NUMERICAL EXPERIMENTATION WITH MAXIMUM LIKELIHOOD IDENTIFICATION IN STATIC DISTRIBUTED SYSTEMS

R. E. Scheid, Jr., and G. Rodriguez
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, CA 91109

1. INTRODUCTION

Many important issues in the control of large space structures are intimately related to the fundamental problem of parameter identification. Very often, a complicated structure can be adequately modeled for certain operations by the fitting of a rather simple model with a number of free parameters. This simple model then can be referenced for necessary control operations. Important applications include the many space station designs which are based on the assembly and joining of discrete module* by crew members. This crew-assisted construction will result in a configuration v nich is a large-scale composite of many structural elements and whose static and dynamic characteristics cannot be adequately modeled in advance. In fact, any modeling will require periodic updating as more modules are added to the system and as the structural properties of the elements slowly change over the lifetime of the station.

One might also ask how well this identification process can be carried out in the presence of noisy data since no sensor system is perfect. With these considerations in mind our algorithms are designed to treat both the case of uncertainties in the modeling and uncertainties in the data.

This paper serves as a companion to [6] where the analytical aspects of maximum likelihood identification are considered in some detail. Here we focus on the questions relevant to the implementation of these schemes, particularly as they apply to models of large space structures. Our emphasis will be on the influence of the infinite-dimensional character of the problem on finite-dimensional implementations of the algorithms. We highlight those areas of current and future analysis which indicate the interplay between error analysis and possible truncations of the state and parameter spaces.

2. MODELS

As in [6], we consider the systems of the form

$$A(\theta)u(\theta) = \sigma_\omega B(\theta)\omega + C(\theta)f$$

$$y(\theta) = H(\theta)u(\theta) + \sigma_\eta \eta$$

(2.1)

Here A is a formally self-adjoint elliptic differential operator defined over the spatial domain Ω; the integral operator Φ is related to A by

$$A\Phi = I \quad ,$$

(2.2)

where I is the identity. B and C are appropriately dimensional operators that model the influence of the process error ω and the input f on the state u. H is an operator that characterizes the state-to-observation map; ω amd η are model errors that form the model error vector

$$\epsilon = [\omega, \eta], \tag{2.3}$$

and f is a deterministic input. Conceptually the error vector ϵ represents spatial white noise and is characterized by the covariance operator

$$E(\epsilon\epsilon^*) = I.$$

σ_ω and σ_η are non-negative scalar weighting parameters that respectively measure the relative importance of the modeling error and the measurement error. Thus, the limit $\sigma_\omega \to 0$ corresponds to the case of perfect modeling while the limit $\sigma_\eta \to 0$ corresponds to the case of perfect measurements.

Θ is the possibly infinite-dimensional parameter which must be estimated. For simplicity we shall generally consider cases where the parameter dependence is restricted to the operator A. Furthermore, we assume as in [6] that the parameter enters linearly into the expression for the potential energy of the system. Thus we assume

$$A(\Theta)u = D^*(\Theta Du) \tag{2.4}$$

where D^* denotes the formal adjoint of D; the corresponding potential energy is given in terms of the appropriate state-space inner-product:

$$\langle A(\Theta)u,u\rangle = \langle \Theta Du,Du\rangle. \tag{2..}$$

And finally, the deterministic and stochastic forcings will be localized to discrete points which might correspond to actuator locations. Similarly, the observation map returns a vector of observations at discrete points which might correspond to sensor locations. We assumed that there are N_s point-sensors at locations $\{\xi_i\}$ and N_a point-actuations at locations $\{\tilde{\xi}_i\}$.

Because of these last assumptions, many of the relevant calculations outlined in [6] reduce to matrix and vector manipulations. In this paper the notation $\vec{g}$ will refer to a finite-dimensional vector whose k-th component is given by $g^{(k)}$. Similarly $\underset{\sim}{G}$ is the notation for a matrix whose (i,j)-component is given by $G^{(i,j)}$. The relevant dimensions of vector and matrix quantities will always be clear from the context.

After taking formal limits in the system (2.1) we have:

$$Au = \sigma_\omega B\vec{\omega} + C\vec{f}$$

$$\vec{y} = Hu + \sigma_\eta \vec{\eta} \tag{2.6}$$

378

where

$$\vec{\omega} = (\omega^{(1)},\ldots,\omega^{(N_a)})^T \tag{2.7a}$$

$$\vec{\eta} = (\eta^{(1)},\ldots,\eta^{(N_s)})^T \tag{2.7b}$$

$$Hu(x) = (u(\xi_1),\ldots,u(\xi_{N_s}))^T \tag{2.7c}$$

$$H^*\vec{g} = \sum_{i=1}^{N_s} \delta(x-\xi_i)g^{(i)} \tag{2.7d}$$

$$C\vec{f} = \sum_{i=1}^{N_a} \delta(x-\tilde{\xi}_i)\, f^{(i)} \tag{2.7e}$$

$$C^*u(x) = (u(\xi_1),\ldots,u(\xi_{N_a}))^T \tag{2.7f}$$

$$B\vec{\omega} = \sum_{i=1}^{N_a} \delta(x-\xi_i)\,\omega^{(i)} \tag{2.7g}$$

$$B^*u(x) = (u(\xi_1),\ldots,u(\xi_{N_a}))^T \tag{2.7h}$$

en in the case of more general sensing and actuating systems, the modeling requirements for the system can be reduced to solving equations of the form

$$Au = f. \tag{2.8}$$

Thus the discussion in this section will focus on how the infinite-dimensional structure of the system (2.8) influences the choice of finite-dimensional approximations which can be made. In this paper we consider two specific structural models: a string under tension and a wrap-rib antenna.

Let $\hat{x}$ be a distance coordinate measured in meters along a string of Length L which is also given in meters. Let $\hat{u}(\hat{x})$ be the displacement in meters and let $\hat{\alpha}(\hat{x})$ be the tension parameter given in units of newtons. The forcing density is given by $\hat{f}(\hat{x})$ in units of newtons/meter. Then the energy potential [3] is given by

$$\hat{V}(\hat{u}) = \frac{1}{2}\int_0^L \hat{\alpha}(\hat{x})(\hat{u}'(\hat{x}))^2\, d\hat{x} - \int_0^L \hat{f}(\hat{x})\,\hat{u}(\hat{x})\, d\hat{x}. \tag{2.9}$$

Of course the energy potential is given in units of newton-meters. The equations of motion can be derived immediately based on the principles of the calculus of variations but it will be convenient to first transform to dimensionless coordinates. Let $\hat{\alpha}_*$ be some characteristic value of the tension parameter. We introduce the dimensionless variables:

$$x = \frac{\hat{x}}{L}$$

$$\alpha(x) = \frac{\hat{\alpha}(\hat{x})}{\hat{\alpha}_*}$$

$$u(x) = \frac{\hat{u}(\hat{x})}{L} \qquad (2.10)$$

$$f(x) = \frac{\hat{f}(\hat{x})L}{\hat{\alpha}_*}$$

$$V(u) = \frac{\hat{V}(\hat{u})}{\hat{\alpha}_* L}$$

and the potential expression becomes

$$V(u) = \frac{1}{2} \int_0^1 \alpha(x)\,(u'(x))^2\,dx - \int_0^1 f(x)u(x)dx. \qquad (2.11)$$

For simplicity we prescribe boundary conditions corresponding to fixed end points:

$$u(o) = u(1) = o \qquad (2.12)$$

Then arguments based on the calculus of variations give the system

$$(\alpha(x)u'(x))' = f(x)$$

$$u(o) = u(1) = 0 \qquad 0 < x < 1 \qquad (2.13)$$

This example has been studied many times in the classical literature but an analogous approach gives comparable expressions for much more complex systems.

