Frequency Domain Identification:
The taming of the condition number

Adhemar Bultheel
Department Computer Science
Numerical Approximation and Linear Algebra Group (NALAG)
K.U.Leuven, Belgium
adhemar.bultheel@cs.kuleuven.ac.be
http://www.cs.kuleuven.ac.be/∼ade/



ICCoS Meeting, VUB, April 26, 2001

Systems

We consider discrete time systems



$$\left.\begin{array}{l} u \text{ is input} \rightarrow U(z) = \sum_k u_k z^{1-k} \\ y \text{ is output} \rightarrow Y(z) = \sum_k y_k z^{1-k} \\ G \text{ transfer function} \end{array}\right\} \quad Y(z) = G(z)U(z)$$

Frequency Domain Identification

Given measurements $\{G(z_j)\}$, (or $\{Y(z_j), U(z_j)\}$) $z_j \in \mathbb{T}$
with variance $\{\sigma_j\}$ on these measurements.
Find estimate $\hat{G}$ such that

$$\min \|G - \hat{G}\|_W = \min \left[\sum_j W_j |G(z_j) - \hat{G}(z_j)|^2 \right]^{1/2}, \quad W_j = \sigma_j^{-2},$$

Frequency Domain Identification(2)

or

$$\begin{aligned} \|Y - \hat{Y}\| &= \|(G - \hat{G})U\| = \|G - \hat{G}\|_W \\ &= \left[\sum_j W_j |G(z_j) - \hat{G}(z_j)|^2\right]^{1/2}, \quad W_j = |U(z_j)|^2/\sigma_j^2 \end{aligned}$$

or

$$\begin{aligned} \|G - \hat{G}\|_{\sigma^{-2}} &= \left\|\frac{Y}{U} - \frac{B}{A}\right\|_{\sigma^{-2}} \\ &= \left\|\frac{YA - BU}{AU}\right\|_{\sigma^{-2}} = \|YA - UB\|_W, \quad W = 1/|\sigma AU|^2 \end{aligned}$$

Numerical problems

Many methods have been proposed.
Serious numerical problems when the number of data points and/or the order of the system is high.

$$A(z) = \sum_k a_k \phi_k(z), \quad B(z) = \sum_k b_k \psi_k(z)$$

The condition number of the Jacobian $([\phi_k(z_j) \mid \psi_k(z_j)])$ grows beyond proportion.
Solutions have been proposed by writing the numerator and denominator as a linear combination of orthogonal polynomials.
Orthogonality should be adapted to the problem.
What orthogonality is required in this case?

Vector Polynomial method

$$\min \left\| \frac{Y}{U} - \frac{B}{A} \right\|^2_{\sigma^{-2}} = \min \|YA - UB\|^2_{|\sigma AU|^{-2}}$$

The weight depends on the solution.

Estimate A so that we have an estimate for the weight $1/|\sigma AU|$,

Set $w = \left[\frac{Y}{\sigma AU} \quad -\frac{U}{\sigma AU} \right], P = \begin{bmatrix} A \\ B \end{bmatrix}^T \Rightarrow \min \|P\|^2_W, \; W = w^H w$

$$\min \|P\|^2_W = \min \sum_{j=0}^{N} P(z_j)^H W_j P(z_j), \quad W_j = w(z_j)^H w(z_j)$$

Hence . . .

Vector/Block orthogonal polynomials

. . . need orthogonality with respect to an inner product

$$\langle f, g \rangle_W = \sum_j f(z_j)^H W_j g(z_j) = \left\{ \begin{array}{l} \Sigma(- \square \, \mathsf{I}) \\ \Sigma(\square \, \square \, \square) \end{array} \right. , \quad W_j \in \mathbb{C}^{2\times 2}$$

Write

$$P = \sum_k \psi_k p_k, \quad \psi_k \in \Pi_k^{2\times 2}, \quad p_k \in \mathbb{C}^{2\times 1}$$

with the ψ_k block orthogonal polynomials $\square \, \square \, \square$
Appropriate column degrees $\Rightarrow$
columns of ψ_k are vector orthogonal polynomials $- \square \, \mathsf{I}$

$$\mathbf{\Psi} = [\psi_0(\mathbf{z}) \cdots \psi_n(\mathbf{z})] \in \mathbb{C}^{2(N+1)\times 2(n+1)}, \quad \mathbf{p} = [p_0 \cdots p_n]^T \Rightarrow P(\mathbf{z}) = \mathbf{\Psi p}.$$

