Master of Science in Mechatronics Engineering

System identification

Professor:
David Macii

Student:
Cardone Andrea 224836

Department of Industrial Engineering
Academic Year 2021 - 2022

1 Load and display data

First of all the data related to the impulse response are loaded from the file `data_impulses.mat`, then the data are scaled according to the project assignment. In figure 1 are reported the loaded data.

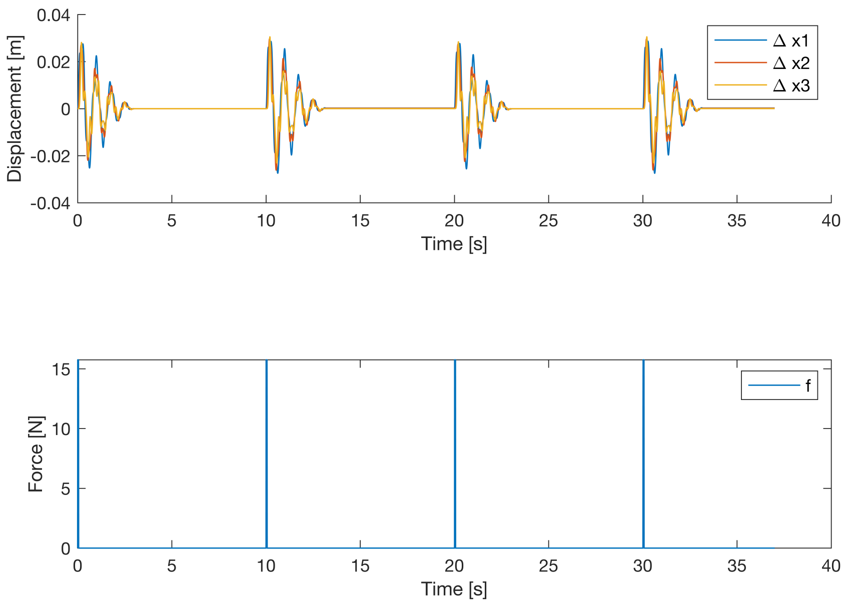

Figure 1: Masses displacement and applied force.

2 Frequency response of the subsystems

First of all the data are splitted in four parts, with N equal to the number of samples of a single segment. There are several ways to compute the frequency response of the signal, considering that the impulse is not ideal, in this project are presented the following methods:

1. The impulse force is characterized by a small number of sample with respect to the duration of the displacement signals. In fact it is a non unitary step with a length of 10/11 samples. So it does not affect qualitatively the frequency response, it only cause an amplification effect of $k = \sum_{n=0}^{N-1} f(n)$, due to $x_1(n) = h_1(n) * f(n)$. Therefore the fft of the signals is computed and then resacaled to obtain the unitary impulse response. This is done for all the splitted part of the signals and then is calculated the average frequency response.

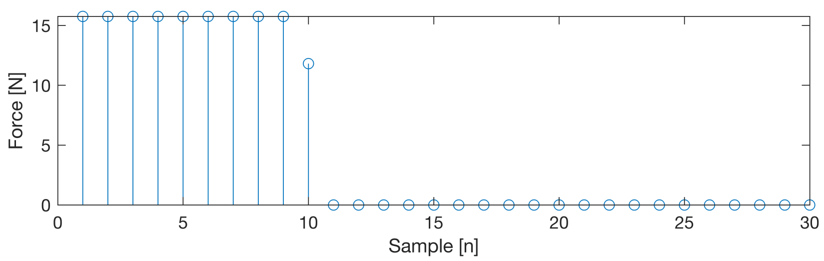

Figure 2: First thirty samples of the first input force.

2. In this case the assumption done before are not valid. The fft of the displacement signals and the input force are computed, then the empirical transfer function of the systems is calculated as $H_1(e^{j\omega}) = X_1(e^{j\omega})F(e^{j\omega})^{-1}$. Then the mean transfer function is calculated.

