

OPTIMALE DÄMPFUNG VON LINEAREN SCHWINGUNGSSYSTEMEN

DISSERTATION

zur

Erlangung des Grades eines Dr. rer. nat.

des Fachbereichs Mathematik

der FernUniversität – Gesamthochschule – in Hagen

vorgelegt von

KATRIN BRABENDER

aus Bochum

Hagen 1998

Eingereicht im Mai 1998

Tag der mündlichen Prüfung: 8. September 1998

Berichterstatter: Prof. Dr. K. Veselić, Hagen

Prof. Dr. F. Locher, Hagen

Prof. Dr. S. J. Cox, Houston, Texas

Meinem Vater

Inhaltsverzeichnis

Einleitung	1
1 Mathematische Grundlagen und physikalischer Hintergrund	7
2 Minimierung über beliebige positiv definite Dämpfungsmatrizen	17
2.1 Bestimmung von lokalen Minima	18
2.2 Experimentelle Untersuchung der lokalen Minima auf Globalität	45
3 Minimierung über die Viskositäten der Dämpfer	51
3.1 Die Wahl eines guten Startwertes	55
3.2 Optimierung beim Gradientenverfahren	60
3.3 Optimierung beim Newtonverfahren	67
4 Abschneidung des Systems	71
4.1 Abschneidung bei einem Dämpfer	75
4.2 Fazit bei einem Dämpfer	107
4.3 Beispiele bei einem Dämpfer	109
4.3 Zum relativen Fehler der optimalen Viskositäten bei einem Dämpfer ...	116
4.4 Abschneidung bei mehreren Dämpfern	118
5 Anhang: Programmpaket zur Optimierung einer Dämpferkonfiguration	133
Literaturverzeichnis	143
Abbildungsverzeichnis	147

Einleitung

Ein gedämpftes lineares Schwingungssystem wird durch die Differentialgleichung

$$M\ddot{x} + D\dot{x} + Kx = 0 \quad (0.1)$$

beschrieben. Dabei sind die Massenmatrix M , die Steifigkeitsmatrix K und die Dämpfungsmatrix D reelle symmetrische Matrizen der Ordnung n , M und K sind positiv definit, während D positiv semidefinit ist.

Die Optimierung einer durch (0.1) gegebenen Schwingungskonstruktion über die Dämpfungsmatrizen D kann unter verschiedenen Gesichtspunkten erfolgen. In der Literatur findet man im wesentlichen drei Gütekriterien, die die Grundlage für die Optimierung bilden. Ein Gütekriterium ist der sog. Stabilitätsgrad d_0 , der den Abstand der imaginären Achse vom nächstgelegenen Eigenwert von (0.1) darstellt, d.h. $d_0 = -\max\{\operatorname{Re} \lambda_i\}$. Der Stabilitätsgrad d_0 ist ein Maß für das Abklingen der Eigenbewegungen des Systems, also ein Maß für den Energieabfall des Systems. Eine analytische Untersuchung des Energieabfalls im Zusammenhang mit der schwingenden Saite findet man z.B. bei Cox und Zuazua in [6]. Wählt man d_0 als Gütekriterium, so erhält man als Optimierungsproblem

$$\max_i \{\operatorname{Re} \lambda_i\} = -d_0 = \min$$

($-d_0$ wird auch oft als Spektralabszisse bezeichnet). Eine analytische Untersuchung der Spektralabszisse der oben angegebenen Schwingungskonstruktion (0.1) für positiv definite Dämpfungsmatrizen betrachtet Cox in [5]. Eine Untersuchung der Minimierung der Spektralabszisse auf numerischer Basis findet man in [4].

Ein zweites Gütekriterium ist der sog. Dämpfungsgrad $D_0 = -\max_i \left\{ \frac{\operatorname{Re} \lambda_i}{|\lambda_i|} \right\}$, der angibt, nach wieviel Schwingungsperioden eine Anfangsauslenkung auf einen bestimmten Bruchteil abgeklungen ist. Ein großer Dämpfungsgrad führt also zu einem Abklingen der Eigenbewegungen des Systems in wenigen Schwingungsperioden. Ein Vergleich zwischen dem Dämpfungsgrad und dem unten angegebenen Energiekriterium bei einem Dämpfer stellen Müller und Gürgöze in [18] vor.

Ein drittes Gütekriterium stellt schließlich das unbestimmte Zeitintegral über die totale Energie $E(t)$ von (0.1) dar. Dieses Gütekriterium, das auf das Optimierungsproblem

$$\int_0^\infty E(t) dt = \min \quad (0.2)$$

führt, wollen wir dieser Arbeit zugrunde legen. Überführt man die Differentialgleichung (0.1) in eine $2n$ -dimensionale Phasenraumdarstellung $\dot{\Psi} = B\Psi$, so wird das Optimierungsproblem (0.2) in das Problem $\Psi_0^T X \Psi_0 = \min$ überführt. Hierbei ist $\Psi = e^{Bt} \Psi_0$ und X die eindeutige Lösung der Ljapunovgleichung $B^T X + X B = -I$, die stets existiert, wenn das Integral in (0.2) konvergent ist. Damit äquivalent (und somit Generalvoraussetzung) ist wiederum, daß B stabil ist (siehe Kapitel 1).

Um eine Unabhängigkeit von Ψ_0 zu erreichen, bildet Cox in [5] das Maximum von $\Psi_0^T X \Psi_0$ über alle Anfangszustände Ψ_0 , die Einheitsenergie besitzen. Dies führt auf ein Optimierungsproblem, bei dem die Ermittlung eines größten Eigenwertes im Vordergrund steht. Cox untersucht in [5] analytisch eine Minimierung des größten Eigenwertes, wobei er über bestimmte Unterklassen von positiv definiten Dämpfungsmatrizen (solche, die proportional zur Massenmatrix M , bzw. solche, die mit M und K simultan diagonalisierbar sind) minimiert. Die Maximierung über die Anfangszustände bilden auch bei Bendsøe, Olhoff und Taylor in [3] die Grundlage für die Minimierung. Veselić umgeht die Abhängigkeit von Ψ_0 , indem er in [24] über alle Ψ_0 mit Einheitsenergie mittelt. Dies führt auf die Minimierung der Spur einer Matrix, genauer auf

$$\text{Tr}(XZ) = \min, \quad (0.3)$$

wobei X die Lösung einer Ljapunovgleichung $B^T X + X B = -I$ und Z eine reelle symmetrische, positiv semidefinite Matrix ist. Auch Gürgöze und Müller wählen z.B. in [10] für ihre Untersuchungen für einen Dämpfer das Zeitintegral über die Energie des Systems mittels der Spur der Lösung einer Ljapunovgleichung.

Diese Mittelung ist auch dieser Arbeit zugrunde gelegt. Wir werden das folgende Problem untersuchen:

Problem 0.1

Gegeben sei $Z \in M_{2n}$, Z positiv semidefinit, $\Omega = \text{diag}(\omega_1, \dots, \omega_n)$ mit $\omega_i > 0$ für alle $i = 1, \dots, n$. Aus der Menge der positiv semidefiniten Matrizen $C \in M_n$, für die

$$A = \begin{pmatrix} 0 & \Omega \\ -\Omega & -C \end{pmatrix} \quad (0.4)$$

stabil ist, wird das C gesucht, für das gilt

$$\begin{aligned} \text{Tr}(XZ) &= \min, \\ X &\text{ eindeutige Lösung von } A^T X + X A = -I. \end{aligned} \quad (0.5)$$

A ist hierbei eine zu B orthogonal ähnliche Matrix, C ist die aus D transformierte Dämpfungsmatrix und ω_i sind die Kreisfrequenzen des ungedämpften Systems. Für

Z wählen wir die Einheitsmatrix I , falls wir die „volle“ Spur minimieren, also das gesamte Frequenzspektrum betrachten, bzw.

$$Z = \begin{pmatrix} I_s & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & I_s & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad (0.6)$$

wobei I_s die Einheitsmatrix der Ordnung s ist, falls wir nur an einem bestimmten Frequenzspektrum interessiert sind und somit nur einen „Teil der Spur“ minimieren.

Nach einer kurzen Erklärung der in der Arbeit benutzten Notationen wird in Kapitel 1 der Arbeit der physikalische Hintergrund beschrieben sowie in Problem 1.3 die Problemstellung der Arbeit genau formuliert. Eine kurze Einführung in die exakte und numerische Lösung einer Ljapunovgleichung schließt sich an. Außerdem wird das Lemma 1.4 formuliert, das uns angibt, unter welchen Bedingungen die Matrix A aus (0.4) stabil ist, was eine Grundvoraussetzung für die Arbeit ist.

Kapitel 2 enthält zwei wesentliche neue Ergebnisse zur Lösung von Problem 0.1 bzw. Problem 1.3. $\text{Tr}(XZ)$ wird über die Menge aller positiv semidefiniten Matrizen C , für die A stabil ist, minimiert. Die Hauptergebnisse werden in den Sätzen 2.1 und 2.6 dargestellt. In Satz 2.1 geben wir eine Matrix C_{\min} an, für die $\text{Tr}(XZ)$ mit $Z = I$ ein lokales Minimum annimmt (und beweisen damit eine Vermutung von K. Veselić). Satz 2.6 beschäftigt sich mit der Minimierung von einem Teil der Spur, d.h. hier ist Z von der Form (0.6). Es wird eine lineare Mannigfaltigkeit \mathcal{M} von Dämpfungsmatrizen angegeben, für die bewiesen wird, daß $\text{Tr}(XZ)$ für jedes C_{\min} aus \mathcal{M} ein lokales Minimum annimmt. Der Funktionswert $\text{Tr}(XZ)$ ist dabei für jedes C_{\min} aus \mathcal{M} gleich. In einem angeschlossenen Abschnitt zeigen wir – wenn auch nur experimentell –, daß sowohl C_{\min} aus Satz 2.1 als auch jedes C_{\min} aus \mathcal{M} in Satz 2.6 nicht nur lokale Minima, sondern sogar globale Minima liefert; bei breit angelegten Experimenten konnte kein Beispiel gefunden werden, das dieser Behauptung widerspräche.

Sind in einem Schwingungssystem, das durch (0.1) gegeben ist, schon Dämpfer an fest vorgegebenen Positionen enthalten, so kann eine Optimierung der Schwingungskonstruktion entweder durch Hinzufügen von Dämpfern oder durch eine Veränderung der Viskositäten der vorhandenen Dämpfer erreicht werden. Mit letzterem beschäftigen wir uns in Kapitel 3. Die Menge der Dämpfungsmatrizen C , über die minimiert wird, wird in diesem Kapitel eingeschränkt auf $C = V\Sigma V^T$ mit $V \in M_{n,g}$, $g \leq n$ (in der Praxis ist g meist viel kleiner als n), $\Sigma = \text{diag}(\varepsilon_1, \dots, \varepsilon_g)$. Dabei enthält die Matrix V die Positionen der g Dämpfer und ist hier fest vorgegeben, während Σ die nichtnegativen Viskositäten der Dämpfer enthält, über die minimiert wird. Veselić hat in [24] das Optimierungsproblem für $g = 1$, d.h. für einen Dämpfer untersucht und gezeigt, daß in diesem Fall die Funktion $\text{Tr}(XZ)$ in Abhängigkeit von der Viskosität ε_1 eine konvexe Funktion ist. In Lemma 3.2 wird

explizit die Gestalt der Diagonalelemente von X für den Fall $n = 2$ mit $g = 2$, d.h. zwei Dämpfer, angegeben. Für beliebige Dimensionen und Dämpferzahlen ist ein Programmpaket entwickelt worden, das $\text{Tr}(XZ)$ numerisch minimiert. Die Programmbeschreibung hierzu findet man im Anhang der Arbeit. Der Abschnitt 3.1 enthält die Beschreibung der Startwertberechnung, die auf den Ergebnissen aus Kapitel 2 basiert, Abschnitt 3.2 und 3.3 beschreiben (nach einer Idee von Veselić) eine Optimierung der Berechnung des Gradienten und der Hesseschen Matrix, die im Programmverlauf zu bestimmen sind. All dies ist im neuen Programmpaket implementiert worden. In diesen Abschnitten werden durchgeführte Tests beschrieben und graphisch verdeutlicht.

Da eine Minimierung von $\text{Tr}(XZ)$ trotz der in Kapitel 3 vorgestellten optimierten Programme für große Dimension n sehr aufwendig ist, beschäftigen wir uns in Kapitel 4 mit der Reduzierung des Minimierungsproblems auf ein Problem niedrigerer Dimension, die in [29] angeregt wurde. Wir untersuchen, ob (und wenn ja, unter welchen Bedingungen) eine Matrix

$$\tilde{A} = \begin{pmatrix} 0 & \tilde{\Omega} \\ -\tilde{\Omega} & -\tilde{C} \end{pmatrix} \in M_{2r}$$

existiert, wobei $\tilde{\Omega}$ und \tilde{C} aus den ersten r Zeilen und Spalten von Ω und C bestehen, so daß $\text{Tr}(XZ) \approx \text{Tr}(\tilde{X}\tilde{Z})$ gilt, wenn \tilde{X} die eindeutige Lösung der Ljapunovgleichung $\tilde{A}^T \tilde{X} + \tilde{X} \tilde{A} = -I$ ist. Mit „ungefähr gleich“ meinen wir dabei, daß der relative Fehler der Spur $\left| \frac{\text{Tr}(XZ) - \text{Tr}(\tilde{X}\tilde{Z})}{\text{Tr}(XZ)} \right|$ möglichst klein wird. Lemma 4.2 zeigt die Gestalt, die C haben muß, damit der relative Fehler der Spur exakt Null ist. In Abschnitt 4.1 wird das Problem der Abschneidung für $g = 1$ explizit untersucht. Eine Reihe dabei auftretender Effekte wird teils aus der Darstellung in Satz 4.4, teils aus der Abschätzung in Satz 4.13 hergeleitet. Satz 4.16 zeigt, daß die Aussagen dieses Kapitels in gewissem Sinne im Grenzwert $n \rightarrow \infty$ erhalten bleiben, Satz 4.17 zeigt, daß der relative Fehler der Spur für $r \rightarrow \infty$ beliebig klein wird. Die – teilweise unerwarteten – Ergebnisse des Abschnitts 4.1 sind in Abschnitt 4.2 noch einmal qualitativ zusammengefaßt und werden in Abschnitt 4.3 anhand von Beispielen verdeutlicht. Die naheliegende Vermutung, daß eine beliebige Verkleinerung der Komponenten der Dämpfungsmatrix $C = (c_{ij})$, die hinter dem Schnitt liegen, d.h. für die $i, j > r$ gilt, zu einer beliebigen Verkleinerung des relativen Fehlers der Spur führt, wird in Korollar 4.7 bzw. Bemerkung 4.8 widerlegt. In Abschnitt 4.4 werden entsprechende Ergebnisse für den relativen Fehler der Minimumstelle festgehalten. Mit der Übertragung einzelner Ergebnisse für einen Dämpfer auf den Fall mehrerer Dämpfer beschäftigen wir uns in Abschnitt 4.5.

Der Anhang schließlich enthält die Programmbeschreibung des in Kapitel 3 vorgestellten Programmpakets.

An dieser Stelle bedanke ich mich bei allen, die mich bei der Erstellung dieser Arbeit durch hilfreiche Diskussionen oder auf andere Weise unterstützt haben. Mein besonderer Dank gilt Herrn Prof. Dr. K. Veselić für das interessante Thema, die wertvollen Anregungen und die konstruktiven Diskussionen, mit denen er den Fortgang der Arbeit unterstützt hat.

1 Mathematische Grundlagen und physikalischer Hintergrund

In diesem Kapitel wollen wir die in der Arbeit verwendeten Notationen festhalten sowie auf den physikalischen Hintergrund eingehen. Des Weiteren werden wir auf die theoretische Lösbarkeit einer Ljapunovgleichung eingehen und numerische Lösungsansätze kurz beschreiben. Außerdem formulieren wir ein Lemma, das uns angibt, unter welchen Bedingungen die in der Arbeit betrachtete Phasenraummatrix A stabil ist. Dies ist eine Grundvoraussetzung für die Arbeit.

Die allgemeinen Bezeichnungen dieser Arbeit orientieren sich im wesentlichen an Golub [8] und Horn/Johnson [13]. Wir betrachten stets nur reelle Matrizen $A = (A)_{ij}$ und bezeichnen den Raum der Matrizen, die n Zeilen und m Spalten besitzen, mit $M_{n,m}$. Den Raum der quadratischen Matrizen kennzeichnen wir mit M_n statt mit $M_{n,n}$. Eine positiv (semi)definite Matrix A beschreiben wir mit $A > 0$ (bzw. mit $A \geq 0$). ($A \geq 0$ oder $A > 0$ ist also stets im Sinne der quadratischen Formen gemeint.) $\sigma(A)$ bezeichnet das Spektrum von A , $\operatorname{Re} \sigma(A)$ bzw. $\operatorname{Im} \sigma(A)$ kennzeichnet den Realteil bzw. den Imaginärteil des Spektrums von A . $\operatorname{Tr}(A)$, $\det(A)$ stehen für die Spur bzw. die Determinante von A , $\operatorname{diag}(a_1, \dots, a_n)$ kennzeichnet eine Diagonalmatrix mit den Diagonalelementen a_1, \dots, a_n . Entsprechend ist $\operatorname{diag}(U, V)$ mit $U \in M_n$ und $V \in M_m$ eine Blockdiagonalmatrix. Die Norm $\|\cdot\|_2$ bezeichnet für einen Vektor stets die Euklidische Norm, Normen von Matrizen sind stets die Frobeniusnorm und werden mit $\|\cdot\|_F$ gekennzeichnet. Weitere in der Arbeit benutzte Begriffe werden im laufenden Text erläutert.

Wir betrachten ein gedämpftes lineares Schwingungssystem, das durch die Differentialgleichung

$$M\ddot{x} + D\dot{x} + Kx = 0 \tag{1.1}$$

gegeben ist. M , D , K sind reelle, symmetrische Matrizen der Ordnung n , wobei die *Massenmatrix* M und die *Steifigkeitsmatrix* K positiv definit sind und die *Dämpfungsmatrix* D positiv semidefinit ist. Die totale Energie (Summe aus kinetischer und potentieller Energie) des Systems ist gegeben durch die Hamilton-Funktion (vgl. [19])

$$E(t) = \frac{1}{2}\dot{x}^T M \dot{x} + \frac{1}{2}x^T K x. \tag{1.2}$$

Wir wollen das lineare Schwingungssystem (1.1) in seine $2n$ -dimensionale Phasenraumdarstellung überführen. Da M und K positiv definite Matrizen sind, existiert eine Cholesky-Zerlegung (s. z.B. [13]), d.h. es gibt nichtsinguläre untere Dreiecksmatrizen L_1 und L_2 mit $K = L_1 L_1^T$, $M = L_2 L_2^T$. Da K und M reell sind, können L_1 und L_2 auch als reell angenommen werden. Mit $y_1 := L_1^T x$ und $y_2 := L_2^T \dot{x}$ erhält man eine zu (1.1) gehörige $2n$ -dimensionale Phasenraumdarstellung der Form

$$\dot{\Psi} = B\Psi \quad (1.3)$$

mit

$$\Psi = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & L_1^T L_2^{-T} \\ -L_2^{-1} L_1 & -L_2^{-1} D L_2^{-T} \end{pmatrix}. \quad (1.4)$$

Das Eigenwertproblem $B\Psi = \lambda\Psi$ entspricht dem zu (1.1) gehörigen Eigenwertproblem $(\lambda^2 M + \lambda D + K)x = 0$. Für die Eigenwerte dieses Problems gilt stets $\operatorname{Re} \lambda \leq 0$ (genauer s. Lemma 1.4). Die Lösung von (1.3) mit der Anfangsbedingung Ψ_0 ist durch $\Psi = e^{Bt} \Psi_0$ gegeben.

Ein grundlegender Begriff bei der Untersuchung von Schwingungsproblemen ist die *Stabilität*:

Definition 1.1 (vgl. [16])

$B \in M_n$ heißt stabil, wenn $\operatorname{Re} \sigma(B) < 0$ gilt.

In der Literatur wird oft auch die Formulierung „ B ist asymptotisch stabil“ gebraucht.

Wir wollen nun die gegebene Schwingungskonstruktion in Abhängigkeit von der Dämpfungsmatrix D optimieren. Hierfür gibt es im wesentlichen drei Gütekriterien (s. [19]), die zur Optimierung einer Schwingungskonstruktion gewählt werden können.

Ein Gütekriterium ist durch den sog. *Stabilitätsgrad*, der den Abstand d_0 der imaginären Achse vom nächstgelegenen Eigenwert von (1.1) bzw. von (1.3) darstellt, gegeben. Der Stabilitätsgrad kann also geschrieben werden als $d_0 = -\max_i \{\operatorname{Re} \lambda_i\}$, wobei λ_i die Eigenwerte von (1.3) sind. Oft bezeichnet man $-d_0$ auch als *Spektralabszisse* (s. [4]). Nach [19] gilt näherungsweise

$$\|\Psi(t)\|_2 < c \cdot e^{-d_0 t},$$

d.h. eine Vergrößerung des Stabilitätsgrad d_0 führt zu einem schnelleren Abklingen der Eigenbewegung des Systems. Damit ist bei dieser Wahl des Gütekriteriums ein Ziel, das Spektrum von B so weit wie möglich in die linke Halbebene zu schieben, d.h. das Optimierungsproblem

$$\max_i \{\operatorname{Re} \lambda_i\} = -d_0 = \min \quad (1.5)$$

zu lösen. Oft wird dieses Optimierungsproblem nur über solche Eigenwerte gelöst, die in einem bestimmten Frequenzstreifen liegen, d.h. für die

$$|\operatorname{Im}\lambda_i| \leq M, \quad M > 0 \quad (1.6)$$

gilt.

Ein zweites Gütekriterium ist der sog. *Dämpfungsgrad* D_0 , der durch

$$D_0 = -\max_i \left\{ \frac{\operatorname{Re} \lambda_i}{|\lambda_i|} \right\} \quad (1.7)$$

gegeben ist, wobei λ_i wieder die Eigenwerte von (1.3) sind. Der Dämpfungsgrad ist ein Maß für den Öffnungswinkel des Sektors (mit Spitze in 0), in dem die Eigenwerte von (1.3) enthalten sind, und gibt an, nach wieviel Schwingungsperioden eine Anfangsauslenkung auf einen bestimmten Bruchteil abgeklungen ist. Ein großer Dämpfungsgrad führt zu einem Abklingen der Eigenbewegungen in wenigen Schwingungsperioden. $D_0 = 1$ entspricht einem rein reellen Spektrum, also dem aperiodischen Grenzfall. Sowohl der Stabilitätsgrad als auch der Dämpfungsgrad sind unabhängig von der Anfangsbedingung des Systems.

Ein weiteres Gütekriterium ist das Zeitintegral über die totale Energie (1.2); diese Wahl führt auf das Optimierungsproblem

$$\int_0^\infty E(t) dt = \min, \quad (1.8)$$

das z.B. auch Veselić in [24] betrachtet (s. dort). Dieses Gütekriterium wollen wir dieser Arbeit zugrunde legen; wir wollen also die Schwingungskonstruktion (1.1) in Abhängigkeit von der Dämpfungsmatrix D optimieren, wobei wir dies im Sinne des Problems (1.8) meinen. Im folgenden wollen wir unsere Wahl präzisieren und sehen, welche Vorteile damit verbunden sind.

Für Ψ aus (1.4) gilt

$$\begin{aligned} \|\Psi\|_2^2 &= \Psi^T \Psi = y_1^T y_1 + y_2^T y_2 = (L_1^T x)^T (L_1^T x) + (L_2^T \dot{x})^T (L_2^T \dot{x}) \\ &= x^T L_1 L_1^T x + \dot{x}^T L_2 L_2^T \dot{x} = x^T K x + \dot{x}^T M \dot{x}, \end{aligned}$$

also (siehe (1.2))

$$\|\Psi\|_2^2 = 2E(t).$$

Damit entspricht das Quadrat der Euklidischen Norm der Lösung von (1.3) der doppelten totalen Energie des Schwingungssystems. Hiermit erhalten wir für das Zeitintegral über die totale Energie $E(t)$

$$\begin{aligned} \int_0^\infty E(t) dt &= \frac{1}{2} \int_0^\infty \Psi^T \Psi dt = \frac{1}{2} \int_0^\infty (e^{Bt} \Psi_0)^T (e^{Bt} \Psi_0) dt = \frac{1}{2} \int_0^\infty \Psi_0^T e^{B^T t} e^{Bt} \Psi_0 dt \\ &= \frac{1}{2} \Psi_0^T \int_0^\infty e^{B^T t} e^{Bt} dt \Psi_0 = \frac{1}{2} \Psi_0^T X \Psi_0 \end{aligned} \quad (1.9)$$

mit

$$X = \int_0^{\infty} e^{B^T t} e^{Bt} dt. \quad (1.10)$$

Dieses Integral ist genau dann absolut konvergent, falls B stabil ist. Unter dieser Voraussetzung ist X gleichzeitig die (symmetrische) eindeutig bestimmte Lösung der Ljapunovgleichung $B^T X + X B = -I$. Dies werden wir später weiter ausführen.

Mit (1.9) ist das Optimierungsproblem (1.8) auf das Problem

$$\Psi_0^T X \Psi_0 = \min \quad (1.11)$$

übertragen worden. Hier liegt allerdings eine Abhängigkeit von der Anfangsbedingung Ψ_0 vor. Um diese zu vermeiden, mitteln wir über alle Anfangszustände Ψ_0 mit derselben Energie, d.h. wir betrachten die Abbildung

$$X \mapsto \int_{\|\Psi_0\|_2=1} \Psi_0^T X \Psi_0 d\mu, \quad (1.12)$$

wobei μ ein beliebiges Maß auf der Einheitskugel $S^{2n} = \{\Psi_0 \in \mathbb{R}^{2n}, \|\Psi_0\|_2 = 1\}$ ist. Diese Abbildung ist ein lineares Funktional auf dem Raum der symmetrischen Matrizen aus M_{2n} . Durch $\langle X, Y \rangle = \text{Tr}(YX)$ wird auf diesem Raum ein Skalarprodukt definiert. Nach dem Satz von Riesz (s. z.B. [15]) existiert eine eindeutige symmetrische Matrix $Z' \in M_{2n}$ mit

$$X \mapsto \int_{\|\Psi_0\|_2=1} \Psi_0^T X \Psi_0 d\mu = \langle X, Z' \rangle = \text{Tr}(Z'X).$$

Z' ist stets positiv semidefinit: Sei $y \in \mathbb{R}^{2n}$ beliebig. Setze $X = yy^T \geq 0$. Dann gilt

$$0 \leq \int_{\|\Psi_0\|_2=1} \Psi_0^T X \Psi_0 d\mu = \langle X, Z' \rangle = \text{Tr}(Z'X) = \text{Tr}(Z'yy^T) = \text{Tr}(y^T Z' y).$$

Damit ist Z' positiv semidefinit.

Unser anfängliches Optimierungsproblem hat sich also auf das Problem

$$\begin{aligned} \text{Tr}(Z'X) &= \min, \\ X &\text{ eindeutige Lösung von } B^T X + X B = -I \end{aligned} \quad (1.13)$$

übertragen, wobei Z' eine symmetrische, positiv semidefinite Matrix ist.

