Passive control of linear systems

K. Veselić[*], K. Brabender[‡] and K. Delinić[§]

April 18, 2000

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 Ljapunov equation in the underlying phase space equipped with the energy norm. We compare our method with other common approaches.

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

[*]Lehrgebiet Mathematische Physik, FernUniversität Hagen, Postf. 940, 58084 Hagen, Germany, email: kresimir.veselic@fernuni-hagen.de

[†]Lehrgebiet Mathematische Physik, FernUniversität Hagen, email: katrin.brabender@fernuni-hagen.de

[‡]Parts of this paper and in particular the proof of Theorem 6.1 in Appendix A are contained in the Ph. D. thesis of the second author under the supervision of the first author.

[§]Siemens AG, Erlangen Germany

in 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. Usual way of doing this is (cf. [8]) to require

$$max \Re \lambda_k = min \ , \tag{1.1}$$

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

Criteria like (1.1) 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 minimise the total energy integral

$$\int_0^\infty E(t)dt = min. \tag{1.2}$$

This criterion leads to the use of the Ljapunov 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 Ljapunov 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 (1.1). On the other hand, it has recently been shown that the solution of the Ljapunov equation provides rigorous bounds to the energy decay of a vibrating system ([13], [14]). The program we propose here has been completely realised 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 Ljapunov 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 repeated solving the Ljapunov equation and must see that the computation time does not become prohibitive. We propose three means to cope with this challenge: (i) using the special structure of our penalty function we derive simple efficient formulae for the gradient and the Hessian which greatly reduce the number of Ljapunov 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.

[1] Sometimes the optimisation 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.

Both (ii) and (iii) are not fully understood as 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 to modal dampings only.

The paper is organized as follows. In Setion 1 we derive the phase space Ljapunov model and the penalty function. We show that the Ljapunov 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 formulae 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 6.1, 6.3 and 6.4). In Section 7 we compare our penalty function with another one, the spectral abscissa. On examples we show that the two may differ significantly. In Section 8 we present a realistic computation on a practical example. The Appendix contains a rather lengthy proof of Theorem 6.1.

The methods presented in this paper have been succesfully implemented to optimise 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

$$M\ddot{x} + C\dot{x} + Kx = 0 \tag{2.1}$$

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 semidefinite. [2] 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* (Fig. 2.1) where

$$M = diag(m_1, m_2, \ldots, m_n) \ , \tag{2.2}$$

[2] In 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 non-singular M at least if the null-space of M is contained in the one of C.

$$K = \begin{pmatrix} k_0+k_1 & -k_1 & & & \\ -k_1 & k_1+k_2 & -k_2 & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & -k_{n-1} \\ & & & -k_{n-1} & k_{n-1}+k_n \end{pmatrix} , \tag{2.3}$$

$$C = \sum_k c_k e_k e_k^T , \tag{2.4}$$

where e_k is the k-th canonical basis vector.

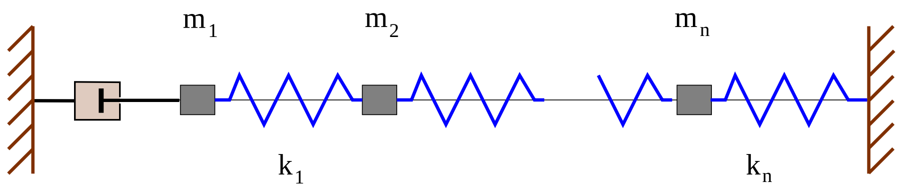

Figure 2.1: The n-mass oscillator with one damper

Here $m_i > 0$ are the masses, $k_i > 0$ the spring constants or stiffnesses and $c_i > 0$ are the damping constant of the damper applied to the mass m_i (on Fig. 2.1 with $k_0 = 0$). Note that in the figure the rank of C is one and we call such damping *one-dimensional.* [3]

To (2.1) there corresponds the eigenvalue equation

$$(\lambda^2 M + \lambda C + K)x = 0 \ . \tag{2.5}$$

Obviously all eigenvalues of (2.5) lie in the left complex plane. We now go over to the $2n$ -dimensional phase space by taking factors

$$K = \Phi_1 \Phi_1^T \ , \ \ M = \Phi_2 \Phi_2^T \tag{2.6}$$

and setting

$$y_1 = \Phi_1^T x \ , \ \ y_2 = \Phi_2^T \dot{x} \ . \tag{2.7}$$

Then (2.1) is directly seen to be equivalent to

$$\dot{\psi} = A\psi \ , \tag{2.8}$$

[3] In 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.

$$\psi = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, \quad A = \begin{pmatrix} 0 & \Phi_1^T \Phi_2^{-T} \\ \\ -\Phi_2^{-1}\Phi_1 & -\Phi_2^{-1} C \Phi_2^{-T} \end{pmatrix}, \tag{2.9}$$

with the solution

$$\psi = e^{At}\psi_0, \quad \psi_0 \quad \text{initial data.}$$

The factors $\Phi_{1,2}$ 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 (2.5). Moreover,

$$\psi^T\psi = \|\psi\|^2 = x^T K x + \dot{x}^T M \dot{x} = 2E(t) \ . \tag{2.10}$$

In the other words, the Euclidian 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":

$$A^T = JAJ \, , J = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}, \tag{2.11}$$

which is directly verified.

Now, (1.2) can be written as

$$\psi_0^T X \psi_0 = min \ , \tag{2.12}$$

where

$$X = \int_0^\infty e^{A^T t} e^{At} dt \tag{2.13}$$

solves the Ljapunov equation

$$A^T X + XA = -I \tag{2.14}$$

and

$$\psi = e^{At}\psi_0 \tag{2.15}$$

is the solution of (2.8), with ψ_0 as initial data. The inconvenience of the criterion (2.12) 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 (2.12) we require

$$\int_{\|\psi_0\|=1} \psi_0^T X \psi_0 d\sigma = min \tag{2.16}$$

where $d\sigma$ is a chosen probability measure on the unit sphere $S^{2n} = \{\psi_0 \in \mathbb{R}^{2n}; \|\psi_0\| = 1\}$. As it is easily seen (2.16) is equivalent to

$$Tr(ZX) = Tr[Z \int_0^\infty e^{A^T t} e^{At} dt] = min \tag{2.17}$$

where Z is a symmetric positive semidefinite matrix (induced by $d\sigma$) which may be normalized to have unit trace. Using (2.11) we obtain another interesting formula for the trace

$$Tr(ZX) = TrY \tag{2.18}$$

where Y is the solution of another Ljapunov equation

$$A^T Y + YA = -Z_1 \ , \ \ Z_1 = JZJ\,.^4 \tag{2.19}$$

Using the well-known "frequency domain formula" ([8])

$$X = \frac{1}{2\pi}\int_{-\infty}^{\infty}(-i\omega - A^T)^{-1}(i\omega - A)^{-1}d\omega \tag{2.20}$$

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

$$x\exp(i\omega t) \ , \ \ x = (i\omega - A)^{-1}f \ . \tag{2.21}$$

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

As a simple illustration consider the one-mass oscillator. Here

$$A = \begin{pmatrix} 0 & \sqrt{k} \\ -\sqrt{k} & -c \end{pmatrix} \ , \ \ k = k_1 \ . \tag{2.22}$$

Now, $TrX = \frac{2}{c} + \frac{c}{2k}$ takes its minimum at $c = 2\sqrt{k}$ and nowhere else. This is the known case of critical damping. Note that here the criterion (1.1) yields the same value of c. For a general system with one dimensional damping $C = cww^T$, $\|w\| = 1$[5] one can show ([12]) that

$$X = X_1 c + X_2/c + X_3 \tag{2.23}$$

where X_1, X_2 can be directly computed from the eigensolution of the undamped problem i.e. from (2.5) with $C = 0$. This formula, which is valid whenever C has rank one, leads to an extremely simple minimisation of $TrZX$ which is always a hyperbola in c. This avoids the use of time consuming algorithms for the Ljapunov equation like the one in [1]. The only eigendecoupling needed is the one for the undamped system and is done only once ([12]).

As is directly seen, the optimal A in (2.22) has no eigenbasis due to nonlinear elementary divisors. This prevents the general use of the eigendecomposition of A to optimise viscous damping. [6] For large systems we have

[4] This is a special case of the general properties: if $A^TX + XA = -B$ and $AY + YA^T = -Z$ then $Tr(ZX) = Tr(BY)$, $JAJ = A^T$, $Tr(Y) = Tr(JYJ)$

[5] This is the case e.g. if in (2.4) only one of c_i is different from zero.

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

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 (2.9) $\Phi_2^{-1}\Phi_1 = \Omega = diag(\omega_1, \ldots, \omega_n)$ 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 (2.5) with $C = 0$. Here ω_i are the undamped circular frequencies and the columns of Φ are the mass-normalised eigenforms:

$$K\Phi = M\Phi\Omega^2 \, , \Phi^T M \Phi = I \, . \tag{2.24}$$

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

$$\psi = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \quad , \quad A = \begin{pmatrix} 0 & \Omega \\ -\Omega & -\hat{C} \end{pmatrix} , \hat{C} = \Phi^T C \Phi \ . \tag{2.25}$$

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 $\omega \le \omega_{max}$ is obtained by choosing

$$Z = Z_s = \begin{pmatrix} I_s & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & I_s & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} , \tag{2.26}$$

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

We will now show that thus obtained X, $\zeta = TrZX$ nicely control the time history of the vibrational system. Indeed, according to [14] for any unit initial data Ψ_0 we have

$$\|e^{At}\Psi_0\|^2 \le \frac{a}{h}\left(1 + \frac{h}{\|X\|}\right)^{2-\frac{t}{h}} \tag{2.27}$$

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

$$a = \Psi_0^T X \Psi_0, \quad \|X\| \ge a \ .$$

For $t > a$ we may set $h = a$ and (2.27) gives

$$\|e^{At}\Psi_0\|^2 \le \left(1 + \frac{a}{\|X\|}\right)^{2-\frac{t}{a}} . \tag{2.28}$$

Let now Ψ_0 be from the subspace spanned by Z i.e. $\Psi_0 = Z\Psi_0$. Then

$$a = \Psi_0^T X \Psi_0 = \Psi_0^T Z X Z \Psi_0 \le \|ZXZ\| \le Tr(ZXZ) = Tr(ZX) = \zeta.$$

Altogether (note that the expression on the right hand side of (2.28) is increasing with a),

$$\|e^{At}\Psi_0\|^2 \le \begin{cases} \left(1+\frac{\zeta}{\|X\|}\right)^{2-\frac{t}{\zeta}} & , \quad t>a \\ 1 & , \quad \text{otherwise} \end{cases} \tag{2.29}$$

Another popular decay criterion is the spectral abscissa

$$\lambda_m = maxRe(\lambda(A)),$$

based on the asymptotic estimate

$$\|e^{At}\Psi_0\| \le C_\varepsilon e^{(\lambda_m+\varepsilon)t}$$

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 Sect. 7 below.