3. To estimate the frequency response of the systems it has been implemented a Welch-Bartlett inspired method. The whole, not splitted, signal is divided into overlapping segments (50% of overlap), then each segment is windowed with the hamming window, then the empirical transfer function is calculated as $H_1(e^{j\omega}) = \alpha X_1(e^{j\omega})F(e^{j\omega})^{-1}$, where α is the correction factor for the applied windows. This is done for all the splitted part of the signals and then is calculated the average frequency response, at the end of the project is described the full implementation in detail.

4. To make a comparison with the previous methods it has been utilized the matlab function `tfestimate` which calculate the empirical transfer function between the input and the output with the Welch's averaged, modified periodogram method. The function is used with default settings, therefore 8 overlapping by 50% segments are windowed with Hamming window, then the power spectral density of x, P_{xx}, and the cross power spectral density of x and y, P_{xy}, is calculated. At the end the empirical tansfer function is obtained: $T_{xy} = P_{xy}P_{xx}^{-1}$.

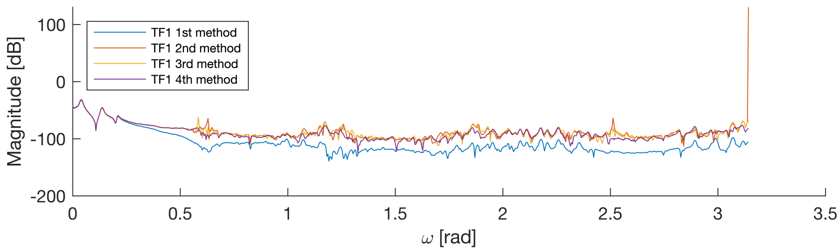

Figure 3: Frequency response of the first system with the four methods.

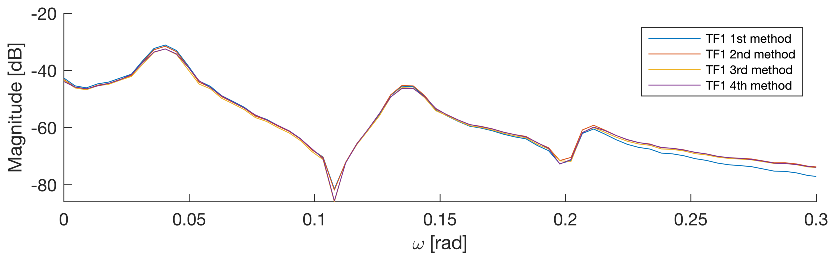

Figure 4: Detail of the notable frequency response of the first system.

The four methods show pretty similar performance for the relevant part related to the system, where we can see the three poles of the systems. For the remaining part the first method present a lower magnitude respect to the others, this is due to the fact that is absent the effect of the ripples of the Fourier transform of the non ideal impulse. In addition the second method present a high peak corresponding to the frequency of π, this is due to the fact that the transform of the non-ideal impulse can have very negative peaks, which, once the inverse is calculated, become positive. This problem was solved in the third method thanks to the overlapping segments and the application of windows. In all methods the fft is computed with N points to permit a good comparison.

3 System model and numerical fitting

3.1 Model analysis

The provided model of the system is shown in the following image. As we can see the model is without damping. Therefore we expect a maximum of three natural frequencies, for which we have an increase of the magnitude frequency response due to the resonance effect. In particular we can observe this behaviour in the empirical transfer function of the three subsystems.

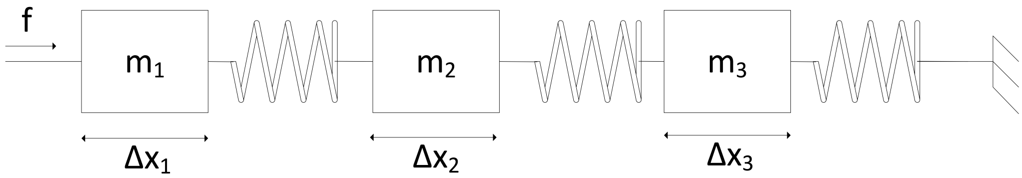

Figure 5: Model of the system without damping.

