OPTIMAL DIGITAL FILTER DESIGN WITH POLE-ZERO SENSITIVITY MINIMIZATION

Gang Li¹ and Michel Gevers²

- 1. School of EEE, Nanyang Technological University, Singapore 2263
- 2. Unite AUTO, Louvain University, B-1348 Louvain la Neuve, Belgium

1. Introduction

Finite Word Length (FWL) effects have been considered as one of the most important issues in digital filter and controller implementations. The performance of a system may be greatly degraded due to these effects. The optimal FWL state-space design has been known as one of the most effective and elegant methods in reducing these effects. It is well known that any linear system can be represented by its state-space model and that this state-space model is not unique. The optimal FWL state-space design is to identify those realizations that minimize the degradation of the filter performance due to the FWL effects. Many classical results in this topic have been collected in [MR87] and a lot of recent developments can be found in [GL92].

The classical results in minimum sensitivity realizations deal with the problem of minimizing some sensitivity measure of the transfer function w.r.t. to the coefficients of all equivalent realizations. In practice it is often the case that it is more important to minimize the errors in the location of certain poles and zeros of a filter w.r.t. the coefficient quantization than the errors in the transfer function. How to minimize the pole and zero sensitivities is the main problem to be addressed in this paper.

The outline of the paper is as follows. In Section 2 we define a global pole-zero sensitivity measure and then formulate our main problem to be dealt with in this paper. In Section 3 we derive the expressions for pole and zero sensitivity functions. The classical pole and zero sensitivity measures are discussed. Our main results in this section is to characterize all the optimal similarity transformations that transform any initial realization into those realizations which have either minimal pole or zero sensitivity.

Section 4 is devoted to finding the optimal realizations minimizing the pole-zero sensitivity measure. A design example is given in Section 5 to show the optimal design procedures. To end this paper, some concluding remarks are given in Section 6.

2. Problem formulation

Consider a discrete-time linear time-invariant filter characterized by its transfer function H(z) and let (A, B, C, d) be a state space realization of this filter, that is

$$x(t+1) = Ax(t) + Bu(t)$$

$$y(t) = Cx(t) + du(t),$$
 (1)

with

$$H(z) = d + C(zI - A)^{-1}B.$$
 (2)

If we denote by $\lambda(M)$ the eigenvalue set of a matrix M, then the poles $\{\lambda_k\}$ of H(z) are the eigenvalues of A, $\{\lambda_k\} = \lambda(A)$, while the zeros $\{v_k\}$ of H(z) are the eigenvalues of

$$Z \triangleq A - d^{-1}BC \tag{3}$$

provided $d \neq 0$. That is $\{v_k\} = \lambda(Z)$ (see [DW78]). Throughout this paper, our zero sensitivity analysis will be limited to the case of systems without delay, i.e. $d \neq 0$.

When the parameters in the matrices A, B, C, d of the state space realization are implemented with error (as is the case in a FWL implementation of the filter) this produces an error in the poles and zeros of the system. The amount of derivation of a pole is approximately proportional to the sensitivity of this pole with respect to each parameter in (A, B, C, d). The same argument applies to a zero. Let $\{\frac{\partial \lambda_k}{\partial A}\}$ and $\{\frac{\partial v_k}{\partial A}, \frac{\partial v_k}{\partial B}, \frac{\partial v_k}{\partial C}, \frac{\partial v_k}{\partial d}\}$ be the sensitivity functions

of individual poles and zeros, respectively, with respect to the matrices A, B, C, d, we shall then adopt the following global pole-zero sensitivity measure, denoted by M_{pz} :

$$M_{pz} = \sum_{k=1}^{n} \{w_{\lambda_k} || \frac{\partial \lambda_k}{\partial A} ||_F^2 + w_{v_k} [|| \frac{\partial v_k}{\partial A} ||_F^2 + || \frac{\partial v_k}{\partial B} ||_F^2 + || \frac{\partial v_k}{\partial C} ||_F^2 + || \frac{\partial v_k}{\partial d} ||_F^2]\}$$

$$(4)$$

where $\{w_{\lambda_k}, k=1,\ldots,n\}$ and $\{w_{v_k}, k=1,\ldots,n\}$ are nonnegative weightings that reflect the relative importance that the designer may care to attach to the sensitivity of individual poles and/or zeros in a sensitivity minimization problem.