Bemerkung 1.2

Man mache sich klar, daß mit dem kanonischen Maß der Gleichverteilung auf S^{2n} dabei die Einheitsmatrix $Z' = I$ korrespondiert. Durch ein Z' , das nicht vollen Rang, also einen nichttrivialen Nullraum hat, erreicht man, daß $d\mu$ auf bestimmten Anfangsbedingungen verschwindet, also z.B. nur über Zustände gemittelt wird, die zu einem bestimmten Frequenzbereich gehören.

Vorteile des Problems (1.13) gegenüber den anderen aufgeführten Ansätzen können in der direkten Verbindung zur Gesamtenergie gesehen werden sowie darin, daß $\text{Tr}(Z'X)$ eine glatte Funktion der Dämpfungsparameter ist, so daß die Minimierung mit Hilfe klassischer Minimierungsverfahren wie Gradienten- oder Newtonverfahren durchgeführt werden kann; gleichzeitig hat Veselić gezeigt, daß auch die Verbindung zur Lösung X der Ljapunovgleichung Abschätzungen für den Energieabfall des Systems erlaubt, z.B. (s. [27], [28])

$$E(t) \leq \frac{\Psi_0^T X \Psi_0}{2h} \left(1 + \frac{h}{\|X\|} \right)^{2-\frac{t}{h}}$$

für $h > 0$ beliebig, $t > h$. $\|X\|$ bezeichnet hier die Spektralnorm von X .

Im allgemeinen ergeben die durch (1.5), (1.7) und (1.8) bzw. (1.13) beschriebenen Gütekriterien zur Optimierung einer Schwingungskonstruktion unterschiedliche optimale Dämpfungsmatrizen D . Es können sich aber auch dieselben optimalen Matrizen D ergeben, wie z.B. beim sog. Ein-Massen-Oszillator, der durch $M = m$, $K = k$ und $D = d$ mit $m, k, d > 0$ gegeben ist. Hier führt (s. [24]) eine Optimierung nach dem Gütekriterium (1.5) und eine Optimierung nach (1.8) zu dem selben optimalen $D = d$.

Eine zu B (s. (1.4)) orthogonal ähnliche Matrix A erhalten wir durch die Singulärwertzerlegung (s. z.B. [13], [8])

$$L_2^{-1}L_1 = W_2\Omega W_1^T. \quad (1.14)$$

Dabei sind die Matrizen W_1 und W_2 reell und orthogonal und die Matrix Ω ist eine Diagonalmatrix mit $\Omega = \text{diag}(\omega_1, \dots, \omega_n)$. O.B.d.A. kann angenommen werden, daß $\omega_1 \leq \omega_2 \leq \dots \leq \omega_n$ gilt. Die ω_i , $1 \leq i \leq n$ sind die positiven Wurzeln der Eigenwerte von $(L_2^{-1}L_1)(L_2^{-1}L_1)^T$, d.h. alle ω_i erfüllen die Gleichung $\det(\omega_i^2 I - (L_2^{-1}L_1 L_1^T L_2^{-T})) = \det(L_2 L_2^T \omega_i^2 - L_1 L_1^T) = \det(M\omega_i^2 - K) = 0$. Damit sind die ω_i genau die ungedämpften Kreisfrequenzen des gegebenen Schwingungssystems (1.1), d.h. die Frequenzen von (1.1) für $D = 0$. Damit sind alle $\omega_i > 0$. Mit $W := \text{diag}(W_1, W_2)$, W_1, W_2 aus (1.14) erhalten wir

$$\begin{aligned} W^T B W &= \begin{pmatrix} W_1^T & 0 \\ 0 & W_2^T \end{pmatrix} \begin{pmatrix} 0 & L_1^T L_2^{-T} \\ -L_2^{-1} L_1 & -L_2^{-1} D L_2^{-T} \end{pmatrix} \begin{pmatrix} W_1 & 0 \\ 0 & W_2 \end{pmatrix} \\ &= \begin{pmatrix} 0 & W_1^T L_1^T L_2^{-T} W_2 \\ -W_2^T L_2^{-1} L_1 W_1 & -W_2^T L_2^{-1} D L_2^{-T} W_2 \end{pmatrix} = \begin{pmatrix} 0 & W_1^T (W_2 \Omega W_1^T)^T W_2 \\ -W_2^T W_2 \Omega W_1^T W_1 & -W_2^T L_2^{-1} D L_2^{-T} W_2 \end{pmatrix} \\ &= \begin{pmatrix} 0 & \Omega \\ -\Omega & -\Phi^T D \Phi \end{pmatrix} = \begin{pmatrix} 0 & \Omega \\ -\Omega & -C \end{pmatrix} =: A \end{aligned}$$

mit

$$\Phi := L_2^{-T} W_2, \quad C := \Phi^T D \Phi. \quad (1.15)$$

Es gilt

$$\begin{aligned} \Phi^T M \Phi &= W_2^T L_2^{-1} L_2 L_2^T L_2^{-T} W_2 = W_2^T W_2 = I, \\ \Phi^T K \Phi &= W_2^T L_2^{-1} L_1 L_1^T L_2^{-T} W_2 = \Omega W_1^{-1} L_1^T L_2^{-T} W_2 \\ &= \Omega W_1^{-1} (W_2^T W_2^{-T} \Omega W_1^{-1})^T = \Omega^2. \end{aligned}$$

Somit haben wir eine Darstellung des Phasenraums erhalten, deren Zustände den ungedämpften Kreisfrequenzen des ursprünglichen Schwingungssystems unmittelbar zugeordnet werden können. Anstelle von (1.13) untersuchen wir somit in der folgenden Arbeit das

Problem 1.3

Gegeben sei $Z \in M_{2n}$, Z positiv semidefinit, $\Omega = \text{diag}(\omega_1, \dots, \omega_n)$ mit $\omega_i > 0$ für alle $i = 1, \dots, n$. Aus der Menge der positiv semidefiniten Matrizen $C \in M_n$, für die

$$A = \begin{pmatrix} 0 & \Omega \\ -\Omega & -C \end{pmatrix} \quad (1.16)$$

stabil ist, wird das C gesucht, für das gilt

$$\begin{aligned} \text{Tr}(XZ) &= \min, \\ X &\text{ eindeutige Lösung von } A^T X + X A = -I. \end{aligned} \quad (1.17)$$

Dabei geht die Matrix Z aus Z' hervor durch die Transformation mit W (s.o.)

$$Z = W^T Z' W.$$

Um dies zu verdeutlichen, bezeichnen wir die eindeutig bestimmte Lösung der Ljapunovgleichung (1.13) mit Y , d.h. Y löst $B^T Y + Y B = -I$. Aus $A^T X + X A = -I$ folgt $W^T B^T W X + X W^T B W = -I$, also $B^T W X W^T + W X W^T B = -W W^T = -I$. Damit ist $W X W^T = Y$, und es folgt $\text{Tr}(Y Z') = \text{Tr}(W X W^T Z') = \text{Tr}(X W^T Z' W) = \text{Tr}(X Z)$. Wenn Z' eine Diagonalmatrix war, ist auch Z eine Diagonalmatrix, insbesondere ist $Z = I$ für $Z' = I$.

Schon bei (1.13) hatten wir die Bedingung (1.10) durch die für stabiles B äquivalente Eigenschaft ersetzt, daß X Lösung der Ljapunovgleichung $B^T X + X B = -I$ ist. Im folgenden sollen daher die grundlegenden Eigenschaften der Ljapunovgleichung betrachtet werden, die im weiteren Verlauf der Arbeit immer wieder eine wesentliche Rolle spielen werden.

Die *Ljapunovgleichung*

$$A^T X + X A = -F \quad (1.18)$$

mit $A \in M_n$, $F \in M_n$ ist ein Spezialfall der Sylvestergleichung $AX - XB = F$. Sie hat für alle $F \in M_n$ eine eindeutig bestimmte Lösung X genau dann, wenn für das Spektrum von A die Bedingung $\sigma(A) \cap \sigma(-A) = \emptyset$ gilt (s. z.B. [14] Kor. 4.4.7). Dies ist jedenfalls erfüllt, falls A stabil ist, d.h. falls $\text{Re}(\sigma(A)) < 0$ gilt. Für alle symmetrischen Matrizen F ergibt sich im Falle der Stabilität von A eine symmetrische Lösung X (s. z.B. [14] Kor. 4.4.10), ist F positiv (semi)definit, so ist dies auch X (s. z.B. [22], [7]). Insbesondere ist X reell für reelle Matrizen A und F . Ist A stabil, dann läßt sich die eindeutig bestimmte Lösung X der Ljapunovgleichung (1.18) als konvergentes uneigentliches Integral

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

darstellen (s. [19]). Gleichzeitig hat X im Falle der Stabilität von A die Darstellung

$$X = \int_0^{\infty} e^{A^T t} F e^{At} dt \quad (1.19)$$

(s. [16]), wobei auch dieses Integral absolut konvergent ist (vgl. mit (1.10)).

Es gibt verschiedene Algorithmen zur numerischen Lösung einer Ljapunovgleichung $A^T X + X A = -F$. Der wohl bekannteste Algorithmus wurde von Bartels und Stewart [2] (1972) entwickelt und liefert die numerische Lösung der allgemeineren Sylvestergleichung $AX - XB = -F$ für beliebige Matrizen $A \in M_m$, $B \in M_n$ und $F \in M_{m,n}$ (falls diese eindeutig lösbar ist, was für $\sigma(A) \cap \sigma(B) = \emptyset$ (s. z.B. [14]) gilt). Der Algorithmus löst die Sylvestergleichung, indem A^T und B mittels der reellen Schur-Zerlegungen $U^T A^T U = \tilde{A}$, $V^T B V = \tilde{B}$, wobei U, V orthogonale Matrizen sind, auf reelle obere quasi-Dreiecksgestalt transformiert werden (eine Matrix hat quasi-Dreiecksgestalt, wenn sie eine Dreiecksmatrix ist, wobei die Diagonalelemente aus Matrizen $\in M_2$ bestehen können). Mit $\tilde{F} = U^T F V$ ergibt sich dann eine neue Sylvestergleichung $\tilde{A}^T \tilde{X} - \tilde{X} \tilde{B} = \tilde{F}$, die direkt gelöst werden kann. X ergibt sich schließlich aus $X = U \tilde{X} V^T$. Die Reduzierung von A^T und B auf ihre reelle Schur-Form erfolgt dabei, indem die Matrizen zuerst auf obere Hessenberg-Form und diese dann durch den QR-Algorithmus auf die reelle Schur-Form gebracht werden (s. hierzu z.B. [8]). Eine Variante zum Algorithmus von Bartels und Stewart liefert der Algorithmus von Hammarling [12] (1981), der die Lösung der Ljapunovgleichung liefert, falls A stabil ist und F nicht negativ definit ist. Dabei wird der Choleskyfaktor von X ermittelt, ohne die Lösung X zu kennen. Auf diesen Algorithmus wollen wir hier nicht näher eingehen. Weitere Algorithmen findet man z.B. in [9], [11], [20]. Algorithmen zur Lösung von Ljapunovgleichungen mit dünn besetzten Matrizen großer Dimension findet man z.B. in [21], [30].

Im folgenden wollen wir untersuchen, unter welchen Bedingungen das uns speziell

interessierende A aus (1.16) stabil ist, d.h. wann die Ljapunovgleichung (1.18) eine eindeutig bestimmte Lösung hat.

Lemma 1.4

Sei

$$A = \begin{pmatrix} 0 & \Omega \\ -\Omega & -C \end{pmatrix},$$

$\Omega = \text{diag}(\omega_1, \dots, \omega_n)$, $\omega_i > 0$ für alle $1 \leq i \leq n$, C positiv semidefinit.

A ist genau dann stabil, wenn auf jedem Teilraum $Y \subset \mathbb{R}^n$, für den $\Omega|_Y = \omega I|_Y$ mit $\omega \in \mathbb{R}$ gilt, C vollen Rang hat.

Beweis

Der Beweis kann auf die Ergebnisse in [19], Kap. 5.2 zurückgeführt werden, soll hier aber unabhängig geführt werden. Für alle $\lambda \in \sigma(A)$ gilt $\text{Re } \lambda \leq 0$, da wegen

$$A = \begin{pmatrix} 0 & \Omega \\ -\Omega & C \end{pmatrix} = \begin{pmatrix} 0 & \Omega \\ -\Omega & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & -C \end{pmatrix}$$

A Summe einer schiefsymmetrischen und einer negativ semidefiniten Matrix ist. Damit ist A genau dann instabil, wenn ein $\lambda \in \sigma(A)$ mit $\text{Re } \lambda = 0$, d.h. $\lambda = i\omega$, $\omega \in \mathbb{R}$ existiert; also genau dann, wenn für einen Vektor $\begin{pmatrix} x \\ y \end{pmatrix} \neq 0$ gilt

$$\begin{pmatrix} 0 & \Omega \\ -\Omega & -C \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = i\omega \begin{pmatrix} x \\ y \end{pmatrix},$$

also genau dann, wenn die beiden Bedingungen

$$\begin{aligned} (i) \quad \Omega y &= i\omega x, \\ (ii) \quad -\Omega x - Cy &= i\omega y \end{aligned}$$

mit $\begin{pmatrix} x \\ y \end{pmatrix} \neq 0$ gelten.

Dies nehmen wir an und führen eine Fallunterscheidung durch:

a) Sei $\omega = 0$. Dann folgt $\Omega y = 0$, also $y = 0$. Damit ergibt sich $\Omega x = 0$, d.h. $x = 0$, was ein Widerspruch zur Instabilität ist.

b) Sei $\omega \neq 0$. Dann ist nach (i) $x = -\frac{i}{\omega}\Omega y$, also eingesetzt in (ii) $\frac{i}{\omega}\Omega^2 y - i\omega y = Cy$ bzw. $i(\Omega^2 - \omega^2)y = \omega Cy$. Es gilt also $i((\Omega^2 - \omega^2)y, y) = \omega(Cy, y)$. Die letzte Gleichung ist nur für $(Cy, y) = 0$, also (da C positiv semidefinit ist) für $Cy = 0$ erfüllt. Daraus wiederum folgt $\Omega^2 y = \omega^2 y$, also $\Omega y = \omega y$, was zusammen mit $Cy = 0$ ein Widerspruch zur Voraussetzung ist, wenn $y \neq 0$ ist. Also gilt $y = 0$ und damit auch $x = 0$, und A ist nicht instabil.

Sei umgekehrt die Voraussetzung nicht erfüllt; d.h. es gibt $\omega \in \mathbb{R}$, $y \neq 0$ mit $\Omega y = \omega y$ und $Cy = 0$. Mit $x = -iy$ gelten dann i) und ii), und $i\omega$ ist Eigenwert von A , d.h. A ist instabil.

Q.E.D.

Bemerkung 1.5

(i) Gilt $\omega_i \neq \omega_{i+1}$ für alle $1 \leq i \leq n-1$, dann ist A genau dann stabil, wenn C keine Nullspalte und Nullzeile hat.

(ii) Gilt $\omega_i = \omega_{i+1}$ für alle $1 \leq i \leq n-1$, dann ist A genau dann stabil, wenn C positiv definit ist.

Enthält C eine Nullspalte, bzw. eine Nullzeile, dann bedeutet dies, daß die Schwingung, die zu der Frequenz ω_i gehört, die dieser Zeile entspricht, ungedämpft ist. Man erkennt dann sofort, daß das Integral (1.19) nicht konvergent ist.

Bemerkung 1.6

Hat C Rang 1, d.h. gilt $C = cc^T$, wobei c ein Vektor ist, dann ist eine notwendige Bedingung für die Stabilität von A , daß die Frequenzen $\omega_1, \dots, \omega_n$ alle voneinander verschieden sind. Ist dies erfüllt und hat C eine Nullzeile und die entsprechende Nullspalte, dann ist nach Lemma 1.4 A nicht stabil. Nach Veselić [26] kann in diesem Fall aber das Problem auf ein kleineres Problem reduziert werden, in dem C keine Nullzeile und Nullspalte hat. Für den Fall $C = cc^T$ bei getrennten Frequenzen können wir also o.B.d.A. annehmen, daß C keine Nullspalte und Nullzeile besitzt.

2 Minimierung über beliebige positiv definite Dämpfungsmatrizen

Cox zeigt in [5], daß das Maximum des Energieintegrals über alle Anfangszustände mit Einheitsenergie (dies ist genau die Norm der Ljapunovlösung X) eine konvexe Funktion auf der Menge

$$\{D : D = D^T, K^{-1}DM^{-1} = M^{-1}DK^{-1}\}$$

(M, K symmetrisch und positiv definit) ist, so daß eine von ihm angegebene Menge von kritischen Punkten ein globales Minimum liefert. Damit berücksichtigt er lediglich eine Menge von Dämpfungsmatrizen, die mit M und K simultan diagonalisierbar sind.

Wir wollen in diesem Kapitel keine derartige Einschränkung an die Dämpfungsmatrizen vornehmen, sondern die Spur der eindeutig bestimmten Lösung X der Ljapunovgleichung über ihr maximales Definitionsgebiet, also über alle positiv semidefiniten Dämpfungsmatrizen, für die die Phasenraummatrix A aus (1.16) stabil ist, minimieren. Veselić und Delinić äußern in [29] die Vermutung, daß das Problem 1.3 mit $Z = I$ – bei dem die Spur der eindeutig bestimmten Lösung X der Ljapunovgleichung $A^T(C)X + XA(C) = -I$ über alle positiv semidefiniten Dämpfungsmatrizen C minimiert wird, für die die Phasenraummatrix $A(C) = A$ aus (1.16) stabil ist – für $C = 2\Omega$ ein lokales Minimum annimmt. (Anschaulich gesehen liefert gerade die Matrix $C = 2\Omega$ eine Matrix $A(C)$, die aus einer direkten Summe von n zweidimensionalen Blöcken besteht, wobei jeder dieser Blöcke ein eindimensionales kritisches Dämpfungssystem beschreibt (genauer siehe Bemerkung 2.4).) In Satz 2.1 beweisen wir diese Vermutung.

Oft betrachtet man in Anlehnung an (1.6) nur ein bestimmtes Frequenzspektrum und minimiert dann nur den entsprechenden Teil der Spur, d.h. es wird (bei geeigneter Sortierung der Frequenzen) das Minimum von $(\sum_{i=1}^s (X)_{ii} + \sum_{i=n+1}^{n+s} (X)_{ii})$ über alle Dämpfungsmatrizen C gebildet. Wir wollen also $\text{Tr}(XZ)$ mit einem entsprechend gewählten Projektor Z minimieren. Hiermit beschäftigen wir uns in Satz 2.6. Wir geben eine lineare Mannigfaltigkeit \mathcal{M} von positiv semidefiniten Dämpfungsmatrizen C an und beweisen, daß die Funktion $\text{Tr}(XZ)$ für diese lineare Mannigfaltigkeit ein striktes lokales Minimum annimmt. Mit anderen Worten, es existiert eine Umgebung \mathcal{O} von \mathcal{M} , so daß die Funktion $\text{Tr}(XZ)$ eingeschränkt auf \mathcal{O} , auf \mathcal{M} konstant ist und dort ihr (absolutes) Minimum annimmt.

Im zweiten Teil des Kapitels wollen wir anhand von breit angelegten Experimenten die Vermutung untermauern, daß $\text{Tr}(X)$ für die in Satz 2.1 vorgestellte Dämpfungsmatrix und $\text{Tr}(XZ)$ für die in Satz 2.6 vorgestellte lineare Mannigfaltigkeit nicht nur ihr lokales, sondern auch ihr globales Minimum annimmt. Andererseits zeigen die Beispiele, daß Konvexität im allgemeinen nicht vorliegt.

2.1 Bestimmung von lokalen Minima

Zuerst betrachten wir den Fall, daß die volle Spur der Lösung X der Ljapunovgleichung $A^T(C)X + XA(C) = -I$ mit $A(C)$ aus (1.16) über alle positiv semidefiniten Dämpfungsmatrizen minimiert wird.

Satz 2.1

Sei

$$A(C) = \begin{pmatrix} 0 & \Omega \\ -\Omega & -C \end{pmatrix} \quad (2.1)$$

eine reelle Matrix der Ordnung $2n$ mit $\Omega = \text{diag}(\omega_1, \dots, \omega_n)$, $\omega_i > 0$ beliebig aber fest für alle $i = 1, \dots, n$, und C eine beliebige positiv semidefinite Matrix der Ordnung n , so daß $A(C)$ stabil ist. Dann nimmt die Spur der eindeutigen Lösung $X(C)$ der Ljapunovgleichung

$$A^T(C)X + XA(C) = -I \quad (2.2)$$

für $C = 2\Omega$ ein lokales Minimum an.

Die bekannten Kriterien für das Vorliegen eines lokalen Minimums einer reellwertigen Funktion in einem endlichdimensionalen Vektorraum formulieren wir für unseren Beweis zunächst etwas um.

Lemma 2.2

Sei f eine stetig differenzierbare reellwertige Funktion um $x = x_0 \in \mathbb{R}^n$.

Dann ist $\text{grad } f(x_0) = 0$ genau dann, wenn

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

Beweis

Aus der Kettenregel folgt

$$\frac{d}{d\mu} f(x_0 + \mu v) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x_0 + \mu v) v_i,$$

insbesondere

$$\left. \frac{d}{d\mu} f(x_0 + \mu v) \right|_{\mu=0} = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x_0) v_i = \langle v, \text{grad } f(x_0) \rangle .$$

Q.E.D.

Lemma 2.3

Sei f eine zweimal stetig differenzierbare reellwertige Funktion um $x = x_0 \in \mathbb{R}^n$. Die Hessesche Matrix $\text{Hess}f(x_0)$ ist positiv definit genau dann, wenn

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

Beweis

$\text{Hess}f(x_0)$ ist symmetrisch, und nach der Kettenregel gilt

$$\begin{aligned} \left. \frac{d^2}{d\mu^2} f(x_0 + \mu v) \right|_{\mu=0} &= \left. \frac{d}{d\mu} \left(\sum_{i=1}^n \frac{\partial f}{\partial x_i}(x_0 + \mu v) v_i \right) \right|_{\mu=0} \\ &= \sum_{i,j} \left. \frac{\partial^2 f}{\partial x_i \partial x_j}(x_0 + \mu v) v_i v_j \right|_{\mu=0} \\ &= v^T \cdot \text{Hess}f(x_0) \cdot v. \end{aligned}$$

Insbesondere ist

$$\left. \frac{d^2}{d\mu^2} f(x_0 + \mu v) \right|_{\mu=0} = v^T \cdot \text{Hess}f(x_0) \cdot v.$$

Q.E.D.

Mit diesen Hilfsmitteln können wir nun Satz 2.1 beweisen.

Beweis von Satz 2.1

Da C eine positiv semidefinite Matrix ist, ist $C_1 = C - 2\Omega$ symmetrisch und kann dargestellt werden in der Form

$$C_1 = \mu V \quad \text{mit } V \text{ symmetrisch und } \mu \in \mathbb{R} .$$

Nach Lemma 2.2 und 2.3 reicht es zu zeigen, daß gilt

$$\begin{aligned} \left. \frac{d}{d\mu} \text{Tr } X(2\Omega + \mu V) \right|_{\mu=0} &= 0, \\ \left. \frac{d^2}{d\mu^2} \text{Tr } X(2\Omega + \mu V) \right|_{\mu=0} &> 0 \end{aligned}$$

für alle symmetrischen $V \neq 0$.

I. Schritt:

Für kleine μ ist $2\Omega + \mu V > 0$, so daß $X(2\Omega + \mu V)$ als Lösung der Ljapunovgleichung (2.2) eindeutig bestimmt ist (s. Lemma 1.4).

Zunächst zeigen wir $\frac{d}{d\mu} \text{Tr } X(2\Omega + \mu V)|_{\mu=0} = 0$ für beliebiges symmetrisches $V = (v_{ij})$ und folgen dabei einem Hinweis von Veselić.

Zuerst wird

$$\frac{d}{d\mu} \text{Tr } X(2\Omega + \mu V) = \text{Tr } \frac{d}{d\mu} X(2\Omega + \mu V)$$

bestimmt. Leitet man die Ljapunovgleichung mit Hilfe der Kettenregel ab, so erhält man (abkürzend schreiben wir A_μ, X_μ statt $A(2\Omega + \mu V), X(2\Omega + \mu V)$)

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

d.h. es ergibt sich folgende neue Ljapunovgleichung

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

Für

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

ist

$$\frac{d}{d\mu} A(2\Omega + \mu V) = \begin{pmatrix} 0 & 0 \\ 0 & -V \end{pmatrix} = \frac{d}{d\mu} A^T(2\Omega + \mu V) =: -\tilde{V}. \quad (2.3)$$

Damit hat die Ljapunovgleichung die Gestalt

$$A_\mu^T \frac{d}{d\mu} X_\mu + \left(\frac{d}{d\mu} X_\mu \right) A_\mu = \tilde{V} X_\mu + X_\mu \tilde{V}. \quad (2.4)$$

$\frac{d}{d\mu} \text{Tr } X_\mu = \text{Tr } \frac{d}{d\mu} X_\mu$ berechnet sich nun auf folgende Weise:

$$\begin{aligned} \text{Tr } \frac{d}{d\mu} X_\mu &= -\text{Tr} \left[\int_0^\infty e^{A_\mu^T t} (\tilde{V} X_\mu + X_\mu \tilde{V}) e^{A_\mu t} dt \right] \\ &= -\text{Tr} \left[\int_0^\infty e^{A_\mu^T t} \tilde{V} X_\mu e^{A_\mu t} dt \right] - \text{Tr} \left[\int_0^\infty e^{A_\mu^T t} X_\mu \tilde{V} e^{A_\mu t} dt \right] \\ &= -\text{Tr} \left[\int_0^\infty \tilde{V} X_\mu e^{A_\mu t} e^{A_\mu^T t} dt \right] - \text{Tr} \left[\int_0^\infty e^{A_\mu t} e^{A_\mu^T t} X_\mu \tilde{V} dt \right] \\ &= -\text{Tr} \left[\tilde{V} X_\mu \int_0^\infty e^{A_\mu t} e^{A_\mu^T t} dt \right] - \text{Tr} \left[\int_0^\infty e^{A_\mu t} e^{A_\mu^T t} dt X_\mu \tilde{V} \right] \\ &= -\text{Tr} \left[\tilde{V} X_\mu Y_\mu \right] - \text{Tr} \left[Y_\mu X_\mu \tilde{V} \right] = -2 \text{Tr} \left[\tilde{V} X_\mu Y_\mu \right] \end{aligned} \quad (2.5)$$

mit

$$Y_\mu = \int_0^\infty e^{A_\mu t} e^{A_\mu^T t} dt,$$

d.h. Y_μ ist die Lösung der Gleichung $A_\mu Y_\mu + Y_\mu A_\mu^T = -I$. Sei

$$J = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}.$$

Dann ist $A_\mu^T = J A_\mu J$ und die Ljapunovgleichung $A_\mu^T X_\mu + X_\mu A_\mu = -I$ geht über in

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

Also ist $Y_\mu = J X_\mu J$ und aus (2.5) ergibt sich

$$\operatorname{Tr} \frac{d}{d\mu} X_\mu = -2 \operatorname{Tr} [\tilde{V} X_\mu J X_\mu J]. \quad (2.6)$$

Für $\mu = 0$ ist die eindeutig bestimmte Lösung X_0 der Ljapunovgleichung

$$\begin{pmatrix} 0 & -\Omega \\ \Omega & -2\Omega \end{pmatrix} \begin{pmatrix} X_{11} & X_{12} \\ X_{12}^T & X_{22} \end{pmatrix} + \begin{pmatrix} X_{11} & X_{12} \\ X_{12}^T & X_{22} \end{pmatrix} \begin{pmatrix} 0 & \Omega \\ -\Omega & -2\Omega \end{pmatrix} = \begin{pmatrix} -I & 0 \\ 0 & -I \end{pmatrix}$$

gegeben durch

$$X_0 = \begin{pmatrix} \frac{3}{2}\Omega^{-1} & \frac{1}{2}\Omega^{-1} \\ \frac{1}{2}\Omega^{-1} & \frac{1}{2}\Omega^{-1} \end{pmatrix}.$$

Damit ist

$$\begin{aligned} X_0 J X_0 J &= \begin{pmatrix} \frac{3}{2}\Omega^{-1} & -\frac{1}{2}\Omega^{-1} \\ \frac{1}{2}\Omega^{-1} & -\frac{1}{2}\Omega^{-1} \end{pmatrix} \begin{pmatrix} \frac{3}{2}\Omega^{-1} & -\frac{1}{2}\Omega^{-1} \\ \frac{1}{2}\Omega^{-1} & -\frac{1}{2}\Omega^{-1} \end{pmatrix} \\ &= \begin{pmatrix} 2\Omega^{-2} & -\frac{1}{2}\Omega^{-2} \\ \frac{1}{2}\Omega^{-2} & 0 \end{pmatrix}, \end{aligned}$$

und wir erhalten

$$\begin{aligned} \operatorname{Tr} \frac{d}{d\mu} X(2\Omega + \mu V) \Big|_{\mu=0} &= -2 \operatorname{Tr} \left[\begin{pmatrix} 0 & 0 \\ 0 & V \end{pmatrix} \begin{pmatrix} 2\Omega^{-2} & -\frac{1}{2}\Omega^{-2} \\ \frac{1}{2}\Omega^{-2} & 0 \end{pmatrix} \right] \\ &= -2 \operatorname{Tr} \begin{pmatrix} 0 & 0 \\ \frac{1}{2}V\Omega^{-2} & 0 \end{pmatrix} = 0. \end{aligned}$$

II. Schritt:

Nun zeigen wir $\frac{d^2}{d\mu^2} \operatorname{Tr} X(2\Omega + \mu V) \Big|_{\mu=0} > 0$ mit beliebigem symmetrischen $V = (v_{ij})$.

