

Passive Control of Linear Systems*

Krešimir Veselić[†], Katrin Brabender[‡] and Krešimir Delinić[§]

Abstract. We propose a method for designing optimal damping viscosities of dampers in order to calm down vibrations of a structure with given mass and stiffness parameters. Our method is based on the minimization of the trace of the Lyapunov equation in the underlying phase space equipped with the energy norm. We compare our method with other common approaches.

AMS subject classification: 70J25, 70Q05

Key words: Lyapunov equation, linear systems, damping

1. Introduction

Dangerous vibrations are frequent practical problem in the industry. Particularly delicate are vibrating systems whose mass and stiffness structure cannot be easily modified. This is the case with piping systems in chemical or power plants. Here pure clamping is forbidden due to the danger of thermic deformations. All this becomes even more serious, if the performance of an already operating system should be improved. A possible solution consists of adding viscous damping by building in a number of dashpots. So, the following question is of interest: *given the masses and the stiffnesses determine the available dampers' viscosities so as to insure an optimal evanescence*¹.

In this paper we will propose a method to this end. We suppose that the conditions are such that the linear model can be applied. Here “optimal evanescence” has to be precisely defined. The usual way of doing this is (cf. [8]) to require

$$\max \operatorname{Re} \lambda_k \rightarrow \min, \quad (1)$$

(here λ_k are the phase space complex eigenfrequencies of the system) to be minimal.

*Parts of this paper and in particular the proof of Theorem 1 in Appendix are contained in the Ph.D. thesis of the second author under the supervision of the first author.

[†]Lehrgebiet Mathematische Physik, FernUniversität Hagen, Postfach 940, 58084 Hagen, Germany, e-mail: kresimir.veselic@fernuni-hagen.de

[‡]Lehrgebiet Mathematische Physik, FernUniversität Hagen, Postfach 940, 58084 Hagen, Germany, e-mail: katrin.brabender@fernuni-hagen.de

[§]Siemens AG, Erlangen, Germany

¹Sometimes the optimization includes also the possible positions of the dampers. This will not be considered here, although some of our methods could be applied in this case as well.

Criteria like (1) concern rather the asymptotic behaviour of the system, and it is not *a priori* clear that they will favourably influence the behaviour of the system at finite times, too.

We propose here to minimize the total energy integral

$$\int_0^{\infty} E(t) dt \rightarrow \min. \quad (2)$$

This criterion leads to the use of the Lyapunov theory. More precisely, our aim is to minimize the average of $E(t)$ over all times t as well as over all initial data of unit total energy and a given frequency range. This average has been shown to be just the trace of the solution of the corresponding Lyapunov equation. The advantages of this quantity are:

- (i) its obvious closeness to the total energy of the vibration, and
- (ii) its smoothness as the function of the damping parameters, which allows standard methods of minimization via gradient or Hessian.

Note that this last property is not shared by the spectral penalty functions (1). On the other hand, it has recently been shown that the solution of the Lyapunov equation provides rigorous bounds to the energy decay of a vibrating system ([13, 14]). The program we propose here has been completely realized in [12] in a special case of one dimensional damping. In this particular case we were able to avoid the tedious solving of the corresponding Lyapunov equation, and provide a simple closed formula for the optimal damping parameter. In applications we are dealing with, the rank of the damping matrix (i.e., the number of the dashpots times their degrees of freedom) is greater than one, but still much smaller than the whole dimension of the vibrating system. Here we cannot avoid solving the Lyapunov equation repeatedly, and must see that the computation time does not become prohibitive. We propose three means to cope with this challenge:

- (i) by using the special structure of our penalty function, we derive simple efficient formulæ for the gradient and the Hessian which greatly reduce the number of Lyapunov equation solvings,
- (ii) we apply a frequency cut-off which allows the reduction of the matrix sizes, and
- (iii) we develop a method to determine the “initial guess”, i.e., the starting point for the trace minimization process.

Both (ii) and (iii) are not fully understood yet, but our existing theory and the experiments indicate their reliability. Our choice for the starting point is to take the so-called “modal critical damping” and approximate it — in the sense of least squares — by the allowed viscosities (note that the damping depends linearly on the viscosity parameters). It is believed that the modal critical damping is the global minimum over *all possible dampings*. Thus far, we only know that at this point the penalty function

has a local minimum. This is one of the key theoretical results of this paper. We show that this minimum is global, if we restrict ourselves only to the modal dampings.

This paper is organized as follows. In Section 2 we derive the phase space Lyapunov model and the penalty function. We show that the Lyapunov solution controls the decay of any oscillation. In Section 3 we describe the frequency cut-off and in Section 4 the so-called ‘‘modal cut-off’’, pointing out the unreliability of the latter. In Section 5 we derive efficient formulæ for the gradient and the Hessian of the penalty function, thus allowing computational savings during the minimization. In Section 6 we define the starting point for the minimization process and give theoretical justifications for our choice. This section contains our main theoretical results (Theorems 1, 2 and 3). In Section 7 we compare our penalty function with another one, the spectral abscissa. On examples we show that the two may differ significantly. The Appendix contains a rather lengthy proof of Theorem 1.

The methods presented in this paper have been successfully implemented to optimize some concrete structures in the course of improving the performance and security of a pipe system in a nuclear power plant. The detailed results will be published elsewhere.

2. Mathematical model. Decay criteria.

We consider a damped linear vibrational system described by the differential equation

$$M\ddot{x} + C\dot{x} + Kx = 0, \quad (3)$$

where M , C , K (called mass, damping, stiffness matrix, respectively) are real, symmetric matrices of order n , with M , K positive definite and C positive semidefinite². Often the matrix C describes few dampers, built in in order to calm down dangerous oscillations. An example is the so-called n -mass oscillator or oscillator ladder (Figure 1), where

$$M = \text{diag}(m_1, m_2, \dots, m_n),$$

$$K = \begin{pmatrix} k_0 + k_1 & -k_1 & & & & & & & & & \\ & -k_1 & k_1 + k_2 & & & & & & & & \\ & & & \ddots & & & & & & & \\ & & & & \ddots & & & & & & \\ & & & & & \ddots & & & & & \\ & & & & & & \ddots & & & & \\ & & & & & & & \ddots & & & \\ & & & & & & & & -k_{n-1} & & \\ & & & & & & & & -k_{n-1} & k_{n-1} + k_n & \end{pmatrix},$$

$$C = \sum_k c_k e_k e_k^T, \quad (4)$$

²In some important applications (e.g., with so-called lumped masses in vibrating structures) M , too, is only semidefinite. This case can be easily reduced to the one with a nonsingular M , at least if the null-space of M is contained in the one of C .

where e_k is the k -th canonical basis vector. Here $m_i > 0$ are the masses, $k_i > 0$ the spring constants or stiffnesses, and $c_i > 0$ the damping constant of the damper applied to the mass m_i (in Figure 1 with $k_0 = 0$). Note that in the figure the rank of C is one and we call such damping *one-dimensional*³.

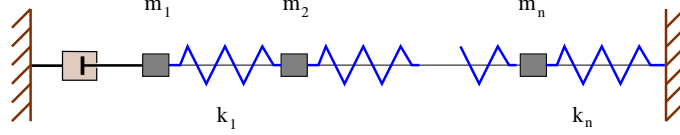


Figure 1. The n -mass oscillator with one damper.

To (3) there corresponds the eigenvalue equation

$$(\lambda^2 M + \lambda C + K)x = 0. \quad (5)$$

All eigenvalues of (5) obviously lie in the left complex plane. We now go over to the $2n$ -dimensional phase space by taking factors

$$K = \Phi_1 \Phi_1^T, \quad M = \Phi_2 \Phi_2^T$$

and setting

$$y_1 = \Phi_1^T x, \quad y_2 = \Phi_2^T \dot{x}.$$

Then (3) is immediately seen to be equivalent to

$$\dot{\psi} = A\psi, \quad (6)$$

$$\psi = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, \quad A = \begin{pmatrix} 0 & \Phi_1^T \Phi_2^{-T} \\ -\Phi_2^{-1} \Phi_1 & -\Phi_2^{-1} C \Phi_2^{-T} \end{pmatrix}, \quad (7)$$

with the solution $\psi = e^{At} \psi_0$, ψ_0 initial data. The factors Φ_1 , Φ_2 may (but need not) be taken as Cholesky factors, i.e., as lower triangular. The eigenvalue problem $A\psi = \lambda\psi$ is obviously equivalent to (5). Moreover,

$$\psi^T \psi = \|\psi\|^2 = x^T K x + \dot{x}^T M \dot{x} = 2E(t).$$

In other words, the Euclidean norm of this phase-space representation equals twice the total energy of the system. From this it follows that *all phase space matrices are unitarily equivalent*. Thus, for all total-energy relevant considerations we may choose any of these representations at our convenience. A further property of any phase matrix A is the so-called “J-symmetry”:

$$A^T = J A J, \quad J = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}, \quad (8)$$

³In general, the rank of C equals the number of the dampers, whereas the range (i.e., the column space) of C determines (and is determined by) their positions.

which is verified directly.

Now, (2) can be written as

$$\psi_0^T X \psi_0 \rightarrow \min, \quad (9)$$

where

$$X = \int_0^\infty e^{A^T t} e^{At} dt$$

solves the Lyapunov equation

$$A^T X + X A = -I, \quad (10)$$

and $\psi = e^{At}\psi_0$ is the solution of (6), with ψ_0 as initial data. The inconvenience of the criterion (9) lies in its dependence on the initial data ψ_0 . Thus, instead of the quantity $\psi_0^T X \psi_0$, we take its mean value over all initial data ψ having the same energy $\|\psi_0\|^2$. Therefore, instead of (9), we require

$$\int_{\|\psi_0\|=1} \psi_0^T X \psi_0 d\sigma \rightarrow \min, \quad (11)$$

where $d\sigma$ is a chosen probability measure on the unit sphere $S^{2n} = \{\psi_0 \in \mathbb{R}^{2n} \mid \|\psi_0\| = 1\}$. As it is easily seen, (11) is equivalent to

$$\text{tr}(ZX) = \text{tr} \left(Z \int_0^\infty e^{A^T t} e^{At} dt \right) \rightarrow \min,$$

where Z is a symmetric positive semidefinite matrix (induced by $d\sigma$) which may be normalized to have unit trace. By using (8) we obtain another interesting formula for the trace

$$\text{tr}(ZX) = \text{tr} Y,$$

where Y is the solution of another Lyapunov equation⁴

$$A^T Y + Y A = -Z_1, \quad Z_1 = JZJ. \quad (12)$$

By using the well-known ‘‘frequency domain formula’’ ([8])

$$X = \frac{1}{2\pi} \int_{-\infty}^{\infty} (-i\omega - A^T)^{-1} (i\omega - A)^{-1} d\omega,$$

we see that averaging over initial conditions is equivalent to averaging over amplitudes f of steady-state solutions

$$x \exp(i\omega t), \quad x = (i\omega - A)^{-1} f,$$

⁴This is a special case of the general properties: if $A^T X + X A = -B$ and $A Y + Y A^T = -Z$ then $\text{tr}(ZX) = \text{tr}(BY)$, $JAJ = A^T$, $\text{tr}(Y) = \text{tr}(JYJ)$.

which are the answer to the harmonic load $f \exp(i\omega t)$. This shows that *in optimizing transient behaviour we also optimize the steady-state one*. Taking the most natural, invariant measure on S^{2n} , we obtain $Z = (1/2n)I$.

As a simple illustration consider the one-mass oscillator, where

$$A = \begin{pmatrix} 0 & \sqrt{k} \\ -\sqrt{k} & -c \end{pmatrix}, \quad k = k_1. \quad (13)$$

Now,

$$\text{tr } X = \frac{2}{c} + \frac{c}{2k}$$

takes its minimum at $c = 2\sqrt{k}$ and nowhere else. This is the known case of critical damping. Note that here the criterion (1) yields the same value of c . For a general system with one dimensional damping⁵ $C = cww^T$, $\|w\| = 1$, one can show ([12]) that

$$X = X_1c + \frac{X_2}{c} + X_3,$$

where X_1 , X_2 can be directly computed from the eigensolution of the undamped problem, i.e., from (5) with $C = 0$. This formula, which is valid whenever C has rank one, leads to an extremely simple minimization of $\text{tr}(ZX)$, which is always a hyperbola in c . This avoids the use of time-consuming algorithms for the Lyapunov equation like the one in [1]. The only eigendecoupling needed is the one for the undamped system, and is done only once ([12]).