It is well known that the realizations (A, B, C, d) satisfying (2) are not unique, they form a realization set, denoted by S_H . In fact, $(T^{-1}AT, T^{-1}B, CT, d)$ is also a realization of H(z) for any real nousingular T. We will show later that the measure M_{pz} depends on the choice of realizations. One of our main contributions in this paper will be to solve the optimal realization problem for the minimization of the polezero sensitivity measure M_{pz} . The computation of M_{pz} in (4) requires all the sensitivity functions of the poles $\{\lambda_k\}$ and of the zeros $\{v_k\}$ concerned in (4). We therefore start our analysis with a derivation of these sensitivity functions in the next section.

3 Pole and zero sensitivities

3.1 pole sensitivity and minimization

The computation of pole sensitivity functions $\{\frac{\partial \lambda_k}{\partial A}\}$ is well known and can be found in e.g. [SS90], [SW84] and [GL92]. These results are based on the following theorem.

Theorem 1: Let $M \in \Re^{n \times n}$ have only simple eigenvalues $\{\lambda_k\} = \lambda(M)$, x_k be a right eigenvalue of M with x_k corresponding to λ_k . Denote $X \triangleq \begin{pmatrix} x_1 & x_2 & \dots & x_n \end{pmatrix}$ and $Y = \begin{pmatrix} y_1 & y_2 & \dots & y_n \end{pmatrix} \triangleq X^{-H}$. Then

$$\left(\frac{\partial \lambda_k}{\partial M}\right)^T = x_k y_k^H, \quad \forall k \tag{5}$$

where y_k is called the reciprocal left eigenvector corresponding to x_k and 'H' denotes the transpose and conjugate operation.

Proof: The result can be proved in several ways. A self-contained derivation can be found in [GL92]. ■

Comment: In the above theorem, it is assumed that

M has only simple eigenvalues. If λ_k is a repeated eigenvalue of M, there exists at least an element of M, say M(i,j), such that $\frac{\partial \lambda_k}{\partial M(i,j)}$ is infinite [GL92]. In the sequel, it is assumed that the filter has only simple poles.

Define $\Psi_k \triangleq ||\frac{\partial \lambda_k}{\partial M}||_F^2$ as the partial eigenvalue sensitivity measure for $\lambda_k(M)$, where $||.||_F$ denotes the Frobenius norm:

$$||M||_F = (\sum_{i,j=1}^{m,n} |M(i,j)|^2)^{1/2}$$

for any matrix $M \in \mathcal{C}^{m \times n}$. Noting that $||M||_F^2 = tr(MM^H) = tr(M^HM)$ where tr(.) denotes the trace operation, we have

$$\Psi_k = tr\{(y_k x_k^H)(y_k x_k^H)^H\}.$$
 (6)

Now, we define the overall eigenvalue sensitivity measure of a matrix $M \in \Re^{n \times n}$ as follows:

$$\Psi(M) \triangleq \sum_{k=1}^{n} \Psi_k. \tag{7}$$

The eigenvalue sensitivity measure as just defined has the following properties:

Theorem 2: Let $M \in \mathbb{R}^{n \times n}$ have distinct eigenvalues; then the overall eigenvalue sensitivity measure is not smaller than n, that is $\Psi(M) \geq n$, and equality holds if and only if M is normal (that is $M^H M = M M^H$).

Proof: A proof of this theorem can be found in [SW84] and [GL92].

All these results above apply to the pole sensitivity problem with M=A.

3.2 zero sensitivity and minimization

We recall that for a realization (A, B, C, d) with $d \neq 0$ the zeros of the system are the eigenvalues of $Z \triangleq A - d^{-1}BC$. Hence they depend not only on the matrix A, but also on B, C and d.

Theorem 3: Let (A, B, C, d) be a realization of H(z). Denote $Z = A - d^{-1}BC$ and $\{v_k\} = \lambda(Z)$, then

$$\frac{\partial v_k}{\partial A} = \frac{\partial v_k}{\partial Z}, \qquad \frac{\partial v_k}{\partial d} = d^{-2}B^T \frac{\partial v_k}{\partial Z}C^T
\frac{\partial v_k}{\partial B} = -\frac{\partial v_k}{\partial Z}(d^{-1}C^T), \frac{\partial v_k}{\partial C} = -d^{-1}B^T \frac{\partial v_k}{\partial Z}. (8)$$

Proof: See the complete version of this paper.