Zuerst wird

$$\frac{d^2}{d\mu^2} \text{Tr} X(2\Omega + \mu V) = \text{Tr} \frac{d^2}{d\mu^2} X(2\Omega + \mu V)$$

bestimmt. Leitet man (2.5) bzw. (2.6) mit Hilfe der Kettenregel ab, so ergibt sich

$$\begin{aligned} \frac{d^2}{d\mu^2} \text{Tr} X_\mu &= \text{Tr} \frac{d^2}{d\mu^2} X_\mu = -2 \text{Tr} \left[\tilde{V} \left(\frac{d}{d\mu} X_\mu \right) J X_\mu J \right] - 2 \text{Tr} \left[\tilde{V} X_\mu J \left(\frac{d}{d\mu} X_\mu \right) J \right] \\ &= -2 \text{Tr} \left[\tilde{V} \left(\frac{d}{d\mu} X_\mu \right) J X_\mu J \right] - 2 \text{Tr} \left[\left(\frac{d}{d\mu} X_\mu \right) J \tilde{V} X_\mu J \right] \\ &= -2 \text{Tr} \left[\tilde{V} \left(\frac{d}{d\mu} X_\mu \right) J X_\mu J \right] - 2 \text{Tr} \left[\left(\frac{d}{d\mu} X_\mu \right) \tilde{V} J X_\mu J \right], \end{aligned}$$

da \tilde{V} blockdiagonal ist.

Wir bestimmen $\tilde{V} \left(\frac{d}{d\mu} X_\mu \right) J X_\mu J \Big|_{\mu=0}$. Sei

$$\frac{d}{d\mu} X_\mu \Big|_{\mu=0} = \begin{pmatrix} \tilde{X}_{11} & \tilde{X}_{12} \\ \tilde{X}_{12}^T & \tilde{X}_{22} \end{pmatrix}.$$

$$\begin{aligned} &\tilde{V} \left(\frac{d}{d\mu} X_\mu \right) J X_\mu J \Big|_{\mu=0} \\ &= \begin{pmatrix} 0 & 0 \\ 0 & V \end{pmatrix} \begin{pmatrix} \tilde{X}_{11} & \tilde{X}_{12} \\ \tilde{X}_{12}^T & \tilde{X}_{22} \end{pmatrix} \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \begin{pmatrix} \frac{3}{2}\Omega^{-1} & \frac{1}{2}\Omega^{-1} \\ \frac{1}{2}\Omega^{-1} & \frac{1}{2}\Omega^{-1} \end{pmatrix} \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 \\ V\tilde{X}_{12}^T & V\tilde{X}_{22} \end{pmatrix} \begin{pmatrix} \frac{3}{2}\Omega^{-1} & -\frac{1}{2}\Omega^{-1} \\ -\frac{1}{2}\Omega^{-1} & \frac{1}{2}\Omega^{-1} \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 \\ \frac{3}{2}V\tilde{X}_{12}^T\Omega^{-1} - \frac{1}{2}V\tilde{X}_{22}\Omega^{-1} & \frac{1}{2}V\tilde{X}_{22}\Omega^{-1} - \frac{1}{2}V\tilde{X}_{12}^T\Omega^{-1} \end{pmatrix}. \end{aligned}$$

Entsprechend ist

$$\begin{aligned} \left(\frac{d}{d\mu} X_\mu \right) \tilde{V} J X_\mu J \Big|_{\mu=0} &= \begin{pmatrix} \tilde{X}_{11} & \tilde{X}_{12} \\ \tilde{X}_{12}^T & \tilde{X}_{22} \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & V \end{pmatrix} \begin{pmatrix} \frac{3}{2}\Omega^{-1} & -\frac{1}{2}\Omega^{-1} \\ -\frac{1}{2}\Omega^{-1} & \frac{1}{2}\Omega^{-1} \end{pmatrix} \\ &= \begin{pmatrix} 0 & \tilde{X}_{12}V \\ 0 & \tilde{X}_{22}V \end{pmatrix} \begin{pmatrix} \frac{3}{2}\Omega^{-1} & -\frac{1}{2}\Omega^{-1} \\ -\frac{1}{2}\Omega^{-1} & \frac{1}{2}\Omega^{-1} \end{pmatrix} \\ &= \begin{pmatrix} -\frac{1}{2}\tilde{X}_{12}V\Omega^{-1} & \frac{1}{2}\tilde{X}_{12}V\Omega^{-1} \\ -\frac{1}{2}\tilde{X}_{22}V\Omega^{-1} & \frac{1}{2}\tilde{X}_{22}V\Omega^{-1} \end{pmatrix}. \end{aligned}$$

Also ist

$$\begin{aligned}
& \frac{d^2}{d\mu^2} \text{Tr} X(2\Omega + \mu V) \Big|_{\mu=0} \\
&= -2 \text{Tr} \left[\frac{1}{2} V \tilde{X}_{22} \Omega^{-1} - \frac{1}{2} V \tilde{X}_{12}^T \Omega^{-1} \right] - 2 \text{Tr} \left[-\frac{1}{2} \tilde{X}_{12} V \Omega^{-1} + \frac{1}{2} \tilde{X}_{22} V \Omega^{-1} \right] \\
&= \text{Tr} \left[V \tilde{X}_{12}^T \Omega^{-1} \right] + \text{Tr} \left[\tilde{X}_{12} V \Omega^{-1} \right] - \text{Tr} \left[V \tilde{X}_{22} \Omega^{-1} \right] - \text{Tr} \left[\tilde{X}_{22} V \Omega^{-1} \right] \\
&= 2 \text{Tr} \left[\tilde{X}_{12} V \Omega^{-1} \right] - 2 \text{Tr} \left[V \tilde{X}_{22} \Omega^{-1} \right]. \tag{2.7}
\end{aligned}$$

Nun müssen wir die Untermatrizen \tilde{X}_{12} und \tilde{X}_{22} der Matrix $\frac{d}{d\mu} X(2\Omega + \mu V) \Big|_{\mu=0}$ bestimmen:

$X'_0 := \left(\frac{d}{d\mu} X_\mu \right) \Big|_{\mu=0}$ ist Lösung der Ljapunovgleichung

$$A_0^T X'_0 + X'_0 A_0 = - \left[\left(\frac{d}{d\mu} A_\mu^T \right) X_\mu + X_\mu \frac{d}{d\mu} A_\mu \right] \Big|_{\mu=0}.$$

Dies bedeutet

$$\begin{aligned}
& \begin{pmatrix} 0 & -\Omega \\ \Omega & -2\Omega \end{pmatrix} \begin{pmatrix} \tilde{X}_{11} & \tilde{X}_{12} \\ \tilde{X}_{12}^T & \tilde{X}_{22} \end{pmatrix} + \begin{pmatrix} \tilde{X}_{11} & \tilde{X}_{12} \\ \tilde{X}_{12}^T & \tilde{X}_{22} \end{pmatrix} \begin{pmatrix} 0 & \Omega \\ -\Omega & -2\Omega \end{pmatrix} \\
&= \begin{pmatrix} 0 & 0 \\ 0 & V \end{pmatrix} \begin{pmatrix} \frac{3}{2}\Omega^{-1} & \frac{1}{2}\Omega^{-1} \\ \frac{1}{2}\Omega^{-1} & \frac{1}{2}\Omega^{-1} \end{pmatrix} + \begin{pmatrix} \frac{3}{2}\Omega^{-1} & \frac{1}{2}\Omega^{-1} \\ \frac{1}{2}\Omega^{-1} & \frac{1}{2}\Omega^{-1} \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & V \end{pmatrix},
\end{aligned}$$

d.h.

$$\begin{aligned}
& \begin{pmatrix} 0 & -\Omega \\ \Omega & -2\Omega \end{pmatrix} \begin{pmatrix} \tilde{X}_{11} & \tilde{X}_{12} \\ \tilde{X}_{12}^T & \tilde{X}_{22} \end{pmatrix} + \begin{pmatrix} \tilde{X}_{11} & \tilde{X}_{12} \\ \tilde{X}_{12}^T & \tilde{X}_{22} \end{pmatrix} \begin{pmatrix} 0 & \Omega \\ -\Omega & -2\Omega \end{pmatrix} \\
&= \begin{pmatrix} 0 & \frac{1}{2}\Omega^{-1}V \\ \frac{1}{2}V\Omega^{-1} & \frac{1}{2}\Omega^{-1}V + \frac{1}{2}V\Omega^{-1} \end{pmatrix}.
\end{aligned}$$

Daraus ergeben sich folgende vier Matrixgleichungen (wobei (ii) die adjungierte Gleichung zu (iii) ist) :

$$-\Omega \tilde{X}_{12}^T - \tilde{X}_{12} \Omega = 0 \tag{i}$$

$$-\Omega \tilde{X}_{22} + \tilde{X}_{11} \Omega - 2\tilde{X}_{12} \Omega = \frac{1}{2}\Omega^{-1}V \tag{ii}$$

$$\Omega \tilde{X}_{11} - 2\Omega \tilde{X}_{12}^T - \tilde{X}_{22} \Omega = \frac{1}{2}V\Omega^{-1} \tag{iii}$$

$$\Omega \tilde{X}_{12} - 2\Omega \tilde{X}_{22} + \tilde{X}_{12}^T \Omega - 2\tilde{X}_{22} \Omega = \frac{1}{2}V\Omega^{-1} + \frac{1}{2}\Omega^{-1}V \tag{iv}.$$

Aus der ersten Gleichung ergibt sich

$$\begin{aligned}\tilde{X}_{12} &= \frac{1}{2}S\Omega^{-1}, \\ \tilde{X}_{12}^T &= -\frac{1}{2}\Omega^{-1}S,\end{aligned}$$

wobei S eine noch zu bestimmende schiefsymmetrische Matrix ist, d.h. $S^T = -S$. Dies zusammen mit der Symmetrie von $\Omega^{-1}V\Omega^{-1}$ ergibt $\text{Tr} [S\Omega^{-1}V\Omega^{-1}] = 0$; somit folgt aus (2.7)

$$\begin{aligned}\frac{d^2}{d\mu^2}\text{Tr} X_\mu \Big|_{\mu=0} &= 2 \text{Tr} \left[\frac{1}{2}S\Omega^{-1}V\Omega^{-1} \right] - 2 \text{Tr} [V\tilde{X}_{22}\Omega^{-1}] \\ &= -2 \text{Tr} [V\tilde{X}_{22}\Omega^{-1}].\end{aligned}\tag{2.8}$$

Damit müssen nur noch die Untermatrix \tilde{X}_{22} sowie die Matrix S bestimmt werden. Setzt man $\tilde{X}_{12} = \frac{1}{2}S\Omega^{-1}$, $\tilde{X}_{12}^T = -\frac{1}{2}\Omega^{-1}S$ in Gleichung (iv) ein, so erhält man

$$\frac{1}{2}\Omega S\Omega^{-1} - 2\Omega\tilde{X}_{22} - \frac{1}{2}\Omega^{-1}S\Omega - 2\tilde{X}_{22}\Omega = \frac{1}{2}V\Omega^{-1} + \frac{1}{2}\Omega^{-1}V.\tag{2.9}$$

Zuerst können die Diagonalelemente der Matrix \tilde{X}_{22} bestimmt werden. Da

$$\frac{1}{2}(\Omega S\Omega^{-1})_{ii} = \frac{1}{2}(\Omega^{-1}S\Omega)_{ii} = 0, \quad (\Omega\tilde{X}_{22})_{ii} = (\tilde{X}_{22}\Omega)_{ii}$$

für $i = 1, \dots, n$, erhält man aus (2.9)

$$\begin{aligned}-4(\Omega\tilde{X}_{22})_{ii} &= \frac{1}{2}(V\Omega^{-1})_{ii} + \frac{1}{2}(\Omega^{-1}V)_{ii} \\ &= \frac{v_{ii}}{\omega_i}.\end{aligned}$$

Also ist

$$\begin{aligned}(\Omega\tilde{X}_{22})_{ii} &= -\frac{v_{ii}}{4\omega_i}, \\ (\tilde{X}_{22})_{ii} &= -\frac{v_{ii}}{4\omega_i^2}.\end{aligned}$$

Als nächstes werden die Außendiagonalelemente bestimmt. (2.9) liefert

$$\begin{aligned}&\frac{1}{2}(\Omega S\Omega^{-1})_{ij} - 2(\Omega\tilde{X}_{22})_{ij} - \frac{1}{2}(\Omega^{-1}S\Omega)_{ij} - 2(\tilde{X}_{22}\Omega)_{ij} \\ &= \frac{1}{2}(V\Omega^{-1})_{ij} + \frac{1}{2}(\Omega^{-1}V)_{ij}.\end{aligned}$$

Mit $S = (s_{ij})$, $s_{ii} = 0$, $\tilde{X}_{22} = (y_{ij})$ folgt

$$\begin{aligned}\frac{\omega_i s_{ij}}{2\omega_j} - \frac{s_{ij}\omega_j}{2\omega_i} - 2\omega_i y_{ij} - 2y_{ij}\omega_j &= \frac{v_{ij}}{2\omega_i} + \frac{v_{ij}}{2\omega_j}, \\ \omega_i^2 s_{ij} - \omega_j^2 s_{ij} - 4\omega_i^2 \omega_j y_{ij} - 4\omega_i \omega_j^2 y_{ij} &= \omega_j v_{ij} + \omega_i v_{ij}.\end{aligned}$$

Damit ist

$$s_{ij}(\omega_i - \omega_j) = v_{ij} + 4\omega_i \omega_j y_{ij}. \quad (2.10)$$

Um eine zweite Abhängigkeit zwischen s_{ij} und y_{ij} zu erhalten, ersetzt man in Gleichung (ii) \tilde{X}_{12} durch $\frac{1}{2}S\Omega^{-1}$ und löst diese komponentenweise:

$$\begin{aligned}-\Omega\tilde{X}_{22} + \tilde{X}_{11}\Omega - 2\left(\frac{1}{2}S\Omega^{-1}\right)\Omega &= \frac{1}{2}\Omega^{-1}V, \\ -\Omega\tilde{X}_{22} + \tilde{X}_{11}\Omega - S &= \frac{1}{2}\Omega^{-1}V, \\ -(\Omega\tilde{X}_{22})_{ij} + (\tilde{X}_{11}\Omega)_{ij} - s_{ij} &= \frac{1}{2}(\Omega^{-1}V)_{ij}.\end{aligned}$$

Mit $\tilde{X}_{11} = (z_{ij})$ ergibt sich

$$-\omega_i y_{ij} + z_{ij}\omega_j - s_{ij} = \frac{v_{ij}}{2\omega_i}.$$

Da die Matrizen \tilde{X}_{22} und \tilde{X}_{11} symmetrisch sind und die Matrix S schiefssymmetrisch ist, d.h. $y_{ij} = y_{ji}$, $z_{ij} = z_{ji}$, $s_{ij} = -s_{ji}$, erhält man daraus die beiden Gleichungen

$$\begin{aligned}-\omega_i y_{ij} + z_{ij}\omega_j - s_{ij} &= \frac{v_{ij}}{2\omega_i}, \\ -\omega_j y_{ij} + z_{ij}\omega_i + s_{ij} &= \frac{v_{ij}}{2\omega_j}.\end{aligned}$$

Multipliziert man die erste Gleichung mit ω_i und die zweite mit ω_j und subtrahiert dann beide Gleichungen, so erhält man

$$-\omega_i^2 y_{ij} + z_{ij}\omega_i \omega_j - s_{ij}\omega_i + \omega_j^2 y_{ij} - z_{ij}\omega_i \omega_j - s_{ij}\omega_j = 0,$$

d.h.

$$\begin{aligned}s_{ij}(\omega_i + \omega_j) &= y_{ij}(\omega_j^2 - \omega_i^2), \\ s_{ij} &= y_{ij}(\omega_j - \omega_i).\end{aligned} \quad (2.11)$$

Setzt man nun (2.11) in (2.10) ein, so ergibt sich

$$\begin{aligned}v_{ij} + 4\omega_i \omega_j y_{ij} &= y_{ij}(\omega_j - \omega_i)(\omega_i - \omega_j) \\ v_{ij} + 4\omega_i \omega_j y_{ij} &= -y_{ij}(\omega_j - \omega_i)^2 \\ y_{ij}(\omega_i + \omega_j)^2 &= -v_{ij}.\end{aligned}$$

Also ist

$$y_{ij} = -\frac{v_{ij}}{(\omega_i + \omega_j)^2}. \quad (2.12)$$

Nun kann $\frac{d^2}{d\mu^2} \text{Tr } X_\mu|_{\mu=0}$ bestimmt werden:

Mit (2.8) ist

$$\frac{d^2}{d\mu^2} \text{Tr } X_\mu \Big|_{\mu=0} = -2 \text{Tr} [V \tilde{X}_{22} \Omega^{-1}] = -2 \sum_{i=1}^n \sum_{j=1}^n \frac{v_{ij} y_{ij}}{\omega_i},$$

und mit (2.12) erhalten wir

$$\begin{aligned} \frac{d^2}{d\mu^2} \text{Tr } X_\mu \Big|_{\mu=0} &= 2 \sum_{i=1}^n \sum_{j=1}^n \frac{v_{ij}^2}{\omega_i (\omega_i + \omega_j)^2} \\ &> 0, \end{aligned}$$

da $V = (v_{ij}) \neq 0$.

Q.E.D.

Bemerkung 2.4

Die Eigenwerte von $A(C)$ aus (2.1) mit $C = 2\Omega$ sind durch $\lambda_i = -\omega_i$ für alle i mit $1 \leq i \leq n$ gegeben, wobei jeder Vielfachheit 2 hat. Der Defekt jedes Eigenwertes ist 1, d.h. es existiert nur ein Eigenvektor. Dies bezeichnet man als *kritische Dämpfung* (s. [24]); wir nennen $C_{\min} = C_{\text{krit}} = 2\Omega$ die zu $A(C)$ *kritische Dämpfungsmatrix*.

Im folgenden betrachten wir den realistischeren und komplizierteren Fall, daß wir nicht die volle Spur von X aus (2.2), sondern nur einen Teil der Spur, d.h. $(\sum_{i=1}^s X_{ii} + \sum_{i=n+1}^{s+n} X_{ii})$, über alle positiv semidefiniten Dämpfungsmatrizen C minimieren. Die Matrix Z aus (1.17) ist dann nicht wie in Satz 2.1 die Einheitsmatrix, sondern hat die Gestalt

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

wobei I_s die Einheitsmatrix der Ordnung s mit $s \leq n$ ist (für $s = n$ ist $Z = I$). Wir bezeichnen diese Matrix im folgenden immer mit Z_s . Minimiert werden soll also $\text{Tr}(XZ_s)$. Dies hat z.B. dann einen Sinn, wenn man die Frequenzen ω_i in der Matrix

$A(C)$ als aufsteigend sortiert betrachtet, was (s. Kapitel 1) o.B.d.A. stets angenommen werden kann. $\text{Tr}(XZ_s)$ zu minimieren bedeutet dann, daß nur ein bestimmter Frequenzunterraum betrachtet wird, und zwar der Unterraum, der durch eine Frequenz ω_{max} begrenzt wird. s ist somit gegeben durch $s = \max\{i : \omega_i \leq \omega_{max}\}$. Es wird sich zeigen, daß wir im Falle der Minimierung von $\text{Tr}(XZ_s)$ eine zusammenhängende Menge von lokalen Minima erhalten, die alle den selben Funktionswert liefern. Bevor wir aber zur Formulierung des Satzes kommen, wollen wir noch ein Lemma beweisen (siehe [29]).

Lemma 2.5

Für alle stabilen Matrizen A aus (2.1) der Ordnung $2n$ und alle diagonalen Matrizen Z der Ordnung $2n$ mit $A^T X + X A = -I$ und $A^T Y + Y A = -Z$ gilt

$$\text{Tr}(XZ) = \text{Tr}(Y).$$

Beweis

Für die eindeutig bestimmte Lösung X der oben angegebenen Ljapunovgleichung gilt $X = \int_0^\infty e^{A^T t} e^{A t} dt$, also

$$\text{Tr}(XZ) = \text{Tr}\left(\int_0^\infty e^{A^T t} e^{A t} dt Z\right) = \text{Tr}\left(\int_0^\infty e^{A t} Z e^{A^T t} dt\right) = \text{Tr}(\tilde{Y}),$$

wobei \tilde{Y} die eindeutig bestimmte Lösung von $A\tilde{Y} + \tilde{Y}A^T = -Z$ ist.

Es gilt $\text{Tr}(\tilde{Y}) = \text{Tr}(Y)$, denn mit $J := \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}$ ist $A^T = JAJ$, $A = JA^T J$, also $JA^T J\tilde{Y} + \tilde{Y}JAJ = -Z$, d.h. $A^T J\tilde{Y}J + J\tilde{Y}JA = -JZJ$. Für diagonales Z ist $JZJ = Z$, also $A^T J\tilde{Y}J + J\tilde{Y}JA = -Z$.

Q.E.D.

Nun können wir Satz 2.6 formulieren. Da eine Sortierung der Frequenzen (s.o.) für die Aussage und den Beweis des Satzes keine Rolle spielt, wird sie in der Voraussetzung des Satzes nicht erwähnt.

Satz 2.6

Sei

$$A(C) = \begin{pmatrix} 0 & \Omega \\ -\Omega & -C \end{pmatrix}$$

eine reelle Matrix der Ordnung $2n$ mit $\Omega = \text{diag}(\omega_1, \dots, \omega_n)$, $\omega_i > 0$ beliebig aber fest für alle $i = 1, \dots, n$, und C eine beliebige positiv semidefinite Matrix der Ordnung

n , so daß $A(C)$ stabil ist. Sei weiterhin die $2n$ -dimensionale Matrix

$$Z_s = \begin{pmatrix} I_s & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & I_s & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad (2.13)$$

gegeben, wobei I_s die Einheitsmatrix der Dimension s ist.

Dann nimmt die Spur der eindeutigen Lösung $Y(C)$ der Ljapunovgleichung

$$A^T(C)Y + YA(C) = -Z_s$$

auf der Menge

$$\mathcal{M} = \left\{ C = 2\tilde{\Omega} = 2 \begin{pmatrix} \hat{\Omega}_s & 0 \\ 0 & H \end{pmatrix} \text{ mit } \hat{\Omega}_s = \text{diag}(\omega_1, \dots, \omega_s), H \in \mathcal{S} \right\}$$

ein strenges lokales Minimum an, wobei \mathcal{S} gegeben ist durch

$$\mathcal{S} = \left\{ H \in M_{n-s}, H \text{ positiv semidefinit: } A \text{ stabil mit } C = 2 \begin{pmatrix} \hat{\Omega}_s & 0 \\ 0 & H \end{pmatrix} \right\};$$

$\text{Tr} Y(C)$ ist dabei konstant auf \mathcal{M} .

Bemerkung 2.7

Der Begriff *strenges lokales Minimum* wird hier sinngemäß auf eine Menge \mathcal{M} erweitert. Dies bedeutet, daß es eine Umgebung \mathcal{O} von \mathcal{M} gibt, so daß die Funktion $\text{Tr} Y(C)|_{\mathcal{O}}$ auf \mathcal{M} und nirgendwo sonst ihr Minimum annimmt.

Beweis von Satz 2.6

Wir beweisen die Aussage des Satzes für die Menge

$$\mathcal{M}_0 = \left\{ C = 2\tilde{\Omega} = 2 \begin{pmatrix} \hat{\Omega}_s & 0 \\ 0 & H \end{pmatrix} \text{ mit } \hat{\Omega}_s = \text{diag}(\omega_1, \dots, \omega_s), H \text{ positiv definit} \right\};$$

\mathcal{M}_0 liegt dicht in \mathcal{M} , so daß die Behauptung des Satzes anschließend aus der Stetigkeit der Lösung der Ljapunovgleichung folgt.

Sei also $2\tilde{\Omega} \in \mathcal{M}_0$. Da C eine positiv semidefinite Matrix ist, kann C wie im Beweis von Satz 2.1 dargestellt werden als $C = 2\tilde{\Omega} + \mu V$, wobei V eine beliebige symmetrische Matrix der Ordnung n ist. Für kleine μ ist $2\tilde{\Omega} + \mu V$ positiv definit, so daß $X(2\tilde{\Omega} + \mu V)$ als Lösung der Ljapunovgleichung (2.2) eindeutig bestimmt ist (s. Lemma 1.4). Nach Lemma 2.2 und 2.3 müssen wir zeigen, daß für beliebiges $2\tilde{\Omega} \in \mathcal{M}_0$

$$\left. \frac{d}{d\mu} \text{Tr} Y(2\tilde{\Omega} + \mu V) \right|_{\mu=0} = 0 \quad (2.14)$$

für alle symmetrischen Matrizen V der Ordnung n gilt und

$$\frac{d^2}{d\mu^2} \text{Tr} Y(2\tilde{\Omega} + \mu V) \Big|_{\mu=0} > 0 \quad (2.15)$$

für alle symmetrischen Matrizen V der Ordnung n , wobei bei (2.15) in der Zerlegung

$$V = \begin{pmatrix} V_{11} & V_{12} \\ V_{12}^T & V_{22} \end{pmatrix} \quad \text{mit } V_{22} \text{ } (n-s) \text{-dimensional} \quad (2.16)$$

V_{11} und V_{12} nicht gleichzeitig verschwinden dürfen (da sonst $2\tilde{\Omega} + \mu V$ für kleines μ in \mathcal{M}_0 liegt).