As can be directly seen, the optimal A in (13) has no eigenbasis due to nonlinear elementary divisors. This prevents the general use of the eigendecomposition of A to optimize viscous damping⁶. For large systems we have to make a choice of Z which takes into account the fact that we are averaging not just over all state vectors, but over those from a certain frequency range.

To this end we choose our factors Φ_1 , Φ_2 , such that in (7)

$$\Phi_2^{-1}\Phi_1 = \Omega = \text{diag}(\omega_1, \dots, \omega_n)$$

is a diagonal matrix with positive diagonal elements ω_i , which we will suppose to be, say, increasingly ordered. This choice is always possible, in fact, it is obtained by the eigensolution of the undamped system (5) with $C = 0$. Here ω_i are the undamped circular frequencies and the columns of Φ are the mass-normalized eigenforms:

$$K\Phi = M\Phi\Omega^2, \quad \Phi^T M\Phi = I.$$

Then by taking $\Phi_1 = \Phi^{-T}\Omega$, $\Phi_2 = \Phi^{-T}$, we obtain the phase matrix

$$\psi = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, \quad A = \begin{pmatrix} 0 & \Omega \\ -\Omega & -\hat{C} \end{pmatrix}, \quad \hat{C} = \Phi^T C \Phi. \quad (14)$$

⁵This is the case e.g., if in (4) only one of c_i is different from zero.

⁶In fact, according to our observations, just "optimally" damped systems tend to have rather highly clustered and defective eigenvalues.

This is the “modal representation” of the phase matrix, and it is in this form that we will use A in the rest of this paper.

Now, averaging over the frequency subspace determined by $\omega \leq \omega_{\max}$ is obtained by choosing

$$Z = Z_s = \begin{pmatrix} I_s & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & I_s & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad (15)$$

where I_s is the identity matrix of dimension s defined by $\omega_s = \omega_{\max}$.

We will now show that thus obtained X , $\zeta = \text{tr}(ZX)$ nicely controls the time history of the vibrational system. Indeed, according to [14], for any unit initial data ψ_0 we have

$$\|e^{At}\psi_0\|^2 \leq \frac{a}{h} \left(1 + \frac{h}{\|X\|}\right)^{2-t/h}, \quad (16)$$

for any $t > 0$ and any h between 0 and t . Here

$$a = \psi_0^T X \psi_0, \quad \|X\| \geq a.$$

For $t > a$ we may set $h = a$, and (16) gives

$$\|e^{At}\psi_0\|^2 \leq \left(1 + \frac{a}{\|X\|}\right)^{2-t/a}. \quad (17)$$

Let now ψ_0 be from the subspace spanned by Z , i.e., $\psi_0 = Z\psi_0$. Then

$$a = \psi_0^T X \psi_0 = \psi_0^T Z X Z \psi_0 \leq \|Z X Z\| \leq \text{tr}(Z X Z) = \text{tr}(Z X) = \zeta.$$

Altogether (note that the expression on the right-hand side of (17) is increasing with a),

$$\|e^{At}\psi_0\|^2 \leq \begin{cases} \left(1 + \frac{\zeta}{\|X\|}\right)^{2-t/\zeta}, & t > a, \\ 1, & \text{otherwise.} \end{cases}$$

Another popular decay criterion is the spectral abscissa

$$\lambda_m = \max \text{Re}(\lambda(A)),$$

based on the asymptotic estimate

$$\|e^{At}\psi_0\| \leq C_\varepsilon e^{(\lambda_m + \varepsilon)t}$$

for any $\varepsilon > 0$. The shortcomings of this estimate are

- (i) there is no control over the constant C_ε , and
- (ii) the quantity λ_m is not a smooth function of the matrix elements.

While the second point is obvious, the first needs an appropriate illustration, which will be given in Section 7.

3. Frequency cut-off

Solving (10) is numerically quite a time-consuming task, especially if in searching for an optimal damping (10) has to be solved many times. According to [1], each solving of a Lyapunov equation of order $2n$ takes about $15(2n)^3 = 120n^3$ operations. This may become very soon prohibitive if, e.g., the number of equations or their dimension get large. Anyhow, since we are interested in lower frequency region — this is expressed in the form of our weight matrix Z — we are heavily tempted to do a similar “cut-off” surgery on the input matrix A as well. This is also suggested by the fact that the rank of the damping matrix is much smaller than n , or, more precisely, we have

$$C = \sum_{i=1}^g c_i d_i d_i^T, \quad (18)$$

where d_i are fixed vectors, determining the dampers’ positions. So, instead of the original

$$A = \begin{pmatrix} 0 & \Omega \\ -\Omega & -\widehat{C} \end{pmatrix}, \quad \widehat{C} = -\Phi_2^{-1} C \Phi_2^{-T} = \sum_{i=1}^g c_i w_i w_i^T, \quad (19)$$

we obtain a “cut-off” matrix

$$A_r = \begin{pmatrix} 0 & \Omega_r \\ -\Omega_r & -\widehat{C}_r \end{pmatrix}, \quad (20)$$

where Ω_r and \widehat{C}_r are obtained by taking the first r rows and columns from Ω and \widehat{C} , respectively; $r \geq s$ has to be conveniently chosen: not too small in order to sufficiently reproduce the true dynamics below ω_s , and not too large in order to prevent prohibitive dimensions. Our experimental evidence suggests that r should be larger not only than s , but also than g , the number of damped dimensions. We still have no rigorous theoretical justification for this “cut-off Ansatz”, particularly for our heuristic ways of choosing r (our r is never much larger than s). When we say that the cut-off r should “sufficiently reproduce the dynamics”, we mean that $\text{tr}(ZX)$ and $\text{tr}(\widetilde{Z}\widetilde{X})$ (corresponding to the cut-off system) are nearly the same, i.e., the relative estimate

$$\left| \frac{\text{tr}(ZX) - \text{tr}(\widetilde{Z}\widetilde{X})}{\text{tr}(ZX)} \right| \quad (21)$$

is small for the whole range of dampings considered, and moreover, that for given dampers positions

$$\max_{i=1, \dots, g} \left| \frac{c_i^{(min)} - \widetilde{c}_i^{(min)}}{c_i^{(min)}} \right| \quad (22)$$

is small, where $c_i^{(min)}$, $\widetilde{c}_i^{(min)}$ represent the viscosities for minimal $\text{tr}(ZX)$, $\text{tr}(\widetilde{Z}\widetilde{X})$, respectively. This assumption is quite strong but not unrealistic. It is exactly fulfilled,

e.g., if

$$\widehat{C} = \sum_{i=1}^g c_i w_i w_i^T = VV^T, \quad \widehat{C}_r = \widetilde{V}\widetilde{V}^T,$$

and VV^T is block-diagonal with

$$VV^T = \begin{pmatrix} \widetilde{V}\widetilde{V}^T & 0 \\ 0 & \widetilde{\widetilde{V}}\widetilde{\widetilde{V}}^T \end{pmatrix}, \quad \text{where } \widetilde{\widetilde{V}}\widetilde{\widetilde{V}}^T \neq 0. \quad (23)$$

Note also that quite often the undamped system is given only through a number of the eigenfrequencies and their eigenforms. This number is usually quite smaller than the total number of degrees of freedom. So, the only way of obtaining *any* result under these circumstances is just to accept the said cut-off approximation idea (of course, our cut-off r is of a more drastic size). All this, however, does not diminish the necessity of theoretical justification, including error estimates.

The conjecture, for instance, that damping matrices \widehat{C} converging to the block-diagonal form (23) lead to an arbitrary small relative estimate (21) and (22), proves to be false. In case $g = 1$, i.e., one damper, we have investigated the cut-off problem exactly (see [2]). One of the important results is, that an arbitrary enlargement of the distance between the frequencies ω_s and ω_{r+1} leads to an arbitrary reduction of the relative estimates. Also, an arbitrary enlargement of a component of w_i in the region before the cut leads to an arbitrary reduction of the relative estimates. In case $g > 1$ we have no exact results. However, experiments show that enlarging of the distance between the frequencies in the region after the cut and before the cut leads to a reduction of the relative estimates, too. The experiments show also that the requirement that the rank g of the damping matrix is less than s , leads to better results (see [2] again).

4. Modal cut-off

Here we would like to briefly discuss another approximation which seems to be very popular, and which reduces the computational effort much more drastically. It is the so-called “modal approximation”, which consists of replacing \widehat{C} in (19) by its diagonal part. Now the Lyapunov equation solution comes for free, once the undamped eigenproblem is solved. However, this may falsify the dynamics dramatically, in particular in cases with small damping rank, which are our main interest.

This phenomenon is certainly not new, but it seems that its importance is not always properly appreciated. Thus, we illustrate it on a three-mass ladder (Figure 1) fixed at both ends and having unit masses and unit spring stiffnesses. The damper is applied at mass #1⁷.

⁷The damper has to be positioned nonsymmetrically, if we want the second mode to be affected as well.

The system at rest is put into vibration by applying unit impulse to mass #1. Our first graph (Figure 2) shows the time history of the position of mass #3 according to the full dynamics (full line), confronted to the one with cut-off dynamics (dashed line). In the same manner, Figure 3 shows the time history of the total energy. We see that neglecting the non-modal part of the damping matrix causes severe underestimation in predicting displacements. The weakness of this approximation is also seen by confronting the true damping matrix

$$C = \begin{pmatrix} c & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

with the modal cut-off damping matrix

$$C_d = \frac{c}{4} \begin{pmatrix} 1 & -\sqrt{2} & 1 \\ -\sqrt{2} & 2 & -\sqrt{2} \\ 1 & -\sqrt{2} & 1 \end{pmatrix},$$

both in Euclidean coordinates. The modal damping matrix shares the symmetry of the undamped system, while the true damping is seriously asymmetric.

Nevertheless, the idea of the “modal Ansatz” is too attractive to be completely abandoned. In fact, it can be very efficiently used in iterative minimization, as will be described later.

5. Gradient and Hessian

In standard minimization procedures the efficiency may be significantly hurt by the costs of computing the gradient, or even the Hessian of the penalty function. Our penalty function reads

$$f(c_1, \dots, c_g) = \text{tr}(ZX),$$

where X is the solution of the Lyapunov equation

$$A^T X + X A = -B, \quad (24)$$

or, in other words

$$X = \int_0^\infty e^{A^T t} B e^{A t} dt. \quad (25)$$

Here X depends on c_1, \dots, c_g through A from (7), whereas the matrices Z, B are symmetric and independent of c_1, \dots, c_g . Writing ∂_i for $\partial/\partial c_i$ and by using (18), (19), we have

$$\partial_i A = -w_i w_i^T, \quad w = \begin{pmatrix} 0 \\ \Phi_2^{-1} d_i \end{pmatrix}.$$

By differentiating (24) we obtain

$$A^T \partial_i X + \partial_i X A = X w_i w_i^T + w_i w_i^T X. \quad (26)$$

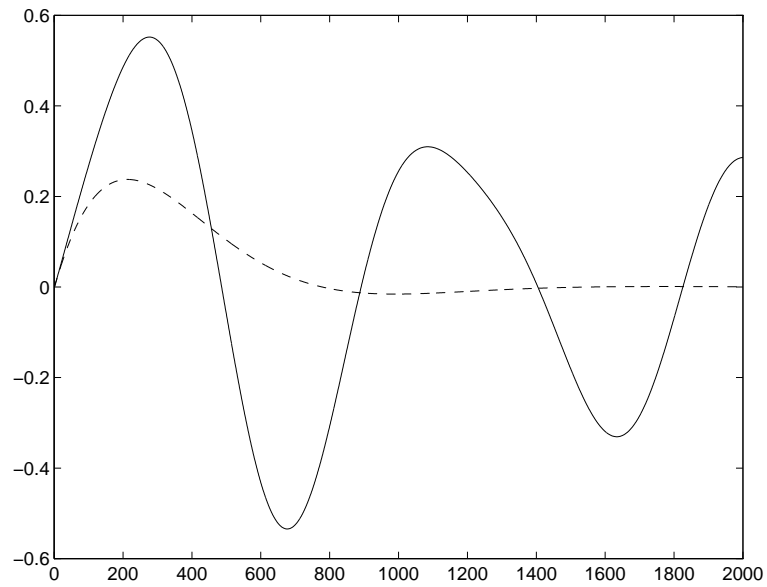


Figure 2. Time history: full line = fully damped, dashed line = modally damped.