We consider now a planar model for a wrap-rib antenna which is used to study out-of-plane vibrations (see Figure 1). The antenna model comprises N gores (subsections) modeled by interconnected ribs and mesh. Since the transformations are similar to those used in the case of the string, we immediately write the potential expression with dimensionless coordinates. Let the vector of rib displacements be $\vec{u}(r)$ where the k-th component of $\vec{u}$ is $u^{(k)}$, the displacement of the k-th rib ($0 < r < 1$). Let the vector of mesh displacements be $\vec{v}(r,\theta)$ where the k-th component of $\vec{v}$ is $v^{(k)}$ the displacement of the k-th mesh sector ($0 < r < 1$, $0 < \theta < 1$).
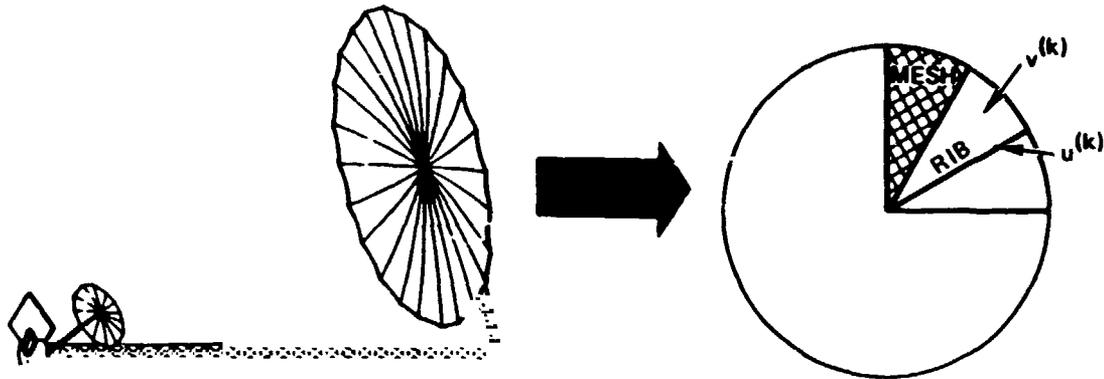
Fig. 1 Simplified model for wrap-rib antenna

Based on analysis of actual antenna designs, our model consists of N identical beams fixed at a central hub. Stretched between the beams are N identical anisotropic membranes. The potential equation is given by

$$V = \frac{1}{2} \int_0^1 G_1 \frac{d^2\vec{u}}{dr^2}^* \frac{d^2\vec{u}}{dr^2} \, dr$$

$$+ \frac{1}{2} \int_0^1 \int_0^1 G_2 \, r \, \frac{\partial\vec{v}}{\partial r}^* \frac{\partial\vec{v}}{\partial r} \, dr \, d\Theta \qquad (2.14)$$

$$+ \frac{1}{2} \int_0^1 \int_0^1 G_3 \, \frac{1}{r} \, \frac{\partial\vec{v}}{\partial\Theta}^* \frac{\partial\vec{v}}{\partial\Theta} \, dr \, d\Theta$$

$$- \int_0^1 \vec{F}_R^* \, \vec{u} \, dr - \int_0^1 \int_0^1 \vec{F}_M^* \, \vec{v} \, r \, dr \, d\Theta$$

Here the coefficients $\{G_i\}$ are related to the physical parameters of the beams and membranes thusly:

$$G_1 = \frac{BI_o}{\sigma L}$$

$$G_2 = \frac{T_r \Theta_o L^2}{\sigma} \qquad (2.15)$$

$$G_3 = \frac{T_\Theta L^2}{\sigma \Theta_o}$$

381

E and I_0 are respectively the Young's modulus and the moment of inertia of the beams. T_r and T_Θ are respectively the radial and circumferential tensions of the membrane. L is the radius of the antenna and Θ_0 is the angular width of a sector; that is, we have

$$\Theta_0 = \frac{2\pi}{N} \quad , \tag{2.16}$$

where N is the number of gores. Finally, σ is some convenient scaling parameter with the dimensions of energy (nt-m). We note that the physical forcing densities $\vec{F}_R$ and $\vec{F}_M$ having respective dimensions nt/m and nt/m^2 were rescaled according to

$$\vec{F}_R = \frac{L^2}{\sigma} \; \hat{\vec{P}}_R$$

$$\vec{F}_M = \frac{L^3}{\sigma} \; \hat{\vec{P}}_M \tag{2.17}$$

Appropriate geometrical boundary conditions follow from fixing the center and attaching each of the ribs to its adjoining membranes:

$$\vec{u}\bigg|_{r=0} = \frac{\partial}{\partial r} \vec{u}\bigg|_{r=0} = 0$$

$$\vec{v}\bigg|_{\Theta=0} = \underset{\sim}{C}\,\vec{v}\bigg|_{\Theta=1} \quad \dashv \vec{u} \tag{2.18}$$

Here $\underset{\sim}{C}$ is an N x N periodic matrix:

$$\underset{\sim}{C} = \begin{bmatrix} 0 & 1 & & & \\ & \cdot & \cdot & & \\ & & \cdot & \cdot & \\ & & & \cdot & 1 \\ 1 & & & & 0 \end{bmatrix} \tag{2.19}$$

382

As in the case of the string, the equations now follow from arguments based on the calculus of variations:

$$\frac{d^2}{dr^2}\left(G_1 \frac{d^2}{dr^2}\vec{u}\right) - G_3\left(\frac{\partial}{\partial\Theta}\vec{v}\Big|_{\Theta=0^+} - \underset{\sim}{C}\frac{\partial}{\partial\Theta}\vec{v}\Big|_{\Theta=1^-}\right) = \vec{F}_R \quad (2.20\text{a})$$

$$-\frac{1}{r}\frac{\partial}{\partial r}\left(G_2 r \frac{\partial\vec{v}}{\partial r}\right) - \frac{1}{r^2}\frac{\partial^2}{\partial\Theta^2}\left(G_3 \vec{v}\right) = \vec{F}_M \qquad (2.20\text{b})$$

with the additional natural boundary conditions:

$$\frac{\partial^2\vec{u}}{\partial r^2}\bigg|_{r=1} = 0$$

$$\frac{\partial^3\vec{u}}{\partial r^3}\bigg|_{r=1} = 0 \qquad (2.21)$$

$$\frac{\partial\vec{v}}{\partial r}\bigg|_{r=1} = 0$$

One of the focuses of this study is the consideration of nonmodal approaches in the finite-dimensional approximation schemes. In practice, this generally will mean directly solving a linear system of equations rather than proceeding from some finite modal synthesis. But the infinite-dimensional structure of the system (2.8) also can influence the particular finite-dimensional approximation schemes used. Our approach is sufficiently general so that any adequate finite e¹ ent model of the system (2.8) should yield adequate numerical approximations. Bu one can often do much better for a particular model or a particular class of models.