Bei Zerlegungen von n -dimensionalen Matrizen ist im folgenden immer das Schema aus (2.16) gemeint, ohne daß zusätzlich darauf hingewiesen wird.

I. Schritt:

Zunächst wird gezeigt, daß $\frac{d}{d\mu} \text{Tr} Y(2\tilde{\Omega} + \mu V) \Big|_{\mu=0} = 0$ für alle symmetrischen Matrizen V gilt.

Zuerst wird

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

bestimmt. Leitet man die Ljapunovgleichung mit Hilfe der Kettenregel ab, so erhält man (abkürzend schreiben wir A_μ, Y_μ statt $A(2\tilde{\Omega} + \mu V), Y(2\tilde{\Omega} + \mu V)$)

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

d.h. es ergibt sich folgende neue Ljapunovgleichung

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

Mit $A_\mu = \begin{pmatrix} 0 & \Omega \\ -\Omega & -(2\tilde{\Omega} + \mu V) \end{pmatrix}$ ergibt sich

$$\frac{d}{d\mu} A_\mu = \begin{pmatrix} 0 & 0 \\ 0 & -V \end{pmatrix} =: -\tilde{V} = \frac{d}{d\mu} A_\mu^T, \quad (2.17)$$

d.h. die Ljapunovgleichung geht über in

$$A_\mu^T \frac{d}{d\mu} Y_\mu + \left(\frac{d}{d\mu} Y_\mu \right) A_\mu = \tilde{V} Y_\mu + Y_\mu \tilde{V}. \quad (2.18)$$

Nun kann $\frac{d}{d\mu} \text{Tr} Y_\mu = \text{Tr} \frac{d}{d\mu} Y_\mu$ bestimmt werden:

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

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

wobei $\tilde{X}_\mu = \int_0^\infty e^{A_\mu t} e^{A_\mu^T t} dt$, d.h Lösung der Ljapunovgleichung $A_\mu \tilde{X}_\mu + \tilde{X}_\mu A_\mu^T = -I$ ist. Sei X_μ Lösung der Ljapunovgleichung $A_\mu^T X_\mu + X_\mu A_\mu = -I$.

Mit $J = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}$ ist $A_\mu^T = J A_\mu J$ und $A_\mu^T X_\mu + X_\mu A_\mu = -I$ geht über in

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

Also ist $\tilde{X}_\mu = J X_\mu J$ und

$$\text{Tr} \frac{d}{d\mu} Y_\mu = -2 \text{Tr} (\tilde{V} Y_\mu J X_\mu J). \quad (2.19)$$

Sei

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

Für $\mu = 0$ ist, wie man leicht nachprüft, die eindeutig bestimmte Lösung Y_0 der Ljapunovgleichung $A_0^T Y_0 + Y_0 A_0 = -Z_s$ gegeben durch

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

mit

$$\Psi_{11} = \begin{pmatrix} \frac{3}{2} \hat{\Omega}_s^{-1} & 0 \\ 0 & 0 \end{pmatrix}, \quad \Psi_{12} = \Psi_{22} = \begin{pmatrix} \frac{1}{2} \hat{\Omega}_s^{-1} & 0 \\ 0 & 0 \end{pmatrix}. \quad (2.20)$$

Ebenso verifiziert man, daß die Lösung X_0 der Ljapunovgleichung $A_0^T X_0 + X_0 A_0 = -I$ von der Form

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

ist mit

$$\Phi_{11} = \begin{pmatrix} \frac{3}{2} \hat{\Omega}_s^{-1} & 0 \\ 0 & W_{11} \end{pmatrix}, \quad \Phi_{12} = \begin{pmatrix} \frac{1}{2} \hat{\Omega}_s^{-1} & 0 \\ 0 & W_{12} \end{pmatrix}, \quad \Phi_{22} = \begin{pmatrix} \frac{1}{2} \hat{\Omega}_s^{-1} & 0 \\ 0 & W_{22} \end{pmatrix}, \quad (2.21)$$

wobei

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

die eindeutige Lösung der Ljapunovgleichung

$$\begin{aligned} & \begin{pmatrix} 0 & -\check{\Omega}_{n-s} \\ \check{\Omega}_{n-s} & -2H \end{pmatrix} \begin{pmatrix} W_{11} & W_{12} \\ W_{12}^T & W_{22} \end{pmatrix} + \begin{pmatrix} W_{11} & W_{12} \\ W_{12}^T & W_{22} \end{pmatrix} \begin{pmatrix} 0 & \check{\Omega}_{n-s} \\ -\check{\Omega}_{n-s} & -2H \end{pmatrix} \\ &= \begin{pmatrix} -I_{n-s} & 0 \\ 0 & -I_{n-s} \end{pmatrix} \end{aligned} \quad (2.22)$$

ist. Damit ergibt sich (siehe (2.19))

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

Somit ist (2.14) gezeigt, und da die Menge \mathcal{M} zusammenhängend und \mathcal{M}_0 dicht in \mathcal{M} ist, folgt $\text{Tr} Y(C) = \text{const}$ auf \mathcal{M} .

II. Schritt:

Als zweites wird $\frac{d^2}{d\mu^2} \text{Tr} (2\tilde{\Omega} + \mu V) \Big|_{\mu=0} > 0$ gezeigt, wobei $V = \begin{pmatrix} V_{11} & V_{12} \\ V_{12}^T & V_{22} \end{pmatrix}$ eine beliebige symmetrische Matrix ist und V_{11}, V_{12} nicht gleichzeitig verschwinden.

Zuerst wird

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

bestimmt. Leitet man die Ljapunovgleichung (2.18) erneut mit Hilfe der Kettenregel ab, so ergibt sich

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

d.h.

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

also (siehe (2.17))

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

Damit ist

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

und

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

mit $A_\mu^T X_\mu + X_\mu A_\mu = -I$ und $A_\mu \tilde{X}_\mu + \tilde{X}_\mu A_\mu^T = -I$ wie oben.

Mit $X_0 = \begin{pmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{12}^T & \Phi_{22} \end{pmatrix}$ und $\frac{d}{d\mu} Y_\mu|_{\mu=0} =: \begin{pmatrix} \tilde{Y}_{11} & \tilde{Y}_{12} \\ \tilde{Y}_{12}^T & \tilde{Y}_{22} \end{pmatrix} =: Y'_0$ folgt

$$\begin{aligned}
& \frac{d^2}{d\mu^2} \text{Tr} Y_\mu \Big|_{\mu=0} \\
&= -2\text{Tr} \left[\begin{pmatrix} 0 & 0 \\ 0 & V \end{pmatrix} \begin{pmatrix} \tilde{Y}_{11} & \tilde{Y}_{12} \\ \tilde{Y}_{12}^T & \tilde{Y}_{22} \end{pmatrix} \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \begin{pmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{12}^T & \Phi_{22} \end{pmatrix} \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \right] \\
& \quad -2\text{Tr} \left[\begin{pmatrix} \tilde{Y}_{11} & \tilde{Y}_{12} \\ \tilde{Y}_{12}^T & \tilde{Y}_{22} \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & V \end{pmatrix} \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \begin{pmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{12}^T & \Phi_{22} \end{pmatrix} \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \right] \\
&= -2\text{Tr} \left[\begin{pmatrix} 0 & 0 \\ V\tilde{Y}_{12}^T & V\tilde{Y}_{22} \end{pmatrix} \begin{pmatrix} \Phi_{11} & -\Phi_{12} \\ -\Phi_{12}^T & \Phi_{22} \end{pmatrix} \right] \\
& \quad -2\text{Tr} \left[\begin{pmatrix} 0 & \tilde{Y}_{12}V \\ 0 & \tilde{Y}_{22}V \end{pmatrix} \begin{pmatrix} \Phi_{11} & -\Phi_{12} \\ -\Phi_{12}^T & \Phi_{22} \end{pmatrix} \right] \\
&= -2\text{Tr} [-V\tilde{Y}_{12}^T\Phi_{12} + V\tilde{Y}_{22}\Phi_{22}] - 2\text{Tr} [-\tilde{Y}_{12}V\Phi_{12}^T + \tilde{Y}_{22}V\Phi_{22}] \\
&= 2\text{Tr} [V\tilde{Y}_{12}^T\Phi_{12}] - 2\text{Tr} [V\tilde{Y}_{22}\Phi_{22}] + 2\text{Tr} [\tilde{Y}_{12}V\Phi_{12}^T] - 2\text{Tr} [\tilde{Y}_{22}V\Phi_{22}] \\
&= 4 \text{Tr} [\tilde{Y}_{12}V\Phi_{12}^T] - 4 \text{Tr} [V\tilde{Y}_{22}\Phi_{22}]. \tag{2.23}
\end{aligned}$$

Als nächstes werden die Untermatrizen \tilde{Y}_{12} und \tilde{Y}_{22} der Matrix $\frac{d}{d\mu} Y(2\tilde{\Omega} + \mu V)|_{\mu=0}$ bestimmt:

$Y'_0 := \left(\frac{d}{d\mu} Y_\mu\right)|_{\mu=0}$ ist (siehe (2.18)) Lösung der Ljapunovgleichung

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

wobei

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

die durch (2.20) gegebene Lösung der Ljapunovgleichung $A_0^T Y_0 + Y_0 A_0 = -Z_s$ ist. Y'_0 muß nun explizit berechnet werden. Dies bedeutet,

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

zu lösen. Daraus ergeben sich folgende vier Matrixgleichungen

$$-\Omega\tilde{Y}_{12}^T - \tilde{Y}_{12}\Omega = 0 \quad (2.24)$$

$$-\Omega\tilde{Y}_{22} + \tilde{Y}_{11}\Omega - 2\tilde{Y}_{12}\tilde{\Omega} = \Psi_{12}V \quad (2.25)$$

$$\Omega\tilde{Y}_{11} - 2\tilde{\Omega}\tilde{Y}_{12}^T - \tilde{Y}_{22}\Omega = V\Psi_{12}^T \quad (2.26)$$

$$\Omega\tilde{Y}_{12} - 2\tilde{\Omega}\tilde{Y}_{22} + \tilde{Y}_{12}^T\Omega - 2\tilde{Y}_{22}\tilde{\Omega} = \Psi_{22}V + V\Psi_{22}, \quad (2.27)$$

wobei Gleichung (2.26) die adjungierte Gleichung zu (2.25) ist. Aus der ersten Gleichung ergibt sich

$$\tilde{Y}_{12} = \frac{1}{2}S\Omega^{-1}, \quad \tilde{Y}_{12}^T = -\frac{1}{2}\Omega^{-1}S, \quad (2.28)$$

wobei S eine schiefsymmetrische Matrix ist, d.h. $S^T = -S$. Setzt man dies in Gleichung (2.23) ein, so ergibt sich

$$\begin{aligned} \frac{d^2}{d\mu^2} \text{Tr} Y_\mu \Big|_{\mu=0} &= 4 \text{Tr} \left[\left(\frac{1}{2}S\Omega^{-1} \right) V \Phi_{12}^T \right] - 4 \text{Tr} [V\tilde{Y}_{22}\Phi_{22}] \\ &= 2 \text{Tr} [S\Omega^{-1}V\Phi_{12}^T] - 4 \text{Tr} [V\tilde{Y}_{22}\Phi_{22}]. \end{aligned} \quad (2.29)$$

Nun müssen noch \tilde{Y}_{22} und S näher bestimmt werden. Seien

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

Wir betrachten zunächst die vier Gleichungen (2.24) - (2.27): Mit (2.20) erhält man

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

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

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

Damit gehen die vier Matrixgleichungen (2.24) - (2.27) über in

$$\begin{aligned}
& - \begin{pmatrix} \hat{\Omega}_s & 0 \\ 0 & \check{\Omega}_{n-s} \end{pmatrix} \begin{pmatrix} Q_{11}^T & Q_{21}^T \\ Q_{12}^T & Q_{22}^T \end{pmatrix} - \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix} \begin{pmatrix} \hat{\Omega}_s & 0 \\ 0 & \check{\Omega}_{n-s} \end{pmatrix} \\
& = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \tag{2.30}
\end{aligned}$$

$$\begin{aligned}
& - \begin{pmatrix} \hat{\Omega}_s & 0 \\ 0 & \check{\Omega}_{n-s} \end{pmatrix} \begin{pmatrix} P_{11} & P_{12} \\ P_{12}^T & P_{22} \end{pmatrix} + \begin{pmatrix} R_{11} & R_{12} \\ R_{12}^T & R_{22} \end{pmatrix} \begin{pmatrix} \hat{\Omega}_s & 0 \\ 0 & \check{\Omega}_{n-s} \end{pmatrix} \\
& - 2 \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix} \begin{pmatrix} \hat{\Omega}_s & 0 \\ 0 & H \end{pmatrix} = \begin{pmatrix} \frac{1}{2}\hat{\Omega}_s^{-1}V_{11} & \frac{1}{2}\hat{\Omega}_s^{-1}V_{12} \\ 0 & 0 \end{pmatrix}, \tag{2.31}
\end{aligned}$$

$$\begin{aligned}
& \begin{pmatrix} \hat{\Omega}_s & 0 \\ 0 & \check{\Omega}_{n-s} \end{pmatrix} \begin{pmatrix} R_{11} & R_{12} \\ R_{12}^T & R_{22} \end{pmatrix} - 2 \begin{pmatrix} \hat{\Omega}_s & 0 \\ 0 & H \end{pmatrix} \begin{pmatrix} Q_{11}^T & Q_{21}^T \\ Q_{12}^T & Q_{22}^T \end{pmatrix} \\
& - \begin{pmatrix} P_{11} & P_{12} \\ P_{12}^T & P_{22} \end{pmatrix} \begin{pmatrix} \hat{\Omega}_s & 0 \\ 0 & \check{\Omega}_{n-s} \end{pmatrix} = \begin{pmatrix} \frac{1}{2}V_{11}\hat{\Omega}_s^{-1} & 0 \\ \frac{1}{2}V_{12}^T\hat{\Omega}_s^{-1} & 0 \end{pmatrix}, \tag{2.32}
\end{aligned}$$

$$\begin{aligned}
& \begin{pmatrix} \hat{\Omega}_s & 0 \\ 0 & \check{\Omega}_{n-s} \end{pmatrix} \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix} - 2 \begin{pmatrix} \hat{\Omega}_s & 0 \\ 0 & H \end{pmatrix} \begin{pmatrix} P_{11} & P_{12} \\ P_{12}^T & P_{22} \end{pmatrix} \\
& + \begin{pmatrix} Q_{11}^T & Q_{21}^T \\ Q_{12}^T & Q_{22}^T \end{pmatrix} \begin{pmatrix} \hat{\Omega}_s & 0 \\ 0 & \check{\Omega}_{n-s} \end{pmatrix} - 2 \begin{pmatrix} P_{11} & P_{12} \\ P_{12}^T & P_{22} \end{pmatrix} \begin{pmatrix} \hat{\Omega}_s & 0 \\ 0 & H \end{pmatrix} \\
& = \begin{pmatrix} \frac{1}{2}\hat{\Omega}_s^{-1}V_{11} + \frac{1}{2}V_{11}\hat{\Omega}_s^{-1} & \frac{1}{2}\hat{\Omega}_s^{-1}V_{12} \\ \frac{1}{2}V_{12}^T\hat{\Omega}_s^{-1} & 0 \end{pmatrix}. \tag{2.33}
\end{aligned}$$

1. Von den Gleichungen (2.30) - (2.33) betrachten wir zunächst nur jeweils den rechten unteren Matrizenblock, d.h. die 2,2 - Blöcke der Matrixgleichungen. Es ergeben sich folgende vier Gleichungen, wobei wieder die zweite und dritte Gleichung zueinander adjungiert sind:

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

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

Dies ist äquivalent mit der Matrixgleichung

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

Da $H \in \mathcal{S}$, ist auch die Matrix $\begin{pmatrix} 0 & \check{\Omega}_{n-s} \\ -\check{\Omega}_{n-s} & -2H \end{pmatrix}$ stabil, und die Matrixgleichung hat nur die triviale Lösung $R_{22} = Q_{22} = P_{22} = 0$. Mit (siehe (2.28))

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

folgt damit auch $S_{22} = 0$.

Setzt man dieses Ergebnis nun in (2.29) ein, so erhalten wir mit (2.21)

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

Da S_{11} eine schiefsymmetrische und V_{11} , $\hat{\Omega}_s^{-1}$ symmetrische Matrizen sind, gilt

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

Daher erhält man

$$\begin{aligned}
\frac{d^2}{d\mu^2} \text{Tr} Y_\mu \Big|_{\mu=0} &= \text{Tr} [S_{12} \check{\Omega}_{n-s}^{-1} V_{12}^T \hat{\Omega}_s^{-1}] - 2 \text{Tr} [S_{12}^T \hat{\Omega}_s^{-1} V_{12} W_{12}^T] \\
&\quad - 2 \text{Tr} [V_{11} P_{11} \hat{\Omega}_s^{-1}] - 2 \text{Tr} [V_{12} P_{12}^T \hat{\Omega}_s^{-1}] - 4 \text{Tr} [V_{12}^T P_{12} W_{22}].
\end{aligned} \tag{2.34}$$

2. Als nächstes betrachten wir die 1,2 - Blöcke der vier Matrixgleichungen (2.30) - (2.33). Man erhält folgende vier Gleichungen:

$$-\hat{\Omega}_s Q_{21}^T - Q_{12} \check{\Omega}_{n-s} = 0 \tag{2.35}$$

$$-\hat{\Omega}_s P_{12} + R_{12} \check{\Omega}_{n-s} - 2Q_{12} H = \frac{1}{2} \hat{\Omega}_s^{-1} V_{12} \tag{2.36}$$

$$\hat{\Omega}_s R_{12} - 2\hat{\Omega}_s Q_{21}^T - P_{12} \check{\Omega}_{n-s} = 0 \tag{2.37}$$

$$\hat{\Omega}_s Q_{12} - 2\hat{\Omega}_s P_{12} + Q_{21}^T \check{\Omega}_{n-s} - 2P_{12} H = \frac{1}{2} \hat{\Omega}_s^{-1} V_{12}. \tag{2.38}$$

Mit (s. (2.28))

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

folgt

$$Q_{12} = \frac{1}{2} S_{12} \check{\Omega}_{n-s}^{-1}, \quad Q_{21} = -\frac{1}{2} S_{12}^T \hat{\Omega}_s^{-1}, \quad Q_{21}^T = -\frac{1}{2} \hat{\Omega}_s^{-1} S_{12}. \tag{2.39}$$

Setzt man dies nun in Gleichung (2.37) ein, so ergibt sich $\hat{\Omega}_s R_{12} + S_{12} - P_{12} \check{\Omega}_{n-s} = 0$, d.h.

$$S_{12} = P_{12} \check{\Omega}_{n-s} - \hat{\Omega}_s R_{12}. \tag{2.40}$$

Nun setzen wir (2.39) in Gleichung (2.36) und (2.38) ein und erhalten

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

Ersetzt man nun in diesen Gleichungen S_{12} durch (2.40), so ergibt sich

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

Subtrahiert man die beiden letzten Gleichungen, so ergibt sich

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

d.h.

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

Diese Gleichung kann auch geschrieben werden als

$$\Delta H + \frac{1}{2} \hat{\Omega}_s \Delta + \frac{1}{2} \hat{\Omega}_s^{-1} \Delta \check{\Omega}_{n-s}^2 = 0 \quad (2.41)$$

mit $\Delta = P_{12} + \hat{\Omega}_s R_{12} \check{\Omega}_{n-s}^{-1}$. Gleichung (2.41) hat nur die triviale Lösung $\Delta = 0$, denn wenn man diese Gleichung mit Δ^T von rechts multipliziert, erhält man

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

Damit gilt auch

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

d.h.

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

Jeder einzelne Summand ist größer oder gleich Null. Da die Matrix $\hat{\Omega}_s$ positiv definit ist, ist die Gleichung nur für $\Delta = 0$ erfüllt.

Damit ist $P_{12} + \hat{\Omega}_s R_{12} \check{\Omega}_{n-s}^{-1} = 0$, d.h. $P_{12} = -\hat{\Omega}_s R_{12} \check{\Omega}_{n-s}^{-1}$. Setzt man dies nun in (2.40) ein, so ergibt sich für S_{12}

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

und damit (siehe (2.39))

$$Q_{12} = \frac{1}{2} S_{12} \check{\Omega}_{n-s}^{-1} = -\hat{\Omega}_s R_{12} \check{\Omega}_{n-s}^{-1} = P_{12} \quad (2.42)$$

$$Q_{21} = -\frac{1}{2} S_{12}^T \hat{\Omega}_s^{-1} = R_{12}^T \hat{\Omega}_s \hat{\Omega}_s^{-1} = R_{12}^T . \quad (2.43)$$

Also erhält man

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

Ersetzt man nun in Gleichung (2.34) P_{12} durch $\frac{1}{2} S_{12} \check{\Omega}_{n-s}^{-1}$ (vgl. (2.42)), so folgt

$$\begin{aligned} \frac{d^2}{d\mu^2} \text{Tr} Y_\mu \Big|_{\mu=0} &= \text{Tr} [S_{12} \check{\Omega}_{n-s}^{-1} V_{12}^T \hat{\Omega}_s^{-1}] - 2\text{Tr} [S_{12}^T \hat{\Omega}_s^{-1} V_{12} W_{12}^T] - 2\text{Tr} [V_{11} P_{11} \hat{\Omega}_s^{-1}] \\ &\quad - 2\text{Tr} [V_{12} (\frac{1}{2} \check{\Omega}_{n-s}^{-1} S_{12}^T) \hat{\Omega}_s^{-1}] - 4\text{Tr} [V_{12}^T (\frac{1}{2} S_{12} \check{\Omega}_{n-s}^{-1}) W_{22}] \\ &= \text{Tr} [V_{12} \check{\Omega}_{n-s}^{-1} S_{12}^T \hat{\Omega}_s^{-1}] - 2\text{Tr} [S_{12}^T \hat{\Omega}_s^{-1} V_{12} W_{12}^T] - 2\text{Tr} [V_{11} P_{11} \hat{\Omega}_s^{-1}] \\ &\quad - \text{Tr} [V_{12} \check{\Omega}_{n-s}^{-1} S_{12}^T \hat{\Omega}_s^{-1}] - 2\text{Tr} [V_{12}^T S_{12} \check{\Omega}_{n-s}^{-1} W_{22}] \\ &= -2\text{Tr} [V_{11} P_{11} \hat{\Omega}_s^{-1}] - 2\text{Tr} [V_{12}^T \hat{\Omega}_s^{-1} S_{12} W_{12}] \\ &\quad - 2\text{Tr} [V_{12}^T S_{12} \check{\Omega}_{n-s}^{-1} W_{22}] . \end{aligned} \quad (2.44)$$

3a. Wir betrachten nun zunächst den ersten Term auf der rechten Seite von (2.44); wir zeigen, daß $-2\text{Tr} [V_{11} P_{11} \hat{\Omega}_s^{-1}] \geq 0$ gilt, und Gleichheit dabei nur für $V_{11} = 0$ möglich ist.

Zuerst bestimmen wir P_{11} . Dazu betrachten wir die 1,1 -Blöcke des Gleichungssystems (2.30) - (2.33). Man erhält die vier Gleichungen

$$-\hat{\Omega}_s Q_{11}^T - Q_{11} \hat{\Omega}_s = 0 \quad (2.45)$$

$$-\hat{\Omega}_s P_{11} + R_{11} \hat{\Omega}_s - 2Q_{11} \hat{\Omega}_s = \frac{1}{2} \hat{\Omega}_s^{-1} V_{11} \quad (2.46)$$

$$\hat{\Omega}_s R_{11} - 2\hat{\Omega}_s Q_{11}^T - P_{11} \hat{\Omega}_s = \frac{1}{2} V_{11} \hat{\Omega}_s^{-1} \quad (2.47)$$

$$\hat{\Omega}_s Q_{11} - 2\hat{\Omega}_s P_{11} + Q_{11}^T \hat{\Omega}_s - 2P_{11} \hat{\Omega}_s = \frac{1}{2} \hat{\Omega}_s^{-1} V_{11} + \frac{1}{2} V_{11} \hat{\Omega}_s^{-1} . \quad (2.48)$$

Aus der ersten Gleichung zusammen mit (2.28) ergibt sich

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

wobei S_{11} eine schiefsymmetrische Matrix ist. Setzt man dies in Gleichung (2.48) ein, so erhält man

$$\frac{1}{2}\hat{\Omega}_s S_{11} \hat{\Omega}_s^{-1} - 2\hat{\Omega}_s P_{11} - \frac{1}{2}\hat{\Omega}_s^{-1} S_{11} \hat{\Omega}_s - 2P_{11} \hat{\Omega}_s = \frac{1}{2}\hat{\Omega}_s^{-1} V_{11} + \frac{1}{2}V_{11} \hat{\Omega}_s^{-1}. \quad (2.49)$$

Zuerst können die Diagonalelemente der Matrix P_{11} bestimmt werden. Da

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

für $i = 1, \dots, s$ ergibt sich aus (2.49)

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

Also ist

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

Nun werden die Außendiagonalelemente bestimmt. Gleichung (2.49) liefert für $1 \leq i < j \leq s$

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

Mit $S_{11} = (f_{ij})$ folgt

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

Also ist

$$f_{ij}(\omega_i - \omega_j) = (V_{11})_{ij} + 4\omega_i \omega_j (P_{11})_{ij}. \quad (2.50)$$

Um eine zweite Abhängigkeit zwischen f_{ij} und $(P_{11})_{ij}$ zu bekommen, ersetzt man in Gleichung (2.46) Q_{11} durch $\frac{1}{2}S_{11} \hat{\Omega}_s^{-1}$ und löst diese komponentenweise:

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

Da die Matrizen P_{11} und R_{11} symmetrisch sind und S_{11} schiefsymmetrisch ist, d.h. $(P_{11})_{ij} = (P_{11})_{ji}$, $(R_{11})_{ij} = (R_{11})_{ji}$, $f_{ij} = -f_{ji}$, erhält man die beiden Gleichungen

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

Multipliziert man die erste Gleichung mit ω_i und die zweite mit ω_j und subtrahiert diese dann, so erhält man

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

also

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

d.h.

$$f_{ij} = (P_{11})_{ij}(\omega_j - \omega_i). \quad (2.51)$$

Setzt man nun (2.51) in (2.50) ein, so erhält man

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

und somit

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

Also ist

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

Damit ist

$$\begin{aligned} -2\text{Tr} [V_{11}P_{11}\hat{\Omega}_s^{-1}] &= -2\sum_{i=1}^s \sum_{j=1}^s \frac{(V_{11})_{ij}(P_{11})_{ji}}{\omega_i} \\ &= 2\sum_{i=1}^s \sum_{j=1}^s \frac{(V_{11})_{ij}^2}{\omega_i(\omega_i + \omega_j)^2} \\ &\geq 0, \end{aligned} \quad (2.52)$$

wobei $-2\text{Tr} [V_{11}P_{11}\hat{\Omega}_s^{-1}] = 0$ nur für $V_{11} = 0$ möglich ist. Schritt 3a. ist damit abgeschlossen.

3b. Als letztes betrachten wir die beiden restlichen Terme auf der rechten Seite von (2.44); wir zeigen, daß

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

gilt, wobei Gleichheit nur für $V_{12} = 0$ möglich ist .