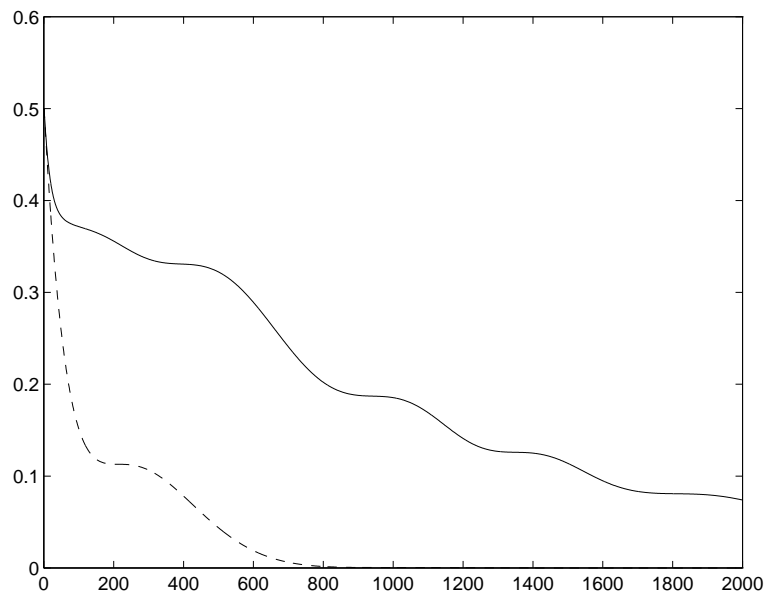


Figure 3. Total energy decay: full line = fully damped, dashed line = modally damped.

Thus, to find

$$\partial_i f = \partial_i \operatorname{tr}(ZX) = \operatorname{tr} \partial_i(ZX),$$

we first have to solve altogether $g + 1$ Lyapunov equations that differ from each other only in their right-hand sides. This can reduce the computational effort considerably, when using the standard Lyapunov solver from [1]. In fact, [1] first reduces A to the upper triangular form (or better, the real upper block-triangular form) which is the same for all equations; it is the backward substitution which has to be repeated each time. Unfortunately, the operational count of this backward substitution is of the same order of magnitude as the Schur reduction⁸, and the computational effort for the gradient may still be much larger than the one for the function value $f(c) = \operatorname{tr}(ZX)$ alone.

However, the special form of the penalty function f allows the computation of its gradient *with solving at most one more Lyapunov equation with the same left-hand side — independently of the number s of the components of the gradient*. Indeed, after applying formula (25) to both the equations (24) and (26), we obtain

$$\begin{aligned} \partial_i f &= - \int_0^\infty \operatorname{tr} \left(Z e^{A^T t} (X w_i w_i^T + w_i w_i^T X) e^{At} \right) dt \\ &= -w_i^T Y X w_i - w_i^T X Y w_i = -2w_i^T X Y w_i, \end{aligned}$$

where $Y = \int_0^\infty \exp(At) Z \exp(A^T t) dt$ solves the “dual Lyapunov equation”

$$AY + Y A^T = -Z.$$

By the J-symmetry (8) we have

$$Y = J X_1 J, \quad \text{with} \quad A^T X_1 + X_1 A = -J Z J.$$

In the special case of $Z = Z_s$, we have $Z_s J = J Z_s$ and

$$Y = J X_1 J, \quad \text{with} \quad A^T X_1 + X_1 A = -Z_s.$$

The last Lyapunov equation has the same left-hand side as (24). Now, the components of the gradient are obtained cheaply:

$$\partial_i f = -2w_i^T J X_1 J X w_i,$$

with some n^2 operations. If $B = Z = I$, then $X_1 = X$ and the Lyapunov solver is needed just once.

A similar formula exists for the $g(g+1)/2$ components $\partial_i \partial_j f$ of the Hessian matrix. After a straightforward calculation we obtain

$$\partial_i \partial_j f = -2w_i^T \partial_i X J X_1 J X w_i - 2w_j^T \partial_i X J X_1 J X w_j,$$

so these can be obtained under the additional cost of computing all s components $\partial_1 X, \dots, \partial_g X$ from (26), which is again a significant saving.

⁸Solver from [1] needs about $20n^3$ operations for the Schur reduction and about $2.5n^3$ operations for the backward substitution.

6. Initial guess

No matter which minimization method is used, the efficiency may be poor if we do not have a good initial guess, i.e., a starting configuration which is not hopelessly away from the optimal one.

This tackles two important questions. Does our penalty function have a global minimum? Are there several local minima? Our experiments indicate a negative answer to the second question, although non-convex penalty functions have been found. The first question remains open, even if we let C vary over the set of *all* positive semidefinite matrices. In this case, however, we know more.

Theorem 1. *Denote by \mathcal{M} the set of all matrices of the form*

$$\widehat{C} = 2\widetilde{\Omega} = 2 \begin{pmatrix} \widehat{\Omega}_s & 0 \\ 0 & H \end{pmatrix}, \quad \widehat{\Omega}_s = \text{diag}(\omega_1, \dots, \omega_s),$$

where H varies over the set of all symmetric positive semidefinite matrices of order $n - s$, such that the corresponding A is stable. Then the set \mathcal{M} gives the function $\text{tr}(Z_s X)$, $A^T X + X A = -I$, a strong local minimum. In other words, in the set of all stable A there is a neighbourhood \mathcal{O} of \mathcal{M} , such that our function, restricted to \mathcal{O} takes its (absolute) minimum on \mathcal{M} and nowhere else. In particular, our function is constant on \mathcal{M} .

For the proof of this theorem we refer the reader to the Appendix at the end of this paper.

In order to understand better the background of this theorem, take a special case $Z = Z_s = I$. Then the set \mathcal{M} reduces to a single matrix 2Ω which gives rise to an A , which is just a direct sum of n two-dimensional blocks, each of them describing a one-dimensional critically damped system. For any such single system the theorem above is known to be true (in this simple case the function is even convex). It seems plausible, and experiments confirm it so far, that 2Ω is generally the best.

Conjecture 1. *The minimum from Theorem 1 is global.*

Another result of a similar kind is given in [3] for the case of a continuous “modally damped” system. Here we present a finite matrix analogue of the aforementioned result.

Modally damped systems are characterized by the generalized commutativity property $CK^{-1}M = MK^{-1}C$. In the modal representation (14) this means just commutativity $\widehat{C}\Omega = \Omega\widehat{C}$. In this case the solution X of the Lyapunov equation

$$\begin{pmatrix} 0 & -\Omega \\ \Omega & -\widehat{C} \end{pmatrix} X + X \begin{pmatrix} 0 & \Omega \\ -\Omega & -\widehat{C} \end{pmatrix} = -I$$

is directly shown to be (cf. Cox [3])

$$X = \begin{pmatrix} \frac{1}{2}\widehat{C}\Omega^{-2} + \widehat{C}^{-1} & \frac{1}{2}\Omega^{-1} \\ \frac{1}{2}\Omega^{-1} & \widehat{C}^{-1} \end{pmatrix}. \quad (27)$$

We now represent Z from (15) as

$$Z = Z_s = \begin{pmatrix} \tilde{Z} & 0 \\ 0 & \tilde{Z} \end{pmatrix}, \quad \tilde{Z} = \tilde{f}(\Omega), \quad \tilde{f}(\lambda) = \begin{cases} 1, & \lambda \leq \omega_{\max}, \\ 0, & \lambda > \omega_{\max}. \end{cases}$$

This \tilde{Z} has the property that any \hat{C} commuting with Ω commutes with \tilde{Z} also.

Theorem 2. *Let Ω be fixed and let \hat{C} vary over the set of all positive definite matrices commuting with Ω . Then in the representation*

$$\tilde{Z} = \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}, \quad \Omega = \begin{pmatrix} \hat{\Omega}_s & 0 \\ 0 & \hat{\Omega}_{n-s} \end{pmatrix}, \quad \hat{C} = \begin{pmatrix} \hat{C}_s & 0 \\ 0 & H \end{pmatrix},$$

the function $\text{tr}(Z_s X)$ can be expressed as a strictly convex function of the variable \hat{C}_s . Consequently, within this set the minimum is global.

Proof. From (27) follows

$$\text{tr}(Z_s X) = f(\hat{C}_s) = 2 \text{tr} \hat{C}_s^{-1} + \frac{1}{2} \text{tr} (\hat{\Omega}_s^{-2} \hat{C}_s).$$

To prove the strict convexity of f , take A, B positive definite and $0 < a < 1$. Let

$$\begin{aligned} U &= aA^{-1} + (1-a)B^{-1} - (aA + (1-a)B)^{-1} \\ &= A^{-1/2} a(1-a) (aI + (1-a)W^{-1})^{-1} (W + W^{-1} - 2I) A^{-1/2}, \end{aligned}$$

where $W = A^{1/2} B^{-1} A^{1/2}$. Then U is positive semidefinite (this follows from the positive definiteness of W and $aI + (1-a)W^{-1}$, and the positive semidefiniteness of $W + W^{-1} - 2I$). So, $\text{tr} U \geq 0$, which is the convexity of f . If this inequality becomes an equality, then $U = 0$ or, equivalently, $A = B$. So, f is strictly convex. ■

Now, our initial guess for minimization methods is obtained as follows: Among all $\hat{C} = c_1 w_1 w_1^T + \dots + c_g w_g w_g^T$, take the one that is closest to the set \mathcal{M} .

This leads to the following least squares problem:

$$\begin{aligned} \text{dist}^2 \left(\mathcal{M}, \sum_{i=1}^g \tilde{c}_i w_i w_i^T \right) &= \min_{c_1, \dots, c_g} \text{dist}^2 \left(\mathcal{M}, \sum_{i=1}^g c_i w_i w_i^T \right) \\ &= \min_{c_1, \dots, c_g} \left(\inf_{H \in \mathcal{S}} \left\| \begin{pmatrix} 2\hat{\Omega}_s & 0 \\ 0 & H \end{pmatrix} - \sum_{i=1}^g c_i w_i w_i^T \right\|_F^2 \right) \end{aligned} \quad (28)$$

where

$$\mathcal{S} = \left\{ H \text{ positive semidefinite of order } n-s \mid A \text{ stable with } C = \begin{pmatrix} 2\hat{\Omega}_s & 0 \\ 0 & H \end{pmatrix} \right\}.$$

Here, $\|\cdot\|_F$ is the Frobenius norm. Now, in general this minimization does not need to yield a positive semidefinite $\tilde{C} = \sum_{i=1}^g \tilde{c}_i w_i w_i^T$. With the assumption $c_i \geq 0$ for

all i , problem (28) leads to the nonnegative least squares problem (see [2]):

$$\min_{c_1, \dots, c_g \geq 0} \left\| X_0 - [A_1 \ A_2 \ \dots \ A_g] \begin{pmatrix} c_1 \\ \vdots \\ c_g \end{pmatrix} \right\|_F^2, \quad (29)$$

with

$$X_0 = \begin{pmatrix} 2\hat{\Omega}_s & 0 \\ 0 & 0 \end{pmatrix},$$

$$A_k = \begin{pmatrix} \hat{V}_s^k (\hat{V}_s^k)^T & 0 \\ 0 & \sqrt{2} \hat{V}_s^k (\hat{V}_{n-s}^k)^T \end{pmatrix}, \quad V = \begin{pmatrix} \hat{V}_s \\ \hat{V}_{n-s} \end{pmatrix} = [w_1 \ \dots \ w_g],$$

where \hat{V}_s^k or \hat{V}_{n-s}^k is the k -th column of \hat{V}_s or of \hat{V}_{n-s} , respectively, with $1 \leq k \leq g$. It is not *a priori* clear that the minimization (29) leads to a \tilde{C} , so that the corresponding A is a stable matrix. However, in all our applications, the obtained \tilde{C} not only led to a stable A , but also to a reasonable starting point for minimization. This cannot be fully explained yet. A partial explanation is given by the following theorem.

Theorem 3. *If $Z_s = I$ and w_i in (19) are mutually orthogonal, then all \tilde{c}_i of the optimal \tilde{C} from the least squares process (28) or (29), respectively, are positive.*

Proof. The initial guess coefficients are given by projections

$$\tilde{c}_i = \frac{2w_i^T \Omega w_i}{\|w_i\|^4}.$$

■

In this case \tilde{C} leads to a stable A , except in the trivial case when $V = [w_1 \ \dots \ w_g]$ has a zero row.

