Bayesian operational modal analysis in time domain using Stan

Javier Cara

Universidad Politécnica de Madrid, Spain

E-mail: javier.cara@upm.es

Abstract. Modal identification consists of determining the natural frequencies, damping ratios and mode shapes of a build structure using measured dynamic data. The Bayesian approach is very appealing when the goal is not only to estimate the modal parameters but also their uncertainty, calculated from their joint probability distribution. This paper presents an example of Bayesian modal identification in time domain using the state space model and the software Stan.

1. Introduction

Operational modal analysis is a technique developed to estimate the modal parameters of a structural or mechanical system using vibration data recorded with sensors. There are many approaches to the problem: parametric and non-parametric; time domain and frequency domain; maximum likelihood and least squares; Bayesian and frequentist estimation,... This work presents the estimation of modal parameters using the Bayesian approach in time domain. A clear advantage of Bayesian formulation is that it provides a rigorous methodology for quantifying the uncertainty of parameters of interest in the presence of measured data and consistent with modeling assumptions.

Recent years have seen many Bayesian formulations in the area of system or modal identification, for example [3], [4]. The method has been successfully applied to bridges [5, 6], high-rise buildings [7], ... Most of these works have been done in the frequency domain. A good example is time domain is [8].

Bayesian estimation is based in three steps:

- Prior distribution: it is the distribution of the modal parameters before the vibration data is recorded.
- Likelihood function: it is the distribution of the recorded data conditional on the modal parameters. In order to calculate the likelihood, a probabilistic model is needed.
- Posterior distribution: it is the distribution of the modal parameters taking into account the recorded data and the model.

Bayesian estimation consists in finding the posterior distributions of the parameters. In this work, the posterior distributions of modal parameters are computed using the following assumptions:

- The likelihood function is derived using the state space model and the Kalman filter. This model is not usual in Bayesian operational modal analysis [3].

- Posterior distributions are computed using Markov Chain Montecarlo sampling by mean of the state-of-the-art software Stan. The usual approach in the bibliography is to adopt the large sample theory and approximate the posterior distributions with normal distributions, so sampling is not needed.

The objective of the paper is to analyze the performance of this approach in Operational Modal Analysis.

2. The model

Consider a single degree of freedom system with equation of motion:

$$\ddot{z}(t) + 2\zeta\omega\dot{z}(t) + \omega^2 z(t) = F(t) \tag{1}$$

where ω is the natural frequency of vibration and ζ is the damping ratio; $z(t)$ is the position of the system at time t and $F(t)$ is the system force. According to the theory of differential equations, this second-order differential equation can be written as first-order matrix differential equation:

$$\begin{bmatrix} \dot{z}(t) \\ \ddot{z}(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\omega^2 & -2\zeta\omega \end{bmatrix} \begin{bmatrix} z(t) \\ \dot{z}(t) \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} F(t) \tag{2}$$

We can rewrite this equation as:

$$\dot{\mathbf{x}}(t) = \mathbf{A_c}\mathbf{x}(t) + \mathbf{B_c}F(t) \tag{3}$$

where

$$\mathbf{x}(t) = \begin{bmatrix} z(t) \\ \dot{z}(t) \end{bmatrix}; \quad \mathbf{A_c} = \begin{bmatrix} 0 & 1 \\ -\omega^2 & -2\zeta\omega \end{bmatrix}; \quad \mathbf{B_c} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \tag{4}$$

Since we work with discrete data, we need the discrete version of this equation. Calling $\mathbf{x}_k = \mathbf{x}(k\Delta t)$ and $F_k = F(k\Delta t)$ we have

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}F_k \tag{5}$$

where

$$\mathbf{A} = \exp(\mathbf{A_c}\Delta t); \quad \mathbf{B} = (\mathbf{A} - \mathbf{I}_2)\mathbf{A_c^{-1}}\mathbf{B_c}. \tag{6}$$

The input force F_k is not measured in OMA, so it is modeled as a zero-mean stationary Gaussian white noise with covariance matrix H:

$$F_k \sim N(0, H). \tag{7}$$

Therefore

$$\mathbf{w}_k = \mathbf{B}F_k \sim N(0, \mathbf{Q}), \tag{8}$$

where $\mathbf{Q} = \mathbf{B}H\mathbf{B}^T$. Then, Equation (5) becomes:

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{w}_k, \quad \mathbf{w}_k \sim N(\mathbf{0}, \mathbf{Q}). \tag{9}$$

In this work we assume we measure the system accelerations. Due to modeling error and measurement noise, the measured acceleration y_k is not the same as the system acceleration $\ddot{z}_k$. It is assumed that

$$y_k = \ddot{z}_k + e_k, \quad e_k \sim N(0, E). \tag{10}$$

Using Equation (1), we have:

$$y_k = \mathbf{C}\mathbf{x}_k + v_k, \quad v_k \sim N(0, R) \tag{11}$$

where

$$\mathbf{C} = \begin{bmatrix} -\omega^2 & -2\zeta\omega \end{bmatrix}; \quad v_k = F_k + e_k; \quad R = H + E \tag{12}$$

It is important to note that the noise processes $\mathbf{w}_k$ and v_k are correlated:

$$Cov[\mathbf{w}_k, v_k] = E[\mathbf{B}F_k(F_k + e_k)] = \mathbf{B}H = \mathbf{S}. \tag{13}$$

The model we will use in this work is composed of equations (9) and (11). This model is called *the state space model*. The parameters of this model are: $\boldsymbol{\theta} = \{f, \zeta, E, H\}$, where f is the system frequency in Hertz ($\omega = 2\pi f$).

3. Bayesian Inference

Given a set of measurements

$$Y = \{y_1, y_2, \ldots, y_N\} \tag{14}$$

the objective is to estimate the parameters $\boldsymbol{\theta} = \{f, \zeta, E, H\}$. In Bayesian inference this is done in a probabilistic framework, that is, we consider these parameters are random variables. Therefore, having observe Y the task is to obtain the posterior distribution of $\boldsymbol{\theta}$. Applying the Bayes' rule:

$$p(\boldsymbol{\theta}|Y) = \frac{p(Y|\boldsymbol{\theta})p(\boldsymbol{\theta})}{p(Y)} \tag{15}$$

where $p(\cdot)$ represents a generic probability density function (PDF):

- $p(Y|\boldsymbol{\theta})$ is known as the likelihood.
- $p(\boldsymbol{\theta})$ is the prior distribution of the parameters.
- $p(Y)$ is the PDF of the data.

In the next sections we describe each element of Equation (15).

3.1. The data Y

We consider acceleration data obtained from a simulated system with f = 4 Hz, ζ = 0.02, H = 1 N^2; the square root of E is 10% of the standard deviation of the noise-free response. The time step used to generate the data is 0.02 s., and the number of data points is N = 1000 (see Figure 1).

3.2. The likelihood $p(Y|\boldsymbol{\theta})$

The starting point is the state space model:

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{w}_k, \quad \mathbf{w}_k \sim N(\mathbf{0}, \mathbf{Q}), \tag{16}$$

$$y_k = \mathbf{C}\mathbf{x}_k + v_k, \quad v_k \sim N(0, R). \tag{17}$$

$$Cov[\mathbf{w}_k, v_k] = \mathbf{S}, \tag{18}$$

where $\mathbf{A}, \mathbf{C}, \mathbf{Q}, R, \mathbf{S}$ have been defined in Section 2. For this model, the likelihood is computed using the Kalman filter. First, we are going to define:

$$\mathbf{x}_{k|k-1} = E[\mathbf{x}_k | y_1, \ldots, y_{k-1}] \tag{19}$$

$$\mathbf{P}_{k|k-1} = E[(\mathbf{x}_k - \mathbf{x}_{k|k-1})(\mathbf{x}_k - \mathbf{x}_{k|k-1})^T | y_1, \ldots, y_{k-1}] \tag{20}$$

The Kalman filter is an algorithm to update the value of $\mathbf{x}_{k|k-1}$ and $\mathbf{P}_{k|k-1}$ when a new value of y_k is available:

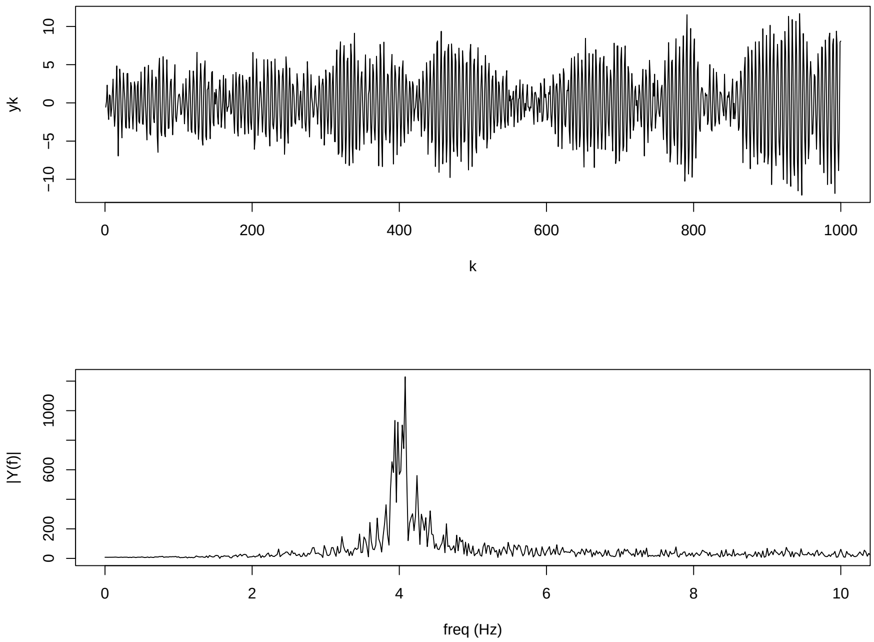

Figure 1. Data obtained from the simulated system. UP: time domain plot; DOWN: absolute value of the Fourier transform.

Property 3.1 (The Kalman filter)**.** *Given* $\mathbf{x}_{1|0}$ *and* $\mathbf{P}_{1|0}$*, for* $k = 1, 2, \ldots, N$

$$\mathbf{x}_{k+1|k} = \mathbf{A}\mathbf{x}_{k|k-1} + \mathbf{K}_k \epsilon_k \tag{21}$$

$$\mathbf{P}_{k+1|k} = \mathbf{A}\mathbf{P}_{k|k-1}\mathbf{A}^T + \mathbf{Q} - \mathbf{K}_k \Sigma_k \mathbf{K}_k^T \tag{22}$$

where

$$\epsilon_k = y_k - \mathbf{C}\mathbf{x}_{k|k-1} \tag{23}$$

$$\Sigma_k = \mathbf{C}\mathbf{P}_{k|k-1}\mathbf{C}^T + R \tag{24}$$

$$\mathbf{K}_k = (\mathbf{A}\mathbf{P}_{k|k-1}\mathbf{C}^T + \mathbf{S})\Sigma_k^{-1} \tag{25}$$

The likelihood is computed using the innovations ϵ_k. When the system force F_k and the noise e_k are Gaussian processes, the innovations also follow a Normal distribution:

$$\epsilon_k \sim N(0, \Sigma_k). \tag{26}$$

Hence we may write the logarithm of the likelihood as:

$$p(Y|\boldsymbol{\theta}) = -\frac{1}{2} N \log(2\pi) - \frac{1}{2} \sum_{k=1}^{N} \log|\Sigma_k| - \frac{1}{2} \sum_{k=1}^{N} \epsilon_k^2 \Sigma_k^{-1}. \tag{27}$$

3.3. Prior distributions $p(\boldsymbol{\theta})$

In this work we consider the parameters are independent random variables, so:

$$p(\boldsymbol{\theta}) = p(f)p(\zeta)p(H)p(E). \tag{28}$$

We use the following distributions:

- $p(f)$: normal distribution.
- $p(\zeta)$: beta distribution.
- $p(H)$: half-Cauchy distribution.
- $p(E)$: half-Cauchy distribution.