Nach (2.42) und (2.43) gilt

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

Setzt man dies in Gleichung (2.36) ein, so ergibt sich

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

d.h.

$$V_{12}^T = -\check{\Omega}_{n-s}^{-1}S_{12}^T\hat{\Omega}_s^2 - \check{\Omega}_{n-s}S_{12}^T - 2H\check{\Omega}_{n-s}^{-1}S_{12}^T\hat{\Omega}_s . \quad (2.53)$$

Damit ist

$$\begin{aligned} \hat{\Delta} &= 2\text{Tr} [\check{\Omega}_{n-s}^{-1}S_{12}^T\hat{\Omega}_sS_{12}W_{12}] + 2\text{Tr} [\check{\Omega}_{n-s}S_{12}^T\hat{\Omega}_s^{-1}S_{12}W_{12}] + 4\text{Tr} [H\check{\Omega}_{n-s}^{-1}S_{12}^T\hat{\Omega}_sS_{12}W_{12}] \\ &\quad + 2\text{Tr} [\check{\Omega}_{n-s}^{-1}S_{12}^T\hat{\Omega}_s^2S_{12}\check{\Omega}_{n-s}^{-1}W_{22}] + 2\text{Tr} [\check{\Omega}_{n-s}S_{12}^T\hat{\Omega}_s^{-1}S_{12}\check{\Omega}_{n-s}^{-1}W_{22}] \\ &\quad + 4\text{Tr} [H\check{\Omega}_{n-s}^{-1}S_{12}^T\hat{\Omega}_sS_{12}\check{\Omega}_{n-s}^{-1}W_{22}] \\ &= 2\text{Tr} [\check{\Omega}_{n-s}^{-1}S_{12}^T\hat{\Omega}_sS_{12}W_{12}] + 2\text{Tr} [\check{\Omega}_{n-s}S_{12}^T\hat{\Omega}_s^{-1}S_{12}W_{12}] + 2\text{Tr} [\check{\Omega}_{n-s}^{-1}S_{12}^T\hat{\Omega}_sS_{12}2W_{12}H] \\ &\quad + 2\text{Tr} [\check{\Omega}_{n-s}^{-1}S_{12}^T\hat{\Omega}_s^2S_{12}\check{\Omega}_{n-s}^{-1}W_{22}] + 2\text{Tr} [\check{\Omega}_{n-s}S_{12}^T\hat{\Omega}_s^{-1}S_{12}\check{\Omega}_{n-s}^{-1}W_{22}] \\ &\quad + \text{Tr} [(2HW_{22} + 2W_{22}H)\check{\Omega}_{n-s}^{-1}S_{12}^T\hat{\Omega}_sS_{12}\check{\Omega}_{n-s}^{-1}] . \end{aligned} \quad (2.54)$$

Die Ljapunovgleichung (2.22) lautet blockweise

$$\begin{aligned} -\check{\Omega}_{n-s}W_{12}^T - W_{12}\check{\Omega}_{n-s} &= -I_{n-s} \\ -\check{\Omega}_{n-s}W_{22} + W_{11}\check{\Omega}_{n-s} - 2W_{12}H &= 0 \\ \check{\Omega}_{n-s}W_{11} - 2HW_{12}^T - W_{22}\check{\Omega}_{n-s} &= 0 \\ \check{\Omega}_{n-s}W_{12} - 2HW_{22} + W_{12}^T\check{\Omega}_{n-s} - 2W_{22}H &= -I_{n-s} . \end{aligned} \quad (2.55)$$

Aus der ersten Gleichung folgt

$$W_{12} = \frac{1}{2}\check{\Omega}_{n-s}^{-1} + \frac{1}{2}G\check{\Omega}_{n-s}^{-1} , \quad (2.56)$$

mit $G^T = -G$, d.h G schiefsymmetrisch. Außerdem ergibt sich aus dem Gleichungssystem (2.55)

$$2W_{12}H = -\check{\Omega}_{n-s}W_{22} + W_{11}\check{\Omega}_{n-s} \quad (2.57)$$

$$2HW_{22} + 2W_{22}H = \check{\Omega}_{n-s}W_{12} + W_{12}^T\check{\Omega}_{n-s} + I_{n-s} . \quad (2.58)$$

Setzt man (2.56) , (2.57) und (2.58) in (2.54) ein, so ergibt sich

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

Da $S_{12}^T \hat{\Omega}_s^{-1} S_{12}$ symmetrisch ist, gilt $\operatorname{Tr} [S_{12}^T \hat{\Omega}_s^{-1} S_{12} G] = \operatorname{Tr} [G^T S_{12}^T \hat{\Omega}_s^{-1} S_{12}]$, und damit erhalten wir

$$\begin{aligned}
\hat{\Delta} &= 2 \operatorname{Tr} [\check{\Omega}_{n-s}^{-1} S_{12}^T \hat{\Omega}_s S_{12} W_{12}] + \operatorname{Tr} [S_{12}^T \hat{\Omega}_s^{-1} S_{12}] \\
&\quad + \operatorname{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 \operatorname{Tr} [(W_{22} \check{\Omega}_{n-s}^{-1} S_{12}^T S_{12} \check{\Omega}_{n-s})^T] \\
&\quad + 2 \operatorname{Tr} [S_{12}^T S_{12} W_{11}] + 2 \operatorname{Tr} [\check{\Omega}_{n-s}^{-1} S_{12}^T \hat{\Omega}_s^2 S_{12} \check{\Omega}_{n-s}^{-1} W_{22}] \\
&\quad + 2 \operatorname{Tr} [\check{\Omega}_{n-s} S_{12}^T S_{12} \check{\Omega}_{n-s}^{-1} W_{22}] + \operatorname{Tr} [\check{\Omega}_{n-s}^{-1} S_{12}^T \hat{\Omega}_s S_{12} W_{12}] \\
&\quad + \operatorname{Tr} [(W_{12}^T S_{12}^T \hat{\Omega}_s S_{12} \check{\Omega}_{n-s}^{-1})^T] + \operatorname{Tr} [\check{\Omega}_{n-s}^{-1} S_{12}^T \hat{\Omega}_s S_{12} \check{\Omega}_{n-s}^{-1}] \\
&= 3 \operatorname{Tr} [\check{\Omega}_{n-s}^{-1} S_{12}^T \hat{\Omega}_s S_{12} W_{12}] + \operatorname{Tr} [S_{12}^T \hat{\Omega}_s^{-1} S_{12}] + \frac{1}{2} \operatorname{Tr} [S_{12}^T \hat{\Omega}_s^{-1} S_{12} G] \\
&\quad - \frac{1}{2} \operatorname{Tr} [S_{12}^T \hat{\Omega}_s^{-1} S_{12} G] - 2 \operatorname{Tr} [\check{\Omega}_{n-s} S_{12}^T S_{12} \check{\Omega}_{n-s}^{-1} W_{22}] + 2 \operatorname{Tr} [S_{12}^T S_{12} W_{11}] \\
&\quad + 2 \operatorname{Tr} [\check{\Omega}_{n-s}^{-1} S_{12}^T \hat{\Omega}_s^2 S_{12} \check{\Omega}_{n-s}^{-1} W_{22}] + 2 \operatorname{Tr} [\check{\Omega}_{n-s} S_{12}^T S_{12} \check{\Omega}_{n-s}^{-1} W_{22}] \\
&\quad + \operatorname{Tr} [\check{\Omega}_{n-s}^{-1} S_{12}^T \hat{\Omega}_s S_{12} W_{12}] + \operatorname{Tr} [\check{\Omega}_{n-s}^{-1} S_{12}^T \hat{\Omega}_s S_{12} \check{\Omega}_{n-s}^{-1}] \\
&= \operatorname{Tr} [\check{\Omega}_{n-s}^{-1} S_{12}^T \hat{\Omega}_s S_{12} \check{\Omega}_{n-s}^{-1}] + \operatorname{Tr} [S_{12}^T \hat{\Omega}_s^{-1} S_{12}] + 4 \operatorname{Tr} [\check{\Omega}_{n-s}^{-1} S_{12}^T \hat{\Omega}_s S_{12} W_{12}] \\
&\quad + 2 \operatorname{Tr} [S_{12}^T S_{12} W_{11}] + 2 \operatorname{Tr} [\check{\Omega}_{n-s}^{-1} S_{12}^T \hat{\Omega}_s^2 S_{12} \check{\Omega}_{n-s}^{-1} W_{22}].
\end{aligned}$$

Da

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

positiv semidefinit sind, erhalten wir

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

Auf Grund der positiven Definitheit von $\hat{\Omega}_s$ kann $\text{Tr} [S_{12}^T \hat{\Omega}_s^{-1} S_{12}]$ nur gleich Null für $S_{12} = 0$, also (siehe (2.53)) für $V_{12} = 0$ sein.

Es bleibt zu zeigen, daß

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

ist. Dazu führt man eine Transformation der (als Lösung von (2.22)) positiv definiten Matrix $\begin{pmatrix} W_{11} & W_{12} \\ W_{12}^T & W_{22} \end{pmatrix}$ durch.

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

ist positiv semidefinit, d.h. insbesondere, daß für alle s -dimensionalen Vektoren x gilt

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

Da $\tilde{W}_{11} + \tilde{W}_{12} + \tilde{W}_{12}^T + \tilde{W}_{22}$ symmetrisch ist, ist diese Matrix auch positiv semidefinit,

d.h.

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

Also haben wir gezeigt, daß

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

ist, wobei Gleichheit nur für $V_{12} = 0$ möglich ist;
damit ist Schritt 3 b. abgeschlossen.

Insgesamt erhält man damit (vgl. (2.44) und (2.52))

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

wobei die Gleichheit nur für $V_{11} = V_{12} = 0$ möglich ist.

Q.E.D.

Bemerkung 2.8

Die Aussage von Satz 2.6 schließt die Aussage von Satz 2.1 ein. Da der Beweis zu Satz 2.1 aber viel leichter ist, wurde er dort explizit vorgeführt.

2.2 Experimentelle Untersuchung der lokalen Minima auf Globalität

In diesem Abschnitt wollen wir experimentell untersuchen, ob das bei den Matrizen C_{\min} aus der Menge \mathcal{M} in Satz 2.6 angenommene lokale Minimum von $\text{Tr} (XZ_s)$ auch globales Minimum ist. Dabei wurden Tests über eine breite Anzahl von Dämpfungsmatrizen mit unterschiedlichem Rang durchgeführt, die nicht nur diese Vermutung

erhärten, sondern auch annehmen lassen, daß für beliebig vorgegebene Dämpferanzahlen und -positionen genau ein lokales Minimum existiert, das auf einer zusammenhängenden Menge angenommen wird und gleichzeitig globales Minimum ist.

Da wir in Kapitel 2 über jede beliebige positiv semidefinite Dämpfungsmatrix C minimiert haben und C in der Form $C = \sum_k \varepsilon_k c_k c_k^T = V \Sigma V^T$ dargestellt werden kann, wobei c_k Vektor der Länge n , $V \in M_{n,g}$, $g \leq n$ und $\Sigma = \text{diag}(\varepsilon_1, \dots, \varepsilon_g)$, $\varepsilon_i \geq 0$ ist, wurden für jeweils feste V Tests durchgeführt, indem die ε_i variiert wurden. Ausgehend von einer Massenmatrix M und einer Steifigkeitsmatrix K aus (1.1) wurde die Dämpfungsmatrix D in der Form $D = \sum_k \varepsilon_k p_k p_k^T = P \Sigma P^T$ mit $\Sigma = \text{diag}(\varepsilon_k)$, $\varepsilon_k \geq 0$ dargestellt, wobei die Vektoren p_k als Euklidische Einheitsvektoren gewählt wurden. Schließlich wurde durch die in Kapitel 1 angegebene Transformation die Matrix A aus (1.16) erzeugt, so daß $\text{Tr}(XZ_s)$ aus (1.17) gelöst werden konnten. Die Matrix C ist dabei aus der Transformationen $C = \Phi^T D \Phi = \Phi^T P \Sigma P^T \Phi =: V \Sigma V^T$ mit $V = \Phi^T P$ hervorgegangen. In diesem Fall repräsentiert g die Anzahl der Dämpfer und ε_i , $1 \leq i \leq g$, die Viskosität des i -ten Dämpfers in einem Schwingungssystem. Abhängig von Φ wird dabei die Menge \mathcal{M} unterschiedlich gut approximiert. Durch (s.u.) die in den meisten Testreihen zufällige Wahl von M und K (also von Φ) ergibt sich aber insgesamt ein repräsentatives Gesamtbild.

Exemplarisch wurde der Fall $n = 6$ in verschiedenen Testreihen untersucht. In einer Testreihe wurde die Steifigkeitsmatrix K von der Form

$$K = \frac{1}{(n+1)^2} \begin{pmatrix} 2 & -1 & & & & 0 \\ -1 & 2 & -1 & & & \\ & -1 & 2 & \ddots & & \\ & & \ddots & \ddots & -1 & \\ 0 & & & -1 & 2 & \end{pmatrix},$$

in allen anderen Testreihen wurde K zufällig, positiv definit gewählt. Die Massenmatrix M war dabei meist die Einheitsmatrix, es wurden aber auch Tests mit beliebig positiv definiten Massenmatrix durchgeführt. Alle Testreihen wurden für verschiedene s , insbesondere für $s = n$ durchgeführt. Der Fall $s = 6 = n$ untersucht also insbesondere, ob das bei $C_{\min} = C_{krit} = 2\Omega$ aus Satz 2.1 angenommene lokale Minimum von $\text{Tr}(X)$ auch globales Minimum ist.

Die Untersuchungen innerhalb einer einzelnen Testreihe fanden für alle möglichen Anzahlen von Dämpfern g mit $1 \leq g \leq n$ statt. Für jede feste Anzahl von Dämpfern g_0 ist jede mögliche Kombination ohne Wiederholung durch Permutation von g_0 verschiedenen der n euklidischen Einheitsvektoren untersucht worden. Diese verschiedenen Kombinationen spiegeln sich in $\binom{n}{g_0}$ verschiedenen Matrizen $P_j^{(g_0)}$, $1 \leq j \leq \binom{n}{g_0}$, wieder. So erhielt man innerhalb einer Testreihe in Abhängigkeit von $g_0 = 1, \dots, n$ und den $\binom{n}{g_0}$ verschiedenen Matrizen $P_j^{(g_0)}$ insgesamt $\sum_{g_0=1}^n \binom{n}{g_0} = 2^n - 1$ Testfälle, wobei in jedem Testfall $T_j^{(g_0)}$ nur noch die Viskositäten $\varepsilon_1, \dots, \varepsilon_{g_0}$ der g_0 Dämpfer zu

varyieren sind. Also ist für jeden der $\binom{n}{g_0}$ Testfälle mit g_0 Dämpfern die Abbildung $h_j^{(g_0)} : (\varepsilon_1, \dots, \varepsilon_{g_0}) \rightarrow \text{Tr}(XZ_s)$, $\varepsilon_i \geq 0$ für alle $1 \leq i \leq g_0$, zu untersuchen. (Im folgenden schreiben wir nur h statt $h_j^{(g_0)}$, falls die Spezialisierung des betrachteten Testfalls unwesentlich ist.) In einem einzelnen Testfall wurde h beschränkt auf einem Gitter Γ_h mit

$$\Gamma_h = \left\{ (\varepsilon_1, \dots, \varepsilon_{g_0}) : \varepsilon_i = \varepsilon_i^{(\min)} + k_i \frac{\varepsilon_i^{(\max)} - \varepsilon_i^{(\min)}}{k_i^{(\max)}}, k_i = 0, \dots, k_i^{(\max)}, i = 1, \dots, g_0 \right\}$$

betrachtet und $\gamma^* \in \Gamma_h$ mit $h(\gamma^*) = \min_{\gamma \in \Gamma} h(\gamma)$ bestimmt.

Dabei konnten (und wurden) $\varepsilon_i^{(\min)}$ und $\varepsilon_i^{(\max)}$ immer so gewählt, daß $0 < k_i^* < k_i^{(\max)}$ für $\gamma_i^* = \varepsilon_i^{(\min)} + k_i^* \frac{\varepsilon_i^{(\max)} - \varepsilon_i^{(\min)}}{k_i^{(\max)}}$, $i = 1, \dots, g_0$ gilt, wobei γ_i die i -te Komponente des Vektors $\gamma \in \Gamma_h$ bezeichnet. D.h. in allen durchgeführten Tests existierte ein Gitter Γ_h zu jedem h , so daß das Minimum von h nicht auf dem Rand des Gitters angenommen wurde.

Innerhalb einer jeden Testreihe galt mit den eindeutigen Minimumstellen $\gamma_j^{*(g_0)}$ die Aussage

$$\min_{1 \leq j \leq \binom{n}{g}} h(\gamma_j^{*(g)}) \leq \min_{1 \leq j \leq \binom{n}{g-1}} h(\gamma_j^{*(g-1)})$$

für alle g mit $2 \leq g \leq n$.

Des weiteren konnte in allen betrachteten Testfällen bei Variation von nur einer Viskosität ε_{i_0} und Festhalten aller anderen Viskositäten ε_i , $i = 1, \dots, i_0 - 1, i_0 + 1, \dots, g_0$ immer ein \tilde{k}_{i_0} bestimmt werden, mit dem für

$$\tilde{\gamma}^{i_0, (k)} := \left(\varepsilon_1, \dots, \varepsilon_{i_0-1}, \varepsilon_{i_0}^{(\min)} + k \frac{\varepsilon_{i_0}^{(\max)} - \varepsilon_{i_0}^{(\min)}}{k_{i_0}^{(\max)}}, \varepsilon_{i_0+1}, \dots, \varepsilon_{g_0} \right)$$

die folgenden zwei Aussagen gelten:

$$h(\tilde{\gamma}^{i_0, (k-1)}) > h(\tilde{\gamma}^{i_0, (k)}) \quad \text{für } k = 2, \dots, \tilde{k}_{i_0} \quad (2.59)$$

$$h(\tilde{\gamma}^{i_0, (k)}) < h(\tilde{\gamma}^{i_0, (k+1)}) \quad \text{für } k = \tilde{k}_{i_0}, \dots, \tilde{k}_{i_0}^{(\max)} - 1. \quad (2.60)$$

Bei Betrachtung der Funktion h – beschränkt auf die Abhängigkeit einer einzelnen Viskosität ε_{i_0} , i_0 beliebig – galt also in allen Testfällen, daß ein eindeutiges lokales Minimum $h(\tilde{\gamma}^{i_0, (\tilde{k}_{i_0})})$ angenommen wurde. Dieses war bezüglich der Einschränkung global, da für alle Punkte des eingeschränkten Gitters, die kleiner als die Minimumstelle sind, die Funktion streng monoton fiel (s. (2.59)) und für alle Punkte des eingeschränkten Gitters, die größer als die Minimumstelle sind, die Funktion streng monoton stieg (s. (2.60)).

Abbildung 2.1 zeigt exemplarisch das Ergebnis eines Programmlaufs für einen Testfall mit $g = 2$ Dämpfern und $s = n$. Die Viskositäten ε_1 und ε_2 sind auf der x -Achse bzw. auf der y -Achse abgetragen. Auf der z -Achse ist die Spur $\text{Tr}(XZ_s) = \text{Tr}(X)$ abgetragen. In diesem Testfall war $\varepsilon_1^{(\min)} = 3$, $\varepsilon_2^{(\min)} = 4$, $\varepsilon_1^{(\max)} = \varepsilon_2^{(\max)} = 100$. Die kleinste in diesem Testfall berechnete Spur war 9,2909. Das lokale Minimum von $\text{Tr}(X)$ für $C_{krit} = 2\Omega$ aus Satz 2.1 war in dieser Testreihe 0,9529.

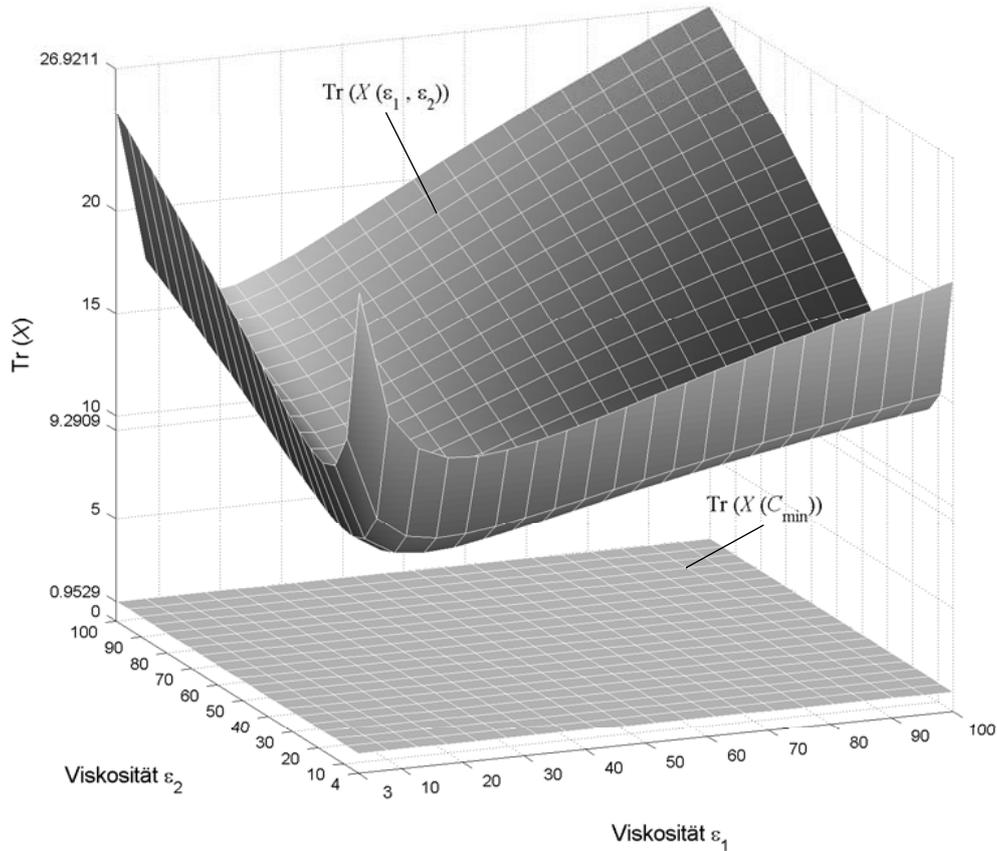


Abb. 2.1: $\text{Tr}(X(\varepsilon_1, \varepsilon_2))$ im Vergleich zu $\text{Tr}(X(C_{\min}))$

Für den Vergleich ist in der Abbildung die Ebene des lokalen Minimums, das durch C_{krit} angenommen wird, eingezeichnet worden. Die obere Fläche zeigt die Größe der Spur der Lösung der Ljapunovgleichung in Abhängigkeit von den zwei Viskositäten ε_1 und ε_2 . Wie man an der Abbildung erkennen kann, ist diese Fläche nicht konvex.

Zusammenfassend konnten also folgende Phänomene in allen Testfällen beobachtet werden:

- In allen Testfällen gab es eine eindeutige Minimumstelle $\gamma^* \in \Gamma$ mit $h(\gamma^*) = \min_{\gamma \in \Gamma} h(\gamma)$.
- In allen Testfällen lag die gefundene Minimumstelle $\gamma^* \in \Gamma$ im Inneren von Γ .
- In keiner Testreihe gab es einen Testfall, bei dem das gefundene Minimum $h(\gamma^*)$ kleiner als $\text{Tr}(X(C_{\min}))$ war ($C_{\min} \in \mathcal{M}$ bzw. $C_{\min} = C_{krit}$).
- In keiner Testreihe wurde bei einer größeren Anzahl der Dämpfer auch eine größere Spur im Gitter angenommen.
- Bei Beschränkung der Funktion h eines jeden Testfalls auf nur eine Viskosität des Gitters wird ein eindeutiges Minimum angenommen.
- Konvexität liegt jedoch nicht immer vor.

Insgesamt legen alle beobachteten Phänomene die Vermutung nahe, daß das bekannte bei C_{\min} angenommene lokale Minimum auch ein globales Minimum ist. Der theoretische Nachweis bleibt allerdings ein offenes Problem.

Alle Tests zur Überprüfung der lokalen Minima auf Globalität wurden auf einem PC Pentium 100 mit 32 MB RAM unter Windows 95 mit Matlab 5.1 durchgeführt.

3 Minimierung über die Viskositäten der Dämpfer

In Kapitel 2 war unser Ziel, eine durch eine Massenmatrix M und eine Steifigkeitsmatrix K gegebene Schwingungskonstruktion (1.1) über alle positiv semidefiniten Dämpfungsmatrizen D im Sinne von Problem 1.3 zu optimieren. Dabei haben wir untersucht, wann $\text{Tr}(XZ_s)$ minimal wird – wobei X eindeutige Lösung der Ljapunovgleichung $A^T X + X A = -I$,

$$A = \begin{pmatrix} 0 & \Omega \\ -\Omega & -C \end{pmatrix}, \quad Z_s = \begin{pmatrix} I_s & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & I_s & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

ist –, wenn über alle positiv semidefiniten Dämpfungsmatrizen C minimiert wird. Es wurde gezeigt, daß die Spur der Lösung der Ljapunovgleichung für alle Matrizen

$$C_{min} = 2 \begin{pmatrix} \hat{\Omega}_s & 0 \\ 0 & H \end{pmatrix}, \quad (3.1)$$

$\hat{\Omega}_s = \text{diag}(\omega_1, \dots, \omega_s)$, $H \in \mathcal{S}$ mit

$$\mathcal{S} = \left\{ H \in M_{n-s}, \quad H \text{ positiv semidefinit} : \quad A \text{ stabil mit } C = 2 \begin{pmatrix} \hat{\Omega}_s & 0 \\ 0 & H \end{pmatrix} \right\}$$

(\mathcal{S} enthält insbesondere alle positiv definiten $H \in M_{n-s}$), ein lokales Minimum annimmt. Für jede Matrix dieser Mannigfaltigkeit der lokalen Minima \mathcal{M} (für $s = n$, d.h. $Z_s = I$ besteht die Menge nur aus dem einen Element $2 \hat{\Omega}_s = 2\Omega$) ergibt sich derselbe Wert für $\text{Tr}(XZ_s)$.

In diesem Kapitel wollen wir nun nicht über alle positiv semidefiniten Dämpfungsmatrizen minimieren, sondern wir wollen den Raum der Dämpfungsmatrizen einschränken. In der Praxis sind die Anzahl und Positionen der Dämpfer einer Schwingungskonstruktion meist fest vorgegeben. Dabei ist im allgemeinen die Anzahl der Dämpfer viel kleiner als die Dimension n des Systems. Eine Optimierung dieser bestehenden Schwingungskonstruktion kann folglich entweder durch Hinzufügen von Dämpfern (was in der Praxis meist nicht der Fall ist) oder durch eine Veränderung der Viskositäten der vorhandenen Dämpfer erreicht werden. Mit letzterem

wollen wir uns in diesem Kapitel beschäftigen. Eine durch die Massenmatrix M , die Steifigkeitsmatrix K sowie die Anzahl und Positionen der Dämpfer gegebene Schwingungskonstruktion $M\ddot{x} + D\dot{x} + Kx = 0$ soll über die Viskositäten der Dämpfer optimiert werden. Die Dämpfungsmatrix D hat die Gestalt $D = P\Sigma P^T$ mit $P \in M_{n,g}$, $\Sigma = \text{diag}(\varepsilon_1, \dots, \varepsilon_g)$, wobei g die Anzahl und ε_i , $1 \leq i \leq g$, die zugehörigen nichtnegativen Viskositäten der Dämpfer sind. Die Matrix P beinhaltet die Positionen der Dämpfer und besteht meistens aus den euklidischen Einheitsvektoren, die den Positionen der Dämpfer entsprechen. Überführen wir dieses Problem wieder in die Form (1.17) und (1.16), so erhalten wir die transformierte Dämpfungsmatrix C der Form $C = \Phi^T D \Phi = \Phi^T P \Sigma P^T \Phi =: V \Sigma V^T$ mit $V = \Phi^T P$. Damit hat A die Gestalt

$$A = \begin{pmatrix} 0 & \Omega \\ -\Omega & -V\Sigma V^T \end{pmatrix}, \quad (3.2)$$

und es ergibt sich das folgende Minimierungsproblem:

Problem 3.1

Gegeben sei Z_s aus (2.13), $\Omega = \text{diag}(\omega_1, \dots, \omega_n)$ mit $\omega_i > 0$ für alle $i = 1, \dots, n$, $V \in M_{n,g}$ fest, $g \leq n$. Bestimme die nichtnegativen Viskositäten $\Sigma = \text{diag}(\varepsilon_1, \dots, \varepsilon_g)$ so, daß $\text{Tr}(XZ_s)$ minimal wird, wobei X die eindeutige Lösung von $A^T X + XA = -I$, A aus (3.2) stabil, ist.