The solution

Set $\mathbf{W} = \mathrm{diag}(W_j) \in \mathbb{C}^{2(N+1)\times 2(N+1)}$ and recall $P(\mathbf{z}) = \mathbf{\Psi p} \Rightarrow$

$$\min \sum_j P(z_j)^H W_j P(z_j) = \min P(\mathbf{z})^H \mathbf{W} P(\mathbf{z}) = \min \mathbf{p}^H \mathbf{\Psi}^H \mathbf{W} \mathbf{\Psi p}.$$

Orthogonal block polynomial $\mathbf{\Psi}^H \mathbf{W} \mathbf{\Psi} = \mathbf{I} \Rightarrow$

$$\min \mathbf{p}^H \mathbf{\Psi}^H \mathbf{W} \mathbf{\Psi p} = \min \mathbf{p}^H \mathbf{p} \Rightarrow P = \psi_n p_n, \quad p_n \neq 0$$

Assume order is $2n$ + appropriate order structure $\Rightarrow p_n = [0\ 1]^T$.
Thus the remaining problem is:

How to construct the vector orthogonal polynomials?

The Hessenberg matrix

Always:

$$z\psi_{k-1}(z) = \psi_0(z)\eta_{0k} + \cdots + \psi_k(z)\eta_{kk}, \quad k = 1, \ldots, n+1$$

Thus with

$$\mathbf{H} = \begin{bmatrix} \eta_{01} & \cdots & \eta_{0n} & \eta_{0,n+1} \\ \eta_{11} & \cdots & \eta_{1n} & \eta_{1,n+1} \\ & \ddots & \vdots & \vdots \\ & & \eta_{nn} & \eta_{n,n+1} \end{bmatrix}$$

$$z[\psi_0(z) \cdots \psi_n(z)] = [\psi_0(z) \cdots \psi_n(z)]\mathbf{H} + [0 \ \cdots \ 0 \ \psi_{n+1}(z)\eta_{n+1,n+1}]$$

Then with $\tilde{\mathbf{Z}} = \mathrm{diag}(z_0 I_2, \ldots, z_N I_2) \in \mathbb{C}^{2(N+1)\times 2(N+1)}$

$$\begin{array}{lcl} \tilde{\mathbf{Z}}[\psi_0(\mathbf{z}) \cdots \psi_n(\mathbf{z})] & = & [\psi_0(\mathbf{z}) \cdots \psi_n(\mathbf{z})]\mathbf{H} + [\mathbf{0} \ \cdots \ \mathbf{0} \ \psi_{n+1}(\mathbf{z})\eta_{n+1,n+1}] \\ \tilde{\mathbf{Z}} \qquad \mathbf{\Psi} & = & \mathbf{\Psi} \qquad \mathbf{H} + [\mathbf{0} \ \cdots \ \mathbf{0} \ \psi_{n+1}(\mathbf{z})\eta_{n+1,n+1}] \end{array}$$

Similarity

$$\tilde{\mathbf{Z}}\mathbf{\Psi} = \mathbf{\Psi}\mathbf{H} + [0 \ \cdots \ 0 \ \psi_{n+1}(\mathbf{z})\eta_{n+1,n+1}]$$

Note that $\mathbf{\Psi}$ contains BOP, but is not an orthogonal matrix.
But $\mathbf{Q} = \mathbf{w}\mathbf{\Psi}$ is with $\mathbf{w} = \mathrm{diag}(w_1, \ldots, w_N) \in \mathbb{C}^{(N+1)\times 2(N+1)}$
(recall $W_j = w_j^H w_j$, thus $\mathbf{W} = \mathbf{w}^H\mathbf{w}$ and $\mathbf{Q}^H\mathbf{Q} = \mathbf{\Psi}^H\mathbf{W}\mathbf{\Psi} = \mathbf{I}$)
Multiply by $\mathbf{w}$ and use $\mathbf{w}\tilde{\mathbf{Z}} = \mathbf{Z}\mathbf{w}$, $\mathbf{Z} = \mathrm{diag}(z_1, \ldots, z_N)$

$$\mathbf{Z}\mathbf{Q} = \mathbf{Q}\mathbf{H} + [0 \ \cdots \ 0 \ \mathbf{w}\psi_{n+1}(\mathbf{z})\eta_{n+1,n+1}]$$