The Figure 2 shows the values for these distributions used in this work:

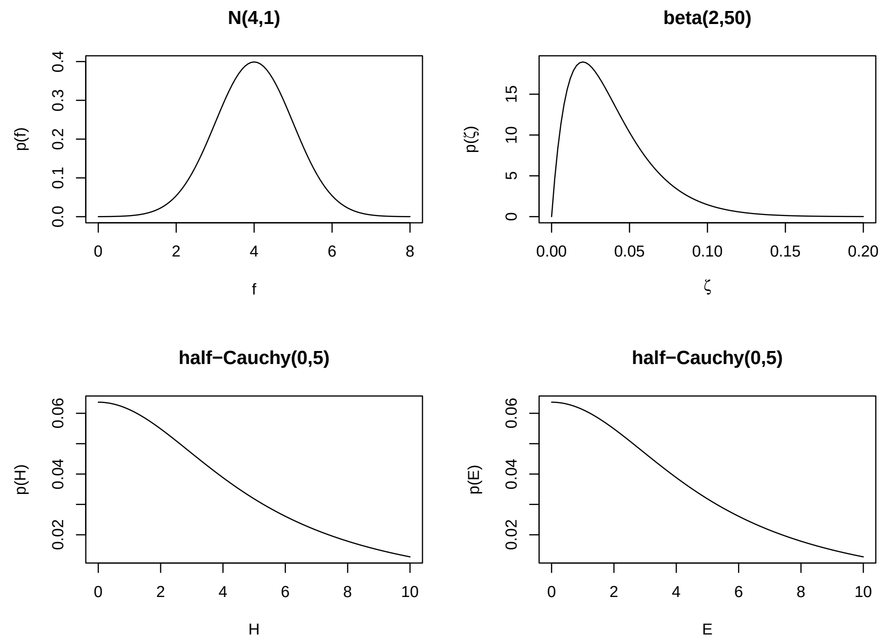

Figure 2. Prior distributions.

3.4. Distribution of the data $p(Y)$

The distribution of the data can be found solving the multiple integral

$$p(Y) = \int_{\boldsymbol{\theta}} p(Y, \boldsymbol{\theta}) d\boldsymbol{\theta} = \int_{\boldsymbol{\theta}} p(Y|\boldsymbol{\theta})p(\boldsymbol{\theta})d\boldsymbol{\theta}. \tag{29}$$

This integral can be only calculated explicitly in simple problems. Furthermore, any approximate numerical scheme that uses a deterministic method to estimate the above integral, for example Gaussian quadrature, will also fail to work. There are two main methods to avoid this difficulty [1]: the first one is to use priors that are conjugate to the likelihood; an alternative approach is to sample from the posterior distribution using Markov chain Monte Carlo (MCMC). There are different MCMC algorithms: Metropolis, Gibbs, Hamiltonian Monte Carlo (HMC), . . . In this work we use HMC by mean of the software Stan [2]. The main reasons for using Stan are:

- Stan implements an improved version of HMC known as the Non-U-Turn Sampler (NUTS) that makes the sampler fast.
- It is very popular, what means we can easily find documentation and examples, and also the program have been extensively tested.
- The language allows us to define our own likelihood functions, like the one shown in Equation (27).

4. Results

4.1. Using Stan

The data and programs used to obtain the posterior distribution of $\boldsymbol{\theta}$ can be found in the Github repository `https://github.com/javiercara/eurodyn2023`, so interested people can replicate the results.

Table 1 shows the mean, standard deviations and quantiles of the posterior distributions. On the other hand, Figure 3 shows the histogram of the posterior distributions $p(f|Y,\zeta,E,H)$ and $p(\zeta|Y,f,E,H)$. At a first look, they are similar to the Gaussian distribution. In fact, we have included (in red) the normal PDF using the mean and standard deviations shown in Table 1.

Table 1. Means, standard deviations and quantiles of the posterior distributions.

parameter	mean	sd	2.5%	25%	50%	75%	97.5%
f	4.025294	0.015608	3.9948e+00	4.0148e+00	4.0249e+00	4.0360e+00	4.0557e+00
ζ	0.016086	0.003608	9.3005e-03	1.3581e-02	1.6028e-02	1.8534e-02	2.3260e-02
H	1.076008	0.062255	9.6330e-01	1.0333e+00	1.0735e+00	1.1164e+00	1.2060e+00
E	0.175564	0.022389	1.3623e-01	1.5954e-01	1.7429e-01	1.9016e-01	2.2279e-01

4.2. Using the Normal approximation to the posterior distribution

In the previous section we have seen that the normal distribution is a good approximation to the posterior distribution. In this section we are going to discuss this approximation.