We use the antenna model to illustrate the point and make some observations that should influence the approximation schemes regardless of which finite element or finite difference scheme is employed. We emphasize that these considerations also apply to much more complicated antenna models which share salient features with the system (2.18) - (2.21). First we note that the structure is periodic in the Θ-direction. This cyclic symmetry leads to considerable savings in the computation of solutions to (2.8). This can be deduced from either the differential equation or the energy expression (2.14). The periodic matrix $\underset{\sim}{C}$ can be diagonalized by means of a finite Fourier transform [1]. That is, let $\underset{\sim}{U}$ be the N x N matrix whose (j,k) component has the form

$$U^{(j,k)} = \frac{\exp\left(i\frac{2\pi}{N}(j-1)(k-1)\right)}{\sqrt{N}} \qquad (2.22)$$

We then have:

$$\underset{\sim}{U}^*\underset{\sim}{U} = \underset{\sim}{I}$$

$$\underset{\sim}{U}^*\underset{\sim}{C}\underset{\sim}{U} = \underset{\sim}{\Lambda} = \underset{j}{\text{diag}} \quad \exp\left(i \ \frac{2\pi}{N} \ (j-1)\right) \tag{2.23}$$

This transformation decouples the system since the potential expression (2.10) with

$$\vec{\tilde{u}} = \underset{\sim}{U}^*u$$

$$\vec{\tilde{v}} = \underset{\sim}{U}^*v$$

$$\vec{\tilde{F}}_R = \underset{\sim}{U}^*F_R \tag{2.24}$$

$$\vec{\tilde{F}}_M = \underset{\sim}{U}^*\vec{F}_m$$

has the same form as the original system except that the matrix $\underset{\sim}{C}$ is replaced by the diagonal matrix $\underset{\sim}{\Lambda}$. The differential system (2.20) is likewise transformed. Thus any particular solution of (2.8) can be expressed in terms of N subsystems each comprising a single rib coupled to a single membrane. Since the cost of solving a m-dimensional linear system is $O(m^3)$ this represents a considerable computational savings.

The balancing of terms in the equation also can influence the choice of discretization. Based on a report by Lockheed on the specifications for a 55-meter wrap-rib antenna with 48 ribs [2], the following nominal parameter ranges were derive :

$$L \sim 27.5m$$

$$\theta_o \sim 1.31 \ 10^{-1}$$

$$I_o \sim 1.31 \ 10^{-6} \ m^4$$

$$\tag{2.25}$$

$$B \sim 9.72 \ 10^{10} \ nt/m^2$$

$$T_R \sim 1.75 \ 10^{-1} \ nt/m$$

$$T_\theta \sim 3.50 \ 10^{-1} \ nt/m$$

This gives the proper scalings in system (2.20). For simplicity we take $\sigma = T_\theta L^2/\theta_o$ and we have

$$G_1 \sim 2.29$$

$$G_2 \sim 8.58 \ 10^{-3} \tag{2.26}$$

$$G_3 \sim 1$$

This means that the radial terms of the mesh potential are comparatively small except when the radial derivatives are large. How this affects the structure of the system is demonstrated in the following example (see Figure 2).
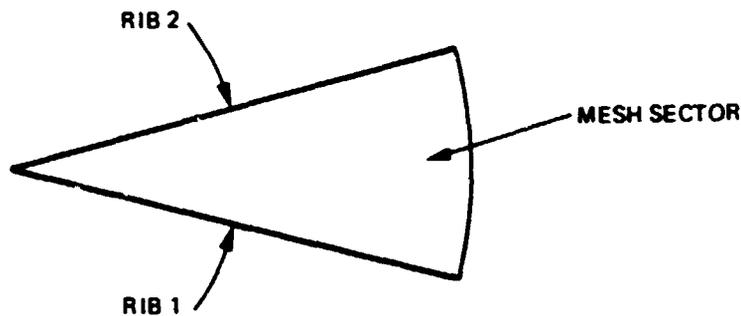
Fig. 2 Single antenna gore

Example:

$$\epsilon \ \frac{1}{r} \ \frac{\partial}{\partial r} \left(r \frac{\partial v}{\partial r} \right) + \ \frac{1}{r^2} \ \frac{\partial^2 v}{\partial \theta^2} \ = 0$$

$$v(r,\theta) \Big|_{\theta=0} = f_1(r) \qquad \begin{array}{l} o < r < 1 \\ o < \theta < 1 \end{array} \tag{2.27}$$

$$v(r,\theta) \Big|_{\theta=1} = f_2(r) \qquad (f_1(o) = f_2(o) = o)$$

$$\frac{\partial v}{\partial r}(r,\theta) \Big|_{r=1} = o$$

In this example we study the equations for a sector of membrane where the prescribed boundary conditions depend on the adjoining rib displacements ($f_1(r)$ and $f_2(r)$). For simplicity we take the forcing on the mesh to be zero although the more

385

general case could be handled in a similar fashion. We are of course interested in the case where

$$0 < \epsilon \ll 1 \qquad (2.28)$$

which corresponds to the parameter ranges (2.26) in (2.20). Physically one expects that the radial terms contribute little to the static behavior except perhaps at the boundary where the gradients may become large. One is also interested in the behavior near the corners $(r,\theta) = (1,0)$ and $(r,\theta) = (1,1)$ since some singular behavior may be possible. Using the techniques of singular perturbations, (see for example [4]), one can show that as ϵ approaches zero we have

$$v(r,\theta) = f_1(r)(1-\theta) + f_2(r)\theta$$

$$+ \epsilon^{1/2} \sum_{n}^{\infty} (-f_1'(1) + f_2'(1)(-1)^n) \frac{2}{(n\pi)^2} \exp\left[\frac{n\pi(r-1)}{\epsilon^{1/2}}\right] \sin(n\pi\theta)$$

$$+ O(\epsilon) \qquad . \qquad (2.29)$$

This expansion could be continued to higher orders, and, as noted before, a more complicated expression would result from forcings on the membrane. One possible approach to the numerical solution of the system (2.20) would by the elimination of the mesh behavior entirely be substituting an expression similar to (2.29) into the beam equations (2.20a). Then one would have only equations along the beams to solve. If higher-order accuracy on the mesh is required, one could then apply finite-element techniques to the system obtained after linearization about the asymptotic expansion for the mesh behavior. Finally we note the appearance of logarithmic singularities in the mesh gradients $(\partial v/\partial v, \partial v/\partial\theta)$ as the corners $(r,\theta) = (1,0)$ and $(r,\theta) = (1,1)$ are approached from the interior of the mesh. This consideration should also influence any finite element approximation of the mesh.

We emphasize that this analysis applies not only to the simplified antenna model we have considered but would hold for more elaborate configurations where a similar structural balance of terms governs the system. Thus, many modeling options can be considered for parameter identification in important classes of structures if one does not insist on a traditional modal characterization of the system.

3. THE LIKELIHOOD FUNCTIONAL

A detailed discussion of the likelihood principle is given in [6]. The functional we consider is the negative logarithm of the likelihood ratio associated with the detection of a Gaussian signal in additive Gaussian noise; this framework is traditional in the theory of communication and signal detection.