Thus

$$\mathbf{Q}^H \cdot \Rightarrow \mathbf{Q}^H\mathbf{Z}\mathbf{Q} = \mathbf{H}$$

$\mathbf{H}$ is unitary Hessenberg: $\mathbf{H}^H\mathbf{H} = \mathbf{Q}^H\mathbf{Z}\mathbf{Q}\mathbf{Q}^H\mathbf{Z}^H\mathbf{Q} = \mathbf{I}$

Furthermore. . .

$$\mathbf{Q}^H \begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_N \end{bmatrix} = \mathbf{\Psi}^H \mathbf{w}^H \mathbf{w} \begin{bmatrix} I_2 \\ I_2 \\ \vdots \\ I_2 \end{bmatrix} = \mathbf{\Psi}^H \mathbf{W} \psi_0(\mathbf{z}) \psi_0^{-1} = \begin{bmatrix} \eta_{0,0} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

because $I_2 = \psi_0 \psi_0^{-1}$, so that by orthonormality $\eta_{0,0} = \psi_0^{-1}$.
Thus

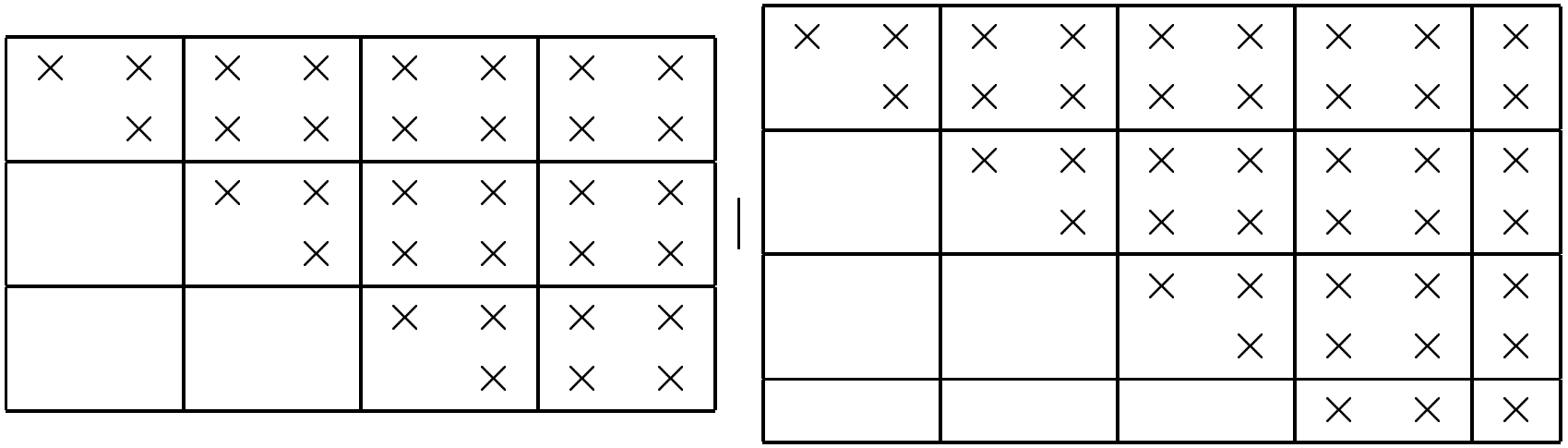
$$\mathbf{Q}^H \left[\begin{array}{c|c} w_0 & \\ \vdots & \mathbf{Z} \\ w_N & \end{array}\right] \left[\begin{array}{c|c} I_2 & \\ \hline & \mathbf{Q} \end{array}\right] = \left[\begin{array}{c|cccc} \eta_{00} & \eta_{01} & \cdots & \eta_{0n} & \eta_{0,n+1} \\ & \eta_{11} & \cdots & \eta_{1n} & \eta_{1,n+1} \\ & & \ddots & \vdots & \vdots \\ & & & \eta_{nn} & \eta_{n,n+1} \end{array}\right]$$