Wäre es möglich, Viskositäten Σ so zu finden, daß für festes V die Beziehung $V\Sigma V^T \in \mathcal{M}$ (s.o.) gilt, dann hätte man optimale Viskositäten erhalten. Dies ist, da V fest vorgegeben und g meist viel kleiner als n ist, im allgemeinen nicht möglich. Veselić hat in [24] das Optimierungsproblem für einen Dämpfer untersucht und analytisch die Lösung X der Ljapunovgleichung und somit $\text{Tr}(XZ_s)$ für $g = 1$ berechnet. Die Funktion $\varepsilon_1 \mapsto \text{Tr}(X)$ ist dabei von der Form

$$\text{Tr}(X) = \frac{a}{\varepsilon_1} + b\varepsilon_1 + \text{const}$$

mit $a, b > 0$. Sie ist also konvex und das Minimum wird bei

$$\varepsilon_1 = \varepsilon_{opt} = \sqrt{\frac{a}{b}}$$

angenommen.

Im Fall mehrerer Dämpfer sind die auftretenden Terme um ein Vielfaches komplexer. Im scheinbar einfachen Fall der Dimension $n = 2$ mit 2 Dämpfern läßt sich die Spur von X mit Hilfe von Lemma 3.2 noch explizit berechnen. Die Dämpfungsmatrix hat dann die Gestalt $C = V\Sigma V^T = \varepsilon_1 cc^T + \varepsilon_2 dd^T$, wobei c und d Vektoren der Länge 2 sind. Zur besseren Übersicht wählen wir im Lemma die Schreibweise $C = (c_{ij})_{i,j=1,2}$.

Lemma 3.2

Sei A eine stabile Matrix der Dimension 4 mit

$$A = \begin{pmatrix} 0 & 0 & \omega_1 & 0 \\ 0 & 0 & 0 & \omega_2 \\ -\omega_1 & 0 & -c_{11} & -c_{12} \\ 0 & -\omega_2 & -c_{12} & -c_{22} \end{pmatrix}, \quad \text{und } X = \begin{pmatrix} X_{11} & X_{12} \\ X_{12}^T & X_{22} \end{pmatrix}$$

sei die eindeutig bestimmten Lösung von $A^T X + X A = -I$. Dann sind die Diagonalelemente von X gegeben durch

$$(X_{11})_{ii} = \frac{c_{ii}}{2\omega_i^2} + \frac{\omega_k^2 \det(C)(c_{11} + c_{22}) + \omega_i^2(c_{11} + c_{22})(c_{kk}^2 + c_{12}^2) + c_{kk}(\omega_1^2 - \omega_2^2)^2}{\det(C)(c_{11}c_{22}\omega_1^2 + c_{22}^2\omega_1^2 + c_{11}c_{22}\omega_2^2 + c_{11}^2\omega_2^2) + c_{11}c_{22}(\omega_1^2 - \omega_2^2)^2},$$

$$(X_{22})_{ii} = \frac{1}{c_{ii}} \left[1 + \frac{c_{12}^2(c_{11} + c_{22})(\omega_1^2 c_{22} + c_{11} \omega_2^2)}{\det(C)(c_{11} + c_{22})(\omega_1^2 c_{22} + \omega_2^2 c_{11}) + c_{11}c_{22}(\omega_1^2 - \omega_2^2)^2} \right]$$

mit $i = 1, 2$ und $k = 3 - i$.

Die Aussage des Lemmas wurde mit Unterstützung des Programms Mathematica bestimmt, welches in der Lage ist, Gleichungssysteme symbolisch zu lösen. Während das Programm den Fall $g = 1$ und $n = 2$ (und auch größere n) ohne nennenswerten Rechenaufwand löst, enden die Berechnungen der Diagonalelemente von X für den Fall $n = g = 2$ mit Mathematica in großen Termen, deren umfangreiche Vereinfachung schließlich zur Aussage von Lemma 3.2 führen. Bereits für den Fall $n = 3$ und $g = 2$ führten Ansätze bei der symbolischen Lösung der Ljapunovgleichung zu so großen Termen, daß die gegebenen Rechnerkapazitäten nicht ausreichten, um mit ihrer Hilfe zu einer Lösung der Ljapunovgleichung zu gelangen. Obwohl $\text{Tr}(X)$ für $n = g = 2$ durch Lemma 3.2 noch explizit berechnet werden kann, scheitert die Berechnung ihres Minimums an dem damit verbundenen hohen Rechenaufwand. Die Berechnung von $\text{Tr}(XZ_s)$ und ihre Minimierung für $g > 1$, d.h. für mehr als einen Dämpfer, kann daher bislang nur numerisch erfolgen. Für feste Positionen und Anzahl der Dämpfer (gegeben durch V) ist $\text{Tr}(XZ_s)$ eine glatte Funktion der Viskositäten Σ . Somit kann das oben gestellte Optimierungsproblem 3.1 mittels klassischer Minimierungsverfahren, wie Gradienten- oder Newtonverfahren, numerisch gelöst werden. In [25] beschreibt Veselić ein von ihm zur Optimierung einer Schwingungskonstruktion entwickeltes Programmpaket. Dieses ermittelt ausgehend von einer diagonalen Massenmatrix M , den Frequenzen ω_i des Systems, der Matrix Φ (siehe (1.15)) sowie den Positionen der Dämpfer, die durch die entsprechenden euklidischen Einheitsvektoren gegeben sind, die Matrix V der Phasenraummatrix A (vgl. Problem 3.1). Des weiteren enthält es die Umsetzung der Algorithmen des Gradienten- und Newtonverfahrens, die eine numerische Lösung von Problem 3.1 ermöglichen. Außerdem ist ein Programm zur Startwertberechnung enthalten. Die

Programme zum Gradientenverfahren und zur Startwertberechnung liegen in doppelter Ausführung vor. Zum einen betrachten sie die Viskositäten $\varepsilon_1, \dots, \varepsilon_g$ als unabhängig von einander, zum anderen werden die Viskositäten durch die typische Einschränkung

$$\varepsilon_{i+1} = \varepsilon_i, \quad \varepsilon_{i+2} = \beta \cdot \varepsilon_i \quad (3.3)$$

mit $i = 1, 4, \dots, (g-2)$ verbunden, wobei g durch 3 teilbar sein muß. Eine Erweiterung auf den Fall $\varepsilon_{i+1} = \alpha \cdot \varepsilon_i, \varepsilon_{i+2} = \beta \cdot \varepsilon_i$ bietet keine technischen Schwierigkeiten und ist in dem im Anhang vorgestellten optimierten Programmpaket enthalten. Dies entspricht der Situation, daß ein einzelner Dämpfer in die drei Raumrichtungen dämpft. Die Minimierung in diesem Programmpaket mittels des Gradienten- und Newtonverfahrens erfordern eine häufige Berechnung der aufwendig zu lösenden Ljapunovgleichung, deren Implementierung sich am Algorithmus von Bartels und Stewart [2] orientiert.

Die folgenden Abschnitte beschreiben die Theorie einer Optimierung dieser Programme. Ein Abschnitt beschäftigt sich mit der Wahl eines guten Startwertes für die Minimierungsverfahren, zwei andere Abschnitte mit der Optimierung der Gradientenberechnung im Gradientenverfahren bzw. mit der Optimierung der Berechnung der Hesseschen Matrix für das Newtonverfahren. Der jeweils angeschlossene Teil in jedem Abschnitt zeigt die Ergebnisse, die in Testreihen mit den optimierten und nichtoptimierten Programmen erzeugt worden sind. Die optimierten Programme sind in modularer Art (gemeinsam mit anderen Programmen) in einem neuen Paket (s. Anhang) zusammengefaßt worden. Dieses Programmpaket enthält neben den optimierten Minimierungsverfahren noch ein Programm, welches aus der Massenmatrix M , der Steifigkeitsmatrix K und der Anzahl und den Positionen der Dämpfer die Frequenzen $\omega_1, \dots, \omega_n$ und die Matrix V (vgl. Problem 3.1) bildet, woraus sich die $2n$ -dimensionale Phasenraummatrix A aus (3.2) zusammensetzt. Die optimierten Programme zur Startwertberechnung und zur Minimierung mittels des Gradientenverfahrens liegen auch hier in doppelter Ausführung vor. Analog zum oben genannten Programmpaket betrachten sie zum einen die Viskositäten $\varepsilon_1, \dots, \varepsilon_g$ als unabhängig voneinander und zum anderen berücksichtigen sie die Abhängigkeit, die durch (3.3) gegeben ist. In dem neuen Programmpaket wurde der implementierte Algorithmus zur Lösung der Ljapunovgleichung aus dem Programmpaket, das in [25] beschrieben wird, übernommen. Damit gibt das Paket eine geschlossene Programmkette her, die ausgehend von der gegebenen Schwingungskonstruktion (1.1) optimale Viskositäten berechnet. Sowohl die in [25] beschriebenen Programme als auch das neu entwickelte Programmpaket sind mit dem Programm-System Matlab realisiert worden (s. Anhang).

3.1 Die Wahl eines guten Startwertes

Um bei iterativen Minimierungsverfahren schnell ein akzeptables Ergebnis zu erzielen, ist die Wahl eines guten Startwertes sehr wichtig. Die Berechnung innerhalb eines Iterationsschrittes sind in Abhängigkeit von der Dimension n sehr aufwendig (s. später). Die Anzahl der Iterationen, die notwendig sind, um eine hinreichend gute Annäherung an das exakte Minimum zu erreichen, ist stark abhängig von einer guten Wahl des Startwertes. Der im folgenden gezeigte Weg zur Berechnung eines Startwertes hat sich in der Praxis als sehr erfolgreich erwiesen und ist im neuen Programmpaket implementiert.

3.1.1 Ansatz zur Bestimmung eines guten Startwertes

Es liegt nahe, den Startwert Σ_0 aus der Menge der nichtnegativen Viskositäten $\{\Sigma\}$ so zu wählen, daß $V\Sigma_0V^T$ möglichst nah an der Menge der lokalen Minima (s. (3.1))

$$\mathcal{M} = \left\{ C_{min} = 2 \begin{pmatrix} \hat{\Omega}_s & 0 \\ 0 & H \end{pmatrix}, H \in \mathcal{S} \text{ beliebig} \right\}$$

liegt. (Eine vergleichbare Überlegung lag der im folgenden Abschnitt 3.1.2 näher erläuterten Wahl des Startwertes in [25] zugrunde.) Zunächst betrachten wir den Fall, daß alle Viskositäten unabhängig voneinander sind und werden später sehen, daß die Startwertberechnung im Falle von abhängigen Viskositäten wie in (3.3) auf diesen Fall übertragbar ist.

Wir berechnen den minimalen Abstand zwischen der Menge \mathcal{M} , die durch die Matrizen $H \in \mathcal{S}$ gegeben ist, und der Menge $\{V\Sigma V^T\}$, die durch die nichtnegativen Viskositäten Σ gegeben ist, d.h. wir wollen das Minimierungsproblem

$$\begin{aligned} \left(\text{dist} \left(\mathcal{M}, V\Sigma_0V^T \right) \right)^2 &= \min_{\varepsilon_1, \dots, \varepsilon_g \geq 0} \left(\text{dist} \left(\mathcal{M}, V\Sigma V^T \right) \right)^2 \\ &= \min_{\varepsilon_1, \dots, \varepsilon_g \geq 0} \left(\inf_{H \in \mathcal{S}} \left\| 2 \begin{pmatrix} \hat{\Omega}_s & 0 \\ 0 & H \end{pmatrix} - V \Sigma V^T \right\|_F^2 \right) \end{aligned} \quad (3.4)$$

lösen.

Hierbei ist $\|\cdot\|_F$ die Frobenius- oder Euklidische Norm mit $\|B\|_F = (\sum_{i,j=1}^n |b_{ij}|^2)^{1/2}$ für $B \in M_n$. Variiert wird also über alle $H \in \mathcal{S}$ und alle positiv semidefiniten diagonalen Matrizen Σ .

Daher gilt mit $V =: \begin{pmatrix} \hat{V}_s \\ \check{V}_{n-s} \end{pmatrix}$ für den minimalen Abstand der Mengen \mathcal{M} und $\{V\Sigma V^T\}$:

$$\begin{aligned}
& \min_{\varepsilon_1, \dots, \varepsilon_g \geq 0} \left(\inf_{H \in \mathcal{S}} \left\| 2 \begin{pmatrix} \hat{\Omega}_s & 0 \\ 0 & H \end{pmatrix} - V \Sigma V^T \right\|_F^2 \right) \\
&= \min_{\varepsilon_1, \dots, \varepsilon_g \geq 0} \left(\inf_{H \in \mathcal{S}} \left\| \begin{pmatrix} 2\hat{\Omega}_s & 0 \\ 0 & 2H \end{pmatrix} - \begin{pmatrix} \hat{V}_s \Sigma \hat{V}_s^T & \hat{V}_s \Sigma \check{V}_{n-s}^T \\ \check{V}_{n-s} \Sigma \hat{V}_s^T & \check{V}_{n-s} \Sigma \check{V}_{n-s}^T \end{pmatrix} \right\|_F^2 \right) \\
&= \min_{\varepsilon_1, \dots, \varepsilon_g \geq 0} \left(\inf_{H \in \mathcal{S}} \left(\|2\hat{\Omega}_s - \hat{V}_s \Sigma \hat{V}_s^T\|_F^2 + \|\hat{V}_s \Sigma \check{V}_{n-s}^T\|_F^2 \right. \right. \\
&\quad \left. \left. + \|\check{V}_{n-s} \Sigma \hat{V}_s^T\|_F^2 + \|2H - \check{V}_{n-s} \Sigma \check{V}_{n-s}^T\|_F^2 \right) \right) \\
&= \min_{\varepsilon_1, \dots, \varepsilon_g \geq 0} \left(\|2\hat{\Omega}_s - \hat{V}_s \Sigma \hat{V}_s^T\|_F^2 + \|\hat{V}_s \Sigma \check{V}_{n-s}^T\|_F^2 \right. \\
&\quad \left. \|\check{V}_{n-s} \Sigma \hat{V}_s^T\|_F^2 + \inf_{H \in \mathcal{S}} \|2H - \check{V}_{n-s} \Sigma \check{V}_{n-s}^T\|_F^2 \right),
\end{aligned}$$

da nur der letzte Term von H abhängt.

Da \mathcal{S} alle positiv definiten Matrizen H enthält, ist $\inf_{H \in \mathcal{S}} \|2H - \check{V}_{n-s} \Sigma \check{V}_{n-s}^T\|_F^2 = 0$.
Damit ist die Minimierung über H bereits erfolgt und wir haben erhalten

$$\begin{aligned}
& \min_{\varepsilon_1, \dots, \varepsilon_g \geq 0} \left(\inf_{H \in \mathcal{S}} \left\| 2 \begin{pmatrix} \hat{\Omega}_s & 0 \\ 0 & H \end{pmatrix} - V \Sigma V^T \right\|_F^2 \right) \\
&= \min_{\varepsilon_1, \dots, \varepsilon_g \geq 0} \left(\|2\hat{\Omega}_s - \hat{V}_s \Sigma \hat{V}_s^T\|_F^2 + \|\hat{V}_s \Sigma \check{V}_{n-s}^T\|_F^2 + \|\check{V}_{n-s} \Sigma \hat{V}_s^T\|_F^2 \right) \\
&= \min_{\varepsilon_1, \dots, \varepsilon_g \geq 0} \left(\|2\hat{\Omega}_s - \hat{V}_s \Sigma \hat{V}_s^T\|_F^2 + 2\|\hat{V}_s \Sigma \check{V}_{n-s}^T\|_F^2 \right) \\
&= \min_{\varepsilon_1, \dots, \varepsilon_g \geq 0} \left\| \begin{pmatrix} 2\hat{\Omega}_s & 0 \\ 0 & 0 \end{pmatrix} - \begin{pmatrix} \hat{V}_s & 0 \\ 0 & \hat{V}_s \end{pmatrix} \begin{pmatrix} \Sigma & 0 \\ 0 & \Sigma \end{pmatrix} \begin{pmatrix} \hat{V}_s^T & 0 \\ 0 & \sqrt{2} \check{V}_{n-s}^T \end{pmatrix} \right\|_F^2 \\
&= \min_{\varepsilon_1, \dots, \varepsilon_g \geq 0} \left\| \begin{pmatrix} 2\hat{\Omega}_s & 0 \\ 0 & 0 \end{pmatrix} - \begin{pmatrix} \hat{V}_s \Sigma \hat{V}_s^T & 0 \\ 0 & \sqrt{2} \hat{V}_s \Sigma \check{V}_{n-s}^T \end{pmatrix} \right\|_F^2 =
\end{aligned}$$

$$\begin{aligned}
&= \min_{\varepsilon_1, \dots, \varepsilon_g \geq 0} \|X_0 - \varepsilon_1 A_1 - \varepsilon_2 A_2 - \dots - \varepsilon_g A_g\|_F^2 \\
&= \min_{\varepsilon_1, \dots, \varepsilon_g \geq 0} \left\| X_0 - [A_1 A_2 \dots A_g] \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_g \end{pmatrix} \right\|_F^2
\end{aligned}$$

mit

$$X_0 = \begin{pmatrix} 2 \hat{\Omega}_s & 0 \\ 0 & 0 \end{pmatrix}, \quad A_k = \begin{pmatrix} \hat{V}_s^k \hat{V}_s^{kT} & 0 \\ 0 & \sqrt{2} \hat{V}_s^k \check{V}_{n-s}^{kT} \end{pmatrix}, \quad (3.5)$$

wobei \hat{V}_s^k bzw. \check{V}_{n-s}^k die k -te Spalte von \hat{V}_s bzw. von \check{V}_{n-s} mit $1 \leq k \leq g$ ist. Zur Lösung des Problems

$$\min_{\varepsilon_1, \dots, \varepsilon_g \geq 0} \left\| X_0 - [A_1 A_2 \dots A_g] \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_g \end{pmatrix} \right\|_F^2, \quad (3.6)$$

das als sog. „Nonnegative Least Square“- Problem bezeichnet wird, wird in [17] ein Algorithmus vorgestellt, der zunächst entlang eines Randes des positiven Sektors $\varepsilon_1, \dots, \varepsilon_g \geq 0$ und dann „in diesen hinein“ optimal gewählte Teilprobleme mit der gewöhnlichen Least-Square-Methode behandelt; wenn dabei die Sektorengrenze übersprungen wird, läuft der Algorithmus (entlang der Verbindungsgerade von der alten zur neuen Lösung) zurück bis auf die Sektorengrenze und setzt von hier das Verfahren fort. Das Verfahren führt nach endlich vielen Schritten zur Lösung (siehe [17]) und ist in der Matlab-Routine `nnls` implementiert, die auch in dem neuen hier vorgestellten Programmpaket verwendet wird. Der Algorithmus und die daraus abgeleitete Routine sind ursprünglich für Vektoren formuliert; da die Frobeniusnorm einer Matrix aber gleich der Summe der Frobeniusnormen ihrer Spaltenvektoren (oder Zeilenvektoren) ist, ist die Umsetzung auf unser Problem ohne Schwierigkeiten durchzuführen.

Bemerkung 3.3

Wie zu Beginn dieses Abschnitts erwähnt, läßt sich das oben beschriebene Verfahren unter Berücksichtigung der in (3.3) gegebenen Abhängigkeiten durchführen. Dabei sind jeweils drei aufeinander folgende Spalten von V gemeinsam zu betrachten. Es ist $V_{(\beta)}$ aus V zu bilden, indem die jeweils $(3k)$ -te Spalte, $1 \leq k \leq \frac{g}{3}$ mit $\sqrt{\beta}$ aus (3.3) multipliziert wird. \hat{V}_s^k ist als $s \times 3$ - Matrix mit der $(3k-2)$ -ten, der $(3k-1)$ -ten und der $(3k)$ -ten Spalte und den ersten s Zeilen von $V_{(\beta)}$ zu bilden. \check{V}_{n-s}^k ist analog zu \hat{V}_s^k zu bilden. Mit diesen Matrizen können nun die Matrizen A_k , $k = 1, \dots, \frac{g}{3}$ wie in (3.5) gebildet werden. Die Lösung des Nonnegative Least Square – Problems liefert nun $\frac{g}{3}$ Viskositäten ε_{3k} , welche die unabhängigen Anteile der g Viskositäten zum Ausgangssystem repräsentieren.

Die Viskositäten $\tilde{\varepsilon}_1, \dots, \tilde{\varepsilon}_g \geq 0$, die sich aus diesem Minimierungsproblem ergeben, dienen als Startwert für die Minimierungsverfahren mit $\Sigma_0 = \text{diag}(\tilde{\varepsilon}_1, \dots, \tilde{\varepsilon}_g)$.

Es ist nicht a priori klar, ob das so gefundene Σ_0 im Inneren oder auf dem Rand von $\{\text{diag}(\varepsilon_1, \dots, \varepsilon_g) \mid \varepsilon_i \geq 0, 1 \leq i \leq g\}$ liegt, d.h. ob alle $\varepsilon_i > 0$ oder ein oder mehrere $\varepsilon_i = 0$ sind, ebenso nicht, ob $V\Sigma_0V^T$ zu einem stabilen A führt. Wenn $\hat{V}_s^k \neq 0$ und orthogonal sind für $1 \leq k \leq g$, gilt $\varepsilon_k > 0$ für alle k , da dann die A_k ein Orthogonalsystem bezüglich des Skalarproduktes $\langle A, B \rangle = \text{Tr} B^T A$ (das $\|\cdot\|_F$ induziert) sind und $\langle \begin{pmatrix} \hat{\Omega}_s & 0 \\ 0 & 0 \end{pmatrix}, A_k \rangle > 0$ gilt. Selbst dann aber könnte $V\Sigma_0V^T$ zu einem instabilen A führen, z.B. durch eine Nullzeile in V (vgl. Lemma 1.4). Jedenfalls ist bei linear unabhängigen V^k die Eigenschaft von A , stabil oder instabil zu sein, invariant unter $\Sigma > 0$, da dann Bild $V\Sigma V^T$ konstant bleibt. Bei Erreichen des Randes dagegen kann Instabilität eintreten, die im Inneren nicht gegeben war. Die Frage, wann garantiert werden kann, daß das vorgestellte Verfahren zu einem stabilen Startwert führt, ist ein offenes Problem. In allen Beispielen, die im folgenden Abschnitt berechnet wurden, lag jeweils Stabilität vor.

3.1.2 Erfahrungen aus der Praxis

Die praktischen Abschnitte in Kapitel 3 sollen verdeutlichen, wie stark sich eine Optimierung der rechenintensiven Vorgänge und/oder die Wahl eines guten Startwertes auf die Verfahren auswirken.

Zur Veranschaulichung des Ergebnisses aus dem vorangegangenen Abschnitt wurde das mit [25] realisierte Verfahren zur Startwertberechnung (im folgenden als Verfahren (A) bezeichnet) dem in Abschnitt 3.1.1 beschriebenen Verfahren zur Startwertberechnung (mit Verfahren (B) bezeichnet) in einem Test gegenübergestellt. Dabei wurden zum einen die Viskositäten als unabhängig von einander betrachtet, zum anderen wurde die in (3.3) beschriebene Abhängigkeit der Viskositäten untersucht. Im Verfahren (A) wird – im Gegensatz zu (3.4) – Σ'_0 mit

$$\|2\Omega - V \Sigma'_0 V^T\|_F = \min_{\Sigma} \|2\Omega - V \Sigma V^T\|_F \quad (3.7)$$

als Startwert berechnet. Im Punkt Σ'_0 wird damit der minimale Abstand zwischen der Menge $\{V\Sigma V^T\}$ und der kritischen Dämpfungsmatrix $C_{krit} = 2\Omega$ (s. Satz 2.1 und Bemerkung 2.4) angenommen. Der Startwert Σ'_0 des Verfahrens (A) entspricht also dem Startwert des Verfahrens (B), falls $n = s$ gilt, das betrachtete Frequenzspektrum also nicht eingeschränkt ist, somit $\text{Tr}(XZ)$ mit $Z = I$ minimiert wird. Bei der durch (3.3) gegebenen Abhängigkeit der Viskositäten wird in Verfahren (A) zunächst $\Sigma'_0 = \text{diag}\{\varepsilon_1, \dots, \varepsilon_g\}$ mittels (3.7) ermittelt und dann je drei aufeinanderfolgende Viskositäten des tatsächlichen Startwertes $\text{diag}\{\tilde{\varepsilon}_1, \dots, \tilde{\varepsilon}_g\}$ durch

$$\tilde{\varepsilon}_{(3k-2)} = \tilde{\varepsilon}_{(3k-1)} = \frac{\tilde{\varepsilon}_{(3k)}}{\beta} := \frac{\varepsilon_{(3k-2)} + \varepsilon_{(3k-1)} + \varepsilon_{(3k)}}{2 + \beta}$$

mit $k = 1, \dots, \frac{g}{3}$, β aus (3.3) definiert. Da sich die Ergebnisse der Tests nicht wesentlich unterscheiden, wird im folgenden nur auf den Test mit unabhängigen Viskositäten eingegangen.