In accordance with the discussion given in [6], and the notation discussed in Section 2 ,the log-likelihood functional is given by,

$$J(\Theta,y) = \frac{1}{2} \text{Tr} \left(\log \left[\sigma_\eta^2 I + \sigma_\omega^2 R \right] \right)$$

$$+ \frac{1}{2} (\vec{y} - \vec{m}(\Theta))^* \left[\sigma_\eta^2 I + \sigma_\omega^2 R \right]^{-1} (\vec{y} - \vec{m}(\Theta)) \tag{3.1}$$

$$- \frac{1}{2} (1/\sigma_\eta)^2 \vec{y}^* \vec{y}$$

where the expected mean and the covariance operator are given by:

$$\vec{m}(\Theta) = H(\Theta) \Phi(\Theta) C(\Theta) \vec{f}$$

$$R(\Theta) = H(\Theta)\Phi(\Theta)B(\Theta)B^*(\Theta)\Phi^*(\Theta)H^*(\Theta) \quad . \tag{3.2}$$

Here $\vec{y}_1^* \vec{y}_2$ indicates the Euclidian inner product in the N_s-dimensional space to which the observators belong.

From the assumptions of Section 2, it is easy to see that R is an $N_s \times N_s$ matrix whose (i,j) component is given by

$$R^{(i,j)} = \sum_{k=1}^{N_a} g(\xi_i | \widetilde{\xi}_k) \, g(\widetilde{\xi}_k | \xi_j)$$

$$= \sum_{k=1}^{N_a} g(\xi_i | \widetilde{\xi}_k) \, g(\xi_j | \widetilde{\xi}_k) \tag{3.3}$$

where $g(x|\xi)$ is the point-source solution of the underlying elliptic system

$$Au = \delta(x - \xi) \tag{3.4}$$

with the appropriate boundary conditions. And likewise the expected observation has the form

$$\vec{m} = G \vec{f}$$

$$\vec{f} = (f^{(1)}, \ldots, f^{(N_a)})^T \tag{3.5}$$

$$G^{(i,j)} = g(\xi_i | \xi_j) \quad ,$$

where G is an $N_s \times N_a$ matrix.

We note that (3.1) also differs from (1.3) of [6] in accordance with the introduction of the positive weighting parameters, σ_ω and σ_η, into the system (2.1). An equivalent form for the likelihood functional follows from a rearrangement of terms.

$$J(\Theta,\vec{y}) = \frac{1}{2} \, \text{Tr log} \, [\underset{\sim}{I} + \mu \underset{\sim}{R}]$$

$$+ (1/\sigma_\eta)^2 \, (\vec{y} - \vec{m}(\Theta))^* \, [\underset{\sim}{I} + \mu \underset{\sim}{R}]^{-1} \, (\vec{y} - \vec{m}(\Theta)) \tag{3.6}$$

$$- \frac{1}{2} \, (1/\sigma_\eta)^2 \, \vec{y}^* \vec{y}$$

where

$$\mu = \frac{\sigma_\omega^2}{\mu_\eta^2} \tag{3.7}$$

This form is useful since one can arrive at the functional given in [6] directly by the substitutions

$$\underset{\sim}{R} \to \mu \underset{\sim}{R} = \hat{\underset{\sim}{R}}$$

$$y \to (1/\sigma_\eta)\vec{y} = \hat{\vec{y}} \tag{3.8}$$

$$\vec{m} \to (1/\sigma_\eta) \, \vec{m} = \hat{\vec{m}}$$

This correspondence allows one to use the algorithms derived in [6] directly on the functional

$$J(\Theta,\hat{\vec{y}}) = \frac{1}{2} \, \text{Tr Log} \, [\underset{\sim}{I} + \hat{\underset{\sim}{R}}]$$

$$+ \frac{1}{2} \, (\hat{\vec{y}} - \hat{\vec{m}}(\Theta))^* \, [\underset{\sim}{I} + \hat{\underset{\sim}{R}}]^{-1} \, (\hat{\vec{y}} - \hat{\vec{m}}) \tag{3.9}$$

$$- \frac{1}{2} \, \hat{\vec{y}}^* \hat{\vec{y}}$$

The goal is to find the parameter value Θ which minimizes the log-likelihood functional; that is, we wish to solve

$$\min_{\Theta} \quad J(\Theta,\hat{\vec{y}}) \qquad , \tag{3.10}$$

where Θ ranges over some appropriate infinite-dimensional space. Assuming that the functional has a Frechet derivative and satisfies an appropriate convexity condition, one can restate the problem (3.10) as

$$\partial J/\partial\Theta \, (\Theta,\hat{\vec{y}}) = 0 \tag{3.11}$$

Both problems (3.10) and (3.11) have been studied in a variety of contexts (See, for example, [5]).

Since both the parameter space and the state space are infinite-dimensional, one must make dual approximations in order to achieve problems that are finite-dimensional and therefore computationally tractable. Thus in practice one solves a sequence of problems of the form

$$\min_{\vec{\hat{\theta}}} \ \hat{j}(\hat{\vec{\theta}}, \vec{\hat{y}}) \qquad\qquad (3.12)$$

or

$$\partial \hat{j}/\partial \vec{\hat{\theta}} \ (\vec{\hat{\theta}}, \vec{\hat{y}}) = 0 \qquad\qquad (3.13)$$

where the state-space and parameter-space have been replaced by finite dimensional spaces. Then the problem reduces to a finite minimization problem which can be treated numerically by a variety of techniques (see, for example, [1,5]).

The state-space can be approximated by a finite-element space which is appropriate for approximating solutions to (3.4), and the parameter-space can be conveniently represented by a spline-based space. Let N_x be the dimension of the finite-dimensional approximation to the state-space and let N_θ be the dimensions of the finite-dimensional parameter space. This leads to the natural substitutions

$$u \to \sum_{k=1}^{N_x} u^{(k)} \psi_k(x)$$

$$\theta \to \sum_{k=1}^{N_\theta} \theta^{(k)} \kappa_k(\theta)$$

$$\theta = [\theta^{(1)},..., \theta^{(N_\theta)}]^T \qquad\qquad (3.14)$$

$$u = [u^{(1)},..., u^{(N_x)}]^T \qquad ,$$

where the sets $\{\psi_k\}$ and $\{\kappa_k\}$ give the basis elements for the state and parameter spaces respectively.

It will also be convenient to consider the state-space inner-product with a weighting given by the basis elements of the parameter space. Thus we define

$$< u,v >_j = <u, \kappa_j(\theta)v>$$

$$(j \in \{1,..., N_\theta\}) \qquad . \qquad\qquad (3.15)$$

In the following we restrict our attention to these finite-dimensional problems, and, when the context is clear, we suppress the ^-notation. Questions concerning the convergence of the numerical schemes and the general relationship between the infinite-dimensional and finite-dimensional problems will be discussed more fully in a future report.

Most nonlinear optimization techniques require solving linearized systems iteratively, and consequently one must solve systems of the form (3.4), where the dimension N_x may be quite large. Since the complexity of solving an m-dimensional linear system is $O(m^3)$, the speed of convergence of the iterates is an important consideration. With this in mind, we emphasize the use of quasi-Newton methods for the solution of (3.13). Consequently much of the resulting effort is directed towards deriving adequate approximations for the N_Θ- dimensional Jacobian vector $\partial J/\partial\Theta$ and the $N_\Theta \times N_\Theta$ Hessian matrix $\partial^2 J/\partial\Theta^2$.

We briefly outline the procedure here; as noted previously, a more complete description is given in [6]. In general for the finite problems, the dimension of the state space (N_x) is much larger than the dimension of the parameter space (N_Θ), the number of sensors (N_s) or the number of actuators (N_a), and so it is preferable to carry out the necessary manipulations in spaces whose dimensions do not depend on the dimension of the state space.