Hessenberg matrix

$$\left[\begin{array}{c|ccc} w_0 & z_0 & & \\ w_1 & & z_1 & \\ & & & \ddots \\ w_N & & & z_N \end{array}\right]$$

unitary similarity $\Rightarrow$ transf.

(Block) Upper Hessenberg



Efficient algorithm

Store the (block upper) unitary Hessenberg matrix in factored form:
$\mathbf{H} = G_1 G_2 \cdots G_m$

$$G_k = \begin{bmatrix} I_{2(k-1)} & & & \\ & -\gamma_k & \hat{\sigma}_k & \\ & \sigma_k & \hat{\gamma}_k & \\ & & & I_{\ldots} \end{bmatrix}, \quad \sigma_k, \gamma_k, \hat{\sigma}_k, \hat{\gamma}_k \in \mathbb{C}^{2\times 2}$$

Chasing the elements to (block) upper Hessenberg form by similarity transformations percolating through the product requires operations on 3×3 or 5×5 blocks $\Rightarrow$ fast algorithm.

Szegő-type recurrence

These BOP satisfy Szegő type recursions

$$\begin{aligned} \phi_n \sigma_n &= z\phi_{n-1} + \phi_{n-1}^* \gamma_n \\ \phi_n^* \hat{\sigma}_n^H &= z\phi_{n-1}\gamma_n^H + \phi_{n-1}^* \end{aligned}$$

generically $\sigma_n, \gamma_n \in \mathbb{C}^{2\times 2}$ (block Schur parameters)

$\sigma_k, \gamma_k, \hat{\sigma}_k, \hat{\gamma}_k$ not independent because $\begin{bmatrix} -\gamma_k & \hat{\sigma}_k \\ \sigma_k & \hat{\gamma}_k \end{bmatrix}$ is a unitary matrix.

Block Schur parameters and Hessenberg

These also show up in the block upper Hessenberg matrix:

$$\left[\begin{array}{c|cccc} \sigma_0 & -\gamma_1 & -\hat{\sigma}_1\gamma_2 & -\hat{\sigma}_1\hat{\sigma}_2\gamma_3 & \hat{\sigma}_1\hat{\sigma}_2\hat{\sigma}_3 \\ & \sigma_1 & -\hat{\gamma}_1\gamma_2 & -\hat{\gamma}_1\hat{\sigma}_2\gamma_3 & -\hat{\gamma}_2\hat{\sigma}_3 \\ & & \sigma_2 & -\hat{\gamma}_2\gamma_3 & -\hat{\gamma}_2\hat{\sigma}_3 \\ & & & \sigma_3 & -\hat{\gamma}_3 \end{array}\right]$$

Note σ_k's on the diagonal are upper triangular.

Nonlinear problem

Initialisation: use OVP with $A = 1$ and solve TLS problem

$$\min \|YA - UB\|_{|\sigma U|^{-2}}.$$

Iterate: use last value of A for OVP and solve GTLS.
Once $P = [A \ B]^T$ has been found, use A to modify the weight $w = 1/(\sigma AU)$ and re-iterate.