3 Frequency cut-off

Solving (2.14) is numerically quite time-consuming task, especially if in searching an optimal damping (2.14) has to be solved many times. According to [1] each solving of a Ljapunov equation of order $2n$ takes some $15(2n)^3 = 120n^3$ operations. This may become very soon prohibitive if e.g. the number of equations or their dimension gets large. Since we are anyhow 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, more precisely, we have

$$C = \sum_{i=1}^{g} c_i d_i d_i^T \tag{3.1}$$

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

$$A = \begin{pmatrix} 0 & \Omega \\ -\Omega & -\hat{C} \end{pmatrix} \ , \ \hat{C} = -\Phi_2^{-1} C \Phi_2^{-T} = \sum_{i=1}^{g} c_i w_i w_i^T \ , \tag{3.2}$$

we obtain a "cut-off" matrix

$$A_r = \begin{pmatrix} 0 & \Omega_r \\ -\Omega_r & -\hat{C}_r \end{pmatrix} \ , \tag{3.3}$$

where Ω_r and $\hat{C}_r$ are obtained by taking the first r rows and columns from Ω, $\hat{C}$, respectively, and $r \ge s$ 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" and in particular 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' then we mean that $Tr(ZX)$ and $Tr(\tilde{Z}\tilde{X})$ (corresponding to the cut-off system) are nearly the same, i.e. the relative estimate

$$\left| \frac{Tr(ZX) - Tr(\tilde{Z}\tilde{X})}{Tr(ZX)} \right| \tag{3.4}$$

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

$$\max_{i=1,\ldots,g} \left| \frac{c_i^{(min)} - \tilde{c}_i^{(min)}}{c_i^{(min)}} \right| \tag{3.5}$$

is small, where $c_i^{(min)}$, $\tilde{c}_i^{(min)}$ represent the viscosities for minimal $Tr(ZX)$, $Tr(\tilde{Z}\tilde{X})$, respectively. This assumption is quite strong but not unrealistic. It is exactly fulfilled, e.g. if $\hat{C} = \sum_{i=1}^{g} c_i w_i w_i^T = VV^T$, $\hat{C}_r = \tilde{V}\tilde{V}^T$ and VV^T is block-diagonal with

$$VV^T = \begin{pmatrix} \tilde{V}\tilde{V}^T & 0 \\ 0 & \check{V}\check{V}^T \end{pmatrix}, \quad \text{where } \check{V}\check{V}^T \neq 0. \tag{3.6}$$

Note also that quite often the undamped system is anyhow given only through a given number of the eigenfrequencies and their eigenforms. This number is usually sensibly less 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 $\hat{C}$ converging to the block-diagonal form (3.6) lead to an arbitrary small relative estimate (3.4) and (3.5) proves to be false. In the 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 the case $g > 1$ we have no exact results. However, experiments show that the enlarging of the distance between the frequencies in the region behind the cut and before the cut lead to a reduction of the relative estimates, too. Also the experiments show, 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 shortly 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 in replacing $\hat{C}$ in (3.2) by its diagonal part. Now the Ljapunov equation solution comes for free, once the undamped eigenproblem is solved. However, this may falsify the dynamics dramatically, in particular, in the 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 (Fig. 2.1) fixed at both ends and having unit masses and unit spring stiffnesses. The damper is applied at mass #1. [7]

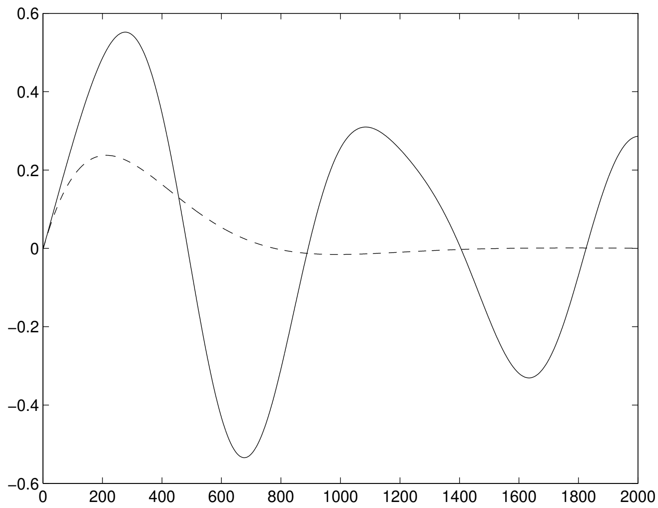

Figure 4.1: Time history, fully, modally damped

The system at rest is put into vibration by applying unit impulse to mass #1. Our first graph (Fig. 4.1) 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 (interruped line). In the same manner Fig. 4.2 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

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

[7] The damper has to be positioned non-symmetrically, if we want the second mode to be affected as well.

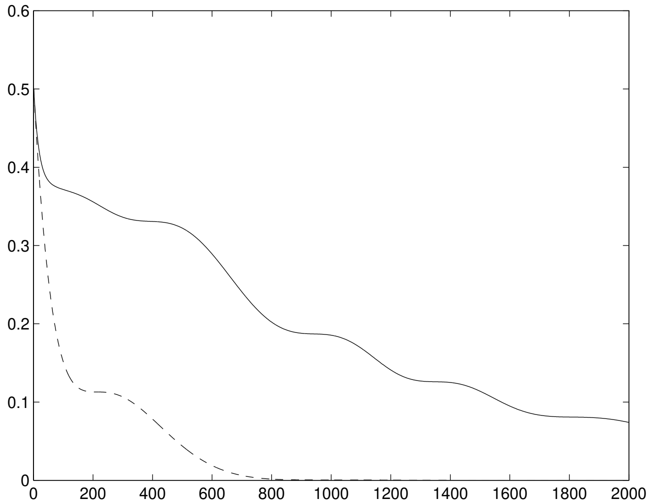

Figure 4.2: Total energy decay, fully, modally damped

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 Euclidian 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

$$f(c_1, \ldots, c_g) = TrZX \tag{5.1}$$

where X is the solution of the Ljapunov equation

$$A^T X + XA = -B\,, \tag{5.2}$$

in other words

$$X = \int_0^\infty e^{A^T t} B e^{At} dt\,. \tag{5.3}$$

Here X depends on $c_1, \ldots, c_g$ through A from (2.9) whereas the matrices Z, B are symmetric and independent of $c_1, \ldots, c_g$. Writing ∂_i for $\partial/\partial c_i$ and using (3.1, 3.2) we have

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

By differentiating (5.2) we obtain

$$A^T \partial_i X + \partial_i X A = X w_i w_i^T + w_i w_i^T X \tag{5.4}$$

Thus, for finding

$$\partial_i f = \partial_i Tr Z X = Tr \partial_i Z X$$

we have to solve altogether $g+1$ Ljapunov equations which differ from each other only in their right hand sides. In using the standard Ljapunov solver from [1] this can be used to reduce the computational effort sensibly. 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 reduction [8] and the computational effort for the gradient may still be much larger than the one for the function value $f(c) = Tr X$ alone.

However, the special form of the penalty function f allows the computation of its gradient *with solving at most one more Ljapunov equation with the same left hand side - independently of the number s of the components of the gradient.* Indeed, applying formula (5.3) to both the equations (5.2) and (5.4) we obtain

$$\partial_i f = -\int_0^\infty Tr Z e^{A^T t}(X w_i w_i^T + w_i w_i^T X) e^{At} dt = -w_i^T Y X w_i - w_i^T X Y w_i =$$

$$= -2 w_i^T X Y w_i \, , \tag{5.5}$$

where $Y = \int_0^\infty \exp(At) Z \exp(A^T t) dt$ solves the 'dual Ljapunov equation'

$$AY + YA^T = -Z \, . \tag{5.6}$$

By the J-symmetry (2.11) we have

$$Y = J X_1 J \, , \text{ with } A^T X_1 + X_1 A = -JZJ \, . \tag{5.7}$$

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

$$Y = J X_1 J \, , \text{ with } A^T X_1 + X_1 A = -Z_s \, . \tag{5.8}$$

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

The last Ljapunov equation has the same left hand side as (5.2). Now, the components of the gradient are obtained cheaply:

$$\partial_i f = -2w_i^T JX_1 JXw_i \,, \tag{5.9}$$

with some n^2 operations. If $B = Z = I$ then $X_1 = X$ and the Ljapunov 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

$$\partial_i\partial_j f = -2w_i^T \partial_i XJX_1 JXw_i - 2w_j^T \partial_i XJX_1 JXw_j \,, \tag{5.10}$$

These are obtained under the additional cost of computing all s components $\partial_1 X, \ldots, \partial_g X$ from (5.4) - which is again a significant saving.

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 mimima? 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 6.1 *Denote by $\mathcal{M}$ the set of all matrices of the form*

$$\hat{C} = 2\tilde{\Omega} = 2\begin{pmatrix} \hat{\Omega}_s & 0 \\ 0 & H \end{pmatrix}, \; \hat{\Omega}_s = diag(\omega_1, \ldots, \omega_s)$$

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 TrZ_sX, $A^TX + XA = -I$, 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 better understand 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 our 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 the best in general.

Conjecture 6.2 *The minimum from Theorem 6.1 is global.*

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

Modally damped systems are characterised by the generalised commutativity property

$$CK^{-1}M = MK^{-1}C. \tag{6.1}$$

In the modal representation (2.25) this means just commutativity

$$\hat{C}\Omega = \Omega\hat{C}.$$

In this case the solution X of the Ljapunov equation

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

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

$$X = \begin{pmatrix} \frac{1}{2}\hat{C}\Omega^{-2} + \hat{C}^{-1} & \frac{1}{2}\Omega^{-1} \\ \frac{1}{2}\Omega^{-1} & \hat{C}^{-1} \end{pmatrix}. \tag{6.2}$$

We now represent Z from (2.26) as

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

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

Theorem 6.3 *Let Ω be fixed and let $\hat{C}$ vary over the set of all positive definite matrices commuting with Ω. Then in the representation*

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

the function TrZ_sX can be expressed as a strictly convex function of the variable $\hat{C}_s$. Consequently, within this set the minimum is global.

Proof From (6.2) is follows

$$TrZ_sX = f(\hat{C}_s) = 2Tr\hat{C}_s^{-1} + \frac{1}{2}Tr\hat{\Omega}_s^{-2}\hat{C}_s. \tag{6.3}$$

To prove the strict convexity of f take $0 < a < 1$. Then the matrix

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

with $W = A^{\frac{1}{2}} B^{-1} A^{\frac{1}{2}}$, is positive semidefinite (this follows from the positive definiteness of $aI + (1-a)W^{-1}$ and the positive semidefiniteness of $W + W^{-1} - 2I$). So,

$$TrU \geq 0$$

and this is the convexity of f. If the above inequality becomes an equality then $U = 0$ or, equivalently, $A = B$. So, f is strictly convex. Q.E.D.

Now, our initial guess is obtained as follows:

Among all $\hat{C} = c_1 w_1 w_1^T + \ldots + c_g w_g^T w_g$ *take the one that is closest to the set* $\mathcal{M}$.

This leads to the following least squares problem:

$$\left(\text{dist}\left(\mathcal{M}, \sum_{i=1}^{g} \tilde{c}_i w_i w_i^T\right)\right)^2 = \min_{c_1,\ldots,c_g} \left(\text{dist}\left(\mathcal{M}, \sum_{i=1}^{g} c_i w_i w_i^T\right)\right)^2$$

$$= \min_{c_1,\ldots,c_g} \left(\inf_{H \in \mathcal{S}} \left\| \begin{pmatrix} 2\,\hat{\Omega}_s & 0 \\ 0 & H \end{pmatrix} - \sum_{i=1}^{g} c_i w_i w_i^T \right\|_F^2 \right) \qquad (6.4)$$

with

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

Here, $\|\cdot\|_F$ is the Frobenius norm. Now, in general this minimisation need not 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 (6.4) leads to the Nonnegative least squares problem (see [2])

$$\min_{c_1,\ldots,c_g \geq 0} \left\| X_0 - [A_1 A_2 \ldots A_g] \begin{pmatrix} c_1 \\ \vdots \\ c_g \end{pmatrix} \right\|_F^2 \qquad (6.5)$$

with

$$X_0 = \begin{pmatrix} 2\,\hat{\Omega}_s & 0 \\ 0 & 0 \end{pmatrix}, \quad A_k = \begin{pmatrix} \hat{V}_s^k\, \hat{V}_s^{kT} & 0 \\ 0 & \sqrt{2}\,\hat{V}_s^k\, \check{V}_{n-s}^{kT} \end{pmatrix},$$

$$V = \begin{pmatrix} \hat{V}_s \\ \check{V}_{n-s} \end{pmatrix} = [w_1 \ldots w_g],$$

where $\hat{V}_s^k$ or $\check{V}_{n-s}^k$ is the k-th column of $\hat{V}_s$ or of $\check{V}_{n-s}$, respectively, with $1 \le k \le g$. It is not a priori clear, that the minimisation (6.5) 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 minimisation. This can not be fully explained as yet. A partial explanation is given by the following

Theorem 6.4 *If $Z_s = I$ and w_i in (3.2) are mutually orthogonal, then all $\tilde{c}_i$ of the optimal $\tilde{C}$ from the least squares process (6.4) or (6.5), respectively, are positive.*

Proof In this case the initial guess coefficients are given by projections

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

Q.E.D.