The models of the undamped systems are derived with the following steps
Newton equations:

$$
\begin{aligned}
m_1\ddot{x}_1 &= f_{imp} - k_1(x_1 - x_2) \\
m_2\ddot{x}_2 &= -k_1(x_2 - x_1) - k_2(x_2 - x_3) \\
m_3\ddot{x}_3 &= -k_2(x_3 - x_2) - k_3x_3
\end{aligned}
$$

Laplace transform and collect the coefficients in the mass matrix and stiffness matrix:

$$
X\left(\begin{bmatrix} m_1 & 0 & 0 \\ 0 & m_2 & 0 \\ 0 & 0 & m_3 \end{bmatrix} s^2 + \begin{bmatrix} k_1 & -k_1 & 0 \\ -k_1 & k_1+k_2 & -k_2 \\ 0 & -k_2 & k_2+k_3 \end{bmatrix}\right) = F
$$

$$
D = \begin{bmatrix} m_1 & 0 & 0 \\ 0 & m_2 & 0 \\ 0 & 0 & m_3 \end{bmatrix} s^2 + \begin{bmatrix} k_1 & -k_1 & 0 \\ -k_1 & k_1+k_2 & -k_2 \\ 0 & -k_2 & k_2+k_3 \end{bmatrix}
$$

$$
\begin{aligned}
X &= D^{-1}F \\
X &= GF
\end{aligned}
$$

Where $G(s)$ is a 3×3 matrix represent the impulse response of the systems, in particular we are interested in the components of the first column of the matrix $G(s)$ which represent the transfer functions of the displacements of the three bodies with respect to a force applied on the first body.
The three transfer functions are characterized by the same denominator $D(s)$ (same natural modes of the systems) and by the numerators, respectively $N_1(s), N_2(s), N_3(s)$.

$$D(s) = a_0 + a_2 s^2 + a_4 s^4 + s^6$$
$$N_1(s) = b_{10} + b_{12} s^2 + b_{14} s^4$$
$$N_2(s) = b_{20} + b_{22} s^2$$
$$N_3(s) = b_{30}$$

From this analysis we can conclude that all the systems are characterized by the same 3 double imaginary poles, in addition the system 1 has two double imaginary zeros, the system 2 has one double imaginary zero and the system 3 doesn't have zeros. This is confirmed by the experimental transfer function of the subsystems.

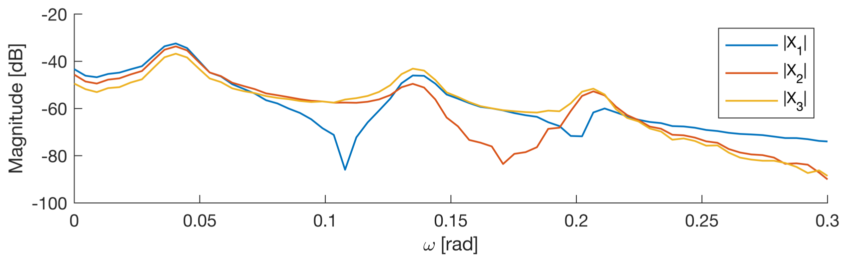

Figure 6: Magnitude of the empirical frequency responses of the three subsystems.

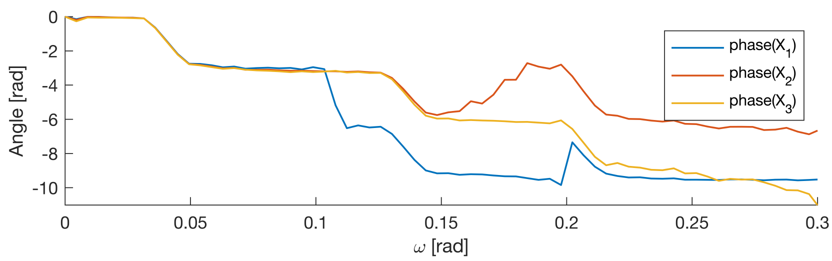

Figure 7: Phase of the empirical frequency responses of the three subsystems.