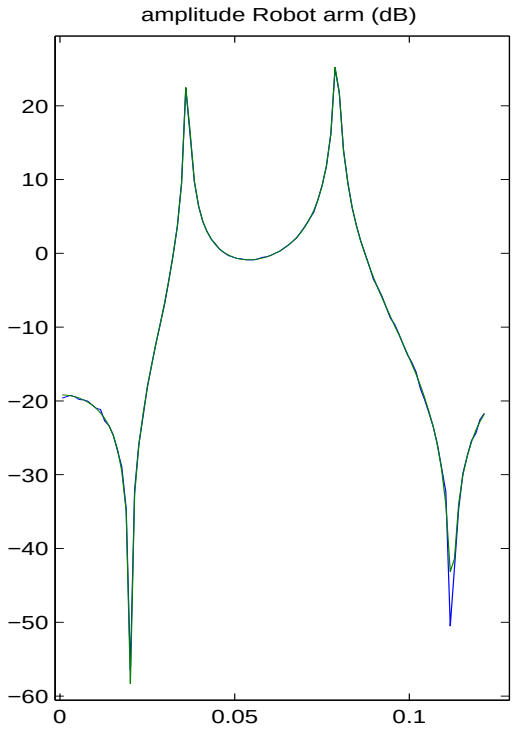
Finally: Do the nonlinear problem: Use nonlinear method to improve the estimate $\sum_{k=0}^{n} \psi_k \lambda_k$. Weight depends on $A \Rightarrow$ loss of orthogonality, but condition number of Jacobian increases only slightly.

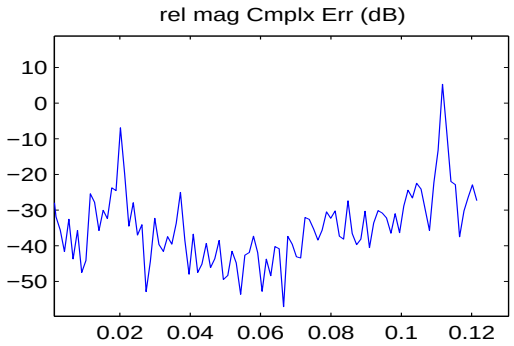
Extras

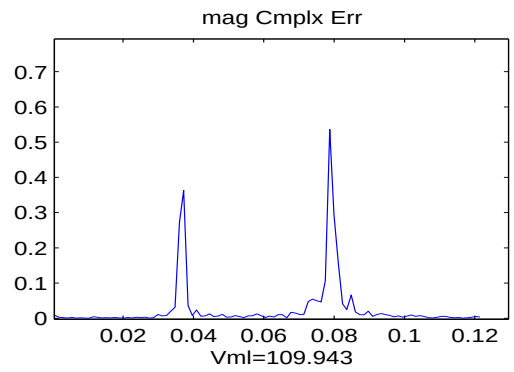
- Similar (but simpler) in the case of continuous time systems
  $\Rightarrow$ (block) tridiagonal instead of (block) Hessenberg
- real computations possible for real problems
- breakdown of the algorithm = exact interpolation
- generalization to MIMO systems possible
- more general degree structure possible

Robot arm

order [6/6]