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

7 Spectral abscissa versus trace

In many applications optimising the spectral abscissa gives about the same results as optimising 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 the optimal - abscissa damping

$$C = \frac{1}{k_1^{1/2} + k_2^{1/2}} \begin{pmatrix} 4k_1^{3/4}k_2^{1/4} & (k_1^{1/2} - k_2^{1/2})^2 \\ (k_1^{1/2} - k_2^{1/2})^2 & 4k_1^{1/4}k_2^{3/4} \end{pmatrix}, \quad k_1 = 1, \ k_2 = 25.$$

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

$$\dot{y} = Ay \quad \dot{y}_d = A_d y$$

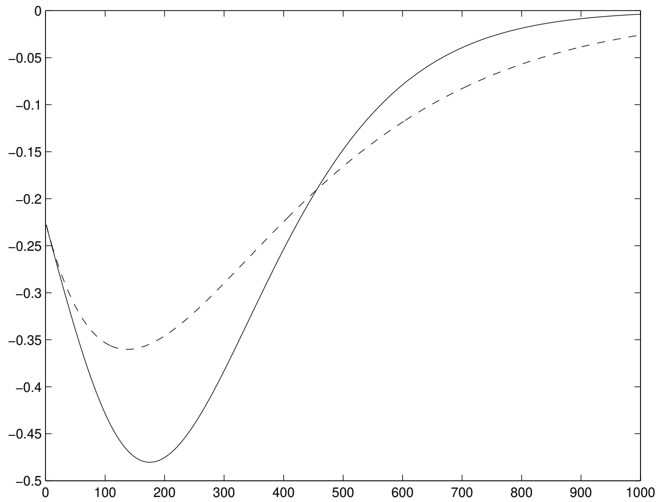

Figure 7.1: full line: optimal abscissa, dashed: optimal trace

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 on Fig. 7.1
The time history of the total energy is displayed on Fig. 7.2

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

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

with the optimal - abscissa damping

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

with $a = \sqrt{385249}$, $b = \sqrt{2585257 + 15776a}$ and $c = \sqrt{103a - 63929}$.
Our best damping in this case is

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

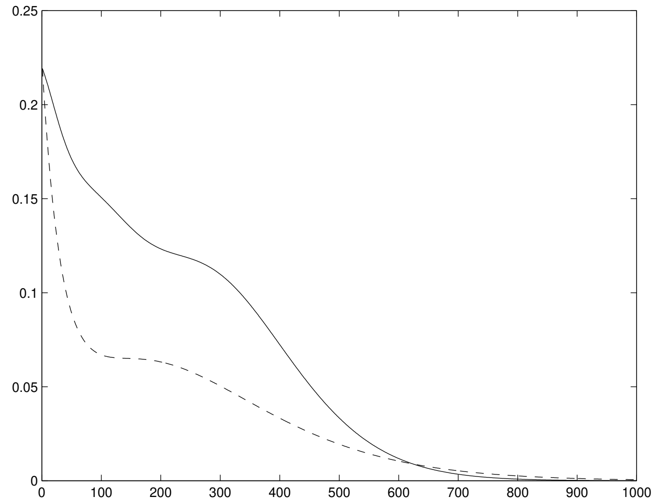

Figure 7.2: full line: optimal abscissa, dashed: optimal trace

The corresponding phase space matrices are A and A_d. For the solutions

$$\dot{y} = Ay \quad \dot{y}_d = A_d y$$

with the common initial data

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

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 on Fig. 7.3, whereas the time history of the total energy is displayed on Fig. 7.4. Here the difference of both criteria is more drastic.

A Appendix

To prove Theorem 6.1 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 A.1 *(i) Let f be a continuously differentiable real-valued function on a neighbourhood of $x = x_0 \in \mathbb{R}^n$. Then grad $f(x_0) = 0$ if and only if*

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

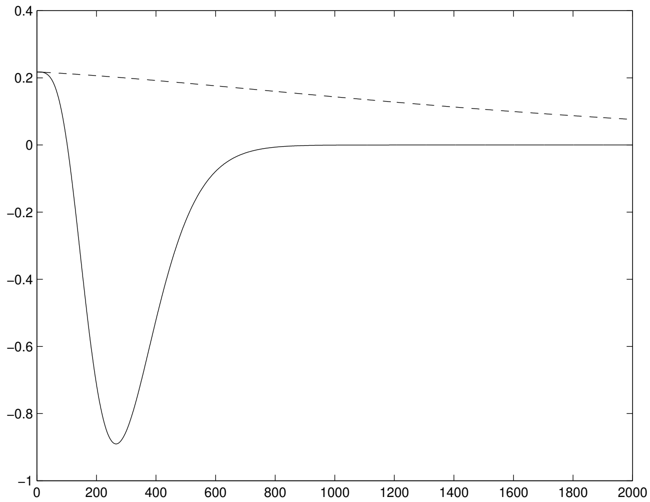

Figure 7.3: full line: optimal abscissa, dashed: optimal trace

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

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

Using this lemma we will now prove Theorem 6.1.

Proof of Theorem 6.1 We prove the statement of the theorem for the set

$$\mathcal{M}_0 = \left\{ C = 2\tilde{\Omega} = 2 \begin{pmatrix} \hat{\Omega}_s & 0 \\ 0 & H \end{pmatrix} \text{ with } \hat{\Omega}_s = diag(\omega_1, \cdots, \omega_s),\ H \text{ positive definite} \right\};$$

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

Let $2\tilde{\Omega} \in \mathcal{M}_0$. Because C is a positive semidefinite matrix, C can be written as $C = 2\tilde{\Omega} + \mu V$, where V is an arbitrary symmetric matrix of order n. For small μ $2\tilde{\Omega} + \mu V$ is positive definite, so that $X(2\tilde{\Omega} + \mu V)$ is uniquely determined as the solution of the Ljapunov equation $A^T X + XA = -I$. According to Lemma A.1 for arbitrary $2\tilde{\Omega} \in \mathcal{M}_0$ we have to show

$$\frac{d}{d\mu} TrY(2\tilde{\Omega} + \mu V)\Big|_{\mu=0} = 0 \tag{A.1}$$

for all symmetric matrices V of order n and

$$\frac{d^2}{d\mu^2} TrY(2\tilde{\Omega} + \mu V)\Big|_{\mu=0} > 0 \tag{A.2}$$

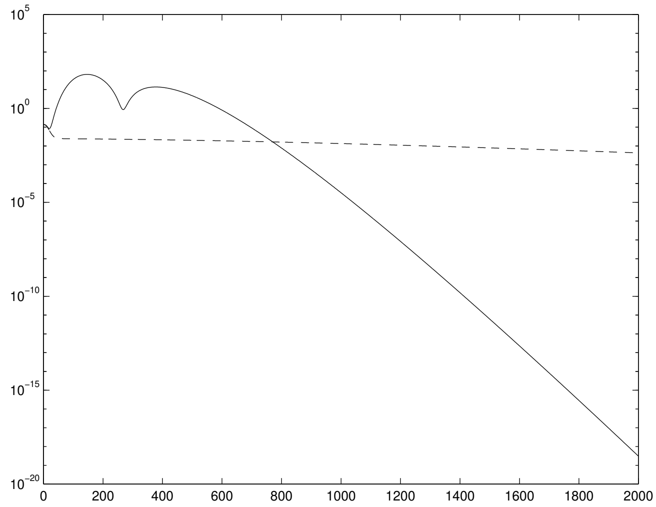

Figure 7.4: full line: optimal abscissa, optimal trace

for all symmetric matrices V of order n, where in (A.2) with

$$V = \begin{pmatrix} V_{11} & V_{12} \\ V_{12}^T & V_{22} \end{pmatrix}, \qquad V_{22} \text{ of order } (n-s) \tag{A.3}$$

the matrices V_{11} and V_{12} may not vanish simultaneously (because in this case $2\tilde{\Omega} + \mu V$ would be in $\mathcal{M}_0$ for small μ).
In the following 2×2- block decompositions of $n \times n$- matrices are always meant as in (A.3) without further notice.

I. Step:
In the first step we show the equality $\frac{d}{d\mu} TrY(2\tilde{\Omega} + \mu V)|_{\mu=0} = 0$ for all symmetric matrices V.

First we determine

$$\frac{d}{d\mu} TrY(2\tilde{\Omega} + \mu V) = Tr \frac{d}{d\mu} Y(2\tilde{\Omega} + \mu V).$$

By differentiating the Ljapunov equation we obtain (during the proof we will use the form A_μ, Y_μ instead of $A(2\tilde{\Omega} + \mu V), Y(2\tilde{\Omega} + \mu V)$)

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

so, we obtain the following new Ljapunov equation

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

With $A_\mu = \begin{pmatrix} 0 & \Omega \\ -\Omega & -(2\tilde{\Omega}+\mu V) \end{pmatrix}$ we obtain

$$\frac{d}{d\mu}A_\mu = \begin{pmatrix} 0 & 0 \\ 0 & -V \end{pmatrix} =: -\tilde{V} = \frac{d}{d\mu}A_\mu^T \ , \tag{A.4}$$

so that the Ljapunov equation takes the form

$$A_\mu^T \frac{d}{d\mu}Y_\mu + \left(\frac{d}{d\mu}Y_\mu\right) A_\mu = \tilde{V}Y_\mu + Y_\mu\tilde{V}. \tag{A.5}$$

Now we can determine $\frac{d}{d_\mu}TrY_\mu = Tr\frac{d}{d\mu}Y_\mu$:

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

where $\tilde{X}_\mu$ is the solution of the Ljapunov equation $A_\mu\tilde{X}_\mu + \tilde{X}_\mu A_\mu^T = -I$, i.e. $\tilde{X}_\mu = \int_0^\infty e^{A_\mu t}e^{A_\mu^T t}dt$. Let X_μ be the solution of the Ljapunov equation $A_\mu^T X_\mu + X_\mu A_\mu = -I$.
By the relation $A_\mu^T = JA_\mu J$, $J = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}$ the Ljapunov equation $A_\mu^T X_\mu + X_\mu A_\mu = -I$ takes the form

$$\begin{aligned}
JA_\mu JX_\mu + X_\mu JA_\mu^T J &= -I \ , \\
A_\mu JX_\mu J + JX_\mu JA_\mu^T &= -I \ .
\end{aligned}$$

So, we obtain $\tilde{X}_\mu = JX_\mu J$ and

$$Tr\frac{d}{d\mu}Y_\mu = -2\,Tr(\tilde{V}Y_\mu JX_\mu J). \tag{A.6}$$

Let

$$\Omega =: \begin{pmatrix} \hat{\Omega}_s & 0 \\ 0 & \check{\Omega}_{n-s} \end{pmatrix} , V = \begin{pmatrix} V_{11} & V_{12} \\ V_{12}^T & V_{22} \end{pmatrix} .$$

In the case $\mu = 0$ the uniquely determined solution Y_0 of the Ljapunov equation $A_0^T Y_0 + Y_0 A_0 = -Z_s$ is easily seen to be

$$Y_0 = \begin{pmatrix} \Psi_{11} & \Psi_{12} \\ \Psi_{12}^T & \Psi_{22} \end{pmatrix}$$

where

$$\Psi_{11} = \begin{pmatrix} \frac{3}{2}\hat{\Omega}_s^{-1} & 0 \\ 0 & 0 \end{pmatrix} \quad , \quad \Psi_{12} = \Psi_{22} = \begin{pmatrix} \frac{1}{2}\hat{\Omega}_s^{-1} & 0 \\ 0 & 0 \end{pmatrix} . \tag{A.7}$$

Likewise we can check that the solution X_0 of the Ljapunov equation $A_0^T X_0 + X_0 A_0 = -I$ has the form

$$X_0 = \begin{pmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{12}^T & \Phi_{22} \end{pmatrix}$$

with

$$\Phi_{11} = \begin{pmatrix} \frac{3}{2}\hat{\Omega}_s^{-1} & 0 \\ 0 & W_{11} \end{pmatrix}, \quad \Phi_{12} = \begin{pmatrix} \frac{1}{2}\hat{\Omega}_s^{-1} & 0 \\ 0 & W_{12} \end{pmatrix}, \quad \Phi_{22} = \begin{pmatrix} \frac{1}{2}\hat{\Omega}_s^{-1} & 0 \\ 0 & W_{22} \end{pmatrix}, \tag{A.8}$$

where

$$\begin{pmatrix} W_{11} & W_{12} \\ W_{12}^T & W_{22} \end{pmatrix}$$

is the uniquely determined solution of the Ljapunov equation

$$\begin{aligned} &\begin{pmatrix} 0 & -\check{\Omega}_{n-s} \\ \check{\Omega}_{n-s} & -2H \end{pmatrix} \begin{pmatrix} W_{11} & W_{12} \\ W_{12}^T & W_{22} \end{pmatrix} + \begin{pmatrix} W_{11} & W_{12} \\ W_{12}^T & W_{22} \end{pmatrix} \begin{pmatrix} 0 & \check{\Omega}_{n-s} \\ -\check{\Omega}_{n-s} & -2H \end{pmatrix} \\ = &\begin{pmatrix} -I_{n-s} & 0 \\ 0 & -I_{n-s} \end{pmatrix} . \end{aligned} \tag{A.9}$$