We assume again that Z has only simple eigenvalues, and we introduce the full rank matrix $X_z \triangleq [x_z(1), \ldots, x_z(n)]$ of right eigenvectors of Z and the matrix $Y_z \triangleq X_z^{-H} = \{y_z(1), \ldots, y_z(n)\}$ of reciprocal left eigenvectors.

We define the sensitivity measure of an arbitrary zero v_k of the system (A, B, C, d) as

$$\Psi_{zk} = ||\frac{\partial v_k}{\partial A}||_F^2 + ||\frac{\partial v_k}{\partial B}||_F^2 + ||\frac{\partial v_k}{\partial C}||_F^2 + ||\frac{\partial v_k}{\partial d}||_F^2. \tag{9}$$

Using expressions (8) and (6) we obtain

$$\Psi_{zk} = tr\{(y_z(k)x_z^H(k))(y_z(k)x_z^H(k))^H\} + tr(\alpha_k^2 y_z(k)y_z^H(k)) + tr(\beta_k^2 x_z(k)x_z^H(k)) + \alpha_k^2 \beta_k^2,$$
(10)

where

$$\alpha_k^2 \triangleq |d^{-1}x_z^H(k)C^T|^2 = |d^{-1}Cx_z(k)|^2, \beta_k^2 \triangleq |d^{-1}B^Ty_z(k)|^2.$$
(11)

The overall unweighted zero sensitivity measure is then given by

$$\Psi_z = \sum_{k=1}^n \Psi_{zk}.\tag{12}$$

Theorem 4: Let (A, B, C, d) be a minimal realization of a system H(z) with $d \neq 0$ and distinct zeros. The unweighted sensitivity measure (12) is lower bounded by

$$\Psi_z \ge n + 2\sum_{k=1}^n |\alpha_k \beta_k| + \sum_{k=1}^n \alpha_k^2 \beta_k^2.$$
 (13)

This lower bound is achieved if and only if Z is normal with its right eigenvector matrix X_z satisfying

$$X_z^H X_z = diag(|\frac{\alpha_1}{\beta_1}|, \dots, |\frac{\alpha_n}{\beta_n}|). \tag{14}$$

Proof: See the complete version of this paper.

3.3 optimal realizations

Consider a matrix M^0 that has a complete set of independent eigenvectors and any nonsingular matrix T, and denote $M = T^{-1}M^0T$. Clearly, $\lambda(M) = \lambda(M^0)$. Let x_k^0 be a right eigenvector of M^0 corresponding to the eigenvalue λ_k and let y_k^0 be its reciprocal left eigenvector, i.e. $y_k^{0H}x_k^0 = 1$. The corresponding eigenvectors of M for the same eigenvalue λ_k are:

$$x_k = T^{-1} x_k^0, \quad y_k = T^T y_k^0.$$
 (15)

It then follows from (6) that the eigenvalue sensitivity measure in the new coordinate system is given by

$$\Psi_k(T) = tr\{ (T^T y_k^0 x_k^{0H} T^{-T}) (T^T y_k^0 x_k^{0H} T^{-T})^H \},$$
(16)

which shows that similar matrices typically have different eigenvalue sensitivity measures.

With M replaced by the state transition matrix A, one can study the pole sensitivity problem. We denote by Ψ_{pk} the partial sensitivity measure of the k-th pole of the realization (A,B,C,d) of a system. The overall unweighted pole sensitivity measure Ψ_p of the realization (A,B,C,d) is defined as

$$\Psi_p \triangleq \sum_{k=1}^n \Psi_{Pk} = \Psi(A). \tag{17}$$

Now, let (A_0, B_0, C_0, d) be some initial realization of a transfer function H(z) and let (A, B, C, d) be obtained from (A_0, B_0, C_0, d) through a similarity transformation T. Let $\{x_p^0(k)\}$ and $\{y_p^0(k)\}$ for $k = 1, \ldots, n$ be the left and right eigenvectors of A_0 . The pole sensitivity measure in the new realization is a function of T obtained directly from (16):

$$\Psi_{pk}(T) = tr\{PH_p(k)P^{-1}H_p^H(k)\}, \qquad (18)$$

where $H_p(k) \triangleq y_p^0(k) x_p^{0H}(k)$ and $P \triangleq TT^T$.