Therefore, as in [6] we represent calculations in terms of the eigen-structure of the $N_s \times N_s$ matrix $\underset{\sim}{R}$.

$$\underset{\sim}{R}\vec{\phi}_k = \lambda_k^2 \vec{\phi}_k$$

$$\lambda_k = \tan \alpha_k \quad (0 \le \alpha_r < \frac{\pi}{2})$$

(3.16)

From the spectral components of $\underset{\sim}{R}$ we define useful quantities as given in [6].

From (3.2) we have the expected observation

$$\vec{m} = H\Phi C\vec{f}$$

$$m_k = \vec{\phi}_k * \vec{m} \qquad ,$$

(3.17)

and also we define the filtered observation

$$\vec{z} = \underset{\sim}{L}\vec{y} + (\underset{\sim}{I} - \underset{\sim}{L})\vec{m}$$

$$z_k = \vec{\phi}_k * \vec{z} \qquad ,$$

(3.18)

where the $N_s \times N_s$ matrix $\underset{\sim}{L}$ is given by

$$\underset{\sim}{L} = \underset{\sim}{I} - (\underset{\sim}{I} + \underset{\sim}{R})^{-1/2}$$

$$= \sum_k (1 - \cos\alpha_k) \vec{\phi}_k \vec{\phi}_k^*$$

(3.19)

and the related matrix $\underset{\sim}{K}$ is given by

390

$$\underset{\sim}{K} = (\underset{\sim}{I} + \underset{\sim}{R})^{1/2} - \underset{\sim}{I}$$

$$= \sum_k (\sec\alpha_k - 1) \vec{\phi}_k \vec{\phi}_k{}^* \qquad (3.20)$$

For algebraic convenience we also define the residual of the process:

$$\vec{e} = \vec{y} - \vec{z}$$

$$e_k = \vec{\phi}_k{}^* \vec{e} \qquad (3.21)$$

The gradient of L is represented by

$$\partial \underset{\sim}{L}/\partial\Theta^{(j)} = \sum_k \sum_{m \neq k} a^j_{km} \vec{\phi}_k \vec{\phi}_m{}^* \qquad (3.22)$$

where the coefficients $\{a^j_{km}\}$ are given by:

$$a^j_{km} = \begin{cases} -(\sin\alpha_k)^2 <Dp_k, Dx_k>_j & , k = m \\[3mm] ((\lambda_m \lambda_k)/(\lambda_k^2 - \lambda_m^2))(\cos\alpha_k - \cos\alpha_m) \cdot \\[2mm] [\lambda_k <Dp_m, Dx_k>_j + \lambda_m <Dp_k, Dx_m>_j] & , k \neq m \end{cases} \qquad (3.23)$$

For later convenience we derive another form for the coefficients $\{a^j_{km}\}$. Using standard trigonometric identities one can easily verify the relation

$$\frac{\cos\alpha_k - \cos\alpha_m}{(\tan\alpha_k)^2 - (\tan\alpha_m)^2} = -\frac{(\cos\alpha_k)^2(\cos\alpha_m)^2}{\cos\alpha_m + \cos\alpha_k} \qquad (3.24)$$

This leads to an alternate form for the coefficients

$$a^j_{km} = \begin{cases} -(\cos\alpha_k)^3 [(\lambda_k)^3 <Dp_k, Dx_k>_j] & , k=m \\[3mm] -(\cos\alpha_k \cos\alpha_m)^2/(\cos\alpha_k + \cos\alpha_m) \cdot \\[2mm] [\lambda_m \lambda_k^2 <Dp_m, Dx_k>_j + \lambda_k \lambda_m^2 <Dp_k, Dx_m>_j] & , k \neq m \end{cases} \qquad (3.25)$$

The point of this last derivation is that the bracketed terms reduce to simpler expressions. From (2.2), (2.7), and (3.4) one can easily show:

$$p_k = \lambda_k^{-1} \, \Phi^* \, H^* \, \phi_k$$

$$= \lambda_k^{-1} \sum_{j=1}^{N_s} g \, (x|\xi_j) \phi_k^{(j)}$$

(3.26)

and also from (2.7) we have:

$$x_m = \lambda_m^{-1} \Phi \, B^* p_m$$

$$= \lambda_m^{-2} \sum_{k=1}^{N_a} \sum_{j=1}^{N_s} g \, (x|\widetilde{\xi}_k) \, g \, (\widetilde{\xi}_k|\xi_j) \, \phi_k^{(j)}$$

(3.27)

By this we have:

$$\lambda_k \, \lambda_m^2 < Dp_k, \, Dx_m >_j = \vec{\phi}_k^* \underset{\sim}{A}_j \vec{\phi}_m$$

(3.28)

where the $N_s \times N_s$ matrix $\underset{\sim}{A}_j$ has the form

$$A_j^{(k,m)} = \sum_{i=1}^{N_a} g \, (\xi_i|\xi_m) <Dg \, (x|\xi_k), \, Dg \, (x|\xi_i) >_j \quad .$$

(3.29)

And similarly we have the useful relation

$$\lambda_k <Dp_k, \, D\overline{u}>_j = \vec{\phi}_k^* \underset{\sim}{B}_j \, \vec{f}$$

$$\overline{u} = \Phi C f$$

(3.30)

where the $N_s \times N_a$ matrix $\underset{\sim}{B}_j$ is given by

$$B_j^{(k,m)} = <Dg \, (x|\xi_k), \, Dg \, (x,\widetilde{\xi}_m)>_j$$

(3.31)

We now give expressions for the gradient and the Hessian in terms of the quantities given above. As in [6] the gradient can be represented as

$$\partial J/\partial \theta^{(j)} = -\sum \sin^2 \alpha_k \tan \alpha_k <Dp_k, \, Dx_k >_j$$

$$-\sum e_k \, (\partial z_k/\partial \theta^{(j)})$$

(3.32)

Here the spectral coefficients $\partial z_k/\partial \theta^{(j)}$ are given by

392

$$\partial z_k / \partial \theta^{(j)} = \sum_m \cos\alpha_m \; a^i_{km} \; e_k$$

$$- (\cos\alpha_k) \; [\lambda_k \; \langle Dp_k, \; D\overline{w} \rangle_j]$$

(3.33)

Exact expressions for the Hessian are given in [6]; in general, however, all terms need not be estimated to give an adequate approximation. In particular, the calculations are much simpler if the terms with second-order derivatives can be ignored. The simplest approximation comes from only keeping those terms which contribute to the expected v$_i$ of the Hessian. Thus, from (1.12) of [6] the (i,j) - component of the $N_s \times N_s$ Hessian approximation $\underset{\sim}{M}$ is given by

$$M^{(i,j)} = \; tr \; [\partial \underset{\sim}{L}/\partial \theta^{(i)} \; (\underset{\sim}{I} + \underset{\sim}{K}) \; \partial \underset{\sim}{L}/\partial \theta^{(j)} \; (\underset{\sim}{I} + \underset{\sim}{K}) \;]$$

$$(\partial \overrightarrow{z}/\partial \theta^{(i)}) * (\partial \overrightarrow{z}/\partial \theta^{(j)})$$

(3.34)

$$= \; tr \; [\underset{\sim}{V_i} \; \underset{\sim}{V_j}] + \sum_k \partial z_k/\partial \theta^{(i)} \; \partial z_k/\partial \theta^{(j)}$$

where by (3.20) and (3.22) we have $V_i^{(m,k)} = a_i^{mk} \sec\alpha_k$.