So, we obtain (see (A.6))

$$\begin{aligned} & Tr\left(\frac{d}{d\mu} Y_\mu\right)\Big|_{\mu=0} \\ = \; & -2\, Tr\left[\begin{pmatrix} 0 & 0 \\ 0 & V \end{pmatrix} \begin{pmatrix} \Psi_{11} & \Psi_{12} \\ \Psi_{12}^T & \Psi_{22} \end{pmatrix} \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \begin{pmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{12}^T & \Phi_{22} \end{pmatrix} \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}\right] \\ = \; & -2\, Tr\left[\begin{pmatrix} 0 & 0 \\ V\Psi_{12}^T & V\Psi_{22} \end{pmatrix} \begin{pmatrix} \Phi_{11} & -\Phi_{12} \\ -\Phi_{12}^T & \Phi_{22} \end{pmatrix}\right] \\ = \; & 2\, Tr(V\Psi_{12}^T\Phi_{12}) - 2\, Tr(V\Psi_{22}\Phi_{22}) \\ = \; & 2\, Tr\left[\begin{pmatrix} V_{11} & V_{12} \\ V_{12}^T & V_{22} \end{pmatrix} \begin{pmatrix} \frac{1}{2}\hat{\Omega}_s^{-1} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{2}\hat{\Omega}_s^{-1} & 0 \\ 0 & W_{12} \end{pmatrix}\right] \\ & -2\, Tr\left[\begin{pmatrix} V_{11} & V_{12} \\ V_{12}^T & V_{22} \end{pmatrix} \begin{pmatrix} \frac{1}{2}\hat{\Omega}_s^{-1} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{2}\hat{\Omega}_s^{-1} & 0 \\ 0 & W_{22} \end{pmatrix}\right] \\ = \; & 2\, Tr\left[\begin{pmatrix} V_{11} & V_{12} \\ V_{12}^T & V_{22} \end{pmatrix} \begin{pmatrix} \frac{1}{4}\hat{\Omega}_s^{-2} & 0 \\ 0 & 0 \end{pmatrix}\right] - 2\, Tr\left[\begin{pmatrix} V_{11} & V_{12} \\ V_{12}^T & V_{22} \end{pmatrix} \begin{pmatrix} \frac{1}{4}\hat{\Omega}_s^{-2} & 0 \\ 0 & 0 \end{pmatrix}\right] \\ = \; & 0 \, . \end{aligned}$$

So, we have proven (A.1) and because the set $\mathcal{M}$ is connected and $\mathcal{M}_0$ is a dense subset in $\mathcal{M}$ we obtain $TrY(C) = const$ on $\mathcal{M}$.

II. Step:

In the second step we show $\frac{d^2}{d\mu^2} Tr(2\tilde{\Omega}+\mu V)|_{\mu=0} > 0$, where $V = \begin{pmatrix} V_{11} & V_{12} \\ V_{12}^T & V_{22} \end{pmatrix}$ is an arbitrary symmetric matrix and $V_{11},\ V_{12}$ do not vanish simultanously.

First we determine

$$\frac{d^2}{d\mu^2} TrY(2\tilde{\Omega}+\mu V) = Tr\frac{d^2}{d\mu^2} Y(2\tilde{\Omega}+\mu V).$$

Differentiating the Ljapunov equation (A.5) we get

$$\begin{aligned} &\left(\frac{d}{d\mu}A_\mu^T\right)\left(\frac{d}{d\mu}Y_\mu\right) + A_\mu^T\left(\frac{d^2}{d\mu^2}Y_\mu\right) + \left(\frac{d^2}{d\mu^2}Y_\mu\right)A_\mu + \left(\frac{d}{d\mu}Y_\mu\right)\left(\frac{d}{d\mu}A_\mu\right) \\ = \;& \tilde{V}\left(\frac{d}{d\mu}Y_\mu\right) + \left(\frac{d}{d\mu}Y_\mu\right)\tilde{V}, \end{aligned}$$

i.e.

$$\begin{aligned} A_\mu^T\left(\frac{d^2}{d\mu^2}Y_\mu\right) + \left(\frac{d^2}{d\mu^2}Y_\mu\right)A_\mu &= -\left(\frac{d}{d\mu}A_\mu^T\right)\left(\frac{d}{d\mu}Y_\mu\right) - \left(\frac{d}{d\mu}Y_\mu\right)\left(\frac{d}{d\mu}A_\mu\right) \\ &\quad + \tilde{V}\left(\frac{d}{d\mu}Y_\mu\right) + \left(\frac{d}{d\mu}Y_\mu\right)\tilde{V}\ , \end{aligned}$$

so (see (A.4))

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

So

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

and

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

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

with $A_\mu^T X_\mu + X_\mu A_\mu = -I$ and $A_\mu \tilde{X}_\mu + \tilde{X}_\mu A_\mu^T = -I$ as before.

With $X_0 = \begin{pmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{12}^T & \Phi_{22} \end{pmatrix}$ and $\frac{d}{d\mu}Y_\mu|_{\mu=0} =: \begin{pmatrix} \tilde{Y}_{11} & \tilde{Y}_{12} \\ \tilde{Y}_{12}^T & \tilde{Y}_{22} \end{pmatrix} =: Y_0'$ it follows

$$
\begin{aligned}
&\left.\frac{d^2}{d\mu^2} Tr Y_\mu\right|_{\mu=0}\\
&= -2Tr\left[\begin{pmatrix} 0 & 0 \\ 0 & V \end{pmatrix}\begin{pmatrix} \tilde{Y}_{11} & \tilde{Y}_{12} \\ \tilde{Y}_{12}^T & \tilde{Y}_{22} \end{pmatrix}\begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}\begin{pmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{12}^T & \Phi_{22} \end{pmatrix}\begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}\right]\\
&\quad -2Tr\left[\begin{pmatrix} \tilde{Y}_{11} & \tilde{Y}_{12} \\ \tilde{Y}_{12}^T & \tilde{Y}_{22} \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 0 & V \end{pmatrix}\begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}\begin{pmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{12}^T & \Phi_{22} \end{pmatrix}\begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}\right]\\
&= -2Tr\left[\begin{pmatrix} 0 & 0 \\ V\tilde{Y}_{12}^T & V\tilde{Y}_{22} \end{pmatrix}\begin{pmatrix} \Phi_{11} & -\Phi_{12} \\ -\Phi_{12}^T & \Phi_{22} \end{pmatrix}\right]\\
&\quad -2Tr\left[\begin{pmatrix} 0 & \tilde{Y}_{12}V \\ 0 & \tilde{Y}_{22}V \end{pmatrix}\begin{pmatrix} \Phi_{11} & -\Phi_{12} \\ -\Phi_{12}^T & \Phi_{22} \end{pmatrix}\right]\\
&= -2Tr[-V\tilde{Y}_{12}^T\Phi_{12} + V\tilde{Y}_{22}\Phi_{22}] - 2Tr[-\tilde{Y}_{12}V\Phi_{12}^T + \tilde{Y}_{22}V\Phi_{22}]\\
&= 2Tr[V\tilde{Y}_{12}^T\Phi_{12}] - 2Tr[V\tilde{Y}_{22}\Phi_{22}] + 2Tr[\tilde{Y}_{12}V\Phi_{12}^T] - 2Tr[\tilde{Y}_{22}V\Phi_{22}]\\
&= 4\,Tr[\tilde{Y}_{12}V\Phi_{12}^T] - 4\,Tr[V\tilde{Y}_{22}\Phi_{22}]. \qquad \text{(A.10)}
\end{aligned}
$$

Now we want to determine the submatrices $\tilde{Y}_{12}$ and $\tilde{Y}_{22}$ of the matrix $\frac{d}{d\mu}Y(2\tilde{\Omega}+\mu V)|_{\mu=0}$:

$Y_0' = \left(\frac{d}{d\mu}Y_\mu\right)|_{\mu=0}$ is (see (A.5)) the solution of the Ljapunov equation

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

where

$$
Y_0 = \begin{pmatrix} \Psi_{11} & \Psi_{12} \\ \Psi_{12}^T & \Psi_{22} \end{pmatrix}
$$

is the solution of the Ljapunov equation $A_0^T Y_0 + Y_0 A_0 = -Z_s$, which is given in (A.7). Y_0' now has to be calculated explicitly. That means, we have to solve

$$\begin{pmatrix} 0 & -\Omega \\ \\ \Omega & -2\tilde{\Omega} \end{pmatrix}\begin{pmatrix} \tilde{Y}_{11} & \tilde{Y}_{12} \\ \\ \tilde{Y}_{12}^T & \tilde{Y}_{22} \end{pmatrix} + \begin{pmatrix} \tilde{Y}_{11} & \tilde{Y}_{12} \\ \\ \tilde{Y}_{12}^T & \tilde{Y}_{22} \end{pmatrix}\begin{pmatrix} 0 & \Omega \\ \\ -\Omega & -2\tilde{\Omega} \end{pmatrix}$$
$$= \begin{pmatrix} \Psi_{11} & \Psi_{12} \\ \\ \Psi_{12}^T & \Psi_{22} \end{pmatrix}\begin{pmatrix} 0 & 0 \\ \\ 0 & V \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ \\ 0 & V \end{pmatrix}\begin{pmatrix} \Psi_{11} & \Psi_{12} \\ \\ \Psi_{12}^T & \Psi_{22} \end{pmatrix}$$
$$= \begin{pmatrix} 0 & \Psi_{12}V \\ \\ V\Psi_{12}^T & \Psi_{22}V + V\Psi_{22} \end{pmatrix}.$$

This leads to the following four matrix equations

$$\begin{aligned} -\Omega\tilde{Y}_{12}^T - \tilde{Y}_{12}\Omega &= 0 & \text{(A.11)} \\ -\Omega\tilde{Y}_{22} + \tilde{Y}_{11}\Omega - 2\tilde{Y}_{12}\tilde{\Omega} &= \Psi_{12}V & \text{(A.12)} \\ \Omega\tilde{Y}_{11} - 2\tilde{\Omega}\tilde{Y}_{12}^T - \tilde{Y}_{22}\Omega &= V\Psi_{12}^T & \text{(A.13)} \\ \Omega\tilde{Y}_{12} - 2\tilde{\Omega}\tilde{Y}_{22} + \tilde{Y}_{12}^T\Omega - 2\tilde{Y}_{22}\tilde{\Omega} &= \Psi_{22}V + V\Psi_{22}, & \text{(A.14)} \end{aligned}$$

where equation (A.13) and (A.12) are adjoint to each other. From the first equation we receive

$$\tilde{Y}_{12} = \frac{1}{2}S\Omega^{-1}\ ,\ \ \tilde{Y}_{12}^T = -\frac{1}{2}\Omega^{-1}S\ , \tag{A.15}$$

where S is a skew-symmetric matrix, i.e. $S^T = -S$. Substituting this into equation (A.10), we get

$$\begin{aligned} \frac{d^2}{d\mu^2}TrY_\mu\Big|_{\mu=0} &= 4\,Tr[(\frac{1}{2}S\Omega^{-1})V\Phi_{12}^T] - 4\,Tr[V\tilde{Y}_{22}\Phi_{22}] \\ &= 2\,Tr[S\Omega^{-1}V\Phi_{12}^T] - 4\,Tr[V\tilde{Y}_{22}\Phi_{22}]. \end{aligned} \tag{A.16}$$