7. Spectral abscissa versus trace

In many applications optimizing the spectral abscissa gives about the same results as optimizing the trace. This is not always so; in particular, the optimal spectral abscissa may fail to control the solution at finite times. We will take two examples from the recent work of Freitas and Lancaster [4]. Both matrices are “optimal” in the sense that they have a “best spectral abscissa” under all symmetric damping matrices. In the first example we have

$$M = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad K = \begin{pmatrix} k_1 & 0 \\ 0 & k_2 \end{pmatrix},$$

with $k_1 = 1$, $k_2 = 25$. The optimal abscissa damping is

$$C = \frac{1}{k_1^{1/2} + k_2^{1/2}} \begin{pmatrix} 4 k_1^{3/4} k_2^{1/4} & (k_1^{1/2} - k_2^{1/2})^2 \\ (k_1^{1/2} - k_2^{1/2})^2 & 4 k_1^{1/4} k_2^{3/4} \end{pmatrix}.$$

Our best damping in this case is

$$C_d = \begin{pmatrix} 2 & 0 \\ 0 & 10 \end{pmatrix}.$$

The corresponding phase space matrices are A and A_d . For the solutions $\dot{y} = Ay$, $\dot{y}_d = A_d y$, with the common initial data

$$y(0) = y_d(0) = [-0.2278 \quad -0.3012 \quad -0.4847 \quad 0.2468]^T,$$

we obtain the following two figures: a typical time history (coordinate #1) can be seen in Figure 4, and the time history of the total energy is displayed in Figure 5.

Optimal abscissa is better for large times, but then, anyhow, everything is calmed down. For finite times, however, optimal abscissa may allow larger oscillations.

Another example from [4] is given by

$$M = \text{diag}(1, 1, 1), \quad K = \text{diag}\left(\frac{1}{100}, 64, 100\right),$$

with the optimal abscissa damping

$$C = \begin{pmatrix} \frac{4(a-493)}{17775} & \frac{(851201-32a)b}{35550(9887+16a)} & \frac{4bc}{395 \cdot 3^{3/2}} \\ \frac{(851201-32a)b}{35550(9887+16a)} & \frac{64(493-a)(a-53818)}{17775(9887+16a)} & \frac{16(a-1204)c}{395 \cdot 3^{3/2}} \\ \frac{4bc}{395 \cdot 3^{3/2}} & \frac{16(a-1204)c}{395 \cdot 3^{3/2}} & \frac{4(53818-a)}{9887+16a} \end{pmatrix},$$

where

$$a = \sqrt{385249}, \quad b = \sqrt{2585257 + 15776a}, \quad c = \sqrt{103a - 63929}.$$

Our best damping in this case is

$$C_d = \begin{pmatrix} 0.2 & 0 & 0 \\ 0 & 16 & 0 \\ 0 & 0 & 20 \end{pmatrix}.$$

The corresponding phase space matrices are A and A_d . For the solutions $\dot{y} = Ay$, $\dot{y}_d = A_d y$, with the common initial data

$$y(0) = y_d(0) = [0.2176 \quad 0.1927 \quad -0.4159 \quad -0.0456 \quad -0.0582 \quad -0.1467]^T,$$

we obtain

$$\max_t \|y(t)\| = 11.4002, \quad \max_t \|y_d(t)\| = 0.5334.$$

A typical time history (coordinate #1) can be seen in Figure 6, whereas the time history of the total energy is displayed in Figure 7. Here the difference of both criteria is more drastic.

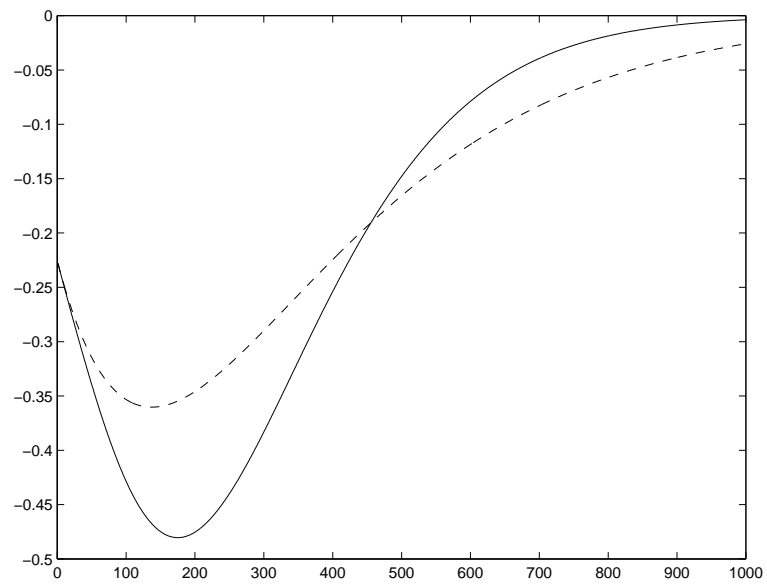


Figure 4. Coordinate #1: full line = optimal abscissa, dashed line = optimal trace.

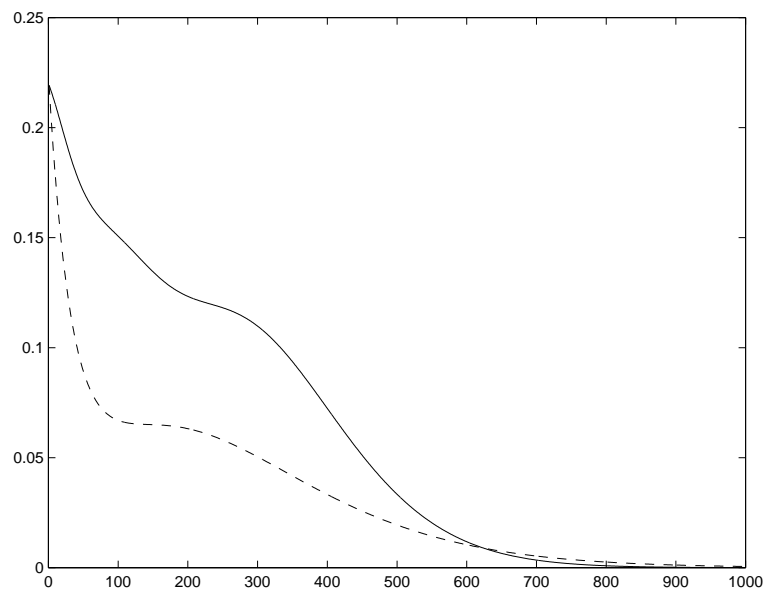


Figure 5. Total energy: full line = optimal abscissa, dashed line = optimal trace.

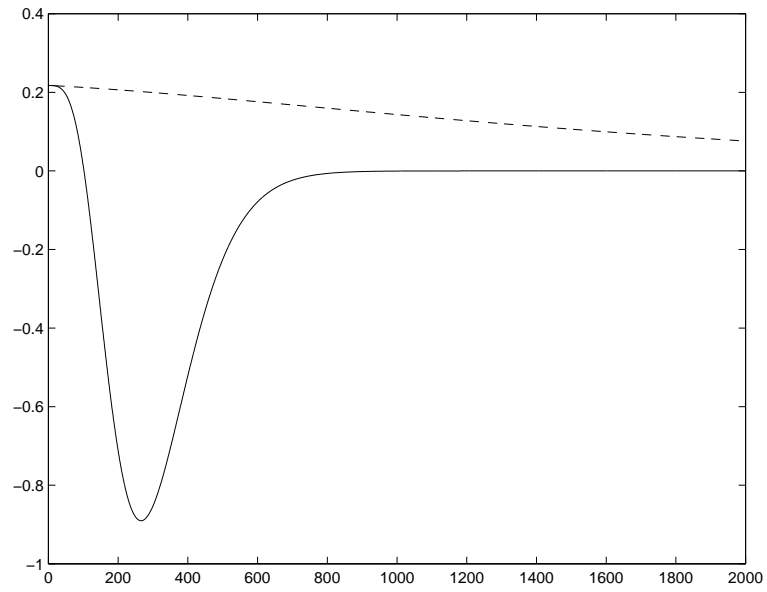


Figure 6. Coordinate #1: full line = optimal abscissa, dashed line = optimal trace.

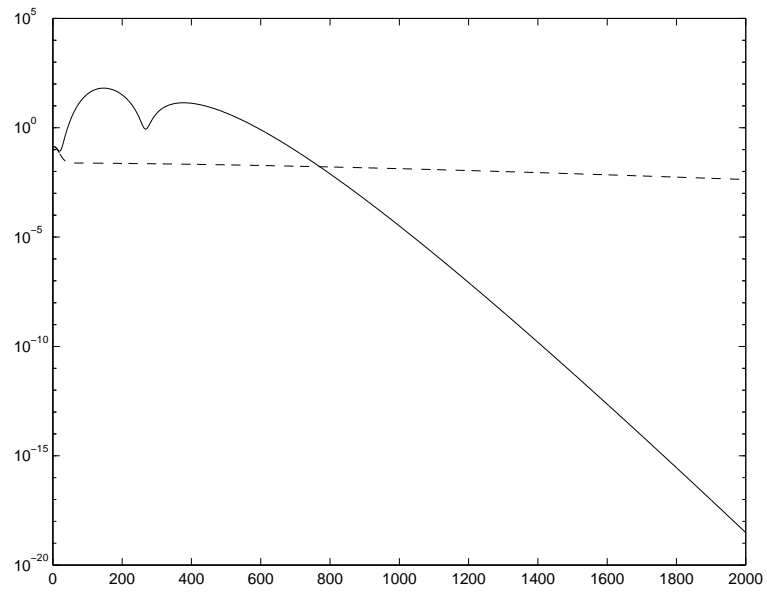


Figure 7. Total energy: full line = optimal abscissa, dashed line = optimal trace.

Appendix

To prove Theorem 1 we first give a suitable formulation of the well-known criteria for a local minimum of a real-valued function in a finite dimensional vector space.

Lemma 1.

(i) Let f be a continuously differentiable real-valued function on a neighbourhood of $x_0 \in \mathbb{R}^n$. Then $\text{grad } f(x_0) = 0$ if and only if

$$\left. \frac{d}{d\mu} f(x_0 + \mu v) \right|_{\mu=0} = 0 \quad \text{for all } v \in \mathbb{R}^n, v \neq 0.$$

(ii) Let f be a twice continuously differentiable real-valued function on a neighbourhood of $x_0 \in \mathbb{R}^n$. The Hessian $\text{Hess } f(x_0)$ is positive definite if and only if

$$\left. \frac{d^2}{d\mu^2} f(x_0 + \mu v) \right|_{\mu=0} > 0 \quad \text{for all } v \in \mathbb{R}^n, v \neq 0.$$

Using this Lemma we will now prove Theorem 1.

Proof of Theorem 1. We prove the statement of the theorem for the set

$$\mathcal{M}_0 = \left\{ \widehat{C}_0 = 2\widetilde{\Omega} = 2 \begin{pmatrix} \widehat{\Omega}_s & 0 \\ 0 & H \end{pmatrix} \text{ with } \widehat{\Omega}_s = \text{diag}(\omega_1, \dots, \omega_s), H \text{ positive definite} \right\}.$$

\mathcal{M}_0 is a dense subset of \mathcal{M} , so that the statement of the theorem follows finally from the continuity of the solution of the Lyapunov equation.

Let C be a positive semidefinite matrix. The corresponding matrix \widehat{C} from (14) is also positive semidefinite. For any $2\widetilde{\Omega} \in \mathcal{M}_0$, the matrix \widehat{C} can be written as $\widehat{C} = 2\widetilde{\Omega} + \mu V$, where V is a symmetric matrix of order n . This representation can be used to define the neighbourhood \mathcal{O} of \mathcal{M}_0 (or \mathcal{M}).

Let V be an arbitrary symmetric matrix of order n and let

$$\widehat{C}_\mu := 2\widetilde{\Omega} + \mu V. \quad (30)$$

For all sufficiently small μ , the matrix \widehat{C}_μ is positive definite, so the Lyapunov equation $A_\mu^T X + X A_\mu = -I$, where

$$A_\mu := A(\widehat{C}_\mu) = \begin{pmatrix} 0 & \Omega \\ -\Omega & -\widehat{C}_\mu \end{pmatrix} = \begin{pmatrix} 0 & \Omega \\ -\Omega & -(2\widetilde{\Omega} + \mu V) \end{pmatrix}, \quad (31)$$

has a unique solution $X_\mu = X(\widehat{C}_\mu)$. In other words, for any given $\widehat{C}_0 = 2\widetilde{\Omega} \in \mathcal{M}_0$ and any V , A_μ is stable and C_μ is in the neighbourhood of \mathcal{M}_0 , for all μ from a certain neighbourhood of zero.

Our function f is $f(\widehat{C}_\mu) = \text{tr } Y_\mu$, where $Y_\mu := Z_s X_\mu$ solves the Lyapunov equation $A_\mu^T Y + Y A_\mu = -Z_s$ (this follows from (12) and the J-symmetry of Z_s).