This estimate is justified when the covariance is small as one might expect if the number of measurements is large. This point will be investigated more rigorously in a future paper.

We now summarize the search procedure for the system (2.6) where the N_s-dimensional observation vector $\overrightarrow{y}$ is given and an initial N_θ-dimensional parameter estimate $\overrightarrow{\theta}_0$ is available.

First the expected observation $\overrightarrow{m}$ and the covariance matrix $\underset{\sim}{R}$ are determined from (3.3) and (3.5). The spectral decomposition of $\underset{\sim}{R}$ as well as the quantities given by (3.17) - (3.33) then can be determined by standard matrix algebra routines. And therefore from (3.32), (3.33) and (3.34) one obtains an N_θ-dimensional gradient approximation $\overrightarrow{g}$ and an $N_s \times N_s$ Hessian approximation $\underset{\sim}{M}$.

The parameter estimate θ_0 can then be updated by making the quasi-Newton correction:

$$\overrightarrow{\theta}_* = \overrightarrow{\theta}_0 - \gamma_0 \; \underset{\sim}{M}^{-1} \overrightarrow{g}$$

(3.35)

Here γ_0 is an appropriate scalar chosen to improve the updated parameter estimate. In accordance with the general theory of Newton iterations in function spaces [5], one can repeat this procedure until the solutions of the linearized problems converge to the solution of the underlying nonlinear problem.

This analysis completes our outline of the maximum likelihood identification process. In Section 4 we give examples which illustrate the successful implementation of these schemes in useful applications.

4. EXAMPLES: In this section we give examples of successful implementations of the previously discussed algorithms.

We first consider the string (cf (2.13)

$$Au = - (\alpha(x) \, u'(x))'$$

$$u(o) = u(1) = 0 \qquad .$$

(4.1)

For the case where the unknown tension parameter is constant, the point-source solution can be explicitly given:

$$u(x|\xi) = \frac{1}{\alpha} (1 - x_>) (x_<)$$

$$x_> = \max \{x, \xi\}$$

$$x_< = \min \{x, \xi\} \qquad .$$

(4.2)

And thus, as outlined in Section 2, all calculations could be given in terms of these quantities, without any truncation of the state space or the parameter space.

In general, however, truncations in both spaces are necessary. For the string problem we consider an N_x - dimensional state space of linear splines; the state variable then becomes the vector of nodal values on the corresponding grid. For simplicity we take the grid to be uniform; thus, since the endpoints $x = 0$ and $x = 1$ are fixed, we have:

$$\Delta x = \frac{1}{N_x + 1} \qquad .$$

(4.3)

The state-space elements are then given by

$$u = \sum_{i=1}^{N_x} u^{(i)} \kappa_i(x)$$

(4.4)

where, as illustrated in Figure 3, the basis elements $\{\kappa_i(x)\}$ have the form

$$\kappa_i(x) = \begin{cases} \dfrac{x - (i-1)\Delta x}{\Delta x} & , \ (i-1) \, \Delta x \, < x \, < (i) \, \Delta x \\[2mm] \dfrac{x - (i)\Delta x}{\Delta x} & , \ (i) \, \Delta x \, < x \, < (i+1) \, \Delta x \\[2mm] 0 & , \quad \text{otherwise} \end{cases}$$

(4.5)
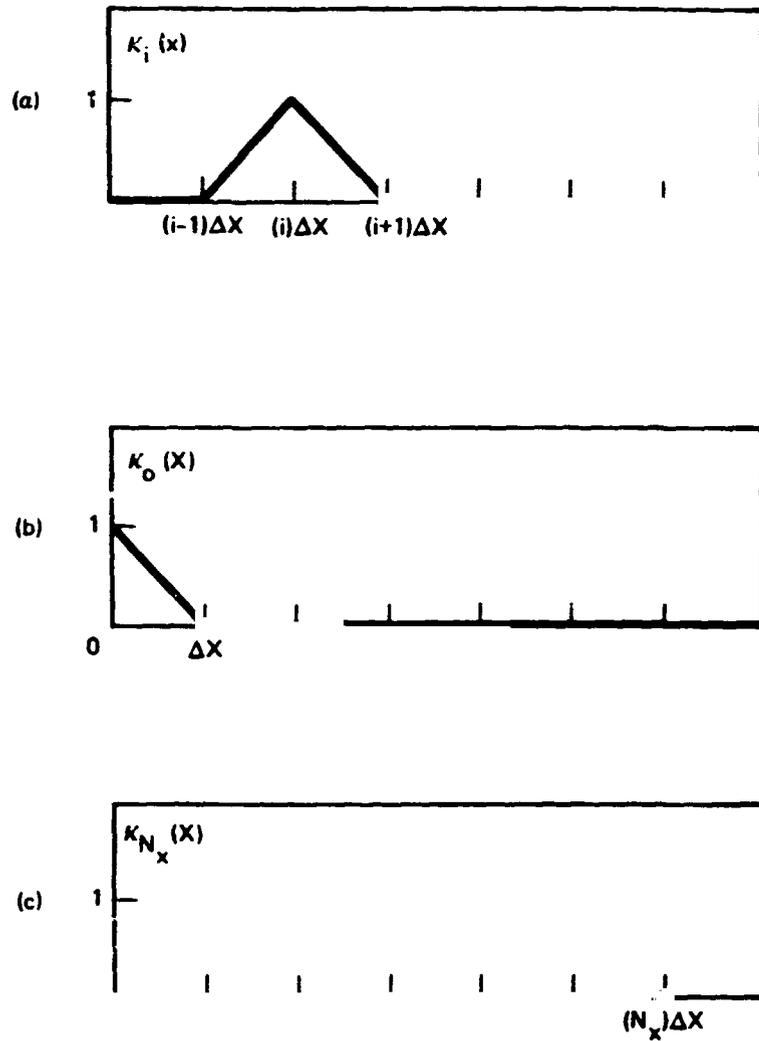
Fig. 3 Linear spline elements

A similar discretization of the parameter space is possible. First we consider the augmented spline space

$$\{\kappa_i(x)\}_{i=0}^{N_x+1} \tag{4.6}$$

where, as illustrated by Figure 3, the endpoint-elements κ_0 and κ_{N_x+1} are given by

$$
\kappa_0(x) = \begin{cases} \dfrac{(\Delta x) - x}{\Delta x} & , \ 0 < x < \Lambda x \\[2em] 0 & , \ \Delta x < x < 1 \end{cases}
$$

$$
\kappa_{N_x+1}(x) = \begin{cases} 0 & , \ 0 < x < N_x (\Delta x) \\[2em] \dfrac{x - N_x (\Delta x)}{\Delta x} & , \ N_x (\Delta x) < x < 1 \end{cases} \tag{4.7}
$$

Thus we have a corresponding parameter element.

$$
\alpha(x) = \sum_{i=0}^{N_x+1} \alpha^{(i+1)} \kappa_i(x) \tag{4.8}
$$

which would give a parameter space $\{\vec{\alpha}\}$ with dimension N_x+2.