Now we have to determine $\tilde{Y}_{22}$ and S more precisely. Let

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

First we look at the four equations (A.11) - (A.14): with (A.7) we obtain

$$\Psi_{12}V = \begin{pmatrix} \frac{1}{2}\hat{\Omega}_s^{-1} & 0 \\ \\ 0 & 0 \end{pmatrix}\begin{pmatrix} V_{11} & V_{12} \\ \\ V_{12}^T & V_{22} \end{pmatrix} = \begin{pmatrix} \frac{1}{2}\hat{\Omega}_s^{-1}V_{11} & \frac{1}{2}\hat{\Omega}_s^{-1}V_{12} \\ \\ 0 & 0 \end{pmatrix},$$

$$V\Psi_{12}^T = (\Psi_{12}V)^T = \begin{pmatrix} \frac{1}{2}V_{11}\hat{\Omega}_s^{-1} & 0 \\ \frac{1}{2}V_{12}^T\hat{\Omega}_s^{-1} & 0 \end{pmatrix},$$

$$\begin{aligned}\Psi_{22}V + V\Psi_{22} &= \begin{pmatrix} \frac{1}{2}\hat{\Omega}_s^{-1} & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} V_{11} & V_{12} \\ V_{12}^T & V_{22} \end{pmatrix} + \begin{pmatrix} V_{11} & V_{12} \\ V_{12}^T & V_{22} \end{pmatrix}\begin{pmatrix} \frac{1}{2}\hat{\Omega}_s^{-1} & 0 \\ 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{2}\hat{\Omega}_s^{-1}V_{11} + \frac{1}{2}V_{11}\hat{\Omega}_s^{-1} & \frac{1}{2}\hat{\Omega}_s^{-1}V_{12} \\ \frac{1}{2}V_{12}^T\hat{\Omega}_s^{-1} & 0 \end{pmatrix}.\end{aligned}$$

So, the four matrix equations (A.11) - (A.14) go over into

$$\begin{aligned}&- \begin{pmatrix} \hat{\Omega}_s & 0 \\ 0 & \check{\Omega}_{n-s} \end{pmatrix}\begin{pmatrix} Q_{11}^T & Q_{21}^T \\ Q_{12}^T & Q_{22}^T \end{pmatrix} - \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix}\begin{pmatrix} \hat{\Omega}_s & 0 \\ 0 & \check{\Omega}_{n-s} \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \end{aligned} \tag{A.17}$$

$$\begin{aligned}&- \begin{pmatrix} \hat{\Omega}_s & 0 \\ 0 & \check{\Omega}_{n-s} \end{pmatrix}\begin{pmatrix} P_{11} & P_{12} \\ P_{12}^T & P_{22} \end{pmatrix} + \begin{pmatrix} R_{11} & R_{12} \\ R_{12}^T & R_{22} \end{pmatrix}\begin{pmatrix} \hat{\Omega}_s & 0 \\ 0 & \check{\Omega}_{n-s} \end{pmatrix} \\ &- 2\begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix}\begin{pmatrix} \hat{\Omega}_s & 0 \\ 0 & H \end{pmatrix} = \begin{pmatrix} \frac{1}{2}\hat{\Omega}_s^{-1}V_{11} & \frac{1}{2}\hat{\Omega}_s^{-1}V_{12} \\ 0 & 0 \end{pmatrix}, \end{aligned} \tag{A.18}$$

$$\begin{aligned}&\begin{pmatrix} \hat{\Omega}_s & 0 \\ 0 & \check{\Omega}_{n-s} \end{pmatrix}\begin{pmatrix} R_{11} & R_{12} \\ R_{12}^T & R_{22} \end{pmatrix} - 2\begin{pmatrix} \hat{\Omega}_s & 0 \\ 0 & H \end{pmatrix}\begin{pmatrix} Q_{11}^T & Q_{21}^T \\ Q_{12}^T & Q_{22}^T \end{pmatrix} \\ &- \begin{pmatrix} P_{11} & P_{12} \\ P_{12}^T & P_{22} \end{pmatrix}\begin{pmatrix} \hat{\Omega}_s & 0 \\ 0 & \check{\Omega}_{n-s} \end{pmatrix} = \begin{pmatrix} \frac{1}{2}V_{11}\hat{\Omega}_s^{-1} & 0 \\ \frac{1}{2}V_{12}^T\hat{\Omega}_s^{-1} & 0 \end{pmatrix}, \end{aligned} \tag{A.19}$$

$$\begin{pmatrix} \hat{\Omega}_s & 0 \\ 0 & \check{\Omega}_{n-s} \end{pmatrix}\begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix} - 2\begin{pmatrix} \hat{\Omega}_s & 0 \\ 0 & H \end{pmatrix}\begin{pmatrix} P_{11} & P_{12} \\ P_{12}^T & P_{22} \end{pmatrix}$$

$$
+ \begin{pmatrix} Q_{11}^T & Q_{21}^T \\ Q_{12}^T & Q_{22}^T \end{pmatrix} \begin{pmatrix} \hat{\Omega}_s & 0 \\ 0 & \check{\Omega}_{n-s} \end{pmatrix} - 2 \begin{pmatrix} P_{11} & P_{12} \\ P_{12}^T & P_{22} \end{pmatrix} \begin{pmatrix} \hat{\Omega}_s & 0 \\ 0 & H \end{pmatrix}
$$

$$
= \begin{pmatrix} \frac{1}{2}\hat{\Omega}_s^{-1}V_{11} + \frac{1}{2}V_{11}\hat{\Omega}_s^{-1} & \frac{1}{2}\hat{\Omega}_s^{-1}V_{12} \\ \frac{1}{2}V_{12}^T\hat{\Omega}_s^{-1} & 0 \end{pmatrix}. \tag{A.20}
$$

1. First we consider the lower right block of the matrices of the equations (A.17) - (A.20) i.e. the 2,2 - block of the matrix equations. We get the following four equations, where once more the second and third equation are adjoint to each other:

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

This is equivalent to the matrix equation

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

Because H is positiv definite, the matrix $\begin{pmatrix} 0 & \check{\Omega}_{n-s} \\ -\check{\Omega}_{n-s} & -2H \end{pmatrix}$ is asymptotically stable too, and the matrix equation has only the trivial solution $R_{22} = Q_{22} = P_{22} = 0$. With (see (A.15))

$$
\begin{aligned}
\tilde{Y}_{12} &= \frac{1}{2}S\Omega^{-1} = \frac{1}{2}\begin{pmatrix} S_{11} & S_{12} \\ -S_{12}^T & S_{22} \end{pmatrix} \begin{pmatrix} \hat{\Omega}_s^{-1} & 0 \\ 0 & \check{\Omega}_{n-s}^{-1} \end{pmatrix} \\
&= \frac{1}{2}\begin{pmatrix} S_{11}\hat{\Omega}_s^{-1} & S_{12}\check{\Omega}_{n-s}^{-1} \\ -S_{12}^T\hat{\Omega}_s^{-1} & S_{22}\check{\Omega}_{n-s}^{-1} \end{pmatrix} = \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & 0 \end{pmatrix}
\end{aligned}
$$

also $S_{22} = 0$ follows.

Putting this result into (A.16) we obtain together with (A.8)

$$
\begin{aligned}
&\frac{d^2}{d\mu^2}TrY_\mu\Big|_{\mu=0} \\
&= 2Tr[S\Omega^{-1}V\Phi_{12}^T] - 4Tr[V\tilde{Y}_{22}\Phi_{22}] \\
&= 2Tr\left[\begin{pmatrix} S_{11} & S_{12} \\ -S_{12}^T & 0 \end{pmatrix} \begin{pmatrix} \hat{\Omega}_s^{-1} & 0 \\ 0 & \check{\Omega}_{n-s}^{-1} \end{pmatrix} \begin{pmatrix} V_{11} & V_{12} \\ V_{12}^T & V_{22} \end{pmatrix} \begin{pmatrix} \frac{1}{2}\hat{\Omega}_s^{-1} & 0 \\ 0 & W_{12}^T \end{pmatrix}\right]
\end{aligned}
$$

$$
-4Tr\left[\begin{pmatrix} V_{11} & V_{12} \\ V_{12}^T & V_{22} \end{pmatrix}\begin{pmatrix} P_{11} & P_{12} \\ P_{12}^T & 0 \end{pmatrix}\begin{pmatrix} \frac{1}{2}\hat{\Omega}_s^{-1} & 0 \\ 0 & W_{22} \end{pmatrix}\right]
$$

$$
\begin{aligned}
&= 2Tr\left[\begin{pmatrix} S_{11}\hat{\Omega}_s^{-1} & S_{12}\check{\Omega}_{n-s}^{-1} \\ -S_{12}^T\hat{\Omega}_s^{-1} & 0 \end{pmatrix}\begin{pmatrix} \frac{1}{2}V_{11}\hat{\Omega}_s^{-1} & V_{12}W_{12}^T \\ \frac{1}{2}V_{12}^T\hat{\Omega}_s^{-1} & V_{22}W_{12}^T \end{pmatrix}\right] \\
&\quad -4Tr\left[\begin{pmatrix} V_{11} & V_{12} \\ V_{12}^T & V_{22} \end{pmatrix}\begin{pmatrix} \frac{1}{2}P_{11}\hat{\Omega}_s^{-1} & P_{12}W_{22} \\ \frac{1}{2}P_{12}^T\hat{\Omega}_s^{-1} & 0 \end{pmatrix}\right] \\
&= 2Tr[\frac{1}{2}S_{11}\hat{\Omega}_s^{-1}V_{11}\hat{\Omega}_s^{-1} + \frac{1}{2}S_{12}\check{\Omega}_{n-s}^{-1}V_{12}^T\hat{\Omega}_s^{-1} - S_{12}^T\hat{\Omega}_s^{-1}V_{12}W_{12}^T] \\
&\quad -4Tr[\frac{1}{2}V_{11}P_{11}\hat{\Omega}_s^{-1} + \frac{1}{2}V_{12}P_{12}^T\hat{\Omega}_s^{-1} + V_{12}^TP_{12}W_{22}] \\
&= Tr[S_{11}\hat{\Omega}_s^{-1}V_{11}\hat{\Omega}_s^{-1}] + Tr[S_{12}\check{\Omega}_{n-s}^{-1}V_{12}^T\hat{\Omega}_s^{-1}] - 2Tr[S_{12}^T\hat{\Omega}_s^{-1}V_{12}W_{12}^T] \\
&\quad -2Tr[V_{11}P_{11}\hat{\Omega}_s^{-1}] - 2Tr[V_{12}P_{12}^T\hat{\Omega}_s^{-1}] - 4Tr[V_{12}^TP_{12}W_{22}].
\end{aligned}
$$

Because S_{11} is a skew- symmetric and V_{11}, $\hat{\Omega}_s^{-1}$ are symmetric matrices, we get

$$
\begin{aligned}
Tr[S_{11}\hat{\Omega}_s^{-1}V_{11}\hat{\Omega}_s^{-1}] &= \frac{1}{2}Tr[S_{11}\hat{\Omega}_s^{-1}V_{11}\hat{\Omega}_s^{-1}] + \frac{1}{2}Tr[(S_{11}\hat{\Omega}_s^{-1}V_{11}\hat{\Omega}_s^{-1})^T] \\
&= \frac{1}{2}Tr[\hat{\Omega}_s^{-1}V_{11}\hat{\Omega}_s^{-1}S_{11}] + \frac{1}{2}Tr[\hat{\Omega}_s^{-1}V_{11}\hat{\Omega}_s^{-1}S_{11}^T] \\
&= \frac{1}{2}Tr[\hat{\Omega}_s^{-1}V_{11}\hat{\Omega}_s^{-1}S_{11}] - \frac{1}{2}Tr[\hat{\Omega}_s^{-1}V_{11}\hat{\Omega}_s^{-1}S_{11}] \\
&= 0\ .
\end{aligned}
$$

So, we get

$$
\begin{aligned}
\frac{d^2}{d\mu^2}TrY_\mu\Big|_{\mu=0} &= Tr[S_{12}\check{\Omega}_{n-s}^{-1}V_{12}^T\hat{\Omega}_s^{-1}] - 2Tr[S_{12}^T\hat{\Omega}_s^{-1}V_{12}W_{12}^T] \\
&\quad -2Tr[V_{11}P_{11}\hat{\Omega}_s^{-1}] - 2Tr[V_{12}P_{12}^T\hat{\Omega}_s^{-1}] - 4Tr[V_{12}^TP_{12}W_{22}].
\end{aligned} \tag{A.21}
$$