Similarly, substituting $x_z(k) = T^{-1}x_z^0(k)$ and $y_z(k) = T^Ty_z^0(k)$ in (10), yields the following expression for the zero sensitivity measure:

$$\Psi_{zk}(T) = Ir\{PH_z(k)P^{-1}H_z^H(k)\} + tr\{\alpha_k^2 PH_y(k)\} + Ir\{\beta_k^2 P^{-1}H_x(k))\} + \alpha_k^2 \beta_k^2,$$
(19)

where $H_z(k) \triangleq y_z^0(k)x_z^{0H}(k)$, $H_x(k) \triangleq x_z^0(k)x_z^{0H}(k)$ and $H_y(k) \triangleq y_z^0(k)y_z^{0H}(k)$. The numbers $\{\alpha_k^2\}$ and $\{\beta_k^2\}$ have been defined in (11). It follows from $C = C_0T$, $B = T^{-1}B_0$ and (15) that these numbers are coordinate independent.

Theorem 2 and 4 show that the realizations whose A and $Z = A - d^{-1}BC$ are normal have their poles and zeros least sensitive to errors in the realization parameters. The question arises as to whether a system H(z) can always be represented in a coordinate space in which the matrix A or Z is in normal form. Stated more simply, can any matrix $M \in \mathbb{R}^{n \times n}$ be transformed to normal form by a real similarity transformation?

Theorem 5: Let (A_0, B_0, C_0, d) be a minimal realization of H(z). Denote $Z_0 = A_0 - d^{-1}B_0C_0$.

Let $X_p^0 \triangleq (x_{p1}^0, \dots, x_{pn}^0)$ and $X_z^0 \triangleq (x_{z1}^0, \dots, x_{zn}^0)$ be a matrix of right eigenvalues of A_0 and Z_0 , respectively. If H(z) has no repeated poles, there exist a set of similarity transformations T_p such that $A = T_p^{-1}A_0T_p$ is normal, and this set is completely characterized by

$$T_p = (X_p^0 D_p X_p^{0H})^{1/2} Q, (20)$$

where D_p is any diagonal positive-definite matrix and Q is an arbitrary orthogonal matrix.

If H(z) has no repeated zeros, there exist a set of similarity transformations T_z such that $Z = T_z^{-1}Z_0T_z$ is normal and satisfies (14), and

this set is completely characterized by

$$T_z = (X_z^0 D_z X_z^{0H})^{1/2} Q_z, \tag{21}$$

where $D_z = diag(\lfloor \frac{\beta_1}{\alpha_1} \rfloor, \ldots, \lfloor \frac{\beta_n}{\alpha_n} \rfloor)$ with $\{\alpha_k\}$ and $\{\beta_k\}$ as defined before, and Q is an arbitrary orthogonal matrix.

Proof: See the complete version of this paper.

For a realization (A, B, C, d) of H(z), generally the normality of A can not guarantee that of $Z = A - d^{-1}BC$, vice-versa. This means that the minimal pole and zero sensitivities can not simultaneously be achieved. A system is uniquely determined by its poles and zeros to within a constant multiplier. The behavior of the system depends on that of the poles as well as of the zeros. That is why we adopt a combined weighted pole-zero sensitivity measure that takes into account of pole sensitivity as well as zero sensitivity. In the next section, we will study this measure.