3.2 Creation of the models for numerical fitting

The models of the systems for the numerical fitting are defined as following:

$$G_1(x_1, s) = \frac{b_{10}(b_{11} + b_{12}s + s^2)(b_{13} + b_{14}s + s^2)}{(a_{11} + a_{12}s + s^2)(a_{21} + a_{22}s + s^2)(a_{31} + a_{32}s + s^2)}$$
$$G_2(x_2, s) = \frac{b_{20}(b_{22} + b_{23}s + s^2)}{(a_{11} + a_{12}s + s^2)(a_{21} + a_{22}s + s^2)(a_{31} + a_{32}s + s^2)}$$
$$G_3(x_3, s) = \frac{b_{30}}{(a_{11} + a_{12}s + s^2)(a_{21} + a_{22}s + s^2)(a_{31} + a_{32}s + s^2)}$$

Where:

$$
\begin{aligned}
x_1 &= [b_{10}, b_{11}, b_{12}, b_{13}, b_{14}] \\
x_2 &= [b_{20}, b_{22}, b_{23}] \\
x_3 &= [b_{30}, a_{11}, a_{12}, a_{21}, a_{22}, a_{31}, a_{32}]
\end{aligned}
$$

x_1 and x_2 doesn't include the denominator values because they will be obtained from the third transfer function, as the optimization can be done with respect to fewer parameters.
This representation is particularly useful because we can approximately derive graphically the position of the zeros and the poles from the empirical transfer functions, which is very useful to set the starting point for the numerical fitting. This models take into account a linear damping of the systems.

3.3 Numerical fitting

First of all is defined the analog frequency vector $\Omega = \omega f_s$, where f_s is the sampling frequency, that defines the interval on which is evaluated the model in order to compute the error function. The frequency vector includes frequencies between 0 and about $60\, rad/s$, where the poles and zeros of the model are located. It is useless to extend beyond the frequency vector, it would lengthen the optimization time.
Then is defined the complex vector $s = i\Omega$.
From the magnitude and phase plot, we can identify the approximate position of poles and zeros, without considering the effect of damping, that modifies the position of the peaks.
Therefore we obtain:

- $\Omega_{p_1} = 8\, rad/s$, $\Omega_{p_2} = 27\, rad/s$, $\Omega_{p_3} = 42\, rad/s$
- $\Omega_{z_{11}} = 21.5\, rad/s$, $\Omega_{z_{12}} = 39.5\, rad/s$
- $\Omega_{z_{21}} = 34\, rad/s$

The error functions are defined as:

$$
\begin{aligned}
e_3(x_3) &= (|H_3(s)| - |G_3(x_3, s)|)^2 + 1\text{e}{-2}(phase(H_3(s)) - phase(G_3(x_3, s)))^2 \\
e_2(x_2) &= (|H_2(s)| - |G_2(x_2, s)|)^2 + 1\text{e}{-5}(phase(H_2(s)) - phase(G_2(x_2, s)))^2 \\
e_1(x_1) &= (|H_1(s)| - |G_1(x_1, s)|)^2 + 1\text{e}{-5}(phase(H_1(s)) - phase(G_1(x_1, s)))^2
\end{aligned}
$$

The weight of the cost part related to the phase is used to balance the relative importance of the magnitude and phase during minimization.

The minimization of $e_3(x_3)$ was done before the other two in order to find the denominator parameters, as mentioned previously.
For the optimization it is used the MATLAB function `fminsearch` (with default options), which uses the Nelder-Mead method and with which good results have been obtained.

The initial guesses for the optimization are:

$$
\begin{aligned}
x_{30} &= [10000, \Omega_{p_1}^2, 1, \Omega_{p_2}^2, 1, \Omega_{p_3}^2] \\
x_{20} &= [500, \Omega_{z_{21}}^2, 1] \\
x_{10} &= [0.1, \Omega_{z_{11}}^2, 0.001, \Omega_{z_{12}}^2, 0.1]
\end{aligned}
$$

The natural frequencies appear squared, as it results in a simple mass, spring, damper system with 1 dof, same reasoning for the position of zeros. The initial values of the other terms were found after a couple of attempts, (assuming a low damping, due to the presence of the increased amplitude of the transfer function module at the natural frequencies of the system).