2. Now we consider the 1,2 - block of the four matrix equations (A.17) - (A.20). We obtain the following four equations:

$$
-\hat{\Omega}_sQ_{21}^T - Q_{12}\check{\Omega}_{n-s} = 0 \tag{A.22}
$$

$$
-\hat{\Omega}_sP_{12} + R_{12}\check{\Omega}_{n-s} - 2Q_{12}H = \frac{1}{2}\hat{\Omega}_s^{-1}V_{12} \tag{A.23}
$$

$$
\hat{\Omega}_sR_{12} - 2\hat{\Omega}_sQ_{21}^T - P_{12}\check{\Omega}_{n-s} = 0 \tag{A.24}
$$

$$
\hat{\Omega}_sQ_{12} - 2\hat{\Omega}_sP_{12} + Q_{21}^T\check{\Omega}_{n-s} - 2P_{12}H = \frac{1}{2}\hat{\Omega}_s^{-1}V_{12}\ . \tag{A.25}
$$

By (see (A.15))

$$\tilde{Y}_{12} = \frac{1}{2} S\Omega^{-1} = \frac{1}{2} \begin{pmatrix} S_{11}\hat{\Omega}_s^{-1} & S_{12}\check{\Omega}_{n-s}^{-1} \\ -S_{12}^T\hat{\Omega}_s^{-1} & 0 \end{pmatrix} = \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & 0 \end{pmatrix}$$

we obtain

$$Q_{12} = \frac{1}{2} S_{12}\check{\Omega}_{n-s}^{-1},\ Q_{21} = -\frac{1}{2} S_{12}^T\hat{\Omega}_s^{-1},\ Q_{21}^T = -\frac{1}{2}\hat{\Omega}_s^{-1} S_{12}\ . \tag{A.26}$$

By putting this into equation (A.24) we get $\hat{\Omega}_s R_{12} + S_{12} - P_{12}\check{\Omega}_{n-s} = 0$, i.e.

$$S_{12} = P_{12}\check{\Omega}_{n-s} - \hat{\Omega}_s R_{12}. \tag{A.27}$$

Now we put (A.26) into equation (A.23) and (A.25) and obtain

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

By replacing S_{12} with (A.27) in these equations we get

$$-\hat{\Omega}_s P_{12} + R_{12}\check{\Omega}_{n-s} - P_{12}H + \hat{\Omega}_s R_{12}\check{\Omega}_{n-s}^{-1} H = \frac{1}{2}\hat{\Omega}_s^{-1} V_{12}\ ,$$

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

Subtracting the last two equations leads to

$$\begin{aligned} & -\hat{\Omega}_s P_{12} + R_{12}\check{\Omega}_{n-s} - P_{12}H + \hat{\Omega}_s R_{12}\check{\Omega}_{n-s}^{-1} H - \frac{1}{2}\hat{\Omega}_s P_{12} \\ + & \frac{1}{2}\hat{\Omega}_s^2 R_{12}\check{\Omega}_{n-s}^{-1} + 2\hat{\Omega}_s P_{12} + \frac{1}{2}\hat{\Omega}_s^{-1} P_{12}\check{\Omega}_{n-s}^2 - \frac{1}{2} R_{12}\check{\Omega}_{n-s} + 2P_{12}H = 0\ , \end{aligned}$$

i.e.

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

This equation can be written as

$$\Delta H + \frac{1}{2}\hat{\Omega}_s \Delta + \frac{1}{2}\hat{\Omega}_s^{-1}\Delta\check{\Omega}_{n-s}^2 = 0 \tag{A.28}$$

with $\Delta = P_{12} + \hat{\Omega}_s R_{12}\check{\Omega}_{n-s}^{-1}$. Equation(A.28) has only the trivial solution $\Delta = 0$, because if we multiply this equation with Δ^T from the right we get

$$\Delta H \Delta^T + \frac{1}{2}\hat{\Omega}_s \Delta\Delta^T + \frac{1}{2}\hat{\Omega}_s^{-1}\Delta\check{\Omega}_{n-s}^2\Delta^T \quad = \quad 0$$

and so

$$Tr[\Delta H\Delta^T] + \frac{1}{2}Tr[\hat{\Omega}_s\Delta\Delta^T] + \frac{1}{2}Tr[\hat{\Omega}_s^{-1}\Delta\check{\Omega}_{n-s}^2\Delta^T] \quad = \quad 0 \ ,$$

i.e.

$$\begin{aligned} & Tr[\Delta H\Delta^T] + \frac{1}{2}Tr[\Delta^T\hat{\Omega}_s\Delta] + \frac{1}{2}Tr[\hat{\Omega}_s^{-\frac{1}{2}}\Delta\check{\Omega}_{n-s}^2\Delta^T\hat{\Omega}_s^{-\frac{1}{2}}] \\ = \ & Tr[(\Delta H^{\frac{1}{2}})(\Delta H^{\frac{1}{2}})^T] + \frac{1}{2}Tr[(\Delta^T\hat{\Omega}_s^{\frac{1}{2}})(\Delta^T\hat{\Omega}_s^{\frac{1}{2}})^T] \\ & + \frac{1}{2}Tr[(\hat{\Omega}_s^{-\frac{1}{2}}\Delta\check{\Omega}_{n-s})(\hat{\Omega}_s^{-\frac{1}{2}}\Delta\check{\Omega}_{n-s})^T] \\ = \ & 0 \ . \end{aligned}$$

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

So, we have $P_{12} + \hat{\Omega}_s R_{12}\check{\Omega}_{n-s}^{-1} = 0$, i.e. $P_{12} = -\hat{\Omega}_s R_{12}\check{\Omega}_{n-s}^{-1}$. Putting this into (A.27), we get for S_{12}

$$S_{12} = P_{12}\check{\Omega}_{n-s} - \hat{\Omega}_s R_{12} = -\hat{\Omega}_s R_{12} - \hat{\Omega}_s R_{12} = -2\hat{\Omega}_s R_{12} \ ,$$

and so (see (A.26))

$$Q_{12} \quad = \quad \frac{1}{2}S_{12}\check{\Omega}_{n-s}^{-1} = -\hat{\Omega}_s R_{12}\check{\Omega}_{n-s}^{-1} = P_{12} \tag{A.29}$$

$$Q_{21} \quad = \quad -\frac{1}{2}S_{12}^T\hat{\Omega}_s^{-1} = R_{12}^T\hat{\Omega}_s\hat{\Omega}_s^{-1} = R_{12}^T. \tag{A.30}$$

So, we obtain

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

Replacing P_{12} in equation (A.21) with $\frac{1}{2}S_{12}\check{\Omega}_{n-s}^{-1}$ (see (A.29)) leads to

$$\begin{aligned} \frac{d^2}{d\mu^2}TrY_\mu\bigg|_{\mu=0} \ = \ & Tr[S_{12}\check{\Omega}_{n-s}^{-1}V_{12}^T\hat{\Omega}_s^{-1}] - 2Tr[S_{12}^T\hat{\Omega}_s^{-1}V_{12}W_{12}^T] - 2Tr[V_{11}P_{11}\hat{\Omega}_s^{-1}] \\ & -2Tr[V_{12}(\frac{1}{2}\check{\Omega}_{n-s}^{-1}S_{12}^T)\hat{\Omega}_s^{-1}] - 4Tr[V_{12}^T(\frac{1}{2}S_{12}\check{\Omega}_{n-s}^{-1})W_{22}] \\ = \ & Tr[V_{12}\check{\Omega}_{n-s}^{-1}S_{12}^T\hat{\Omega}_s^{-1}] - 2Tr[S_{12}^T\hat{\Omega}_s^{-1}V_{12}W_{12}^T] - 2Tr[V_{11}P_{11}\hat{\Omega}_s^{-1}] \\ & -Tr[V_{12}\check{\Omega}_{n-s}^{-1}S_{12}^T\hat{\Omega}_s^{-1}] - 2Tr[V_{12}^T S_{12}\check{\Omega}_{n-s}^{-1}W_{22}] \\ \\ = \ & -2Tr[V_{11}P_{11}\hat{\Omega}_s^{-1}] - 2Tr[V_{12}^T\hat{\Omega}_s^{-1}S_{12}W_{12}] \\ & -2Tr[V_{12}^T S_{12}\check{\Omega}_{n-s}^{-1}W_{22}] \ . \end{aligned} \tag{A.31}$$

3a. First we look at the first term of the right side of (A.31); we show that

$-2Tr[V_{11}P_{11}\hat{\Omega}_s^{-1}] \geq 0$, where identity is possible only in the case $V_{11} = 0$.

First we evaluate P_{11}. For that we have a look at the 1,1 -block of the equation system (A.17) - (A.20). We obtain the four equations

$$\begin{aligned}
-\hat{\Omega}_s Q_{11}^T - Q_{11}\hat{\Omega}_s &= 0 && \text{(A.32)}\\
-\hat{\Omega}_s P_{11} + R_{11}\hat{\Omega}_s - 2Q_{11}\hat{\Omega}_s &= \frac{1}{2}\hat{\Omega}_s^{-1}V_{11} && \text{(A.33)}\\
\hat{\Omega}_s R_{11} - 2\hat{\Omega}_s Q_{11}^T - P_{11}\hat{\Omega}_s &= \frac{1}{2}V_{11}\hat{\Omega}_s^{-1} && \text{(A.34)}\\
\hat{\Omega}_s Q_{11} - 2\hat{\Omega}_s P_{11} + Q_{11}^T\hat{\Omega}_s - 2P_{11}\hat{\Omega}_s &= \frac{1}{2}\hat{\Omega}_s^{-1}V_{11} + \frac{1}{2}V_{11}\hat{\Omega}_s^{-1} \; . && \text{(A.35)}
\end{aligned}$$

The first equation together with (A.15) leads to

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

where S_{11} is a skew-symmetric matrix. Putting this into equation (A.35), we obtain

$$\frac{1}{2}\hat{\Omega}_s S_{11}\hat{\Omega}_s^{-1} - 2\hat{\Omega}_s P_{11} - \frac{1}{2}\hat{\Omega}_s^{-1}S_{11}\hat{\Omega}_s - 2P_{11}\hat{\Omega}_s = \frac{1}{2}\hat{\Omega}_s^{-1}V_{11} + \frac{1}{2}V_{11}\hat{\Omega}_s^{-1} \; . \tag{A.36}$$

First we determine the diagonal elements of the matrix P_{11}. Because of

$$\frac{1}{2}(\hat{\Omega}_s S_{11}\hat{\Omega}_s^{-1})_{ii} - \frac{1}{2}(\hat{\Omega}_s^{-1}S_{11}\hat{\Omega}_s)_{ii} = 0, (\hat{\Omega}_s P_{11})_{ii} = (P_{11}\hat{\Omega}_s)_{ii}$$

for $i = 1, \cdots, s$ we get with (A.36)

$$(-4\hat{\Omega}_s P_{11})_{ii} = \frac{1}{2}(\hat{\Omega}_s^{-1}V_{11})_{ii} + \frac{1}{2}(V_{11}\hat{\Omega}_s^{-1})_{ii} = \frac{(V_{11})_{ii}}{\omega_i}.$$

So, we obtain

$$\begin{aligned}
(\hat{\Omega}_s P_{11})_{ii} &= -\frac{(V_{11})_{ii}}{4\omega_i}, \text{ d.h.}\\
(P_{11})_{ii} &= -\frac{(V_{11})_{ii}}{4\omega_i^2} \; .
\end{aligned}$$

Now we determine the off-diagonal elements. For $1 \leq i < j \leq s$ equation (A.36) yields

$$\begin{aligned}
&\frac{1}{2}(\hat{\Omega}_s S_{11}\hat{\Omega}_s^{-1})_{ij} - 2(\hat{\Omega}_s P_{11})_{ij} - \frac{1}{2}(\hat{\Omega}_s^{-1}S_{11}\hat{\Omega}_s)_{ij} - 2(P_{11}\hat{\Omega}_s)_{ij}\\
= \; &\frac{1}{2}(\hat{\Omega}_s^{-1}V_{11})_{ij} + \frac{1}{2}(V_{11}\hat{\Omega}_s^{-1})_{ij}.
\end{aligned}$$