According to Lemma 1, for arbitrary $\widehat{C}_0 = 2\widetilde{\Omega} \in \mathcal{M}_0$ we have to show

$$\frac{d}{d\mu} \operatorname{tr} Y_\mu \Big|_{\mu=0} = 0 \quad (32)$$

for all symmetric matrices V of order n , and

$$\frac{d^2}{d\mu^2} \operatorname{tr} Y_\mu \Big|_{\mu=0} > 0 \quad (33)$$

for all symmetric matrices V of order n , such that $\widehat{C}_\mu \notin \mathcal{M}_0$ for small μ . From (30), if we write V as a 2×2 block-matrix

$$V = \begin{pmatrix} V_{11} & V_{12} \\ V_{12}^T & V_{22} \end{pmatrix}, \quad V_{22} \text{ of order } (n-s), \quad (34)$$

this is equivalent to the requirement that V_{11} and V_{12} should not vanish simultaneously (otherwise, $\widehat{C}_\mu \in \mathcal{M}_0$ for small μ).

In the rest of the proof, 2×2 block decompositions of $n \times n$ matrices always have the same form as in (34) without further notice.

I. Step:

In the first step we show that (32) holds for all symmetric matrices V .

First we determine

$$\frac{d}{d\mu} \operatorname{tr} Y_\mu = \operatorname{tr} \frac{d}{d\mu} Y_\mu.$$

By differentiating the Lyapunov equation for Y_μ

$$A_\mu^T Y_\mu + Y_\mu A_\mu = -Z_s, \quad (35)$$

we obtain

$$\left(\frac{d}{d\mu} A_\mu^T \right) Y_\mu + A_\mu^T \frac{d}{d\mu} Y_\mu + \left(\frac{d}{d\mu} Y_\mu \right) A_\mu + Y_\mu \frac{d}{d\mu} A_\mu = 0. \quad (36)$$

From (31), we have

$$-\widetilde{V} := \frac{d}{d\mu} A_\mu = \begin{pmatrix} 0 & 0 \\ 0 & -V \end{pmatrix} = \frac{d}{d\mu} A_\mu^T. \quad (37)$$

Now, (36) can be written as the new Lyapunov equation

$$A_\mu^T \frac{d}{d\mu} Y_\mu + \left(\frac{d}{d\mu} Y_\mu \right) A_\mu = \widetilde{V} Y_\mu + Y_\mu \widetilde{V}, \quad (38)$$

with the solution

$$\frac{d}{d\mu} Y_\mu = - \int_0^\infty e^{A_\mu^T t} (\widetilde{V} Y_\mu + Y_\mu \widetilde{V}) e^{A_\mu t} dt.$$

Taking the trace yields

$$\begin{aligned}
\operatorname{tr} \frac{d}{d\mu} Y_\mu &= -\operatorname{tr} \left[\int_0^\infty e^{A_\mu^T t} (\tilde{V} Y_\mu + Y_\mu \tilde{V}) e^{A_\mu t} dt \right] \\
&= -\operatorname{tr} \left[\int_0^\infty e^{A_\mu^T t} \tilde{V} Y_\mu e^{A_\mu t} dt \right] - \operatorname{tr} \left[\int_0^\infty e^{A_\mu^T t} Y_\mu \tilde{V} e^{A_\mu t} dt \right] \\
&= -\operatorname{tr} \left[\int_0^\infty \tilde{V} Y_\mu e^{A_\mu t} e^{A_\mu^T t} dt \right] - \operatorname{tr} \left[\int_0^\infty e^{A_\mu t} e^{A_\mu^T t} Y_\mu \tilde{V} dt \right] \\
&= -\operatorname{tr} \left[\tilde{V} Y_\mu \int_0^\infty e^{A_\mu t} e^{A_\mu^T t} dt \right] - \operatorname{tr} \left[\left(\int_0^\infty e^{A_\mu t} e^{A_\mu^T t} dt \right) Y_\mu \tilde{V} \right] \\
&= -\operatorname{tr}(\tilde{V} Y_\mu \tilde{X}_\mu) - \operatorname{tr}(\tilde{X}_\mu Y_\mu \tilde{V}) = -2 \operatorname{tr}(\tilde{V} Y_\mu \tilde{X}_\mu),
\end{aligned}$$

where

$$\tilde{X}_\mu = \int_0^\infty e^{A_\mu t} e^{A_\mu^T t} dt \quad (39)$$

is the solution of the Lyapunov equation $A_\mu \tilde{X}_\mu + \tilde{X}_\mu A_\mu^T = -I$.

By assumption, X_μ is the solution of the Lyapunov equation $A_\mu^T X_\mu + X_\mu A_\mu = -I$. Since A_μ is J -symmetric (8), this equation can be written as

$$J A_\mu J X_\mu + X_\mu J A_\mu^T J = -I,$$

or

$$A_\mu J X_\mu J + J X_\mu J A_\mu^T = -I.$$

It follows that $\tilde{X}_\mu = J X_\mu J$ and

$$\operatorname{tr} \frac{d}{d\mu} Y_\mu = -2 \operatorname{tr}(\tilde{V} Y_\mu J X_\mu J). \quad (40)$$

Finally, we have to find explicit forms for Y_0 and X_0 . Let

$$\Omega =: \begin{pmatrix} \hat{\Omega}_s & 0 \\ 0 & \hat{\hat{\Omega}}_{n-s} \end{pmatrix}.$$

The solution Y_0 of (35) for $\mu = 0$ is easily seen to be

$$Y_0 = \begin{pmatrix} \Psi_{11} & \Psi_{12} \\ \Psi_{12}^T & \Psi_{22} \end{pmatrix}, \quad (41)$$

where

$$\Psi_{11} = \begin{pmatrix} \frac{3}{2} \hat{\Omega}_s^{-1} & 0 \\ 0 & 0 \end{pmatrix}, \quad \Psi_{12} = \Psi_{22} = \begin{pmatrix} \frac{1}{2} \hat{\Omega}_s^{-1} & 0 \\ 0 & 0 \end{pmatrix}. \quad (42)$$

Likewise, the solution X_0 of the Lyapunov equation $A_0^T X_0 + X_0 A_0 = -I$ has the form

$$X_0 = \begin{pmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{12}^T & \Phi_{22} \end{pmatrix}, \quad (43)$$

with

$$\Phi_{11} = \begin{pmatrix} \frac{3}{2}\widehat{\Omega}_s^{-1} & 0 \\ 0 & W_{11} \end{pmatrix}, \quad \Phi_{12} = \begin{pmatrix} \frac{1}{2}\widehat{\Omega}_s^{-1} & 0 \\ 0 & W_{12} \end{pmatrix}, \quad \Phi_{22} = \begin{pmatrix} \frac{1}{2}\widehat{\Omega}_s^{-1} & 0 \\ 0 & W_{22} \end{pmatrix}, \quad (44)$$

where

$$W := \begin{pmatrix} W_{11} & W_{12} \\ W_{12}^T & W_{22} \end{pmatrix} \quad (45)$$

of order $2(n-s)$ is the uniquely determined solution of the Lyapunov equation

$$\begin{pmatrix} 0 & -\widehat{\Omega}_{n-s} \\ \widehat{\Omega}_{n-s} & -2H \end{pmatrix} W + W \begin{pmatrix} 0 & \widehat{\Omega}_{n-s} \\ -\widehat{\Omega}_{n-s} & -2H \end{pmatrix} = \begin{pmatrix} -I_{n-s} & 0 \\ 0 & -I_{n-s} \end{pmatrix}. \quad (46)$$

From (40), by using (37) and (41)–(44), we obtain

$$\begin{aligned} \operatorname{tr} \left(\frac{d}{d\mu} Y_\mu \right) \Big|_{\mu=0} &= -2 \operatorname{tr}(\widetilde{V} Y_0 J X_0 J) = -2 \operatorname{tr} \left[\begin{pmatrix} 0 & 0 \\ V \Psi_{12}^T & V \Psi_{22} \end{pmatrix} \begin{pmatrix} \Phi_{11} & -\Phi_{12} \\ -\Phi_{12}^T & \Phi_{22} \end{pmatrix} \right] \\ &= 2 \operatorname{tr}(V \Psi_{12}^T \Phi_{12}) - 2 \operatorname{tr}(V \Psi_{22} \Phi_{22}) \\ &= 2 \operatorname{tr} \left[V \begin{pmatrix} \frac{1}{4}\widehat{\Omega}_s^{-2} & 0 \\ 0 & 0 \end{pmatrix} \right] - 2 \operatorname{tr} \left[V \begin{pmatrix} \frac{1}{4}\widehat{\Omega}_s^{-2} & 0 \\ 0 & 0 \end{pmatrix} \right] = 0. \end{aligned}$$

We have just proved (32). Furthermore, the set \mathcal{M} is connected and \mathcal{M}_0 is a dense subset in \mathcal{M} , so we obtain that $\operatorname{tr} Y(\widehat{C}_0) = \text{const}$ on \mathcal{M} .

II. Step:

In the second step we show that (33) holds for all symmetric matrices V , such that V_{11}, V_{12} do not vanish simultaneously in (34).

First we determine

$$\frac{d^2}{d\mu^2} \operatorname{tr} Y_\mu = \operatorname{tr} \frac{d^2}{d\mu^2} Y_\mu.$$

By differentiating the Lyapunov equation (38) and using (37), we obtain the Lyapunov equation

$$A_\mu^T \left(\frac{d^2}{d\mu^2} Y_\mu \right) + \left(\frac{d^2}{d\mu^2} Y_\mu \right) A_\mu = 2\widetilde{V} \left(\frac{d}{d\mu} Y_\mu \right) + 2 \left(\frac{d}{d\mu} Y_\mu \right) \widetilde{V},$$

with the solution

$$\frac{d^2}{d\mu^2} Y_\mu = - \int_0^\infty e^{A_\mu^T t} \left[2\widetilde{V} \left(\frac{d}{d\mu} Y_\mu \right) + 2 \left(\frac{d}{d\mu} Y_\mu \right) \widetilde{V} \right] e^{A_\mu t} dt.$$

As in the first step, by using (39) and $\tilde{X}_\mu = JX_\mu J$, the trace can be written as

$$\begin{aligned} \operatorname{tr} \frac{d^2}{d\mu^2} Y_\mu &= -\operatorname{tr} \left[\int_0^\infty e^{A_\mu^T t} \left[2\tilde{V} \left(\frac{d}{d\mu} Y_\mu \right) + 2 \left(\frac{d}{d\mu} Y_\mu \right) \tilde{V} \right] e^{A_\mu t} dt \right] \\ &= -2 \operatorname{tr} \left[\int_0^\infty \tilde{V} \left(\frac{d}{d\mu} Y_\mu \right) e^{A_\mu t} e^{A_\mu^T t} dt \right] - 2 \operatorname{tr} \left[\int_0^\infty \left(\frac{d}{d\mu} Y_\mu \right) \tilde{V} e^{A_\mu t} e^{A_\mu^T t} dt \right] \\ &= -2 \operatorname{tr} \left[\tilde{V} \left(\frac{d}{d\mu} Y_\mu \right) JX_\mu J \right] - 2 \operatorname{tr} \left[\left(\frac{d}{d\mu} Y_\mu \right) \tilde{V} JX_\mu J \right]. \end{aligned}$$

Let Y'_0 be defined as

$$Y'_0 := \begin{pmatrix} \tilde{Y}_{11} & \tilde{Y}_{12} \\ \tilde{Y}_{12}^T & \tilde{Y}_{22} \end{pmatrix} := \frac{d}{d\mu} Y_\mu \Big|_{\mu=0}. \quad (47)$$

With X_0 from (43), it follows

$$\begin{aligned} \operatorname{tr} \left(\frac{d^2}{d\mu^2} Y_\mu \right) \Big|_{\mu=0} &= -2 \operatorname{tr} (\tilde{V} Y'_0 JX_0 J) - 2 \operatorname{tr} (Y'_0 \tilde{V} JX_0 J) \\ &= -2 \operatorname{tr} \left[\begin{pmatrix} 0 & 0 \\ V\tilde{Y}_{12}^T & V\tilde{Y}_{22} \end{pmatrix} \begin{pmatrix} \Phi_{11} & -\Phi_{12} \\ -\Phi_{12}^T & \Phi_{22} \end{pmatrix} \right] \\ &\quad - 2 \operatorname{tr} \left[\begin{pmatrix} 0 & \tilde{Y}_{12} V \\ 0 & \tilde{Y}_{22} V \end{pmatrix} \begin{pmatrix} \Phi_{11} & -\Phi_{12} \\ -\Phi_{12}^T & \Phi_{22} \end{pmatrix} \right] \\ &= -2 \operatorname{tr} (-V\tilde{Y}_{12}^T \Phi_{12} + V\tilde{Y}_{22} \Phi_{22}) - 2 \operatorname{tr} (-\tilde{Y}_{12} V \Phi_{12}^T + \tilde{Y}_{22} V \Phi_{22}) \\ &= 4 \operatorname{tr} (\tilde{Y}_{12} V \Phi_{12}^T) - 4 \operatorname{tr} (V\tilde{Y}_{22} \Phi_{22}). \end{aligned} \quad (48)$$