Im Test wurden verschiedene Dimensionen n zwischen 5 und 70, sowie für jedes n unterschiedliche Spurlängen s betrachtet. Die Anzahl der Dämpfer lag bei $g = 4$. In allen Fällen wurden zu zufällig ermittelten positiv definiten Matrizen M und K , welche die Massenmatrix und Steifigkeitsmatrix der gegebenen Differentialgleichung (1.1) repräsentieren, der Startwert aus beiden Verfahren ermittelt. Anschließend wurde ausgehend von diesen Startwerten mit dem Gradientenverfahren ein lokales Minimum bis zu einer festen vorgegebenen Genauigkeit berechnet. Die jeweiligen Anzahlen der vom Gradientenverfahren benötigten Iterationen wurden letztendlich als Vergleichskriterium für die unterschiedlichen Startviskositäten Σ_0 und Σ'_0 herangezogen. Um große Ausreißer aus den Testergebnissen zu neutralisieren, wurde jeder Test für festes n und feste Spurenlänge s mehrfach durchgeführt und der Durchschnittswert über alle Iterationen betrachtet. Die Ergebnisse des Tests sind in der Abbildung 3.1 dargestellt. Die Abbildung 3.1 zeigt den Faktor, um den das Verfahren

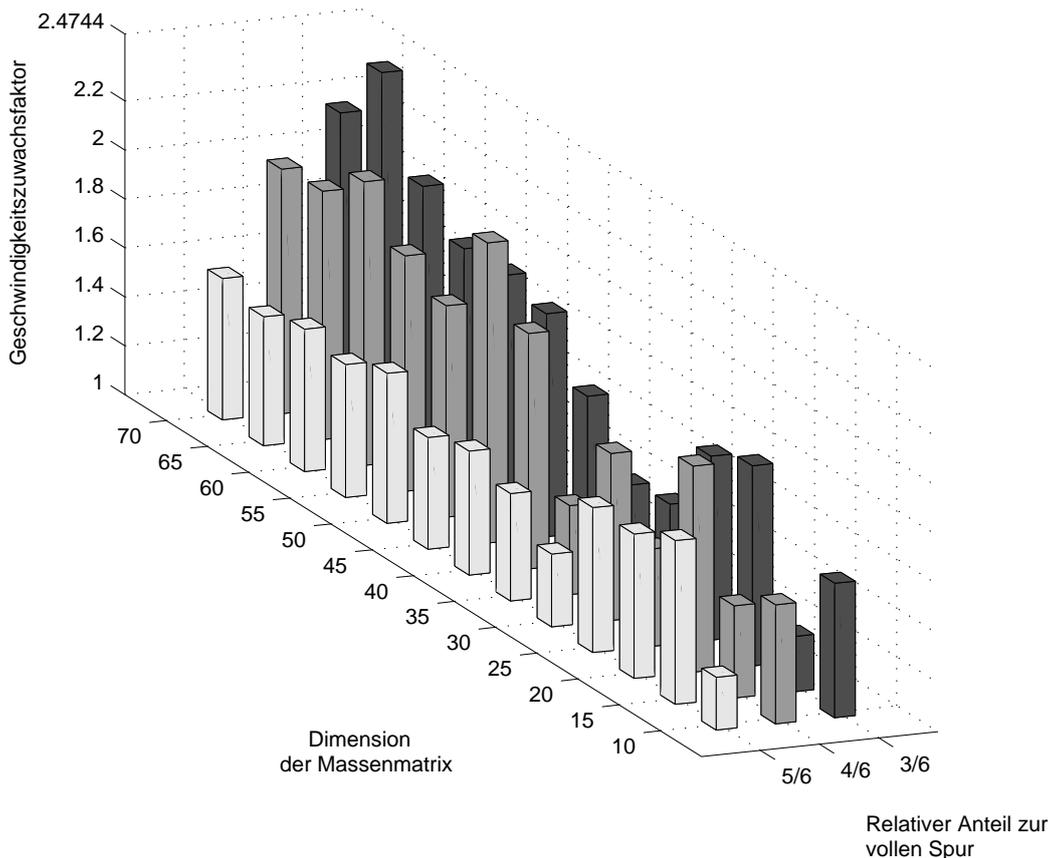


Abb. 3.1: Vergleich der Startwertberechnung

(B) weniger Iterationen als das Verfahren (A) im einzelnen Testfall im Durchschnitt benötigt. Auf der z-Achse ist die Anzahl der Iterationen des Verfahrens (A) dividiert durch die Anzahl der Iterationen des Verfahrens (B), auf der y-Achse sind die verschiedenen betrachteten Dimensionen n von 10 bis 70 in Fünferschritten abgetragen. Die drei verschiedenfarbigen Kolonnen, abgetragen auf der x-Achse, stellen die Berechnung unter Berücksichtigung einer jeweils unterschiedlichen relativen Spurlänge ca. $\frac{s}{n}$ dar. Die Anzahlen der Iterationen bewegten sich mit Startwerten aus Verfahren (B) zwischen 8 und 45 und für Startwerte aus Verfahren (A) zwischen 14 und 69. Es ist leicht zu sehen, daß alle Faktoren größer als 1 sind, d.h. in allen Fällen das Gradientenverfahren mit dem Startwert aus Verfahren (B) eher zum Ziel führte. Die Faktoren werden umso größer, je höher die Dimension der Massenmatrix ist. Für eine kleinere Länge der Spur scheint bei großer Dimension der Faktor auch noch einmal zu steigen. Die volle Länge der Spur ist in dieser Abbildung nicht als eigene Kolonne abgetragen, da in diesem Fall der Faktor konstant 1 ist.

Die klassischen Verfahren wie Gradienten- (Methode des steilsten Abstiegs) und Newtonverfahren zur Minimierung von $F(\varepsilon) = F(\varepsilon_1, \dots, \varepsilon_g) = \text{Tr}(XZ_s)$ erfordern die Berechnung des Gradienten bzw. beim Newtonverfahren zusätzlich die Berechnung der Hesseschen Matrix für jeden Iterationsschritt. Die Größe des Gradienten und der Hesseschen Matrix ist dabei von der Anzahl g der Dämpfer abhängig. Es gilt

$$\text{grad}(F(\varepsilon)) = \left[\frac{\partial \text{Tr}(XZ_s)}{\partial \varepsilon_i} \right] \in M_{1,g}, \quad \text{Hess}(F(\varepsilon)) = \left[\frac{\partial^2 (\text{Tr}(XZ_s))}{\partial \varepsilon_i \partial \varepsilon_j} \right] \in M_g,$$

wobei zur Berechnung einer jeden Komponente die Lösung einer Ljapunovgleichung erforderlich ist. Dies ist für eine hohe Dimension n und mehrere Dämpfer sehr aufwendig. In [29] wird gezeigt, wie der Gradient beim Gradientenverfahren – unabhängig von der Anzahl der Dämpfer – mit der Lösung höchstens einer zusätzlichen Ljapunovgleichung berechnet werden kann. Außerdem wird gezeigt, wie beim Newtonverfahren die Hessesche Matrix mit der Lösung von lediglich g zusätzlichen Ljapunovgleichungen bestimmt werden kann, wodurch gleichzeitig der Gradient bestimmt ist. Mit diesen Ergebnissen läßt sich der Rechenaufwand je Iterationsschritt – abhängig von der Dämpferzahl – erheblich verringern. Dies wollen wir im folgenden verdeutlichen.

3.2 Optimierung beim Gradientenverfahren

Das Gradientenverfahren (Methode des steilsten Abstiegs) ermittelt zur Minimierung der differenzierbaren Funktion $F(\varepsilon) = F(\varepsilon_1, \dots, \varepsilon_g) = (\text{Tr}(XZ_s))(\varepsilon)$ in jedem

Iterationsschritt einen neuen Viskositätenvektor nach folgender Vorschrift:

$$\varepsilon^{(i+1)} = \varepsilon^{(i)} - \lambda \cdot \text{grad}(F(\varepsilon^{(i)})).$$

Dabei bezeichnet $\varepsilon^{(i)}$ den Viskositätenvektor ε nach dem i -ten Iterationsschritt. $-\text{grad}(F(\varepsilon^{(i)}))$ gibt die Richtung des steilsten Gefälles der Funktion im Punkt $\varepsilon^{(i)}$ an. Für die Berechnung des Gradienten sind (s.o.) g verschiedene Ljapunovgleichungen zu lösen. Insgesamt erfordert jeder Iterationsschritt des Gradientenverfahrens die Lösung von $g + 1$ verschiedenen Ljapunovgleichungen (eine zur Berechnung von X und g zur Berechnung des Gradienten). Unter Berücksichtigung der in (3.3) genannten Abhängigkeit sind entsprechend $\frac{g}{3}$ verschiedene Ljapunovgleichungen für die Berechnung des Gradienten zu lösen. Der folgenden Abschnitt zeigt (s. [29]), wie der Gradient unabhängig von der Anzahl der Dämpfer berechnet werden kann und damit die Anzahl der zu lösenden Ljapunovgleichungen pro Iterationsschritt reduziert wird.

3.2.1 Berechnung des Gradienten

Wir wollen den Gradienten $\text{grad}(\text{Tr}(XZ_s))$ berechnen. Im folgenden werden wir statt $\frac{\partial}{\partial \varepsilon_i}$ die Schreibweise ∂_i wählen. Zunächst betrachten wir die Gradientenberechnung unter der Voraussetzung, daß die Dämpfer und somit die Viskositäten unabhängig voneinander sind (s. [29]). Es gilt $\partial_i \text{Tr}(XZ_s) = \text{Tr}(\partial_i X \cdot Z_s)$. Leiten wir die Ljapunovgleichung $A^T X + XA = -I$ einmal partiell ab, so erhalten wir

$$\partial_i A^T X + A^T \partial_i X + \partial_i X A + X \partial_i A = 0,$$

d.h. wir bekommen eine neue Ljapunovgleichung der Form

$$A^T \partial_i X + \partial_i X A = -X \partial_i A - \partial_i A^T X. \quad (3.8)$$

Mit

$$\partial_i A = \partial_i A^T = \begin{pmatrix} 0 & 0 \\ 0 & -V^i V^{iT} \end{pmatrix} =: -\tilde{V}^i \tilde{V}^{iT}$$

erhalten wir aus (3.8)

$$A^T \partial_i X + \partial_i X A = X \tilde{V}^i \tilde{V}^{iT} + \tilde{V}^i \tilde{V}^{iT} X. \quad (3.9)$$

Damit ergibt sich für die i -te Komponente $\text{Tr}(\partial_i X Z_s)$ des Gradienten

$$\begin{aligned} \text{Tr}(\partial_i X Z_s) &= -\text{Tr} \left(\int_0^\infty e^{A^T t} (X \tilde{V}^i \tilde{V}^{iT} + \tilde{V}^i \tilde{V}^{iT} X) e^{At} dt Z_s \right) \\ &= -\text{Tr} \left(\int_0^\infty \tilde{V}^{iT} e^{At} Z_s e^{A^T t} X \tilde{V}^i dt \right) - \text{Tr} \left(\int_0^\infty \tilde{V}^{iT} X e^{At} Z_s e^{A^T t} \tilde{V}^i dt \right) \\ &= -\left(\tilde{V}^{iT} \int_0^\infty e^{At} Z_s e^{A^T t} dt X \tilde{V}^i \right) - \left(\tilde{V}^{iT} X \int_0^\infty e^{At} Z_s e^{A^T t} dt \tilde{V}^i \right) \\ &= -\tilde{V}^{iT} Y X \tilde{V}^i - \tilde{V}^{iT} X Y \tilde{V}^i = -2 \tilde{V}^{iT} Y X \tilde{V}^i, \end{aligned}$$

wobei Y die eindeutige Lösung der Ljapunovgleichung $AY + YA^T = -Z_s$ ist. Die Berechnung des Gradienten erfordert also neben der schon durch $\text{Tr}(XZ_s)$ bekannten Lösung X nur die Lösung Y einer zusätzlichen Ljapunovgleichung. Damit werden im Gradientenverfahren in jedem Iterationsschritt genau zwei verschiedene Ljapunovgleichungen gelöst, egal wieviel Dämpfer das System enthält. Gilt sogar $s = n$, d.h. $Z_s = I$, dann erfolgt die Berechnung des Gradienten ganz ohne Lösung einer zusätzlichen Ljapunovgleichung, denn dann gilt $Y = JXJ$ mit $J := \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}$.

(Es gilt $JA^TJ = A$, d.h. aus $A^TX + XA = -I$ folgt $JAJX + XJA^TJ = -I$ also $JJAJXJ + JXJA^TJJ = -JJ$, d.h. $AJXJ + JXJA^T = -I$.) Pro Iterationsschritt wird in diesem Fall also nur eine Ljapunovgleichung gelöst.

Unter Berücksichtigung der durch (3.3) gegebenen Abhängigkeit der Viskositäten verläuft die Gradientenberechnung auf ähnliche Weise. Die $\frac{g}{3}$ unabhängigen Komponenten des Gradienten werden durch $\text{Tr}(\partial_k XZ_s) = -2 \text{Tr}(\tilde{V}_{(\beta)}^{kT} Y X \tilde{V}_{(\beta)}^k)$ mit $k = 1, \dots, \frac{g}{3}$ gebildet, wobei $\tilde{V}_{(\beta)}^k$ die $(2n \times 3)$ -Matrix $\tilde{V}_{(\beta)}^k = \begin{pmatrix} 0 \\ V_{(\beta)}^k \end{pmatrix}$ ist.

$V_{(\beta)}^k$ ist dabei die $n \times 3$ -Matrix, die aus der $(3k-2)$ -ten, der $(3k-1)$ -ten und der mit $\sqrt{\beta}$ multiplizierten $(3k)$ -ten Spalte von V besteht, wobei β durch (3.3) gegeben ist. Die so erhaltenen $\frac{g}{3}$ Komponenten des Gradienten repräsentieren die unabhängigen Anteile der g Komponenten des Gradienten des Ausgangssystems.

3.2.2 Erfahrungen aus der Praxis

Um zu verdeutlichen, wie stark sich die im vorangegangenen Abschnitt beschriebenen Berechnungen des Gradienten $\text{grad}(\text{Tr}(XZ_s))$ im Aufwand der numerischen Minimierung bemerkbar macht, ist das in [25] realisierte Gradientenverfahren ($P_{Grad I}$), welches die Berechnung des Gradienten auf nicht optimierte Weise enthält (d.h. in jedem Iterationsschritt werden $g + 1$ Ljapunovgleichungen gelöst), dem Programm ($P_{Grad II}$) aus dem neuen Programmpaket gegenüber gestellt. Es berechnet den Gradienten auf die im Abschnitt 3.2.1 beschriebenen Weise, d.h. es kommt mit der Lösung von zwei Ljapunovgleichungen, im Fall der vollen Spur sogar mit der Lösung nur einer Ljapunovgleichung pro Iterationsschritt aus. (Vorgestellt werden hier nur die Ergebnisse im Falle unabhängiger Viskositäten; bei einer Abhängigkeit nach (3.3) sind in ($P_{Grad I}$) noch $1 + \frac{g}{3}$ Ljapunovgleichungen zu lösen, und die Unterschiede im Aufwand sind entsprechend geringer.) In den Tests wurde die Dimension n und die Anzahl der Dämpfer g variiert und in jedem der beiden Programme die Anzahl der Fließkommaoperationen (Flops) einer Iteration in MFlops (Megaflops) gemessen. Zwei Tests betrachten die Dimensionen n von 1 bis 20 bei allen möglichen Anzahlen der Dämpfer $g \leq n$ jeweils unter Berücksichtigung der Minimierung der vollen Spur

(Test (G1a)) oder nur eines Teils der Spur (Test (G1b)). Genau genommen sind bei Test (G1b) (und auch bei allen anderen Testreihen unter Berücksichtigung eines Teils der Spur) zwei Testreihen durchgeführt worden: eine mit $s \approx \frac{n}{2}$ und eine mit $s \approx \frac{2n}{3}$. Wie erwartet hat sich jedoch herausgestellt, daß die Ergebnisse für jeweils zwei solcher Tests nahezu identisch sind. Im folgenden werden alle Testreihen unter gleichen Rahmenbedingungen dieser Art, aber verschieden langen Anteilen der vollen Spur, als ein Test unter Berücksichtigung eines Teils der Spur bezeichnet. In zwei weiteren Tests wurde in Anlehnung an die in der Praxis vorkommenden Daten die Dimension n von 5 bis 100 in Fünferschritten deutlich größer als die Anzahl der Dämpfer g von 1 bis 10 gewählt. Auch hier ist je ein Test unter Berücksichtigung der Minimierung der vollen Spur (Test (G2a)) und einer zur Minimierung unter Berücksichtigung eines Teils der Spur (Test (G2b)) durchgeführt worden. Tabelle 3.1 gibt einen Überblick über die durchgeführten Tests.

Test	Minimierung unter Berücksichtigung der Spur	Variation von Dimension und Dämpferzahl
(G1a)	voll	$n = 1, 2, \dots, 20$ $g = 1, 2, \dots, n$
(G1b)	teilweise	$n = 1, 2, \dots, 20$ $g = 1, 2, \dots, n$
(G2a)	voll	$n = 5, 10, \dots, 100$ $g = 1, 2, \dots, \min(n, 10)$
(G2b)	teilweise	$n = 5, 10, \dots, 100$ $g = 1, 2, \dots, \min(n, 10)$

Tabelle 3.1: Durchgeführte Tests mit dem Gradientenverfahren

Abbildung 3.2 zeigt die Einzelergebnisse von Test (G1a), d.h. Minimierung unter Berücksichtigung der vollen Spur. Das obere Achsensystem stellt den Aufwand in MFlops je Iteration des Programms ($P_{Grad I}$) in Abhängigkeit der Dimension und der Anzahl der Dämpfer, das untere Achsensystem den Aufwand des Programms ($P_{Grad II}$), dar. Zur Verdeutlichung der Krümmung und der Oberflächenstruktur wurde jeweils ein Lichteffect eingesetzt und die z-Achse in Abhängigkeit des höchsten vorkommenden Ergebnisses skaliert. Die Graphiken lassen sofort erkennen, daß der Aufwand beim Programm ($P_{Grad I}$) in Abhängigkeit von der Anzahl der Dämpfer nicht unwesentlich linear ansteigt, während der Aufwand für das Programm ($P_{Grad II}$) bei fester Dimension nahezu konstant, also unabhängig von der Anzahl der Dämpfer ist. Die Eckdaten des höchsten Aufwands, der hier für Dimension 20 bei 20 Dämpfern auftritt, lassen mit ca. 3,2 MFlops je Iteration bei Berechnung mit Programm ($P_{Grad II}$) und ca. 43 MFlops je Iteration bei Berechnung mit Programm

($P_{Grad I}$) die Effizienzsteigerung bei Nutzung der Berechnung des Gradienten nach Abschnitt 3.2.1 erkennen.

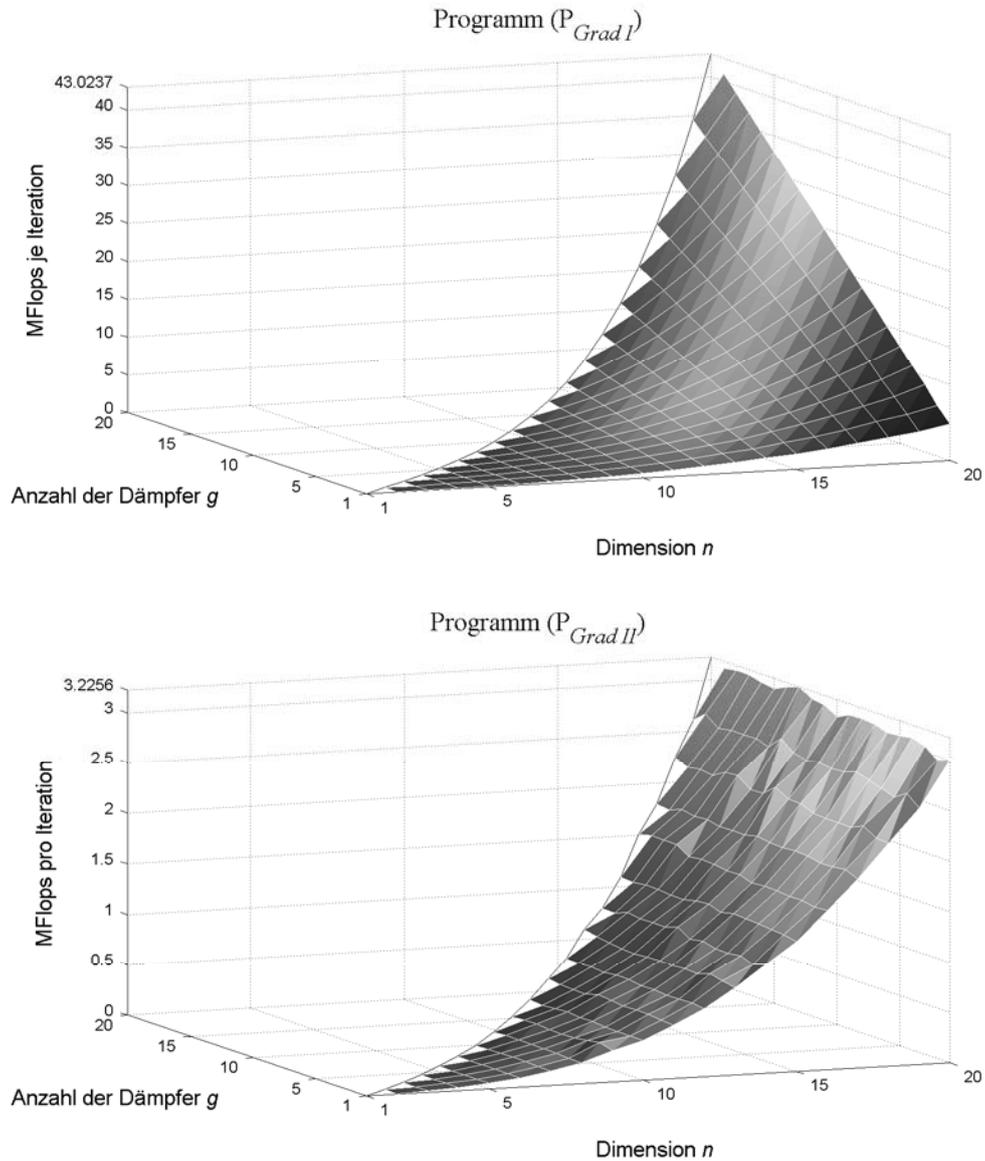


Abb. 3.2: Test (G1a), Vergleich von ($P_{Grad I}$) und ($P_{Grad II}$) bei voller Spur

Abbildung 3.3 auf der nächsten Seite zeigt den Aufwand je Iteration der Programme ($P_{Grad I}$) und ($P_{Grad II}$) bei Minimierung unter Berücksichtigung eines Teils der Spur für Dimensionen zwischen 5 und 100. Auch hier ist im oberen Koordinatensystem das Ergebnis von Programm ($P_{Grad I}$), d.h. mit der nicht optimierten Berechnung des Gradienten und im unteren Koordinatensystem das Ergebnis von Programm

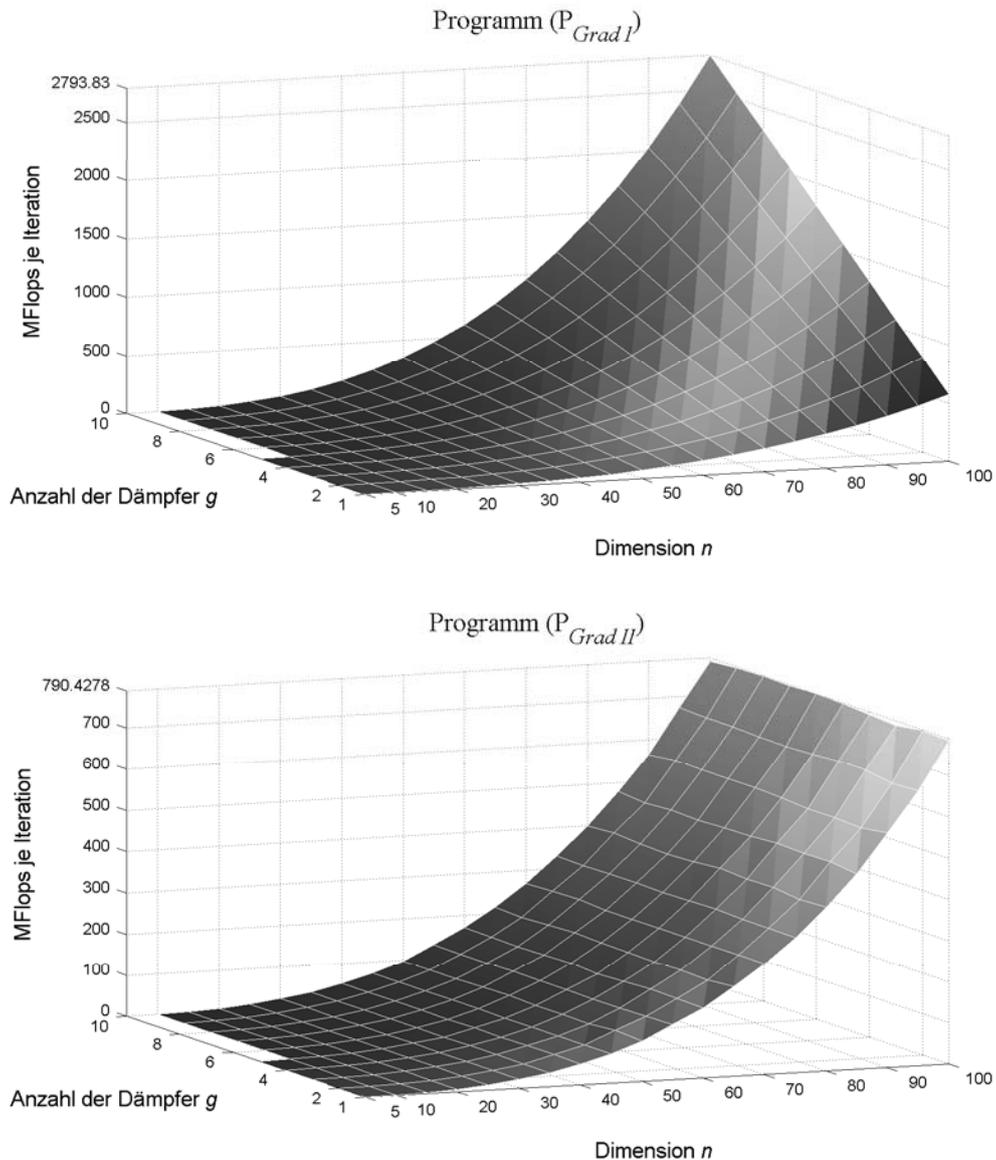


Abb. 3.3: Test (G2b), Vergleich von ($P_{Grad I}$) und ($P_{Grad II}$) bei nichtvoller Spur

($P_{Grad II}$) dargestellt. Hier ist, wie in der Praxis üblich, die Anzahl der Dämpfer zwischen 1 und 10 erheblich kleiner als die Dimensionen. Die Ergebnisse von ($P_{Grad I}$) lassen erkennen, daß sich mit höherer Dimension die Hinzunahme eines weiteren Dämpfers zwar weiterhin nur linear, jedoch mit einem grösseren Faktor im Aufwand bemerkbar macht. Der Aufwand je Iteration bei Programm ($P_{Grad II}$) bleibt, wie erwartet, auch bei großen Dimensionen n , von der Anzahl der Dämpfer g unabhängig, nahezu konstant. Der höchste Aufwand fällt für Dimension $n = 100$ bei höchster

Dämpferanzahl $g = 10$ an. Er liegt für Programm $(P_{Grad I})$ bei ca. 2794 und für Programm $(P_{Grad II})$ bei ca. 790 MFlops je Iteration.

Abbildung (3.4) zeigt die im Test (G2a) berechneten Aufwände je Iteration beider Programme in einem gemeinsamen Achsensystem. Zur besseren Übersicht wurden die x-Achse und die y-Achse vertauscht und die Perspektive geändert.