With $S_{11} = (f_{ij})$ we get

$$\begin{aligned}
\frac{\omega_i f_{ij}}{2\omega_j} - \frac{f_{ij}\omega_j}{2\omega_i} - 2\omega_i(P_{11})_{ij} - 2(P_{11})_{ij}\omega_j &= \frac{(V_{11})_{ij}}{2\omega_i} + \frac{(V_{11})_{ij}}{2\omega_j}, \text{ i.e.}\\
\omega_i^2 f_{ij} - f_{ij}\omega_j^2 - 4\omega_i^2\omega_j(P_{11})_{ij} - 4\omega_i\omega_j^2(P_{11})_{ij} &= (V_{11})_{ij}\omega_j + (V_{11})_{ij}\omega_i \; .
\end{aligned}$$

So, we have

$$f_{ij}(\omega_i - \omega_j) = (V_{11})_{ij} + 4\omega_i\omega_j(P_{11})_{ij}. \tag{A.37}$$

To get a second dependence between f_{ij} and $(P_{11})_{ij}$, we replace Q_{11} with $\frac{1}{2}S_{11}\hat{\Omega}_s^{-1}$ in equation (A.33) and solve it componente wise.

$$\begin{aligned}
-\hat{\Omega}_s P_{11} + R_{11}\hat{\Omega}_s - S_{11} &= \frac{1}{2}\hat{\Omega}_s^{-1}V_{11}, \\
-(\hat{\Omega}_s P_{11})_{ij} + (R_{11}\hat{\Omega}_s)_{ij} - f_{ij} &= \frac{1}{2}(\hat{\Omega}_s^{-1}V_{11})_{ij}, \text{ d.h.} \\
-\omega_i(P_{11})_{ij} + (R_{11})_{ij}\omega_j - f_{ij} &= \frac{(V_{11})_{ij}}{2\omega_i} .
\end{aligned}$$

Because the matrices P_{11} and R_{11} are symmetric and S_{11} is skew-symmetric, i.e. $(P_{11})_{ij} = (P_{11})_{ji}$,$(R_{11})_{ij} = (R_{11})_{ji}$, $f_{ij} = -f_{ji}$, we obtain the two equations

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

Multiplying the first equation with ω_i and the second equation with ω_j und then subtracting this, leads to

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

therefore

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

i.e.

$$f_{ij} = (P_{11})_{ij}(\omega_j - \omega_i) . \tag{A.38}$$

Now we put (A.38) into (A.37) and obtain

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

and therefore

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

Thus, we get

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

Therefore we have

$$\begin{aligned}
-2Tr[V_{11}P_{11}\hat{\Omega}_s^{-1}] &= -2\sum_{i=1}^{s}\sum_{j=1}^{s}\frac{(V_{11})_{ij}(P_{11})_{ji}}{\omega_i} \\
&= 2\sum_{i=1}^{s}\sum_{j=1}^{s}\frac{(V_{11})_{ij}^2}{\omega_i(\omega_i+\omega_j)^2} \\
&\geq 0\ , \qquad \text{(A.39)}
\end{aligned}$$

where $-2Tr[V_{11}P_{11}\hat{\Omega}_s^{-1}] = 0$ is only possible for $V_{11} = 0$.
So, step 3a. is completed.

3b. At last we consider the both last terms on the right side of (A.31); we show that

$$\hat{\Delta} := -2Tr[V_{12}^T\hat{\Omega}_s^{-1}S_{12}W_{12}] - 2Tr[V_{12}^T S_{12}\check{\Omega}_{n-s}^{-1}W_{22}] \geq 0$$

yields, where equality is only possible for $V_{12} = 0$.

By (A.29) and (A.30) we have

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

Putting this into equation (A.23) we get

$$-\hat{\Omega}_s(\frac{1}{2}S_{12}\check{\Omega}_{n-s}^{-1}) + (-\frac{1}{2}\hat{\Omega}_s^{-1}S_{12})\check{\Omega}_{n-s} - 2(\frac{1}{2}S_{12}\check{\Omega}_{n-s}^{-1})H = \frac{1}{2}\hat{\Omega}_s^{-1}V_{12}\ ,$$

i.e.

$$V_{12}^T = -\check{\Omega}_{n-s}^{-1}S_{12}^T\hat{\Omega}_s^2 - \check{\Omega}_{n-s}S_{12}^T - 2H\check{\Omega}_{n-s}^{-1}S_{12}^T\hat{\Omega}_s\ . \qquad \text{(A.40)}$$

Hence

$$\begin{aligned}
\hat{\Delta} &= 2Tr[\check{\Omega}_{n-s}^{-1}S_{12}^T\hat{\Omega}_s S_{12}W_{12}] + 2Tr[\check{\Omega}_{n-s}S_{12}^T\hat{\Omega}_s^{-1}S_{12}W_{12}] + 4Tr[H\check{\Omega}_{n-s}^{-1}S_{12}^T S_{12}W_{12}] \\
&\quad +2Tr[\check{\Omega}_{n-s}^{-1}S_{12}^T\hat{\Omega}_s^2 S_{12}\check{\Omega}_{n-s}^{-1}W_{22}] + 2Tr[\check{\Omega}_{n-s}S_{12}^T S_{12}\check{\Omega}_{n-s}^{-1}W_{22}] \\
&\quad +4Tr[H\check{\Omega}_{n-s}^{-1}S_{12}^T\hat{\Omega}_s S_{12}\check{\Omega}_{n-s}^{-1}W_{22}] \\
&= 2Tr[\check{\Omega}_{n-s}^{-1}S_{12}^T\hat{\Omega}_s S_{12}W_{12}] + 2Tr[\check{\Omega}_{n-s}S_{12}^T\hat{\Omega}_s^{-1}S_{12}W_{12}] + 2Tr[\check{\Omega}_{n-s}^{-1}S_{12}^T S_{12}2W_{12}H] \\
&\quad 2Tr[\check{\Omega}_{n-s}^{-1}S_{12}^T\hat{\Omega}_s^2 S_{12}\check{\Omega}_{n-s}^{-1}W_{22}] + 2Tr[\check{\Omega}_{n-s}S_{12}^T S_{12}\check{\Omega}_{n-s}^{-1}W_{22}] \\
&\quad +Tr[(2HW_{22} + 2W_{22}H)\check{\Omega}_{n-s}^{-1}S_{12}^T\hat{\Omega}_s S_{12}\check{\Omega}_{n-s}^{-1}]. \qquad \text{(A.41)}
\end{aligned}$$

The block decomposition of the Ljapunov equation (A.9) yields

$$\begin{aligned}
-\check{\Omega}_{n-s}W_{12}^T - W_{12}\check{\Omega}_{n-s} &= -I_{n-s} \\
-\check{\Omega}_{n-s}W_{22} + W_{11}\check{\Omega}_{n-s} - 2W_{12}H &= 0 \qquad \text{(A.42)} \\
\check{\Omega}_{n-s}W_{11} - 2HW_{12}^T - W_{22}\check{\Omega}_{n-s} &= 0 \\
\check{\Omega}_{n-s}W_{12} - 2HW_{22} + W_{12}^T\check{\Omega}_{n-s} - 2W_{22}H &= -I_{n-s}\ .
\end{aligned}$$

The first equation leads to

$$W_{12} = \frac{1}{2}\check{\Omega}^{-1}_{n-s} + \frac{1}{2}G\check{\Omega}^{-1}_{n-s} \ , \tag{A.43}$$

where $G^T = -G$, i.e. G is skew-symmetric. Furthermore the equation system (A.42) leads to

$$\begin{aligned} 2W_{12}H &= -\check{\Omega}_{n-s}W_{22} + W_{11}\check{\Omega}_{n-s} && \text{(A.44)} \\ 2HW_{22} + 2W_{22}H &= \check{\Omega}_{n-s}W_{12} + W_{12}^T\check{\Omega}_{n-s} + I_{n-s} \ . && \text{(A.45)} \end{aligned}$$

Placing (A.43) , (A.44) and (A.45) into (A.41), we obtain

$$\begin{aligned} \hat{\Delta} &= 2\ Tr[\check{\Omega}^{-1}_{n-s}S_{12}^T\hat{\Omega}_s S_{12}W_{12}] + 2\ Tr[\check{\Omega}_{n-s}S_{12}^T\hat{\Omega}_s^{-1}S_{12}(\frac{1}{2}\check{\Omega}^{-1}_{n-s} + \frac{1}{2}G\check{\Omega}^{-1}_{n-s})] \\ &\quad +2\ Tr[\check{\Omega}^{-1}_{n-s}S_{12}^T S_{12}(-\check{\Omega}_{n-s}W_{22} + W_{11}\check{\Omega}_{n-s})] \\ &\quad +2\ Tr[\check{\Omega}^{-1}_{n-s}S_{12}^T\hat{\Omega}_s^2 S_{12}\check{\Omega}^{-1}_{n-s}W_{22}] + 2\ Tr[\check{\Omega}_{n-s}S_{12}^T S_{12}\Omega^{-1}_{n-s}W_{22}] \\ &\quad +Tr[(\check{\Omega}_{n-s}W_{12} + W_{12}^T\check{\Omega}_{n-s} + I_{n-s})\check{\Omega}^{-1}_{n-s}S_{12}^T\hat{\Omega}_s S_{12}\check{\Omega}^{-1}_{n-s}] \\[2ex] &= 2\ Tr[\check{\Omega}^{-1}_{n-s}S_{12}^T\hat{\Omega}_s S_{12}W_{12}] + Tr[\check{\Omega}_{n-s}S_{12}^T\hat{\Omega}_s^{-1}S_{12}\check{\Omega}^{-1}_{n-s}] \\ &\quad +Tr[\check{\Omega}_{n-s}S_{12}^T\hat{\Omega}_s^{-1}S_{12}G\check{\Omega}^{-1}_{n-s}] - 2\ Tr[\check{\Omega}^{-1}_{n-s}S_{12}^T S_{12}\check{\Omega}_{n-s}W_{22}] \\ &\quad +2\ Tr[\check{\Omega}^{-1}_{n-s}S_{12}^T S_{12}W_{11}\check{\Omega}_{n-s}] + 2\ Tr[\check{\Omega}^{-1}_{n-s}S_{12}^T\hat{\Omega}_s^2 S_{12}\check{\Omega}^{-1}_{n-s}W_{22}] \\ &\quad +2\ Tr[\check{\Omega}_{n-s}S_{12}^T S_{12}\check{\Omega}^{-1}_{n-s}W_{22}] + Tr[\check{\Omega}_{n-s}W_{12}\check{\Omega}^{-1}_{n-s}S_{12}^T\hat{\Omega}_s S_{12}\check{\Omega}^{-1}_{n-s}] \\ &\quad +Tr[W_{12}^T\check{\Omega}_{n-s}\check{\Omega}^{-1}_{n-s}S_{12}^T\hat{\Omega}_s S_{12}\check{\Omega}^{-1}_{n-s}] + Tr[\check{\Omega}^{-1}_{n-s}S_{12}^T\hat{\Omega}_s S_{12}\check{\Omega}^{-1}_{n-s}] \\[2ex] &= 2\ Tr[\check{\Omega}^{-1}_{n-s}S_{12}^T\hat{\Omega}_s S_{12}W_{12}] + Tr[S_{12}^T\hat{\Omega}_s^{-1}S_{12}] + Tr[S_{12}^T\hat{\Omega}_s^{-1}S_{12}G] \\ &\quad -2\ Tr[W_{22}\check{\Omega}^{-1}_{n-s}S_{12}^T S_{12}\check{\Omega}_{n-s}] + 2\ Tr[S_{12}^T S_{12}W_{11}] + 2\ Tr[\check{\Omega}^{-1}_{n-s}S_{12}^T\hat{\Omega}_s^2 S_{12}\check{\Omega}^{-1}_{n-s}W_{22}] \\ &\quad +2\ Tr[\check{\Omega}_{n-s}S_{12}^T S_{12}\check{\Omega}^{-1}_{n-s}W_{22}] + Tr[W_{12}\check{\Omega}^{-1}_{n-s}S_{12}^T\hat{\Omega}_s S_{12}] + Tr[W_{12}^T S_{12}^T\hat{\Omega}_s S_{12}\check{\Omega}^{-1}_{n-s}] \\ &\quad +Tr[\check{\Omega}^{-1}_{n-s}S_{12}^T\hat{\Omega}_s S_{12}\check{\Omega}^{-1}_{n-s}]. \end{aligned}$$