Now we have to determine the submatrices \tilde{Y}_{12} and \tilde{Y}_{22} of the matrix Y'_0 , which is the solution of the Lyapunov equation (38) for $\mu = 0$. In other words, we have to explicitly solve the Lyapunov equation

$$A_0^T Y'_0 + Y'_0 A_0 = \tilde{V} Y_0 + Y_0 \tilde{V},$$

where Y_0 is given by (41) and (42). In accordance with (47), this leads to the following four matrix equations

$$-\Omega \tilde{Y}_{12}^T - \tilde{Y}_{12} \Omega = 0 \quad (49)$$

$$-\Omega \tilde{Y}_{22} + \tilde{Y}_{11} \Omega - 2\tilde{Y}_{12} \tilde{\Omega} = \Psi_{12} V \quad (50)$$

$$\Omega \tilde{Y}_{11} - 2\tilde{\Omega} \tilde{Y}_{12}^T - \tilde{Y}_{22} \Omega = V \Psi_{12}^T \quad (51)$$

$$\Omega \tilde{Y}_{12} - 2\tilde{\Omega} \tilde{Y}_{22} + \tilde{Y}_{12}^T \Omega - 2\tilde{Y}_{22} \tilde{\Omega} = V \Psi_{22} + \Psi_{22} V, \quad (52)$$

where equations (50) and (51) are adjoint to each other. The first equation yields

$$\tilde{Y}_{12} = \frac{1}{2} S \Omega^{-1}, \quad \tilde{Y}_{12}^T = -\frac{1}{2} \Omega^{-1} S, \quad (53)$$

where S is a skew-symmetric matrix, i.e., $S^T = -S$. Substituting this into (48), we get

$$\operatorname{tr} \left(\frac{d^2}{d\mu^2} Y_\mu \right) \Big|_{\mu=0} = 2 \operatorname{tr}(S\Omega^{-1}V\Phi_{12}^T) - 4 \operatorname{tr}(V\tilde{Y}_{22}\Phi_{22}). \quad (54)$$

Now we have to determine \tilde{Y}_{22} and S more precisely. Let the block-matrix representations of S and the blocks in Y'_0 be

$$\begin{aligned} \tilde{Y}_{11} &= \begin{pmatrix} R_{11} & R_{12} \\ R_{12}^T & R_{22} \end{pmatrix}, & \tilde{Y}_{12} &= \frac{1}{2} S\Omega^{-1} = \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix}, \\ \tilde{Y}_{22} &= \begin{pmatrix} P_{11} & P_{12} \\ P_{12}^T & P_{22} \end{pmatrix}, & S &= \begin{pmatrix} S_{11} & S_{12} \\ -S_{12}^T & S_{22} \end{pmatrix}. \end{aligned}$$

With these definitions and (42), each of the equations (49)–(52) can be written as a 2×2 block-matrix equation. Altogether, this leads to 16 block-matrix equations in terms of the smaller blocks that have just been defined. We group them in accordance with the 2×2 block-form of (49)–(52).

First we consider the $(2, 2)$ -block of the block-matrix equations corresponding to (49)–(52). We get the following four equations, where once more the second and third equation are adjoint to each other:

$$\begin{aligned} -\widehat{\Omega}_{n-s}Q_{22}^T - Q_{22}\widehat{\Omega}_{n-s} &= 0 \\ -\widehat{\Omega}_{n-s}P_{22} + R_{22}\widehat{\Omega}_{n-s} - 2Q_{22}H &= 0 \\ \widehat{\Omega}_{n-s}R_{22} - 2HQ_{22}^T - P_{22}\widehat{\Omega}_{n-s} &= 0 \\ \widehat{\Omega}_{n-s}Q_{22} - 2HP_{22} + Q_{22}^T\widehat{\Omega}_{n-s} - 2P_{22}H &= 0. \end{aligned}$$

This system is equivalent to the matrix equation

$$\begin{pmatrix} 0 & -\widehat{\Omega}_{n-s} \\ \widehat{\Omega}_{n-s} & -2H \end{pmatrix} \begin{pmatrix} R_{22} & Q_{22} \\ Q_{22}^T & P_{22} \end{pmatrix} + \begin{pmatrix} R_{22} & Q_{22} \\ Q_{22}^T & P_{22} \end{pmatrix} \begin{pmatrix} 0 & \widehat{\Omega}_{n-s} \\ -\widehat{\Omega}_{n-s} & -2H \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Positive definiteness of H implies that the matrix

$$\begin{pmatrix} 0 & \widehat{\Omega}_{n-s} \\ -\widehat{\Omega}_{n-s} & -2H \end{pmatrix}$$

is asymptotically stable too, and the matrix equation has only the trivial solution $R_{22} = Q_{22} = P_{22} = 0$. From (53), we obtain

$$\tilde{Y}_{12} = \frac{1}{2} S\Omega^{-1} = \frac{1}{2} \begin{pmatrix} S_{11}\widehat{\Omega}_s^{-1} & S_{12}\widehat{\Omega}_{n-s}^{-1} \\ -S_{12}^T\widehat{\Omega}_s^{-1} & S_{22}\widehat{\Omega}_{n-s}^{-1} \end{pmatrix} = \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & 0 \end{pmatrix}, \quad (55)$$

so $S_{22} = 0$ follows, as well.

Putting this result into (54), together with (44), we obtain

$$\begin{aligned} \operatorname{tr} \left(\frac{d^2}{d\mu^2} Y_\mu \right) \Big|_{\mu=0} &= 2 \operatorname{tr} \left[\begin{pmatrix} S_{11} \widehat{\Omega}_s^{-1} & S_{12} \widehat{\Omega}_{n-s}^{-1} \\ -S_{12}^T \widehat{\Omega}_s^{-1} & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{2} V_{11} \widehat{\Omega}_s^{-1} & V_{12} W_{12}^T \\ \frac{1}{2} V_{12}^T \widehat{\Omega}_s^{-1} & V_{22} W_{12}^T \end{pmatrix} \right] \\ &\quad - 4 \operatorname{tr} \left[\begin{pmatrix} V_{11} & V_{12} \\ V_{12}^T & V_{22} \end{pmatrix} \begin{pmatrix} \frac{1}{2} P_{11} \widehat{\Omega}_s^{-1} & P_{12} W_{22} \\ \frac{1}{2} P_{12}^T \widehat{\Omega}_s^{-1} & 0 \end{pmatrix} \right] \\ &= \operatorname{tr}(S_{11} \widehat{\Omega}_s^{-1} V_{11} \widehat{\Omega}_s^{-1}) + \operatorname{tr}(S_{12} \widehat{\Omega}_{n-s}^{-1} V_{12}^T \widehat{\Omega}_s^{-1}) - 2 \operatorname{tr}(S_{12}^T \widehat{\Omega}_s^{-1} V_{12} W_{12}^T) \\ &\quad - 2 \operatorname{tr}(V_{11} P_{11} \widehat{\Omega}_s^{-1}) - 2 \operatorname{tr}(V_{12} P_{12}^T \widehat{\Omega}_s^{-1}) - 4 \operatorname{tr}(V_{12}^T P_{12} W_{22}). \end{aligned}$$

Since S_{11} is a skew-symmetric matrix and V_{11} , $\widehat{\Omega}_s^{-1}$ are symmetric matrices, we have

$$\begin{aligned} \operatorname{tr}(S_{11} \widehat{\Omega}_s^{-1} V_{11} \widehat{\Omega}_s^{-1}) &= \frac{1}{2} \operatorname{tr}(S_{11} \widehat{\Omega}_s^{-1} V_{11} \widehat{\Omega}_s^{-1}) + \frac{1}{2} \operatorname{tr}((S_{11} \widehat{\Omega}_s^{-1} V_{11} \widehat{\Omega}_s^{-1})^T) \\ &= \frac{1}{2} \operatorname{tr}(\widehat{\Omega}_s^{-1} V_{11} \widehat{\Omega}_s^{-1} S_{11}) + \frac{1}{2} \operatorname{tr}(\widehat{\Omega}_s^{-1} V_{11} \widehat{\Omega}_s^{-1} S_{11}^T) \\ &= \frac{1}{2} \operatorname{tr}(\widehat{\Omega}_s^{-1} V_{11} \widehat{\Omega}_s^{-1} S_{11}) - \frac{1}{2} \operatorname{tr}(\widehat{\Omega}_s^{-1} V_{11} \widehat{\Omega}_s^{-1} S_{11}) = 0. \end{aligned}$$

So, we get

$$\begin{aligned} \operatorname{tr} \left(\frac{d^2}{d\mu^2} Y_\mu \right) \Big|_{\mu=0} &= \operatorname{tr}(S_{12} \widehat{\Omega}_{n-s}^{-1} V_{12}^T \widehat{\Omega}_s^{-1}) - 2 \operatorname{tr}(S_{12}^T \widehat{\Omega}_s^{-1} V_{12} W_{12}^T) - 2 \operatorname{tr}(V_{11} P_{11} \widehat{\Omega}_s^{-1}) \\ &\quad - 2 \operatorname{tr}(V_{12} P_{12}^T \widehat{\Omega}_s^{-1}) - 4 \operatorname{tr}(V_{12}^T P_{12} W_{22}). \end{aligned} \quad (56)$$

Now we consider the (1, 2)-block of the block-matrix equations corresponding to (49)–(52). We obtain the following four equations:

$$-\widehat{\Omega}_s Q_{21}^T - Q_{12} \widehat{\Omega}_{n-s} = 0 \quad (57)$$

$$-\widehat{\Omega}_s P_{12} + R_{12} \widehat{\Omega}_{n-s} - 2Q_{12} H = \frac{1}{2} \widehat{\Omega}_s^{-1} V_{12} \quad (58)$$

$$\widehat{\Omega}_s R_{12} - 2\widehat{\Omega}_s Q_{21}^T - P_{12} \widehat{\Omega}_{n-s} = 0 \quad (59)$$

$$\widehat{\Omega}_s Q_{12} - 2\widehat{\Omega}_s P_{12} + Q_{21}^T \widehat{\Omega}_{n-s} - 2P_{12} H = \frac{1}{2} \widehat{\Omega}_s^{-1} V_{12}. \quad (60)$$

From (55), we conclude that

$$Q_{12} = \frac{1}{2} S_{12} \widehat{\Omega}_{n-s}^{-1}, \quad Q_{21} = -\frac{1}{2} S_{12}^T \widehat{\Omega}_s^{-1}, \quad Q_{21}^T = -\frac{1}{2} \widehat{\Omega}_s^{-1} S_{12}, \quad (61)$$

so (57) is satisfied. By putting this into (59), we get $\widehat{\Omega}_s R_{12} + S_{12} - P_{12} \widehat{\Omega}_{n-s} = 0$, i.e.,

$$S_{12} = P_{12} \widehat{\Omega}_{n-s} - \widehat{\Omega}_s R_{12}. \quad (62)$$

Now we put (61) into equations (58) and (60), and obtain

$$\begin{aligned} -\widehat{\Omega}_s P_{12} + R_{12} \widehat{\widehat{\Omega}}_{n-s} - S_{12} \widehat{\widehat{\Omega}}_{n-s}^{-1} H &= \frac{1}{2} \widehat{\Omega}_s^{-1} V_{12}, \\ \frac{1}{2} \widehat{\Omega}_s S_{12} \widehat{\widehat{\Omega}}_{n-s}^{-1} - 2\widehat{\Omega}_s P_{12} - \frac{1}{2} \widehat{\Omega}_s^{-1} S_{12} \widehat{\widehat{\Omega}}_{n-s} - 2P_{12} H &= \frac{1}{2} \widehat{\Omega}_s^{-1} V_{12}. \end{aligned}$$