However, as previously noted, the resolution of the parameter space often does not need to be as fine as the resolution of the state-space. We consider then the use of a piecewise linear parameter space of lower dimension where the only requirement is that the nodal points must be a subset of the nodal points of the state-space. The new parameter space is then a subset of the (N_x+2)-dimensional space given by (4.6). Let α be an N_Θ-dimensional parameter element $(N_\Theta \leq N_x+2)$. Then α identifies with an element $\vec{\alpha}$ of the larger (N_x+2)-dimensional space and the relationship is given by

$$
\vec{\tilde{\alpha}} = \underline{B}\vec{\alpha} \tag{4.9}
$$

where $\underline{B}$ is an $(N_x+2) \times N_\Theta$ matrix. And correspondingly, we have

$$
\partial\vec{\tilde{\alpha}}/\partial\vec{\alpha} = \underline{B} \quad . \tag{4.10}
$$

This relationship simplifies the algorithms as described below since B is easy to construct, and the more cumbersome calculations which are needed to determine partial derivatives with respect to the parameter space are then specified in terms of the grid associated with the state space. Thus we have:

$$
\partial/\partial\vec{\alpha} = \underline{B} \ \partial/\partial\vec{\tilde{\alpha}} \tag{4.11}
$$

We illustrate these points with a sample calculation (see Fig. 4). We consider the case where there are seven sensors at the locations

$$\xi \in \{.125, .25, .375, .5, .625, .75, .85\} \qquad (4.12)$$
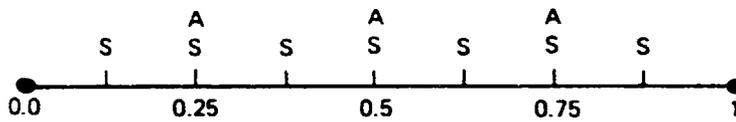
S: SENSOR LOCATION
A: ACTUATOR LOCATION

Fig. 4 String tension identification: sensor and actuator locations

and three actuators at the locations

$$\tilde{\xi} \in \{.25, .5, .75\}. \qquad (4.13)$$

The data vector was derived from a plant with specifications

$$\alpha_{plant}(x) = 3 + x$$

$$C \vec{f} = \delta(x-.25) + \delta(x-.5) + \delta(x-.75)$$

$$\qquad (4.14)$$

$$\sigma_{\omega} = .001$$

$$\sigma_{\eta} = .001$$

For the state space we take the seven-dimensional space of linear splines $\{\kappa_i(x)\}$ with nodes corresponding to the sensor locations (4.12), and for the parameter space we take the five-dimensional subset of linear splines with nodes corresponding to the set

$$\{0, .25, .5, .75, 1.\} \qquad (4.15)$$

The relaxation parameter γ_0 in (3.35) was chosen to speed up to the convergence of the iteration; these issues will be discussed more fully in a future report but we give the results of the calculations in Fig. 5. These numerical experiments appear to be very encouraging although with a crude approximation to the Hessian the convergence can be very slow.
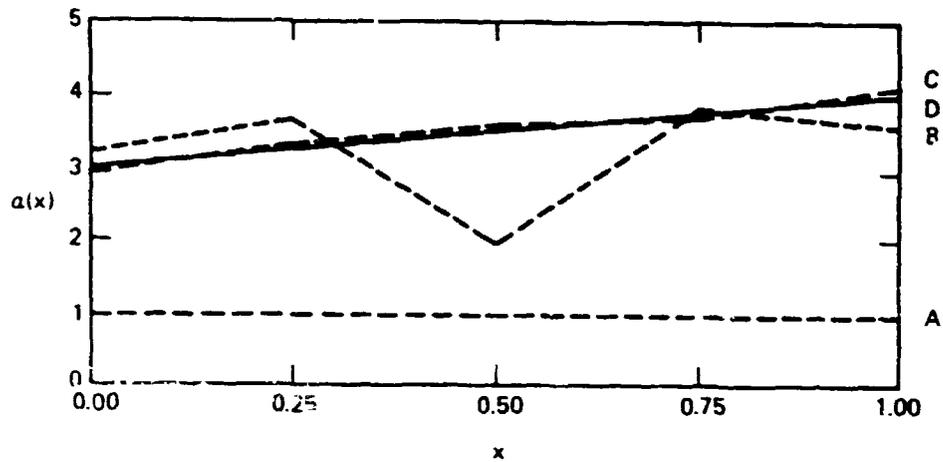
ITERATE	$a^{(1)}$	$a^{(2)}$	$a^{(3)}$	$a^{(4)}$	$a^{(5)}$
A INITIAL	1.00	1.00	1.00	1.00	1.00
B INTERMEDIATE	3.19	3.64	1.98	3.81	3.62
C FINAL	2.9R	3.29	3.55	3.68	4.05
D TRUE	3.00	3.25	3.50	3.75	4.00

Fig. 5 Distributed parameter identification via spline analysis

In a similar way, the algorithm was successfully applied to the wrap-rib antenna model (2.4). To simplify the calculations, we assumed here that the stiffness parameters were scalers although one could introduce a spline-based space as in the previous example.

Again for simplicity we consider the case where there are six gores ($N = 6$), where a sensor is placed on the outer endpoint of each rib ($r = 1$), and where an actuator is placed at the midpoint of each rib ($r = .5$). This scheme is outlined in Figure 6.

We introduce the set of N-dimensional unit vectors

$$\left\{ \vec{\hat{e}}_k \right\}_{k=1}^{N} \qquad (4.16)$$

where the components of each $\vec{\hat{e}}_k$ are determined by

398

● POINT — SENSOR LOCATION
✱ POINT — ACTUATOR LOCATION
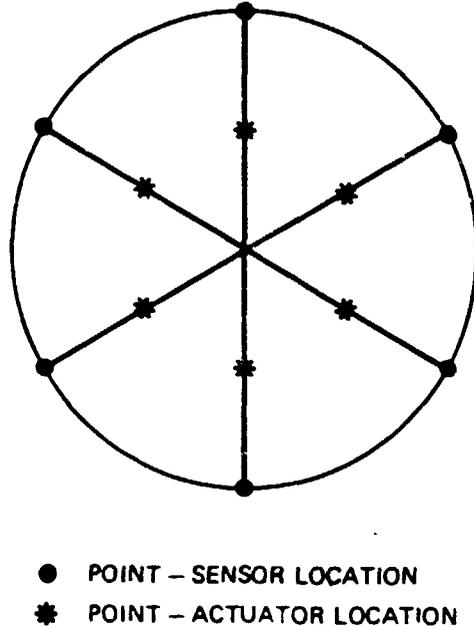
Fig. 6 Rib stiffness identification: sensor and actuator locations

$$\widetilde{e}_k^{(i)} = \begin{cases} 1, & i = k \\ 0, & i \neq k \end{cases}$$ (4.17)

The parameters of the likelihood functional are then given by

$$\sigma_\omega = .001$$

$$\sigma_\eta = .001$$ (4.18)

$$\vec{Cf} = \sum \vec{\widetilde{e}}_k \, \delta \, (r - .5)$$

And the stiffness parameters of the plate are given by

$$EI = 1.25 \cdot 10^5 \ nt\text{-}m^2$$ (4.19a)

$$T_R = 1 \cdot 75 \cdot 10^{-1} \ nt/m$$ (4.19b)

$$T_\Theta = 3.5 \cdot 10^{-1} \ nt/m$$ (4.19c)

$$L = 2.75 \cdot 10^{-1} \ m$$ (4.19d)

We applied the algorithm then to the case where the unknown parameter was EI while the other stiffness parameters were assumed to be known.