Let's called $\hat{\boldsymbol{\theta}}$ the mode of the posterior distribution $P(Y|\boldsymbol{\theta})$. The Taylor expansion of $P(Y|\boldsymbol{\theta})$ up to the quadratic term of the log posterior density centered at $\hat{\boldsymbol{\theta}}$ is:

$$log(p(\boldsymbol{\theta}|Y)) \approx log(p(\hat{\boldsymbol{\theta}}|Y)) + \mathbf{G}(\hat{\boldsymbol{\theta}})^T(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}) + \frac{1}{2}(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}})^T\mathbf{H}(\hat{\boldsymbol{\theta}})(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}})$$

where $\mathbf{G}(\hat{\boldsymbol{\theta}})$ is the gradient of $log(p(\boldsymbol{\theta}|Y))$ evaluated at $\hat{\boldsymbol{\theta}}$, $\mathbf{H}(\hat{\boldsymbol{\theta}})$ is the Hessian of $log(p(\boldsymbol{\theta}|Y))$ evaluated at $\hat{\boldsymbol{\theta}}$. Since $\hat{\boldsymbol{\theta}}$ is the mode ot the distribution, the gradient $\mathbf{G}(\hat{\boldsymbol{\theta}})$ is equal to zero. Therefore:

$$log(p(\boldsymbol{\theta}|Y)) \approx log(p(\hat{\boldsymbol{\theta}}|Y)) + \frac{1}{2}(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}})^T\mathbf{H}(\hat{\boldsymbol{\theta}})(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}) \Rightarrow$$

$$p(\boldsymbol{\theta}|Y) \approx exp\left(const + \frac{1}{2}(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}})^T\mathbf{H}(\hat{\boldsymbol{\theta}})(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}})\right) \Rightarrow$$

$$p(\boldsymbol{\theta}|Y) \approx N(\hat{\boldsymbol{\theta}}, -\mathbf{H}(\hat{\boldsymbol{\theta}})^{-1}) \tag{30}$$

To apply this approximation we need to find the mode of the posterior density, $\hat{\boldsymbol{\theta}}$. Due to the complexity of the posterior distribution we have to use a numerical optimization algorithm. In

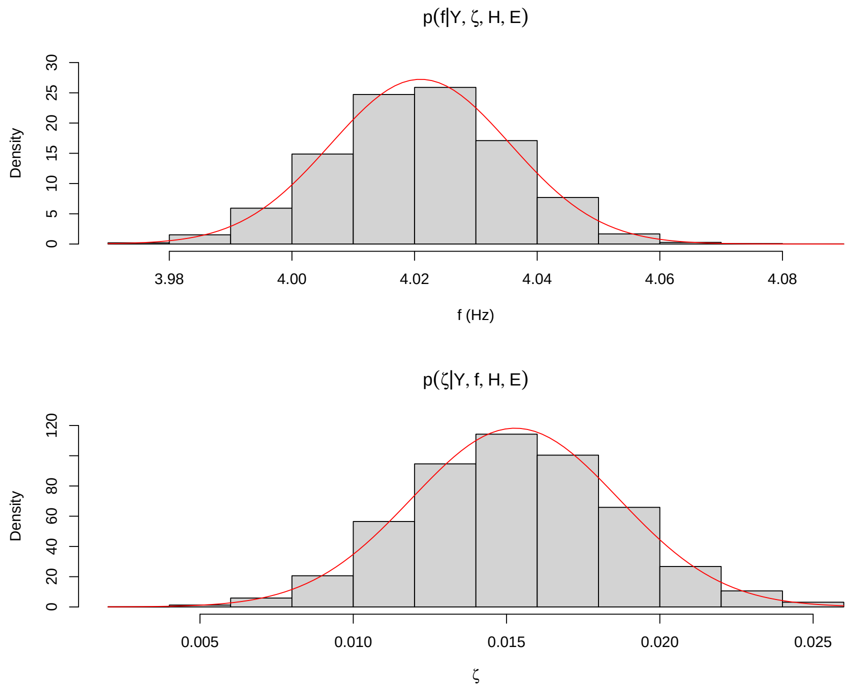

Figure 3. Posterior distributions $p(f|Y, \zeta, E, H)$ and $p(\zeta|Y, f, E, H)$.