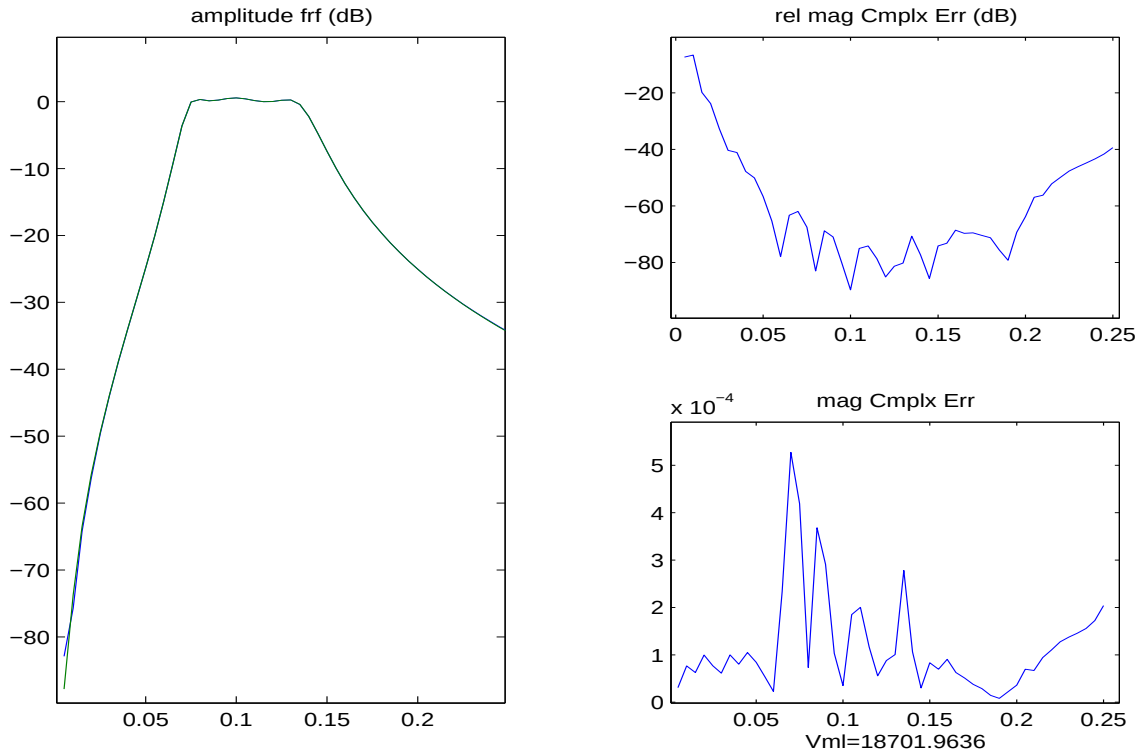






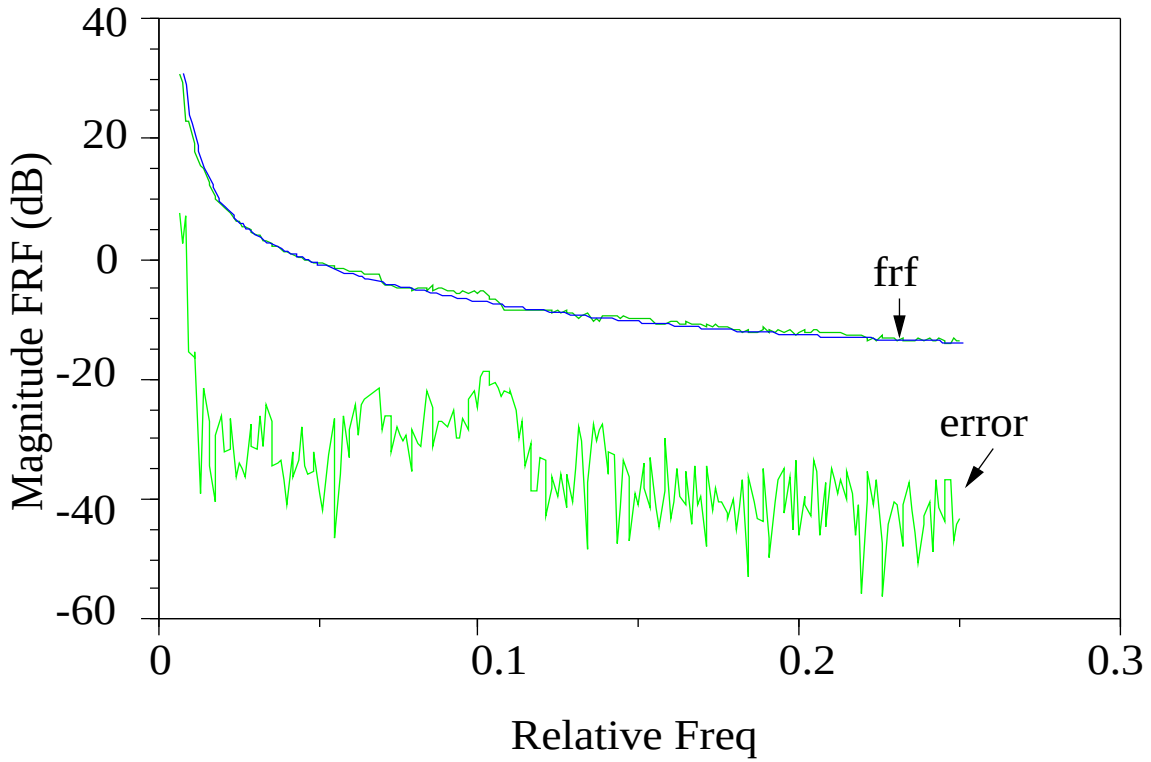
Bandpass filter

order [6/6]