4. Pole-zero sensitivity minimization

It follows from (18) and (19) that the weighted measure M_{pz} introduced in (4) can now be written as:

$$M_{pz} = \sum_{k=1}^{2n} tr(PH_kP^{-1}H_k^H) + tr(PM_y) + tr(P^{-1}M_x) + c \triangleq R(P)$$
 (22)

where $P = TT^T$ and

$$H_{k} = w_{\lambda_{k}}^{1/2} y_{p}^{0}(k) x_{p}^{0H}(k), \quad k = 1, ..., n$$

$$= w_{\nu_{k}}^{1/2} y_{z}^{0}(k-n) x_{z}^{0H}(k-n), \quad k = n+1, ..., 2n$$

$$M_{y} = \sum_{k=1}^{n} w_{\nu_{k}} \alpha_{k}^{2} y_{z}^{0}(k) y_{z}^{0H}(k)$$

$$= Y_z^0 diag(w_{v_1}\alpha_1^2, \dots, w_{v_n}\alpha_n^2) Y_z^{0H}$$

$$M_x = \sum_{k=1}^n w_{v_k} \beta_k^2 x_z^0(k) x_z^{0H}(k)$$

$$= X_z^0 diag(w_{v_1}\beta_1^2, \dots, w_{v_n}\beta_n^2) X_z^{0H}$$

$$c = \sum_{k=1}^n w_{v_k} \alpha_k^2 \beta_k^2.$$
(23)

So, the minimization problem can be formulated as

$$\min_{\substack{(A,B,C,d)\in S_H}} M_{pz} \Longleftrightarrow \min_{P>0} R(P). \tag{24}$$

This problem is solved by the following theorem.

Theorem 6: Let H(z) be a discrete time transfer function with a set of distinct poles and zeros, and let (A_0, B_0, C_0, d_0) be a minimal realization of H(z). Furthermore, let $X_p^0 \triangleq [x_p^0(1) \dots x_p^0(n)]$ and $X_z^0 \triangleq [x_z^0(1) \dots x_z^0(n)]$ be right eigenvector matrices of A_0 and $Z_0 \triangleq A_0 - d_0^{-1}B_0C_0$, respectively, let $Y_p^0 \triangleq [y_p^0(1) \dots y_p^0(n)]$ and $Y_z^0 \triangleq [y_z^0(1) \dots y_z^0(n)]$ be the matrices of reciprocal left eigenvectors and let $\{w_{v_k} > 0\}$ for $k = 1, \dots, n$. Then, with $P = TT^T$, the minimum of (22) exists and is achieved by a nonsingular P. The minimum of R(P) is unique and is obtained as the solution of the following equation:

$$P[M_y + \sum_{k=1}^{2n} H_k P^{-1} H_k^H] P = M_x + \sum_{k=1}^{2n} H_k^H P H_k$$
 (25)

Comment: The optimal solution to the minimization of $M_{P^{\pm}}$ is given as the unique solution P of (25). This means that the optimal realization problem has a set of optimal solutions characterized by

$$\{A, B, C, d = T_{opt}^{-1} A_0 T_{opt}, T_{opt}^{-1} B_0, C_0 T_{opt}, d_0\}$$
 (26)

where $T_{opt} = \tilde{T}V$, \tilde{T} is any square root of P ($P = \tilde{T}\tilde{T}^T$, and V is any orthogonal matrix.

It seems difficult to have an explicit expression of the solutions to (25). One can always obtain the solutions by an iterative procedure using a gradient algorithm:

$$P(k+1) = P(k) - \mu \frac{\partial R(P)}{\partial P}|_{P=P(k)}$$
 (27)

where μ is the step-size and

$$\frac{\partial R(P)}{\partial P} = \sum_{k=1}^{2n} \{ H_k P^{-1} H_k^H P^{-1} H_k^H P H_k P^{-1} \} + M_y - P^{-1} M_x P^{-1}.$$
 (28)

Since there is no local minimum (or maximum either), this algorithm will converge to the unique solution to the equation (25) as long as μ is small enough.

5. Numerical example

We now illustrate our theoretical results with a numerical example. We consider a filter that is initially described in direct form R_r :

$$A_{c} = \begin{pmatrix} 3.6147 & -4.9242 & 2.9971 & -0.6881 \\ 1.0000 & 0 & 0 & 0 \\ 0 & 1.0000 & 0 & 0 \\ 0 & 0 & 1.0000 & 0 \end{pmatrix}$$

$$C_{c} = \begin{pmatrix} 0.5640 & 0.0588 & 0.5284 & 0.0179 \end{pmatrix} \times 10^{-3}$$

$$B_{c} = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix}^{T}, d = 0.1578.$$

The corresponding poles and zeros are given, respectively , by the following vectors V_p and V_z : $V_p = (0.9550 + j0.0953 \ 0.9550 - j0.0953 \ 0.8524 + j0.1432 \ 0.8524 - j0.1432)^T$ and $V_z = (1.0818 + j0.2556 \ 1.0818 - j0.2556 \ 0.7238 + j0.1819)^T$.