The results of the optimizations are:

$$
\begin{aligned}
x3_{opt} &= [236953.30, 68.19, 1.46, 748.93, 1.83, 1755.38, 1.64] \\
x2_{opt} &= [310.75, 1133.83, 2.05] \\
x1_{opt} &= [0.61, 471.10, 0.007, 1619.26, 0.59]
\end{aligned}
$$

And the minimum costs obtained are:

$$
\begin{aligned}
e3_{opt} &= 2.68\text{e}{-2} \\
e2_{opt} &= 5.94\text{e}{-5} \\
e1_{opt} &= 1.06\text{e}{-4}
\end{aligned}
$$

In the following images are reported the empirical transfer function amplitude and phase compared with the fitted one.

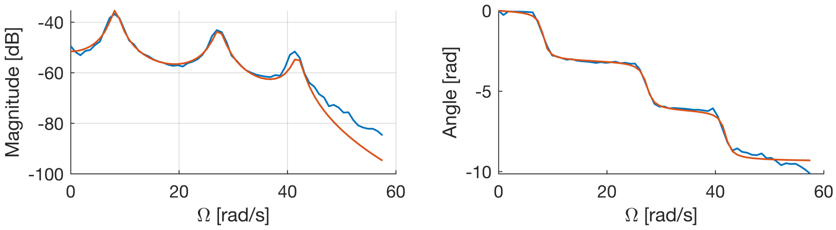

Figure 8: H_3(blue) vs $G_{3_{opt}}$(red).

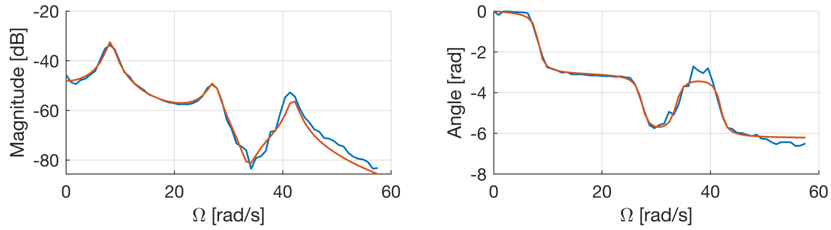

Figure 9: H_2(blue) vs $G_{2_{opt}}$(red).

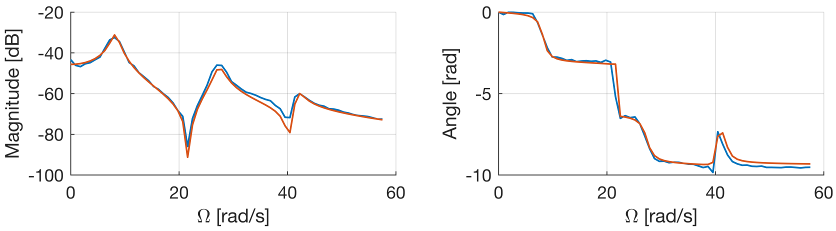

Figure 10: H_1(blue) vs $G_{1_{opt}}$(red).

4 Systems step response and comparison

4.1 Fitted models step response

First of all a new analog frequency vector is created to evaluate the fitted models up to the Nyquist frequency (it has been used the same sampling frequency of the provided data).
Then the step signal is defined as:

$$step(n) = \begin{cases} 1 & \text{if } 0 \leq n \leq 1000 \\ 0 & \text{otherwise} \end{cases}$$

At the end the step response of the signal is obtained with the convolution between the step an the ifft of the three systems.

4.2 Experimental step response

The step response is loaded from the file `data_steps.mat`, then the signals are re scaled as done before. Then the signals are normalized by a factor `k` equal to the inverse of the maximum amplitude of the re scaled input signal. At the end the average experimental step response is obtained splitting the signals every step (considering also the release step) and then is computed the average.