By replacing S_{12} with (62) in these equations and subtracting them, we get

$$\frac{1}{2} \widehat{\Omega}_s P_{12} + \frac{1}{2} R_{12} \widehat{\widehat{\Omega}}_{n-s} + P_{12} H + \widehat{\Omega}_s R_{12} \widehat{\widehat{\Omega}}_{n-s}^{-1} H + \frac{1}{2} \widehat{\Omega}_s^2 R_{12} \widehat{\widehat{\Omega}}_{n-s}^{-1} + \frac{1}{2} \widehat{\Omega}_s^{-1} P_{12} \widehat{\widehat{\Omega}}_{n-s}^2 = 0.$$

This equation can be written as

$$\Delta H + \frac{1}{2} \widehat{\Omega}_s \Delta + \frac{1}{2} \widehat{\Omega}_s^{-1} \Delta \widehat{\widehat{\Omega}}_{n-s}^2 = 0, \quad (63)$$

where $\Delta = P_{12} + \widehat{\Omega}_s R_{12} \widehat{\widehat{\Omega}}_{n-s}^{-1}$. Equation (63) has only the trivial solution $\Delta = 0$. To see this, postmultiply (63) by Δ^T and take the trace:

$$\text{tr}(\Delta H \Delta^T) + \frac{1}{2} \text{tr}(\widehat{\Omega}_s \Delta \Delta^T) + \frac{1}{2} \text{tr}(\widehat{\Omega}_s^{-1} \Delta \widehat{\widehat{\Omega}}_{n-s}^2 \Delta^T) = 0,$$

i.e.,

$$\begin{aligned} \text{tr}(\Delta H \Delta^T) + \frac{1}{2} \text{tr}(\Delta^T \widehat{\Omega}_s \Delta) + \frac{1}{2} \text{tr}(\widehat{\Omega}_s^{-1/2} \Delta \widehat{\widehat{\Omega}}_{n-s}^2 \Delta^T \widehat{\Omega}_s^{-1/2}) \\ = \text{tr}[(\Delta H^{1/2})(\Delta H^{1/2})^T] + \frac{1}{2} \text{tr}[(\Delta^T \widehat{\Omega}_s^{1/2})(\Delta^T \widehat{\Omega}_s^{1/2})^T] \\ + \frac{1}{2} \text{tr}[(\widehat{\Omega}_s^{-1/2} \Delta \widehat{\widehat{\Omega}}_{n-s})(\widehat{\Omega}_s^{-1/2} \Delta \widehat{\widehat{\Omega}}_{n-s})^T] = 0. \end{aligned}$$

Every single term is greater or equal zero. Because of the positive definiteness of the matrix $\widehat{\Omega}_s$, this equation is valid only for $\Delta = 0$.

So, we have $P_{12} + \widehat{\Omega}_s R_{12} \widehat{\widehat{\Omega}}_{n-s}^{-1} = 0$, or $P_{12} = -\widehat{\Omega}_s R_{12} \widehat{\widehat{\Omega}}_{n-s}^{-1}$. Putting this into (62), for S_{12} we get

$$S_{12} = P_{12} \widehat{\widehat{\Omega}}_{n-s} - \widehat{\Omega}_s R_{12} = -\widehat{\Omega}_s R_{12} - \widehat{\Omega}_s R_{12} = -2\widehat{\Omega}_s R_{12}.$$

From (61), we have

$$Q_{12} = \frac{1}{2} S_{12} \widehat{\widehat{\Omega}}_{n-s}^{-1} = -\widehat{\Omega}_s R_{12} \widehat{\widehat{\Omega}}_{n-s}^{-1} = P_{12} \quad (64)$$

$$Q_{21} = -\frac{1}{2} S_{12}^T \widehat{\widehat{\Omega}}_{n-s}^{-1} = R_{12}^T \widehat{\Omega}_s \widehat{\widehat{\Omega}}_{n-s}^{-1} = R_{12}^T, \quad (65)$$

and

$$\widetilde{Y}_{12} = \begin{pmatrix} Q_{11} & P_{12} \\ R_{12}^T & 0 \end{pmatrix}.$$

Substitution $P_{12} = \frac{1}{2}S_{12}\widehat{\Omega}_{n-s}^{-1}$ from (64) into (56) leads to

$$\begin{aligned} \operatorname{tr} \left(\frac{d^2}{d\mu^2} Y_\mu \right) \Big|_{\mu=0} &= \operatorname{tr}(V_{12}\widehat{\Omega}_{n-s}^{-1}S_{12}^T\widehat{\Omega}_s^{-1}) - 2\operatorname{tr}(S_{12}^T\widehat{\Omega}_s^{-1}V_{12}W_{12}^T) - 2\operatorname{tr}(V_{11}P_{11}\widehat{\Omega}_s^{-1}) \\ &\quad - \operatorname{tr}(V_{12}\widehat{\Omega}_{n-s}^{-1}S_{12}^T\widehat{\Omega}_s^{-1}) - 2\operatorname{tr}(V_{12}^T S_{12}\widehat{\Omega}_{n-s}^{-1}W_{22}) \\ &= -2\operatorname{tr}(V_{11}P_{11}\widehat{\Omega}_s^{-1}) - 2\operatorname{tr}(V_{12}^T\widehat{\Omega}_s^{-1}S_{12}W_{12}) \\ &\quad - 2\operatorname{tr}(V_{12}^T S_{12}\widehat{\Omega}_{n-s}^{-1}W_{22}). \end{aligned} \quad (66)$$

To complete the proof, we analyze the terms on the right-hand side of (66). For the first term we will show that

$$-2\operatorname{tr}(V_{11}P_{11}\widehat{\Omega}_s^{-1}) \geq 0,$$

and that equality is possible only in the case $V_{11} = 0$.

We have to evaluate P_{11} , so we consider the $(1,1)$ -block of the block-matrix equations corresponding to (49)–(52). We obtain the following four equations:

$$-\widehat{\Omega}_s Q_{11}^T - Q_{11}\widehat{\Omega}_s = 0 \quad (67)$$

$$-\widehat{\Omega}_s P_{11} + R_{11}\widehat{\Omega}_s - 2Q_{11}\widehat{\Omega}_s = \frac{1}{2}\widehat{\Omega}_s^{-1}V_{11} \quad (68)$$

$$\widehat{\Omega}_s R_{11} - 2\widehat{\Omega}_s Q_{11}^T - P_{11}\widehat{\Omega}_s = \frac{1}{2}V_{11}\widehat{\Omega}_s^{-1} \quad (69)$$

$$\widehat{\Omega}_s Q_{11} - 2\widehat{\Omega}_s P_{11} + Q_{11}^T\widehat{\Omega}_s - 2P_{11}\widehat{\Omega}_s = \frac{1}{2}\widehat{\Omega}_s^{-1}V_{11} + \frac{1}{2}V_{11}\widehat{\Omega}_s^{-1}. \quad (70)$$

Note that (68) and (69) are adjoint to each other. Furthermore, from (53)

$$Q_{11} = \frac{1}{2}S_{11}\widehat{\Omega}_s^{-1},$$

where S_{11} is a skew-symmetric matrix, so (67) is satisfied. Substitution of Q_{11} into equations (68) and (70) yields

$$-\widehat{\Omega}_s P_{11} + R_{11}\widehat{\Omega}_s - S_{11} = \frac{1}{2}\widehat{\Omega}_s^{-1}V_{11} \quad (71)$$

$$\frac{1}{2}(\widehat{\Omega}_s S_{11}\widehat{\Omega}_s^{-1} - \widehat{\Omega}_s^{-1}S_{11}\widehat{\Omega}_s) - 2(\widehat{\Omega}_s P_{11} + P_{11}\widehat{\Omega}_s) = \frac{1}{2}(\widehat{\Omega}_s^{-1}V_{11} + V_{11}\widehat{\Omega}_s^{-1}). \quad (72)$$

We solve them componentwise, using that P_{11} , R_{11} and V_{11} are symmetric matrices, S_{11} is skew-symmetric, and $\widehat{\Omega}_s$ is diagonal, i.e., $(P_{11})_{ij} = (P_{11})_{ji}$, $(R_{11})_{ij} = (R_{11})_{ji}$, $(V_{11})_{ij} = (V_{11})_{ji}$, $(S_{11})_{ij} = -(S_{11})_{ji}$, and $(\widehat{\Omega}_s)_{ij} = \omega_i\delta_{ij}$, for $i, j = 1, \dots, s$. From (71), we obtain the following equations in positions (i, j) and (j, i) , respectively:

$$\begin{aligned} -\omega_i(P_{11})_{ij} + (R_{11})_{ij}\omega_j - (S_{11})_{ij} &= \frac{(V_{11})_{ij}}{2\omega_i}, \\ -\omega_j(P_{11})_{ij} + (R_{11})_{ij}\omega_i + (S_{11})_{ij} &= \frac{(V_{11})_{ij}}{2\omega_j}. \end{aligned}$$

Multiplying the first equation with ω_i and the second one with ω_j and then subtracting them, gives

$$(\omega_j^2 - \omega_i^2)(P_{11})_{ij} - (\omega_i + \omega_j)(S_{11})_{ij} = 0,$$

or, since $\omega_i > 0$,

$$(S_{11})_{ij} = (\omega_j - \omega_i)(P_{11})_{ij}. \quad (73)$$

From (72), we have the following equation in position (i, j) :

$$\frac{1}{2} \left(\frac{\omega_i}{\omega_j} - \frac{\omega_j}{\omega_i} \right) (S_{11})_{ij} - 2(\omega_i + \omega_j)(P_{11})_{ij} = \frac{1}{2} \left(\frac{1}{\omega_i} + \frac{1}{\omega_j} \right) (V_{11})_{ij}.$$

Multiplication by $-2\omega_i\omega_j/(\omega_i + \omega_j)$ yields

$$(\omega_j - \omega_i)(S_{11})_{ij} + 4\omega_i\omega_j(P_{11})_{ij} = -(V_{11})_{ij}.$$

By using (73), we get

$$(P_{11})_{ij} = -\frac{(V_{11})_{ij}}{(\omega_i + \omega_j)^2} = (P_{11})_{ji}.$$

Therefore, we have

$$-2 \operatorname{tr}(V_{11}P_{11}\widehat{\Omega}_s^{-1}) = -2 \sum_{i=1}^s \sum_{j=1}^s \frac{(V_{11})_{ij}(P_{11})_{ji}}{\omega_i} = 2 \sum_{i=1}^s \sum_{j=1}^s \frac{(V_{11})_{ij}^2}{\omega_i(\omega_i + \omega_j)^2} \geq 0, \quad (74)$$

where $-2 \operatorname{tr}(V_{11}P_{11}\widehat{\Omega}_s^{-1}) = 0$ is possible only for $V_{11} = 0$.

Finally, we consider the last two terms on the right-hand side of (66); we will show that

$$\widehat{\Delta} := -2 \operatorname{tr}(V_{12}^T \widehat{\Omega}_s^{-1} S_{12} W_{12}) - 2 \operatorname{tr}(V_{12}^T S_{12} \widehat{\Omega}_{n-s}^{-1} W_{22}) \geq 0,$$

and that equality is possible only when $V_{12} = 0$.