399

To discretize the state-space eight equal subdivisions were made in the radial direction on each rib and in each mesh sector; in the circumferential direction five equal subdivisions were made in each mesh sector. The shape functions on the ribs were given by Hermite cubics while on the mesh the shape functions were given by splines linear in r and Θ. In test cases this discretization produced at least three digits of accuracy in solving problems of the form (3.4). In all calculations the principle of cyclic symmetry (cf. (2.24)) was exploited to reduce the number of calculations.

Convergence of the likelihood algorithm was very fast (see Fig. 7) when the relaxation parameter was taken to be

$$\gamma_o = 2.5 \ .$$

LIKELIHOOD FUNCTIONAL

RIB STIFFNESS X 10^{-6} nt-m^2
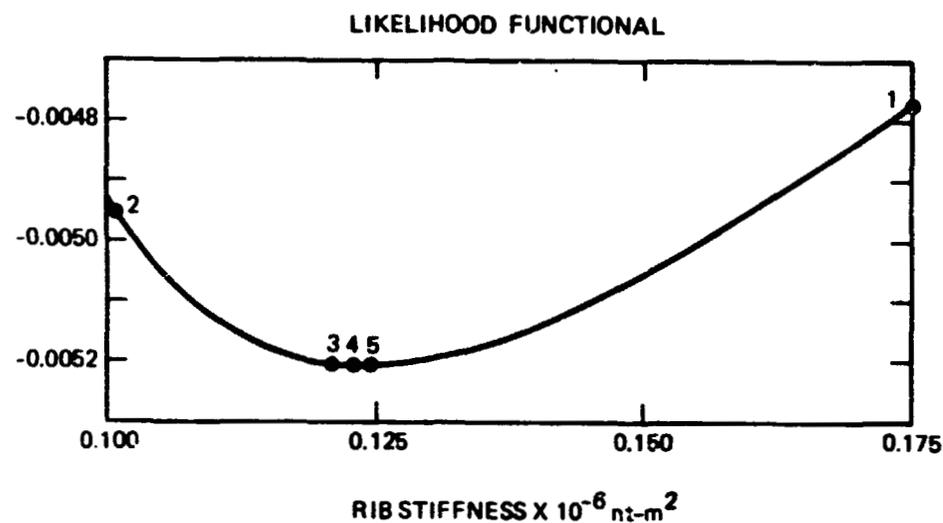
ITERATION N. .IBER N	ESTIMATED PARAMETER EI_N	RELATIVE ERROR $\|(EI_N - EI_{EST})/EI_{EST})\|$
1	$1.75000 \cdot 10^5$	$4. \cdot 10^{-1}$
2	$1.01930 \cdot 10^5$	$2. \cdot 10^{-1}$
3	$1.20484 \cdot 10^5$	$3. \cdot 10^{-2}$
4	$1.23866 \cdot 10^5$	$2. \cdot 10^{-4}$
5	$1.23889 \cdot 10^5$	—

$$EI_{PLANT} = 1.25000 \cdot 10^5$$

$$EI_{EST} = 1.23889 \cdot 10^5$$

Fig. 7 Distributed parameter identification of beam stiffness parameter

400

Numerical experiments also demonstrated an improvement in the sensitivity of the identification schemes as the number of measurements was increased. Thus, Figure 8 illustrates how, for the antenna problem considered, an increase in the number of sensors led to a steepening of the likelihood functional. Here the curves were shifted transversely for illustrative purposes. We note that no corresponding improvement in the parameter estimate occurred in these trials, possibly because of the less favorable signal-to-noise ratio which corresponds to sensing in the interior of the ribs.
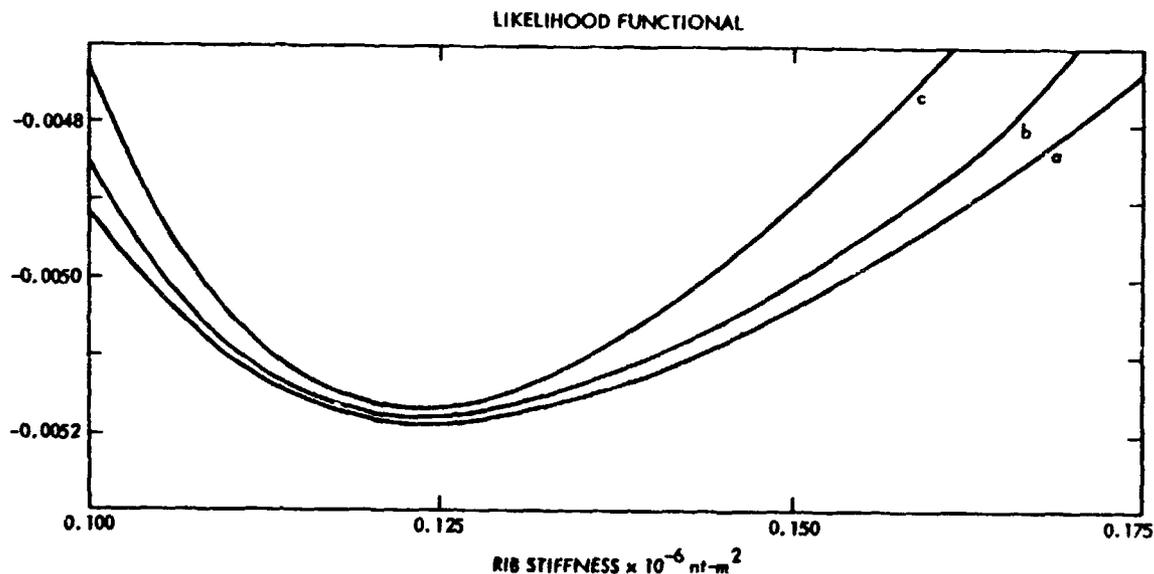
Fig. 8   Sensitivity of antenna stiffness - parameter identification according to number of sensors

(a)   One sensor per rib at r = 1.0

(b)   Two sensors per rib at r = 0.5, 1.0

(c)   Three sensors per rib at r = 0.5, 0.75, 1.0

More detailed numerical experiments with distributed antenna stiffness parameters will be given in a future report. But the results outlined in this report demonstrate already the great potential for these algorithms.

REFERENCES

1.   Dahlquist, G. and Bjork, A., Numerical Methods, Prentice Hall, 1974.

2.   El-Raheb, M. and Wagner, P., "Static and Dynamic Characteristics of Large Deployable Space Reflectors" AIAA Conference Paper 81-0503-PC, AIAA Dynamics Specialists Conference, Atlanta, Georgia, 1981.

3.   Gelfand, I. M., and Fomin, S. V., Calculus of Variations, Prentice-Hall, 1963.

4. Kevorkian, J. and Cole, J., *Perturbation Methods in Applied Mathematics*, Springer-Verlag, New York, 1980.

5. Luenburger, D. G., *Optimization by Vector Space Methods*, Prentice-Hall, 1971.

6. Rodriguez, G., "A Residuals Approach to Filtering, Smoothing and Identification for Elliptic Systems", Proceedings of NASA/ACC Workshop on Iden⁺ification and Control of Large Flexible Structures, 1984.