4.3 Final comparison

In the following images are reported the step responses for the first five seconds.

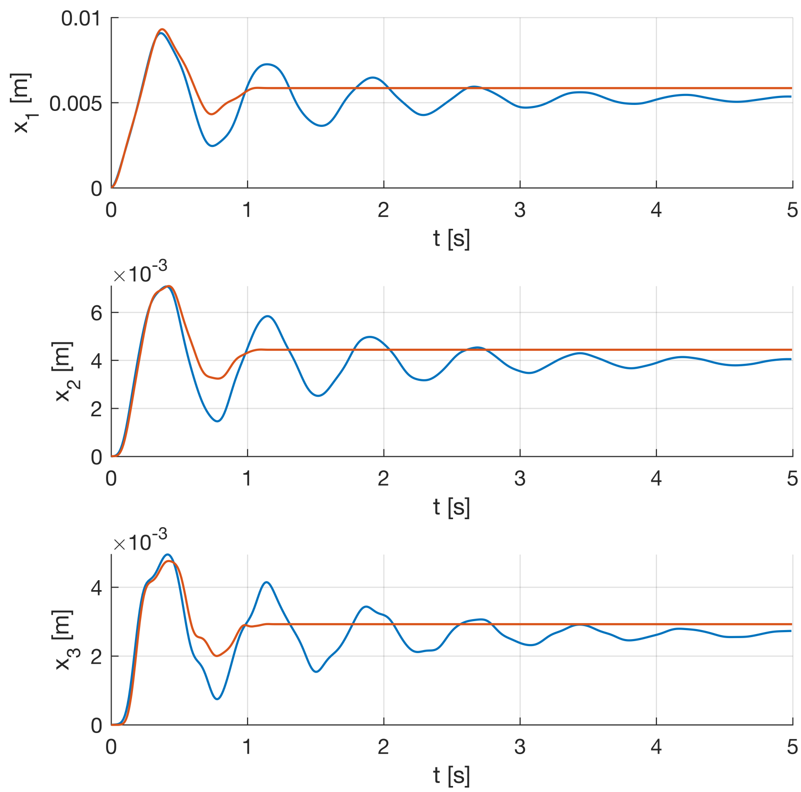

Figure 11: Model (blue) and experimental (red) step response.

The model step response follows quite well the experimental step response for the first oscillation, after which the two trajectories differ quite a lot. This is probably at the limits of the linear model to describe the analyzed system. After the first oscillation it seems that damping effects not described by the model become dominant, moreover the difference of the two curves at steady state is probably due to static friction.

5 Welch modified periodogram inspired method

To obtain a better estimation of the frequency response of the systems, the signals are divided into overlapping by 50% segments, and then the frequencies responses obtained are averaged together. The implementation of this method is shown in the following steps.

- Merge the data splitted before to get the complete signal.
- 7 overlapping windows showed the best performance.
- Perform circular shifting to have a better performance, in particular the signal of the force creates problems, such as high peaks at high frequencies, if the non ideal impulse is located at the center or at the extremities of a window. Considering the number of windows, the signal is shifted by 1/5 of its length.
- The signals are divided into overlapping (50%) segments.
- Each segment is windowed with a hamming window
- Then the fft of the windowed signals is computed.
- It is calculated the amplitude correction factor of the windows, which is defined as $k_f = \sum_{n=0}^{N-1} f(n) \left(\sum_{n=0}^{N-1} f(n)w(n) \right)^{-1}$. For the masses displacement signals it can be calculated as $k_x = N \left(\sum_{n=0}^{N-1} w(n) \right)^{-1}$, because there are lots of nonzero samples.
- It is calculated $H_{mn}(e^{j\omega}) = X_{mn}(e^{j\omega})(k_f F_{mn}(e^{j\omega}))^{-1}$, where m, n indicate the signal and the segment, k_f is the amplitude correction factor.
- The segments are averaged together and then is applied the amplitude correction factor k_x.