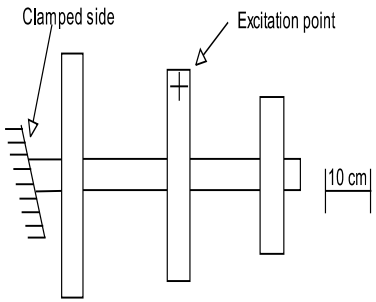
CD's radial servo system

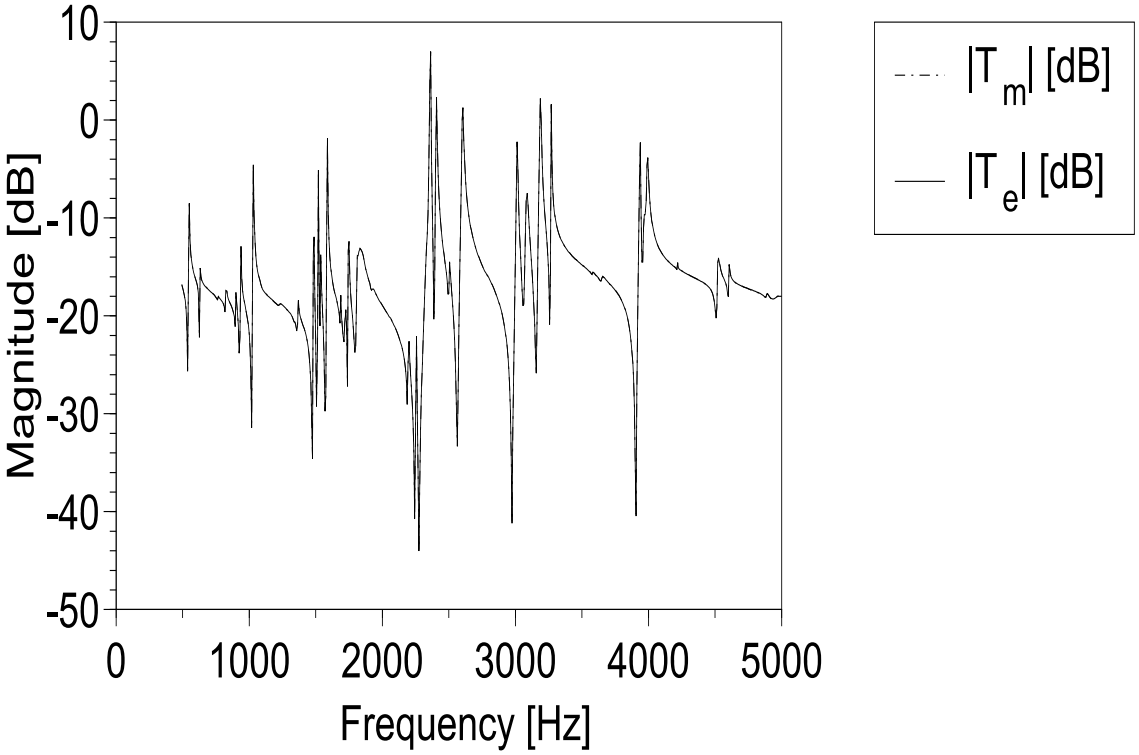
sampling frequency 9.7 kHz, order [5/5]



Mechanical structure

order [120/120]





References OPV

[1] A. Bultheel, M. Van Barel, Y. Rolain. Robust rational approximation for identification, 2001, Submitted

[2] M. Van Barel, A. Bultheel, Discrete linearized least squares approximation on the unit circle, *J. Comput. Appl. Math.* 50 (1994) 965-972.

[3] A. Bultheel, M. Van Barel. Vector orthogonal polynomials and least squares approximation, *SIAM J. Matrix Anal. Appl.* 16 (1995) 863-885.

[4] M. Van Barel, A. Bultheel, Orthogonal polynomial vectors and least squares approximation for a discrete inner product, *Electron. Trans. Numer. Anal.* 3 (1995) 1-23.