Using Theoem 5, we obtained a T_p and a T_z that give two optimal realizations R_p and R_z in terms of pole and zero sensitivities, respectively. We computed the pole and zero sensitivity measures Ψ_p and Ψ_z for R_c , R_p and R_z , respectively. The results are given in the following list:

	Ψ_p	Ψ.
R_c	4.469×10^{6}	9.5477×10^4
R_p	4	3.7684×10^{6}
R_z	70.2677	8.3889

We note that the zero sensitivity measure of R_p , 3.7684×10^6 is far from the minimal value, 8.3889, while the pole sensitivity measure of R_z , 70.2677, is far higher than the minimal value, 4. Thus, these two optimal realizations are significantly different. It therefore makes sense to use our combined measure M_{pz} in order to achieve a tradeoff between the two apparently conflicting design criterion Ψ_p and Ψ_z .

¿From a stability point of view, one would like to implement a realization for which the two poles closest to the unit circle, numbered 1 and 2, have a smaller sensitivity. Their sensitivities in the three realizations examined so far are: $\Psi_p(1) = \Psi_p(2) = 1.6142 \times 10^6$ for R_c , $\Psi_p(1) = \Psi_p(2) = 1$ for R_p , and $\Psi_p(1) = \Psi_p(2) = 23.3233$ for R_z .

We therefore choose the following weighting fac-

tors: $\left(\begin{array}{cccc} w_{\lambda_1} & w_{\lambda_2} & w_{\lambda_3} & w_{\lambda_4} \end{array}\right) = \left(\begin{array}{cccc} 20 & 20 & 1 & 1 \end{array}\right)$ and $\left(\begin{array}{cccc} w_{v_1} & w_{v_2} & w_{v_3} & w_{v_4} \end{array}\right) = \left(\begin{array}{ccccc} 1 & 1 & 1 & 1 \end{array}\right)$. Using algorithm (27), one can find the corresponding optimal realization R_{pz} for which the pole and zero sensitivity measures are 7.4555 and 27.0285, and $\Psi_p(1) = \Psi_p(2) = 1.8564$. Clearly, this realization is a tradeoff of R_p and R_z . These theoretical results have been confirmed by some simulations.

6. Conclusions

We have devoted this paper to the study of pole and zero sensitivities. One of our contributions is to characterize the optimal similarity transformations that transform any initial realization into the one that has a minimal pole or zero sensitivity. By defining a combined weighted pole-zero sensitivity measure, the otimal realization has been formulated. A necessary and sufficient condition that any optimal realization minimizing this measure has to satisfy has been given. An algorithm ffor solving this optimization problem has been proposed. A design example has been given to illustrate the optimal design procedure.

References:

[DW78] E. J. Davison and S. H. Wang, 'An Algorithm for the Calculation of Transmission Zeros of the System (C,A,B,D) Using High Gain Output Feedback', *IEEE Trans. on Automatic Control*, Vol. AC-23, No.4, pp. 738-741, August 1978.

[GL92] Michel Gevers and Gang Li, Parametrizations in Control, Estimation and Filtering Problems: Accuracy Aspects, Springer Verlag New York, Communication and Control Engineering Series, 1992.

[MR87] C.T. Mullis and R.A. Roberts, *Digital Signal Processing*, Addison-Wesley, 1987.

[SS90] G.W. Stewart and Ji-guan Sun, Matrix Pertubation Theory, Academic Press, 1990.

[SW84] R. E. Skelton and D. A. Wagie, 'Minimal root sensitivity in linear systems,' *J. Guidance Contr.*, Vol. 7, pp. 570-574, September-October 1984.

[Wil86] D. Williamson, 'Roundoff noise minimization and pole-zero sens itivity in fixed-point digital filters using residue feedback,' *IEEE Trans. on Acoust.*, Speech, and Signal Processing, Vol. ASSP-34, No.5, pp.1210-1220, Oct. 1986.