this work we have used the BFGS algorithm. The objective function is obtained from Equation (15):

$$p(\boldsymbol{\theta}|Y) = \frac{p(Y|\boldsymbol{\theta})p(\boldsymbol{\theta})}{p(Y)} \Rightarrow log[p(\boldsymbol{\theta}|Y)] = log[p(Y|\boldsymbol{\theta})] + log[p(\boldsymbol{\theta})] - log[p(Y)]$$

The term $log[p(Y)]$ does not depend of $\boldsymbol{\theta}$, so we find the mode of:

$$f(\boldsymbol{\theta}) = log[p(Y|\boldsymbol{\theta})] + log[p(\boldsymbol{\theta})]. \tag{31}$$

It is important to remark that $log[p(Y|\boldsymbol{\theta})]$ is computed using the Kalman filter, Property 3.1. On the other hand, $log[p(\boldsymbol{\theta})]$ is computed using Section 3.3. The results are shown in Table 2. We can check that the agreement with Table 1 is quite good.

Table 2. Estimates and their standard deviation computed from the Normal approximation to the posterior distribution.

parameter	estimate	sd
f	4.02502640	0.015301990
ζ	0.01593136	0.003653876
$\sqrt{H}$	1.03365917	0.029914477
$\sqrt{E}$	0.41400722	0.026655830

5. Conclusions

Bayesian methodology has a long tradition in Operational Modal Analysis. The usual approach is to work in the frequency domain and to use the normal approximation so, in practice, this approach is equivalent to maximize the likelihood.

On the other hand, the bayesian community has develop numerical tools to find posterior distributions in different and general problems. For example, the Hamiltonian Monte Carlo sampler. In this context, the Stan software is possibly the reference software. In this work we have shown that it is possible to estimate the modal parameters of structures using this approach. More specifically:

- The probabilistic model for the likelihood is the state space model.
- The posterior distribution is obtained using sampling (in this case, using Hamiltonian Monte Carlo by mean of Stan).

In future works we will try to extend this methodology to the case of multiple degrees of freedom.

References

[1] Gelman A, Carlin J, Stern H, Dunson D, Vehtari A and Rubin D 2013 *Bayesian Data Analysis* (Chapman and Hall/CRC, 3rd ed.)

[2] Stan Development Team 2023 *Stan Modeling Language Users Guide and Reference Manual* (version 2.31. `https://mc-stan.org`)

[3] Yuen KV *Bayesian methods for structural dynamics and civil engineering* (John Wiley & Sons)

[4] Au SK, Zhang FL, Ni YC 2013 Bayesian operational modal analysis: theory, computation, practice. *Computers & Structures*, Volume 126, Pages 3-14.

[5] Brownjohn JMW, Au SK, Zhu Y, Sun Z, Li B, Bassitt J, Hudson E, Sun H 2018 Bayesian operational modal analysis of Jiangyin Yangtze River Bridge, *Mechanical Systems and Signal Processing*, Volume 110, Pages 210-230.

[6] Yan-Chun N, Zhang QW, Liu JF 2022 Dynamic performance investigation of a long-span suspension bridge using a Bayesian approach. *Mechanical Systems and Signal Processing*, Volume 168.

[7] Ni Y, Lu X, Lu W 2017 Operational modal analysis of a high-rise multi-function building with dampers by a Bayesian approach. *Mechanical Systems and Signal Processing*, Volume 86, Pages 286-307.

[8] Yuen KV, Katafygiotis LS 2001 Bayesian time-domain approach for modal updating using ambient data, *Probabilistic Engineering Mechanics*, Volume 16, Pages 219-231.