Because $S_{12}^T\hat{\Omega}_s^{-1}S_{12}$ is symmetric, we get $Tr[S_{12}^T\hat{\Omega}_s^{-1}S_{12}G] = Tr[G^T S_{12}^T\hat{\Omega}_s^{-1}S_{12}]$, and so, we obtain

$$\begin{aligned} \hat{\Delta} &= 2\ Tr[\check{\Omega}^{-1}_{n-s}S_{12}^T\hat{\Omega}_s S_{12}W_{12}] + Tr[S_{12}^T\hat{\Omega}_s^{-1}S_{12}] \\ &\quad +Tr[\frac{1}{2}S_{12}^T\hat{\Omega}_s^{-1}S_{12}G + \frac{1}{2}G^T S_{12}^T\hat{\Omega}_s^{-1}S_{12}] - 2\ Tr[(W_{22}\check{\Omega}^{-1}_{n-s}S_{12}^T S_{12}\check{\Omega}_{n-s})^T] \\ &\quad +2\ Tr[S_{12}^T S_{12}W_{11}] + 2\ Tr[\check{\Omega}^{-1}_{n-s}S_{12}^T\hat{\Omega}_s^2 S_{12}\check{\Omega}^{-1}_{n-s}W_{22}] \\ &\quad +2\ Tr[\check{\Omega}_{n-s}S_{12}^T S_{12}\check{\Omega}^{-1}_{n-s}W_{22}] + Tr[\check{\Omega}^{-1}_{n-s}S_{12}^T\hat{\Omega}_s S_{12}W_{12}] \\ &\quad +Tr[(W_{12}^T S_{12}^T\hat{\Omega}_s S_{12}\check{\Omega}^{-1}_{n-s})^T] + Tr[\check{\Omega}^{-1}_{n-s}S_{12}^T\hat{\Omega}_s S_{12}\check{\Omega}^{-1}_{n-s}] \end{aligned}$$

$$
\begin{aligned}
= \;& 3\,Tr[\check{\Omega}_{n-s}^{-1}S_{12}^T\hat{\Omega}_s S_{12}W_{12}] + Tr[S_{12}^T\hat{\Omega}_s^{-1}S_{12}] + \frac{1}{2}Tr[S_{12}^T\hat{\Omega}_s^{-1}S_{12}G] \\
& -\frac{1}{2}Tr[S_{12}^T\hat{\Omega}_s^{-1}S_{12}G] - 2\,Tr[\check{\Omega}_{n-s}S_{12}^T S_{12}\check{\Omega}_{n-s}^{-1}W_{22}] + 2\,Tr[S_{12}^T S_{12}W_{11}] \\
& +2\,Tr[\check{\Omega}_{n-s}^{-1}S_{12}^T\hat{\Omega}_s^2 S_{12}\check{\Omega}_{n-s}^{-1}W_{22}] + 2\,Tr[\check{\Omega}_{n-s}S_{12}^T S_{12}\check{\Omega}_{n-s}^{-1}W_{22}] \\
& +Tr[\check{\Omega}_{n-s}^{-1}S_{12}^T\hat{\Omega}_s S_{12}W_{12}] + Tr[\check{\Omega}_{n-s}^{-1}S_{12}^T\hat{\Omega}_s S_{12}\check{\Omega}_{n-s}^{-1}]
\end{aligned}
$$

$$
\begin{aligned}
= \;& Tr[\check{\Omega}_{n-s}^{-1}S_{12}^T\hat{\Omega}_s S_{12}\check{\Omega}_{n-s}^{-1}] + Tr[S_{12}^T\hat{\Omega}_s^{-1}S_{12}] + 4\,Tr[\check{\Omega}_{n-s}^{-1}S_{12}^T\hat{\Omega}_s S_{12}W_{12}] \\
& +2\,Tr[S_{12}^T S_{12}W_{11}] + 2\,Tr[\check{\Omega}_{n-s}^{-1}S_{12}^T\hat{\Omega}_s^2 S_{12}\check{\Omega}_{n-s}^{-1}W_{22}]\;.
\end{aligned}
$$

Because

$$
\begin{aligned}
(\check{\Omega}_{n-s}^{-1}S_{12}^T\hat{\Omega}_s S_{12}\check{\Omega}_{n-s}^{-1}) &= (\hat{\Omega}_s^{1/2}S_{12}\check{\Omega}_{n-s}^{-1})^T(\hat{\Omega}_s^{1/2}S_{12}\check{\Omega}_{n-s}^{-1}) \text{ and} \\
(S_{12}^T\hat{\Omega}_s^{-1}S_{12}) &= (\hat{\Omega}_s^{-\frac{1}{2}}S_{12})^T(\hat{\Omega}_s^{-\frac{1}{2}}S_{12})
\end{aligned}
$$

is positive semidefinite, we get

$$
\begin{aligned}
Tr[S_{12}^T\hat{\Omega}_s^{-1}S_{12}] &\geq 0\;, \\
Tr[\check{\Omega}_{n-s}^{-1}S_{12}^T\hat{\Omega}_s S_{12}\check{\Omega}_{n-s}^{-1}] &\geq 0\;.
\end{aligned}
$$

Because of the positive definitness of $\hat{\Omega}_s$, $Tr[S_{12}^T\hat{\Omega}_s^{-1}S_{12}]$ can only equal zero if $S_{12} = 0$, i.e. (see (A.40)) for $V_{12} = 0$.

It remains to prove

$$
\hat{\hat{\Delta}} :=
4Tr[\check{\Omega}_{n-s}^{-1}S_{12}^T\hat{\Omega}_s S_{12}W_{12}] + 2Tr[S_{12}^T S_{12}W_{11}] + 2Tr[\check{\Omega}_{n-s}^{-1}S_{12}^T\hat{\Omega}_s^2 S_{12}\check{\Omega}_{n-s}^{-1}W_{22}] \geq 0\,.
$$

For this we consider a transformation of the (as a solution of (A.9)) positive definite matrix $\begin{pmatrix} W_{11} & W_{12} \\ W_{12}^T & W_{22} \end{pmatrix}$.

$$
\begin{aligned}
& \begin{pmatrix} S_{12} & 0 \\ 0 & \hat{\Omega}_s S_{12}\check{\Omega}_{n-s}^{-1} \end{pmatrix} \begin{pmatrix} W_{11} & W_{12} \\ W_{12}^T & W_{22} \end{pmatrix} \begin{pmatrix} S_{12} & 0 \\ 0 & \hat{\Omega}_s S_{12}\check{\Omega}_{n-s}^{-1} \end{pmatrix}^T \\
= \;& \begin{pmatrix} S_{12}W_{11} & S_{12}W_{12} \\ \hat{\Omega}_s S_{12}\check{\Omega}_{n-s}^{-1}W_{12}^T & \hat{\Omega}_s S_{12}\check{\Omega}_{n-s}^{-1}W_{22} \end{pmatrix} \begin{pmatrix} S_{12}^T & 0 \\ 0 & \check{\Omega}_{n-s}^{-1}S_{12}^T\hat{\Omega}_s \end{pmatrix} \\
= \;& \begin{pmatrix} S_{12}W_{11}S_{12}^T & S_{12}W_{12}\check{\Omega}_{n-s}^{-1}S_{12}^T\hat{\Omega}_s \\ \hat{\Omega}_s S_{12}\check{\Omega}_{n-s}^{-1}W_{12}^T S_{12}^T & \hat{\Omega}_s S_{12}\check{\Omega}_{n-s}^{-1}W_{22}\check{\Omega}_{n-s}^{-1}S_{12}^T\hat{\Omega}_s \end{pmatrix} \\
=: \;& \begin{pmatrix} \tilde{W}_{11} & \tilde{W}_{12} \\ \tilde{W}_{12}^T & \tilde{W}_{22} \end{pmatrix}
\end{aligned}
$$

is positive semidefinite, i.e. particularly, that we have for all s - dimensional vectors x

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

Because $\tilde{W}_{11}+\tilde{W}_{12}+\tilde{W}_{12}^T+\tilde{W}_{22}$ is symmetric, this matrix is positive semidefinite too, i.e.,

$$\begin{aligned}
& Tr[\tilde{W}_{11} + \tilde{W}_{12}^T + \tilde{W}_{12} + \tilde{W}_{22}] \\
= \; & Tr[S_{12}W_{11}S_{12}^T] + Tr[S_{12}W_{12}\check{\Omega}_{n-s}^{-1}S_{12}^T\hat{\Omega}_s] \\
& + Tr[\hat{\Omega}_s S_{12}\check{\Omega}_{n-s}^{-1}W_{12}^T S_{12}^T] + Tr[\hat{\Omega}_s S_{12}\check{\Omega}_{n-s}^{-1}W_{22}\check{\Omega}_{n-s}^{-1}S_{12}^T\hat{\Omega}_s] \\
= \; & Tr[S_{12}^T S_{12}W_{11}] + Tr[\check{\Omega}_{n-s}^{-1}S_{12}^T\hat{\Omega}_s S_{12}W_{12}] + Tr[(S_{12}W_{12}\check{\Omega}_{n-s}^{-1}S_{12}^T\hat{\Omega}_s)^T] \\
& + Tr[\check{\Omega}_{n-s}^{-1}S_{12}^T\hat{\Omega}_s^2 S_{12}\check{\Omega}_{n-s}^{-1}W_{22}] \\
= \; & Tr[S_{12}^T S_{12}W_{11}] + 2Tr[\check{\Omega}_{n-s}^{-1}S_{12}^T\hat{\Omega}_s S_{12}W_{12}] + Tr[\check{\Omega}_{n-s}^{-1}S_{12}^T\hat{\Omega}_s^2 S_{12}\check{\Omega}_{n-s}^{-1}W_{22}] \\
= \; & 2\,\hat{\hat{\Delta}} \\
\geq \; & 0 \ .
\end{aligned}$$

So, we proved that

$$\hat{\Delta} = -2Tr[V_{12}^T\hat{\Omega}_s^{-1}S_{12}W_{12}] - 2Tr[V_{12}^T S_{12}\check{\Omega}_{n-s}^{-1}W_{22}] \geq 0$$

where equality is only possible for $V_{12} = 0$;
so, step 3 b. is completed.

Altogether we obtain (see (A.31) and (A.39))

$$\frac{d^2}{d\mu^2} Tr Y_\mu \Big|_{\mu=0} \geq 0 \ ,$$

where equality is only possible for $V_{11} = V_{12} = 0$.

Q.E.D.

References

[1] Bartels, R.H., Stewart, G.W., A solution of the matrix equation AX + XB = C, Comm. ACM **15** (1972) 820-826.

[2] Brabender, K., Optimale Dämpfung von linearen Schwingungssystemen, PhD Thesis, FernUniversität Hagen, 1998.

[3] Cox, S.J., Designing of optimal engergy absorption, I: lumped parameter systems, J. Vib. and Acoustics, 1996 (to appear).

[4] Freitas, P., Lancaster,P., Optimal decay of energy for a system of linear oscillators, preprint December 1997.

[5] Gohberg, I., Lancaster, P., Rodman, L., Matrices and indefinite scalar products, Birkhäuser, Basel 1983.

[6] Golub, G. H., van Loan, Ch. F., Matrix computations, J. Hopkins University Press, Baltimore 1989.

[7] Lancaster, P., Lambda-matrices and vibrating systems, Pergamon Press Oxford 1966

[8] Müller, P.C., Schiehlen, W.O., Lineare Schwingungen, Akademische Verlagsgesellschaft Wiesbaden 1976.

[9] Veselić, K., Global Jacobi method for a symmetric indefinite problem $Sx = \lambda Tx$, Comp. Meth. Appl. Mech. Engrg. **38** (1983) 273-290.

[10] Veselić, K., Some convergent Jacobi-like procedures for diagonalizing J-symmetric matrices, Numer. Math. **27** (1976) 67-75.

[11] Veselić, K., On linear vibrational systems with one dimensional damping, Appl. Anal. **29** (1988) 1-18.

[12] Veselić, K., On linear vibrational systems with one dimensional damping II, Integral Eq. Operator Th. **13** (1990) 883-897.

[13] Veselić, K., Exponential decay of semigroups in Hilbert space, Semigroup Forum, **55** (1997) 325-331.

[14] K. Veselić, Estimating the operator exponential, Linear Algebra Appl. **280** (1998) 241-244.