By (64) and (65) we have

$$P_{12} = Q_{12} = \frac{1}{2} S_{12} \widehat{\Omega}_{n-s}^{-1}, \quad R_{12} = -\frac{1}{2} \widehat{\Omega}_s^{-1} S_{12}.$$

Putting this into equation (58), we get

$$V_{12}^T = -\widehat{\Omega}_{n-s}^{-1} S_{12}^T \widehat{\Omega}_s^2 - \widehat{\Omega}_{n-s} S_{12}^T - 2H \widehat{\Omega}_{n-s}^{-1} S_{12}^T \widehat{\Omega}_s. \quad (75)$$

Hence

$$\begin{aligned} \widehat{\Delta} &= 2 \operatorname{tr}(\widehat{\Omega}_{n-s}^{-1} S_{12}^T \widehat{\Omega}_s S_{12} W_{12}) + 2 \operatorname{tr}(\widehat{\Omega}_{n-s} S_{12}^T \widehat{\Omega}_s^{-1} S_{12} W_{12}) + 4 \operatorname{tr}(H \widehat{\Omega}_{n-s}^{-1} S_{12}^T S_{12} W_{12}) \\ &\quad + 2 \operatorname{tr}(\widehat{\Omega}_{n-s}^{-1} S_{12}^T \widehat{\Omega}_s^2 S_{12} \widehat{\Omega}_{n-s}^{-1} W_{22}) + 2 \operatorname{tr}(\widehat{\Omega}_{n-s} S_{12}^T S_{12} \widehat{\Omega}_{n-s}^{-1} W_{22}) \\ &\quad + 4 \operatorname{tr}(H \widehat{\Omega}_{n-s}^{-1} S_{12}^T \widehat{\Omega}_s S_{12} \widehat{\Omega}_{n-s}^{-1} W_{22}) \\ &= 2 \operatorname{tr}(\widehat{\Omega}_{n-s}^{-1} S_{12}^T \widehat{\Omega}_s S_{12} W_{12}) + 2 \operatorname{tr}(\widehat{\Omega}_{n-s} S_{12}^T \widehat{\Omega}_s^{-1} S_{12} W_{12}) + 2 \operatorname{tr}[\widehat{\Omega}_{n-s}^{-1} S_{12}^T S_{12} (2W_{12}H)] \\ &\quad + 2 \operatorname{tr}(\widehat{\Omega}_{n-s}^{-1} S_{12}^T \widehat{\Omega}_s^2 S_{12} \widehat{\Omega}_{n-s}^{-1} W_{22}) + 2 \operatorname{tr}(\widehat{\Omega}_{n-s} S_{12}^T S_{12} \widehat{\Omega}_{n-s}^{-1} W_{22}) \\ &\quad + \operatorname{tr}[(2HW_{22} + 2W_{22}H) \widehat{\Omega}_{n-s}^{-1} S_{12}^T \widehat{\Omega}_s S_{12} \widehat{\Omega}_{n-s}^{-1}]. \end{aligned} \quad (76)$$

The block decomposition of the Lyapunov equation (46) for W is

$$-\widehat{\Omega}_{n-s}W_{12}^T - W_{12}\widehat{\Omega}_{n-s} = -I_{n-s} \quad (77)$$

$$-\widehat{\Omega}_{n-s}W_{22} + W_{11}\widehat{\Omega}_{n-s} - 2W_{12}H = 0 \quad (78)$$

$$\widehat{\Omega}_{n-s}W_{11} - 2HW_{12}^T - W_{22}\widehat{\Omega}_{n-s} = 0 \quad (79)$$

$$\widehat{\Omega}_{n-s}W_{12} - 2HW_{22} + W_{12}^T\widehat{\Omega}_{n-s} - 2W_{22}H = -I_{n-s}, \quad (80)$$

where (78) and (79) are adjoint to each other. From (77), we can write W_{12} as

$$W_{12} = \frac{1}{2}\widehat{\Omega}_{n-s}^{-1} + G\widehat{\Omega}_{n-s}^{-1}, \quad (81)$$

where G is skew-symmetric, i.e., $G^T = -G$. Furthermore, from (78) and (80), we get

$$2W_{12}H = -\widehat{\Omega}_{n-s}W_{22} + W_{11}\widehat{\Omega}_{n-s} \quad (82)$$

$$2HW_{22} + 2W_{22}H = \widehat{\Omega}_{n-s}W_{12} + W_{12}^T\widehat{\Omega}_{n-s} + I_{n-s}. \quad (83)$$

Placing (81), (82) and (83) into (76), we obtain

$$\begin{aligned} \widehat{\Delta} &= 2 \operatorname{tr}(\widehat{\Omega}_{n-s}^{-1}S_{12}^T\widehat{\Omega}_sS_{12}W_{12}) + \operatorname{tr}(\widehat{\Omega}_{n-s}S_{12}^T\widehat{\Omega}_s^{-1}S_{12}\widehat{\Omega}_{n-s}^{-1}) \\ &\quad + 2 \operatorname{tr}(\widehat{\Omega}_{n-s}S_{12}^T\widehat{\Omega}_s^{-1}S_{12}G\widehat{\Omega}_{n-s}^{-1}) - 2 \operatorname{tr}(\widehat{\Omega}_{n-s}^{-1}S_{12}^T\widehat{\Omega}_sS_{12}\widehat{\Omega}_{n-s}W_{22}) \\ &\quad + 2 \operatorname{tr}(\widehat{\Omega}_{n-s}^{-1}S_{12}^T\widehat{\Omega}_sS_{12}W_{11}\widehat{\Omega}_{n-s}) + 2 \operatorname{tr}(\widehat{\Omega}_{n-s}^{-1}S_{12}^T\widehat{\Omega}_s^2S_{12}\widehat{\Omega}_{n-s}^{-1}W_{22}) \\ &\quad + 2 \operatorname{tr}(\widehat{\Omega}_{n-s}S_{12}^T\widehat{\Omega}_s\widehat{\Omega}_{n-s}^{-1}W_{22}) + \operatorname{tr}(\widehat{\Omega}_{n-s}W_{12}\widehat{\Omega}_{n-s}^{-1}S_{12}^T\widehat{\Omega}_sS_{12}\widehat{\Omega}_{n-s}^{-1}) \\ &\quad + \operatorname{tr}(W_{12}^T\widehat{\Omega}_{n-s}\widehat{\Omega}_{n-s}^{-1}S_{12}^T\widehat{\Omega}_sS_{12}\widehat{\Omega}_{n-s}^{-1}) + \operatorname{tr}(\widehat{\Omega}_{n-s}^{-1}S_{12}^T\widehat{\Omega}_sS_{12}\widehat{\Omega}_{n-s}^{-1}) \\ &= 4 \operatorname{tr}(\widehat{\Omega}_{n-s}^{-1}S_{12}^T\widehat{\Omega}_sS_{12}W_{12}) + \operatorname{tr}(S_{12}^T\widehat{\Omega}_s^{-1}S_{12}) + 2 \operatorname{tr}(S_{12}^T\widehat{\Omega}_s^{-1}S_{12}G) \\ &\quad + 2 \operatorname{tr}(S_{12}^T\widehat{\Omega}_sS_{12}W_{11}) + 2 \operatorname{tr}(\widehat{\Omega}_{n-s}^{-1}S_{12}^T\widehat{\Omega}_s^2S_{12}\widehat{\Omega}_{n-s}^{-1}W_{22}) + \operatorname{tr}(\widehat{\Omega}_{n-s}^{-1}S_{12}^T\widehat{\Omega}_sS_{12}\widehat{\Omega}_{n-s}^{-1}). \end{aligned}$$

Since G is skew-symmetric, and $S_{12}^T\widehat{\Omega}_s^{-1}S_{12}$ is symmetric, we have

$$\operatorname{tr}(S_{12}^T\widehat{\Omega}_s^{-1}S_{12}G) = \operatorname{tr}(G^T S_{12}^T\widehat{\Omega}_s^{-1}S_{12}) = -\operatorname{tr}(S_{12}^T\widehat{\Omega}_s^{-1}S_{12}G) = 0,$$

As $\widehat{\Omega}_s$ is positive semidefinite, so are $S_{12}^T\widehat{\Omega}_s^{-1}S_{12}$ and $\widehat{\Omega}_{n-s}^{-1}S_{12}^T\widehat{\Omega}_sS_{12}\widehat{\Omega}_{n-s}^{-1}$, and we get

$$\operatorname{tr}(S_{12}^T\widehat{\Omega}_s^{-1}S_{12}) \geq 0, \quad \operatorname{tr}(\widehat{\Omega}_{n-s}^{-1}S_{12}^T\widehat{\Omega}_sS_{12}\widehat{\Omega}_{n-s}^{-1}) \geq 0,$$

with equality if and only if $S_{12} = 0$, i.e., (see (75)) $V_{12} = 0$. It remains to prove

$$\widehat{\Delta} := 4 \operatorname{tr}(\widehat{\Omega}_{n-s}^{-1}S_{12}^T\widehat{\Omega}_sS_{12}W_{12}) + 2 \operatorname{tr}(S_{12}^T\widehat{\Omega}_sS_{12}W_{11}) + 2 \operatorname{tr}(\widehat{\Omega}_{n-s}^{-1}S_{12}^T\widehat{\Omega}_s^2S_{12}\widehat{\Omega}_{n-s}^{-1}W_{22}) \geq 0.$$

The matrix W from (45) is positive definite, as a solution of (46). We consider the following congruence transformation of W :

$$\widetilde{W} := \begin{pmatrix} \widetilde{W}_{11} & \widetilde{W}_{12} \\ \widetilde{W}_{12}^T & \widetilde{W}_{22} \end{pmatrix} := \begin{pmatrix} S_{12} & 0 \\ 0 & \widehat{\Omega}_s S_{12} \widehat{\Omega}_{n-s}^{-1} \end{pmatrix} \begin{pmatrix} W_{11} & W_{12} \\ W_{12}^T & W_{22} \end{pmatrix} \begin{pmatrix} S_{12} & 0 \\ 0 & \widehat{\Omega}_s S_{12} \widehat{\Omega}_{n-s}^{-1} \end{pmatrix}^T$$

\widetilde{W} is positive semidefinite, and, in particular, for all s -dimensional vectors x , we have

$$(x^T \ x^T) \begin{pmatrix} \widetilde{W}_{11} & \widetilde{W}_{12} \\ \widetilde{W}_{12}^T & \widetilde{W}_{22} \end{pmatrix} \begin{pmatrix} x \\ x \end{pmatrix} = x^T (\widetilde{W}_{11} + \widetilde{W}_{12}^T + \widetilde{W}_{12} + \widetilde{W}_{22}) x \geq 0.$$

So, the matrix $\widetilde{W}_{11} + \widetilde{W}_{12} + \widetilde{W}_{12}^T + \widetilde{W}_{22}$ is symmetric, positive semidefinite, and has a nonnegative trace. We conclude that

$$\begin{aligned} 0 &\leq \text{tr}(\widetilde{W}_{11} + \widetilde{W}_{12}^T + \widetilde{W}_{12} + \widetilde{W}_{22}) = \text{tr} \widetilde{W}_{11} + 2 \text{tr} \widetilde{W}_{12} + \text{tr} \widetilde{W}_{22} \\ &= \text{tr}(S_{12} W_{11} S_{12}^T) + 2 \text{tr}(S_{12} W_{12} \widehat{\Omega}_{n-s}^{-1} S_{12}^T \widehat{\Omega}_s) + \text{tr}(\widehat{\Omega}_s S_{12} \widehat{\Omega}_{n-s}^{-1} W_{22} \widehat{\Omega}_{n-s}^{-1} S_{12}^T \widehat{\Omega}_s) = \frac{\widehat{\Delta}}{2}. \end{aligned}$$

So, we have proved that $\widehat{\Delta} \geq 0$, where equality is possible only for $V_{12} = 0$. Together with (74), this proves (33). \blacksquare

References

- [1] R. H. BARTELS AND G. W. STEWART, *A solution of the matrix equation $AX + XB = C$* , Comm. ACM, 15 (1972), pp. 820–826.
- [2] K. BRABENDER, *Optimale Dämpfung von Linearen Schwingungs-Systemen*, Ph.D. thesis, Fern-Universität Hagen, 1998.
- [3] S. J. COX, *Designing of optimal energy absorption, I: lumped parameter systems*, J. Vib. and Acoustics, to appear, (1996).
- [4] P. FREITAS AND P. LANCASTER, *Optimal decay of energy for a system of linear oscillators*, preprint, December 1997.
- [5] I. GOHBERG, P. LANCASTER, AND L. RODMAN, *Matrices and Indefinite Scalar Products*, Birkhäuser, Basel, 1983.
- [6] G. H. GOLUB AND CH. F. VAN LOAN, *Matrix Computations*, 2nd ed., Johns Hopkins University Press, Baltimore, 1989.
- [7] P. LANCASTER, *Lambda-matrices and vibrating systems*, Pergamon Press, Oxford, 1966.
- [8] P. C. MÜLLER AND W. O. SCHIEHLEN, *Lineare Schwingungen*, Akademische Verlagsgesellschaft, Wiesbaden, 1976.
- [9] K. VESELIĆ, *Some convergent Jacobi-like procedures for diagonalizing J -symmetric matrices*, Numer. Math., 27 (1976), pp. 67–75.
- [10] K. VESELIĆ, *Global Jacobi method for a symmetric indefinite problem $Sx = \lambda Tx$* , Comput. Methods Appl. Mech. Engrg., 38 (1983), pp. 273–290.
- [11] K. VESELIĆ, *On linear vibrational systems with one dimensional damping*, Appl. Anal., 29 (1988), pp. 1–18.
- [12] K. VESELIĆ, *On linear vibrational systems with one dimensional damping II*, Integral Equations Operator Theory, 13 (1990), pp. 883–897.
- [13] K. VESELIĆ, *Exponential decay of semigroups in Hilbert space*, Semigroup Forum, 55 (1997), pp. 325–331.
- [14] K. VESELIĆ, *Estimating the operator exponential*, Linear Algebra Appl., 280 (1998), pp. 241–244.