\alpha}_r$. Define a diagonal matrix A by:

$$A = \begin{bmatrix} \gamma_1 & & & 0 \\ & \gamma_2 & & \\ & & \ddots & \\ 0 & & & \gamma_N \end{bmatrix} \tag{B.3}$$

then from (3.4.9) and (3.4.12):

$$\Phi^t \, B^t \, B \Phi = A^2 \tag{B.4}$$

Using the fact that Φ is unitary:

$$\Phi^t B^t B = A^2 \Phi^t$$

which, together with (3. 4. 8) leads to:

$$A^2 \Phi^t \underline{\rho} = \Phi^t B^t B \underline{\rho} = \Phi^t B^t \underline{\rho}$$

Thus, from (3. 4. 11):

$$\gamma_r (\underline{\varphi}^{(r)})^t \underline{\rho} = (\underline{\tilde{\varphi}}^{(r)})^t \underline{\rho}$$

or
$$\tilde{\alpha}_r = \gamma_r \alpha_r \qquad \text{(B. 5)}$$

since from (3. 2. 8), (3. 4. 5) and (3. 4. 10):

$$\alpha_r = \sum_{k=1}^{N} \rho_k \varphi_k^{(r)} = (\underline{\varphi}^{(r)})^t \underline{\rho} \qquad \text{(B. 6)}$$

with a similar expression for $\tilde{\alpha}_r$. Substituting (B. 5) and (B. 2) into (B. 1) leads to:

$$\tilde{\beta}_i^{(r)} = m_i^{-\frac{1}{2}} \varphi_i^{(r)} \alpha_r = \beta_i^{(r)}, \ \forall\ r=1,\ldots,N \text{ and } \forall i \in \mathcal{I} \qquad \text{(B. 7)}$$

(b) It is now shown that f is onto, that is, given any model in $\mathcal{L}_N$, there exists B in $\mathcal{B}_N$ such that the model is equal to f(B).

Let the model in $\mathcal{L}_N$ have a transformed modeshape matrix $\tilde{\Phi}$, then $\forall$ r=1,...,N and $\forall i \in \mathcal{I}$:

$$m_i^{-\frac{1}{2}} \tilde{\varphi}_i^{(r)} \tilde{\alpha}_r = \beta_i^{(r)} = m_i^{-\frac{1}{2}} \varphi_i^{(r)} \alpha_r \qquad \text{(B. 8)}$$

Define
$$\gamma_r = \tilde{\alpha}_r / \alpha_r \qquad \text{(B. 9)}$$

and define a matrix A by (B. 3). Each ratio γ_r is well-defined since $\alpha_r \neq 0$, otherwise $\beta_i^{(r)} = 0$, $\forall i \in \mathcal{I}$, which contradicts an original

hypothesis. Similarly, $\gamma_r \neq 0$ and so the matrix A is nonsingular.

Define $$B = \tilde{\Phi} A \Phi^t \tag{B. 10}$$

then the aim is to show that B belongs to $\mathcal{B}_N$ and that f(B) is equal to the given model in $\mathcal{L}_N$. First, B is real and nonsingular. Also from (B. 8) and (B. 9)

$$\tilde{\alpha}_r = \gamma_r \alpha_r \quad \text{and} \quad \varphi_i^{(r)} = \gamma_r \varphi_i^{(r)}$$

These results can be written in vector form using (B. 6):

$$(\tilde{\underline{\varphi}}^{(r)})^t \underline{\rho} = \gamma_r (\underline{\varphi}^{(r)})^t \underline{\rho} \quad \text{and} \quad (\underline{\varphi}^{(r)})^t \underline{e}_i = \gamma_r (\tilde{\underline{\varphi}}^{(r)})^t \underline{e}_i$$

In matrix form, they become:

$$\tilde{\Phi}^t \underline{\rho} = A \Phi^t \underline{\rho} \quad \text{and} \quad \Phi^t \underline{e}_i = A \tilde{\Phi}^t \underline{e}_i \,, \; \forall i \in \mathcal{J} \tag{B. 11}$$

In view of the definition of B in Eq. (B. 10), and the fact that Φ and $\tilde{\Phi}$ are unitary, (B. 11) leads to (3. 4. 7) and (3. 4. 8). Furthermore, from the definition of B:

$$\tilde{\Phi} = B \Phi A^{-1}$$

or $$\tilde{\underline{\varphi}}^{(r)} = \frac{1}{\gamma_r} B \underline{\varphi}^{(r)} \tag{B. 12}$$

Applying the unitary property (3. 4. 6) for the $\tilde{\underline{\varphi}}^{(r)}$:

$$\frac{1}{\gamma_r \gamma_s} (B \underline{\varphi}^{(r)})^t (B \underline{\varphi}^{(s)}) = \delta_{rs} \tag{B. 13}$$

This shows (3. 4. 9) holds and, together with the earlier results, proves that B belongs to $\mathcal{B}_N$.

Finally, from (B. 12) and (B. 13), the matrix B and the transformed modeshapes $\tilde{\underline{\varphi}}^{(r)}$ of the given model in $\mathcal{L}_N$ satisfy (3. 4. 11)

and (3.4.12). Thus, f(B) is equal to the given model.

(c) The final result is to show that f is one-to-one, that is, if B_1, B_2 in $\mathcal{B}_N$ are such that $f(B_1) = f(B_2)$, then $B_1 = B_2$.

Since $f(B_1)$ and $f(B_2)$ are the same model in $\mathcal{L}_N$, the transformed modeshapes corresponding to B_1 and B_2 can be made equal by an appropriate choice of sign for $\gamma_r^{(1)}$ and $\gamma_r^{(2)}$. Thus:

$$\frac{1}{\gamma_r^{(1)}} B_1 \underline{\varphi}^{(r)} = \underline{\tilde{\varphi}}^{(r)} = \frac{1}{\gamma_r^{(2)}} B_2 \underline{\varphi}^{(r)} \tag{B.14}$$

where the $\underline{\tilde{\varphi}}^{(r)}$ are the transformed modeshapes of the model $f(B_1) = f(B_2)$. The results in (a) show that:

$$\gamma_r^{(1)} \alpha_r = \tilde{\alpha}_r = \gamma_r^{(2)} \alpha_r$$

and so $\gamma_r^{(1)} = \gamma_r^{(2)}$ because $\alpha_r \neq 0$. Simplifying (B.14):

$$B_1 \underline{\varphi}^{(r)} = B_2 \underline{\varphi}^{(r)} \tag{B.15}$$

or

$$B_1 \Phi = B_2 \Phi$$

Since Φ is nonsingular, $B_1 = B_2$.