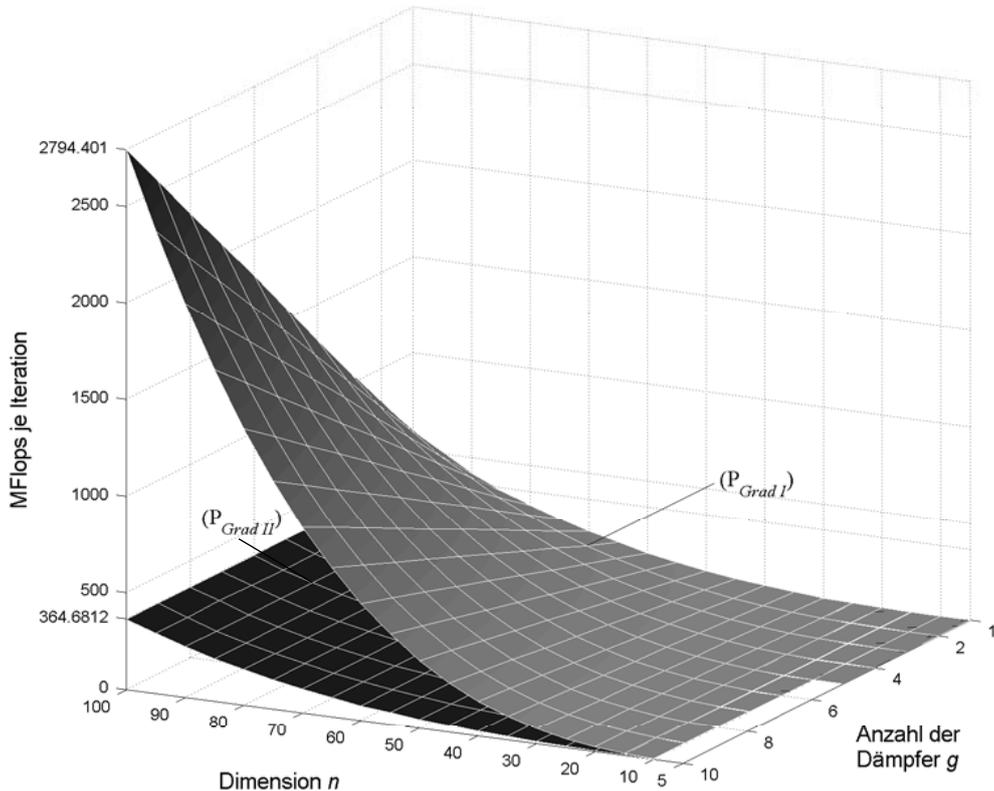


Abb. 3.4: Test (G2a), Vergleich von $(P_{Grad I})$ und $(P_{Grad II})$ bei voller Spur

Wie im Test (G2b) sind die Aufwände für die Dimension n zwischen 5 und 100 in Fünferschritten und die Anzahlen der Dämpfer g zwischen 1 und 10 berechnet worden. Im Gegensatz zu Test (G2b) ist hier jedoch unter Berücksichtigung der vollen Spur minimiert worden. Wie erwartet halbiert sich im Test (G2a) noch einmal der Aufwand für Programm $(P_{Grad II})$ gegenüber Test (G2b), da bei Minimierung unter Berücksichtigung der vollen Spur nur noch eine Ljapunovgleichung pro Iteration zu lösen ist. Die Ergebnisfläche des Programms $(P_{Grad I})$ mit der Berechnung des Gradienten ohne Optimierung liegt vollständig oberhalb der Ergebnisfläche von Programm $(P_{Grad II})$, wobei sich der Abstand, d.h. der Mehraufwand von Programm $(P_{Grad I})$ gegenüber Programm $(P_{Grad II})$, mit steigender Anzahl der Dämpfer linear und mit steigender Anzahl der Dimension n etwa mit n^3 vergrößert.

3.3 Optimierung beim Newtonverfahren

Das Newtonverfahren ermittelt zur Minimierung der zweimal stetig differenzierbaren Funktion $F(\varepsilon) = F(\varepsilon_1, \dots, \varepsilon_g) = (\text{Tr}(XZ_s))(\varepsilon)$ in jedem Iterationsschritt einen neuen Viskositätenvektor nach folgender Vorschrift (vgl. z.B. [23]):

$$\varepsilon^{(i+1)} = \varepsilon^{(i)} - \lambda \cdot (\text{Hess}(F(\varepsilon^{(i)})))^{-1} \cdot \text{grad}(F(\varepsilon^{(i)})).$$

Dabei bezeichnet $\varepsilon^{(i)}$ den Viskositätenvektor ε nach dem i -ten Iterationsschritt. Für die Berechnung des Gradienten sind (s.o.) g verschiedene Ljapunovgleichungen, für die Hessesche Matrix (auf Grund ihrer Symmetrie) $\frac{g+g^2}{2}$ verschiedene Ljapunovgleichungen zu lösen. Insgesamt sind damit in jedem Iterationsschritt beim Newtonverfahren im zu [25] gehörenden Programm $1 + \frac{3g+g^2}{2}$ (eine für X , g für den Gradienten und $(g+g^2)/2$ für die Hessesche Matrix) verschiedene Ljapunovgleichungen zu lösen. Der nächste Abschnitt zeigt (s. [29]), wie die Hessesche Matrix unabhängig von der Anzahl der Dämpfer berechnet werden kann, wenn der Gradient bekannt ist.

3.3.1 Berechnung der Hesseschen Matrix

Wir wollen die Hessesche Matrix unabhängig von der Anzahl der vorhandenen Dämpfer berechnen (s. [29]). Wir wählen wieder die Schreibweise ∂_{ij}^2 an Stelle von $\frac{\partial^2}{\partial \varepsilon_i \partial \varepsilon_j}$. Leiten wir (3.8) noch einmal partiell ab, so erhalten wir die Gleichung

$$\partial_j A^T \partial_i X + A^T \partial_{ij}^2 X + \partial_{ij}^2 X A + \partial_i X \partial_j A = -\partial_j X \partial_i A - \partial_i A^T \partial_j X.$$

Damit ergibt sich eine neue Ljapunovgleichung der Form

$$\begin{aligned} A^T \partial_{ij}^2 X + \partial_{ij}^2 X A &= -\partial_j A^T \partial_i X - \partial_i X \partial_j A - \partial_j X \partial_i A - \partial_i A^T \partial_j X \\ &= \tilde{V}^j \tilde{V}^{jT} \partial_i X + \partial_i X \tilde{V}^j \tilde{V}^{jT} + \partial_j X \tilde{V}^i \tilde{V}^{iT} + \tilde{V}^i \tilde{V}^{iT} \partial_j X. \end{aligned}$$

Nun können wir die (i, j) -te Komponente der Hesseschen Matrix berechnen:

$$\begin{aligned} &\text{Tr}(\partial_{ij}^2 X \cdot Z_s) \\ &= -\text{Tr} \left(\int_0^\infty e^{A^T t} \left(\tilde{V}^j \tilde{V}^{jT} \partial_i X + \partial_i X \tilde{V}^j \tilde{V}^{jT} + \partial_j X \tilde{V}^i \tilde{V}^{iT} + \tilde{V}^i \tilde{V}^{iT} \partial_j X \right) e^{At} dt \cdot Z_s \right) \\ &= -\text{Tr} \int_0^\infty \tilde{V}^{jT} \partial_i X e^{AT} Z_s e^{AT} \tilde{V}^j dt - \text{Tr} \int_0^\infty \tilde{V}^{jT} e^{AT} Z_s e^{AT} \partial_i X \tilde{V}^j dt \\ &\quad - \text{Tr} \int_0^\infty \tilde{V}^{iT} e^{AT} Z_s e^{AT} \partial_j X \tilde{V}^i dt - \text{Tr} \int_0^\infty \tilde{V}^{iT} \partial_j X e^{AT} Z_s e^{AT} \tilde{V}^i dt = \end{aligned}$$

$$\begin{aligned}
&= -\tilde{V}^{jT} \partial_i X \int_0^\infty e^{AT} Z_s e^{A^T t} dt \tilde{V}^j - \tilde{V}^{jT} \int_0^\infty e^{AT} Z_s e^{A^T t} dt \partial_i X \tilde{V}^j \\
&\quad - \tilde{V}^{iT} \int_0^\infty e^{AT} Z_s e^{A^T t} dt \partial_j X \tilde{V}^i - \tilde{V}^{iT} \partial_j X \int_0^\infty e^{AT} Z_s e^{A^T t} dt \tilde{V}^i \\
&= -2 \tilde{V}^{jT} \partial_i X Y \tilde{V}^j - 2 \tilde{V}^{iT} Y \partial_j X \tilde{V}^i,
\end{aligned}$$

wobei Y wiederum die eindeutige Lösung der Ljapunovgleichung $AY + YA^T = -Z_s$ ist. Gilt $s = n$, d.h. $Z_s = I$, dann ist $Y = JXJ$ und für die Hessesche Matrix ergibt sich

$$\text{Tr}(\partial_{ij} X \cdot Z_s) = -2 \tilde{V}^{jT} \partial_i X JXJ \tilde{V}^j - 2 \tilde{V}^{iT} JXJ \partial_j X \tilde{V}^i.$$

Die Berechnung der Hesseschen Matrix erfordert also für den Fall $Z_s \neq I$ die Lösung von g verschiedenen Ljapunovgleichungen (3.9) für die Berechnung des Vektors $(\partial_i X)$, $1 \leq i \leq g$, sowie einer Ljapunovgleichung für die Berechnung von Y . Gilt $Z_s = I$, dann muß Y nicht mittels einer Ljapunovgleichung berechnet werden. Damit müssen insgesamt pro Iterationsschritt beim Newtonverfahren $g+2$ (eine für X , eine für Y , g für $(\partial_i X)$) Ljapunovgleichungen für $Z_s \neq I$ und $g+1$ Ljapunovgleichungen für $Z_s = I$ gelöst werden.

3.3.2 Erfahrungen aus der Praxis

Nutzt man bei der numerischen Berechnung der Hesseschen Matrix im programmierten Newtonverfahren die Erkenntnisse aus Abschnitt 3.3.1, so kann man im Vergleich zur Berechnung der Hesseschen Matrix nach [25] einen erheblichen Geschwindigkeitsgewinn erzielen. Dies ist in einer zu Abschnitt 3.2.2 analogen Testreihe mit dem Vergleich der Programme ($P_{Newton I}$) und ($P_{Newton II}$) bestätigt worden, wobei nur der Fall unabhängiger Viskositäten betrachtet wurde. Dabei bezeichnet

Test	Minimierung unter Berücksichtigung der Spur	Variation von Dimension und Dämpferzahl
(N1a)	voll	$n = 1, 2, \dots, 20$ $g = 1, 2, \dots, n$
(N1b)	teilweise	$n = 1, 2, \dots, 20$ $g = 1, 2, \dots, n$
(N2a)	voll	$n = 5, 10, \dots, 75$ $g = 1, 2, \dots, \min(n, 10)$
(N2b)	teilweise	$n = 5, 10, \dots, 75$ $g = 1, 2, \dots, \min(n, 10)$

Tabelle 3.2: Durchgeführte Tests mit dem Newtonverfahren

Programm ($P_{Newton I}$) die Version mit der nicht optimierten Berechnung der Hesseschen Matrix, d.h. mit $1 + \frac{3g+g^2}{2}$ Lösungen von Ljapunovgleichungen je Iteration. Das Programm ($P_{Newton II}$) enthält die Berechnung der Hesseschen Matrix in der nach Abschnitt 3.3.1 optimierten Variante, d.h. es kommt mit der Lösung von $g + 2$ Ljapunovgleichungen je Iteration bei Minimierung unter Berücksichtigung eines Teils der Spur, oder nur $g + 1$ Lösungen von Ljapunovgleichungen bei Minimierung unter Berücksichtigung der vollen Spur aus. Die Rahmenbedingungen beim Vergleich der beiden Programme ($P_{Newton I}$) und ($P_{Newton II}$) waren die gleichen wie in Abschnitt 3.2.2. Es wurde also vier Tests (N1a), (N1b), (N2a) und (N2b) bei Variation der Anzahl der Dämpfer und der Dimension gemäß Tabelle 3.2 durchgeführt.

Abbildung 3.5 enthält die Ergebnisse von Test (N2a). Sie zeigt die MFlops je Iteration des Programms ($P_{Newton I}$) und des Programms ($P_{Newton II}$) in einem Achsen-system.

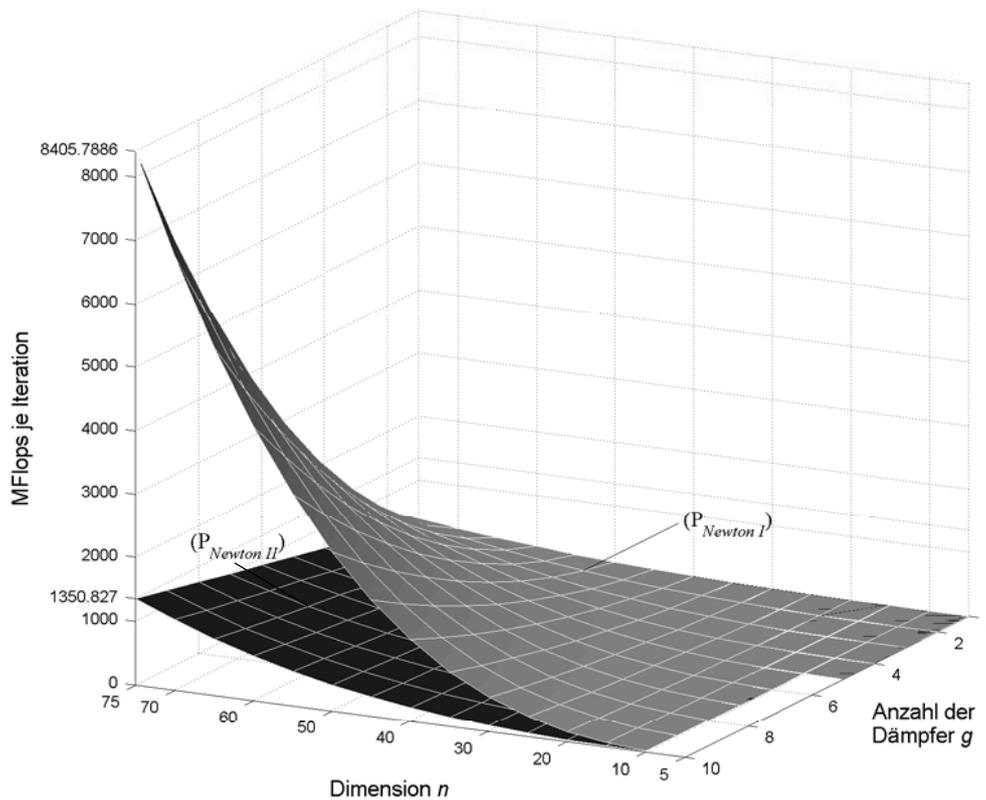


Abb. 3.5: Test (N2a), Vergleich von ($P_{Newton I}$) und ($P_{Newton II}$) bei voller Spur

Die hellere obere Ergebnisfläche repräsentiert die Werte des Programms ($P_{Newton I}$), d.h. des Programms mit der Berechnung der Hesseschen Matrix nach nicht optimierter Art. Der Höchstwert bei Dimension 75 und zehn Dämpfern von über 8,4 Milliarden Flops für eine Iteration zeigt, warum sich hier eine Optimierung lohnt. Das Ergebnis des Programms ($P_{Newton II}$) mit der optimierten Berechnung der Hesseschen Matrix benötigt unter den gleichen Voraussetzungen nur noch ca. 1,35 Milliarden Flops. Im Gegensatz zu den Abbildungen beim Test der Optimierung des Gradientenverfahrens ist hier der Rechenaufwand des optimierten Verfahrens zwar nicht mehr unabhängig von der Anzahl der Dämpfer, aber er steigt nur linear, während er beim nicht optimierten Verfahren zur Berechnung der Hesseschen Matrix quadratisch steigt.

Die Tests zum Vergleich der Startwertberechnung sowie die Tests (N1) und (N2) wurden auf einer IBM RS6000 -530 H mit 96 MB RAM unter AIX ausgeführt, die Tests (G1) und (G2) auf einem PC Pentium 100 mit 32 MB RAM unter Windows 95. Wie alle Tests gezeigt haben, wirkt sich sowohl eine optimierte Startwertauswahl (vgl. Abschnitt 3.1) als auch eine Optimierung in den Berechnungen innerhalb einer Iteration, d.h. bei der Berechnung des Gradienten (vgl. Abschnitt 3.2) und der Berechnung der Hesseschen Matrix (vgl. Abschnitt 3.3) in starkem Maße auf die Laufzeit in der Benutzung numerischer Näherungsverfahren zur Ermittlung optimaler Viskositäten aus.

4 Abschneidung des Systems

In Kapitel 3 haben wir gesehen, wie aufwendig eine Minimierung von $\text{Tr}(XZ_s)$ über die Menge der nichtnegativen Viskositäten Σ ist, wobei X die eindeutige Lösung von $A^T X + XA = -I$ mit

$$A = \begin{pmatrix} 0 & \Omega \\ -\Omega & -C \end{pmatrix} \in M_{2n}, \quad C = V\Sigma V^T, \quad Z_s = \begin{pmatrix} I_s & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & I_s & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad (4.1)$$

ist. Optimierungen, wie sie in Kapitel 3 vorgestellt worden sind, verkürzen zwar die Aufwände einer Minimierung teils erheblich, indem sie die Anzahl der benötigten Iterationen verringern und innerhalb jedes Iterationsschrittes die Anzahl der MFlops durch ein (zumindest beim Gradientenverfahren) von der Anzahl der Dämpfer unabhängiges Verfahren senken. Jedoch zeigen die Testreihen in Kapitel 3 auch, daß trotz einer Optimierung die Aufwände für große Dimensionen n immer noch sehr hoch sind (s. Abbildungen im Kapitel 3). Dies liegt an der Tatsache, daß die Ljapunovgleichung mit der $(2n)$ -dimensionalen Matrix A mehrfach pro Iterationsschritt gelöst werden muß. Dies ist sehr aufwendig, wählt man z.B. hierfür den Algorithmus von Bartels und Stewart [2] (s. Kapitel 1), so erfordert eine Lösung etwa $120n^3$ Operationen. In [29] wurde daher die Frage formuliert, ob die Matrix A – besonders für große Dimensionen n – auf eine Matrix \tilde{A} mit kleinerer Dimension $2r$, $r < n$, aber gleicher Anzahl der Dämpfer g reduziert werden kann, so daß die Spur der Lösung der Ljapunovgleichung mit der Matrix A in etwa der Spur der Lösung der Ljapunovgleichung mit der Matrix \tilde{A} entspricht. Wäre dies möglich, dann könnte eine numerische Lösung von Problem 3.1 für große Dimensionen n durch eine numerische Lösung des Problems für eine kleinere Dimension r ersetzt werden. Dies würde, falls r bedeutend kleiner als n ist, zu einer erheblichen Verkleinerung des Aufwandes führen und nahezu dieselben Ergebnisse wie die beim Originalsystem liefern. Damit stellt sich für uns das folgende Problem:

Problem 4.1

Existiert (und wenn ja, unter welchen Bedingungen) eine Matrix \tilde{A} der Dimension $2r$ ($r < n$) der Form

$$\tilde{A} = \begin{pmatrix} 0 & \tilde{\Omega} \\ -\tilde{\Omega} & -\tilde{C} \end{pmatrix}, \quad (4.2)$$

wobei $\tilde{\Omega}$ durch die ersten r Zeilen und Spalten von Ω und \tilde{C} durch die ersten r Zeilen und Spalten von C gegeben ist, so daß

$$\text{Tr}(XZ_s) \approx \text{Tr}(\tilde{X}\tilde{Z}_s)$$

gilt, wenn X eindeutige Lösung von $A^T X + X A = -I$ und \tilde{X} eindeutige Lösung von $\tilde{A}^T \tilde{X} + \tilde{X} \tilde{A} = -I$ ist? \tilde{Z}_s ist dabei die Z_s entsprechende $2r$ - dimensionale Matrix.

Mit \approx meinen wir dabei, daß der relative Fehler

$$\left| \frac{\text{Tr}(XZ_s) - \text{Tr}(\tilde{X}\tilde{Z}_s)}{\text{Tr}(XZ_s)} \right| \quad (4.3)$$

möglichst klein wird.

Mit der Untersuchung dieses Problems wollen wir uns in diesem Kapitel beschäftigen. Es ist klar, daß Problem 4.1 nicht für jede beliebige Matrix A und jedes beliebige r möglich ist. Auf jeden Fall muß aber $g, s \leq r < n$ gelten, da sonst Informationen, die in der Systemmatrix A enthalten sind, in der Matrix \tilde{A} verloren gehen. Wir werden untersuchen, welche Voraussetzungen die Frequenzen $\omega_1, \dots, \omega_n$ und die Dämpfungsmatrix C erfüllen müssen, damit A im Sinne von Problem 4.1 durch r reduziert werden kann.

Wie schon in Kapitel 2 angesprochen und o.B.d.A. möglich, werden wir im gesamten Kapitel 4 voraussetzen, daß die Frequenzen ω_i aufsteigend geordnet sind. Im „vollen“ Problem wird also der zu $\omega_i \leq \omega_s$ gehörende Frequenzunterraum betrachtet, und im „geschnittenen“ Problem werden die zu den Frequenzen $\omega_i > \omega_r$ gehörenden Informationen außer acht gelassen.

Zuerst wollen wir ein Lemma formulieren, das uns angibt, unter welchen Bedingungen $\text{Tr}(XZ_s)$ und $\text{Tr}(\tilde{X}\tilde{Z}_s)$ exakt gleich sind. O.B.d.A. nehmen wir an, daß die Dämpfungsmatrix C die Gestalt $C = VV^T$, $V \in M_{n,g}$ hat. Entsprechend ist $\tilde{C} = \tilde{V}\tilde{V}^T$, wobei \tilde{V} aus den ersten r Zeilen von V besteht. (V enthält also hier die Dämpferpositionen sowie die zugehörigen Viskositäten).

Lemma 4.2

Seien

$$A = \begin{pmatrix} 0 & \Omega \\ -\Omega & -VV^T \end{pmatrix} \in M_{2n} \quad \text{und} \quad \tilde{A} = \begin{pmatrix} 0 & \tilde{\Omega} \\ -\tilde{\Omega} & -\tilde{V}\tilde{V}^T \end{pmatrix} \in M_{2r}$$

stabile Matrizen mit $s \leq r < n$ und $\Omega = \text{diag}(\omega_1, \dots, \omega_n)$, $\tilde{\Omega} = \text{diag}(\omega_1, \dots, \omega_r)$, $0 < \omega_1 \leq \dots \leq \omega_n$, $V \in M_{n,g}$, $g \leq s$ mit

$$V = \begin{pmatrix} \tilde{V} \\ \tilde{V} \end{pmatrix},$$

wobei \tilde{V} die ersten r Zeilen von V und \check{V} die restlichen $n - r$ Zeilen von V bilden. Sei weiter X die eindeutige Lösung von $A^T X + X A = -I$ und \tilde{X} die eindeutige Lösung von $\tilde{A}^T \tilde{X} + \tilde{X} \tilde{A} = -I$. Dann gilt mit Z_s und \tilde{Z}_s (s.o.)

$$\text{Tr}(X Z_s) = \text{Tr}(\tilde{X} \tilde{Z}_s),$$

falls $\tilde{V} \check{V}^T = 0$ ist und zusätzlich $\check{V} \check{V}^T \neq 0$ gilt, d.h. mit anderen Worten, falls $V V^T$ eine blockdiagonale Matrix der Form

$$V V^T = \begin{pmatrix} \tilde{V} \tilde{V}^T & 0 \\ 0 & \check{V} \check{V}^T \end{pmatrix} \quad (4.4)$$

ist.

Beweis

Zur besseren Übersicht wählen wir eine andere Darstellung von A und \tilde{A} . Seien die Elemente von V durch v_{ij} , $i = 1, \dots, n$, $j = 1, \dots, g$ gegeben. A hat die Form

$$A = \begin{pmatrix} 0 & \Omega \\ -\Omega & 0 \end{pmatrix} - V_0 V_0^T \quad \text{mit} \quad V_0 = \begin{pmatrix} 0 \\ V \end{pmatrix} \in M_{2n, g},$$

und entsprechend hat \tilde{A} die Gestalt

$$\tilde{A} = \begin{pmatrix} 0 & \tilde{\Omega} \\ -\tilde{\Omega} & 0 \end{pmatrix} - \tilde{V}_0 \tilde{V}_0^T \quad \text{mit} \quad \tilde{V}_0 = \begin{pmatrix} 0 \\ \tilde{V} \end{pmatrix} \in M_{2s, g}.$$

Wir permutieren A und \tilde{A} so, daß sich aus A die Matrix

$$A_\tau = \text{diag}(\Omega_1, \dots, \Omega_n) - V_{0\tau} V_{0\tau}^T$$

mit

$$\Omega_i = \begin{pmatrix} 0 & \omega_i \\ -\omega_i & 0 \end{pmatrix}, \quad V_{0\tau} = \begin{pmatrix} 0 & 0 & \dots & 0 \\ v_{11} & v_{12} & \dots & v_{1g} \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & 0 \\ v_{n1} & v_{n2} & \dots & v_{ng} \end{pmatrix}$$

und aus \tilde{A} die Matrix

$$\tilde{A}_\tau = \text{diag}(\Omega_1, \dots, \Omega_r) - \tilde{V}_{0\tau} \tilde{V}_{0\tau}^T$$

ergibt. Durch diese Permutationen der Matrizen von A und \tilde{A} ergibt sich für die Lösung $A^T X + X A = -I$ und $A_\tau^T X_\tau + X_\tau A_\tau = -I$ (und entsprechend für \tilde{A}) der folgende Zusammenhang:

Sei $X = \begin{pmatrix} X_{11} & X_{12} \\ X_{12}^T & X_{22} \end{pmatrix}$, dann sind die Diagonalelemente von X_τ gegeben durch

$$(X_\tau)_{ii} = \begin{cases} (X_{11})_{kk} & \text{für } i \text{ ungerade} \quad \text{und } k = \frac{i+1}{2} \\ (X_{22})_{kk} & \text{für } i \text{ gerade} \quad \text{und } k = \frac{i}{2} \end{cases},$$

es gilt also

$$\text{Tr}(XZ_s) = \sum_{i=1}^s (X)_{ii} + \sum_{i=n+1}^{s+n} (X)_{ii} = \sum_{i=1}^{2s} (X_\tau)_{ii}$$

und entsprechend

$$\text{Tr}(\tilde{X}\tilde{Z}_s) = \sum_{i=1}^{2s} (\tilde{X}_\tau)_{ii}.$$

Mit der Schreibweise $V_{0\tau} := \begin{pmatrix} \hat{V}_{0\tau} \\ \check{V}_{0\tau} \end{pmatrix}$, wobei $\hat{V}_{0\tau}$ die ersten $2r$ Zeilen von $V_{0\tau}$ und $\check{V}_{0\tau}$ die restlichen $2(n-r)$ Zeilen von $V_{0\tau}$ bezeichnet und $\hat{\Omega}^{(r)} := \text{diag}(\Omega_1, \dots, \Omega_r)$, $\check{\Omega}^{(n-r)} := \text{diag}(\Omega_{r+1}, \dots, \Omega_n)$, haben dann \tilde{A}_τ und A_τ die Gestalt

$$\tilde{A}_\tau = \hat{\Omega}^{(r)} - \hat{V}_{0\tau}\hat{V}_{0\tau}^T, \quad A_\tau = \begin{pmatrix} \hat{\Omega}^{(r)} - \hat{V}_{0\tau}\hat{V}_{0\tau}^T & -\hat{V}_{0\tau}\check{V}_{0\tau}^T \\ -\check{V}_{0\tau}\hat{V}_{0\tau}^T & \check{\Omega}^{(n-r)} - \check{V}_{0\tau}\check{V}_{0\tau}^T \end{pmatrix}. \quad (4.5)$$

Mit (4.5) kann man nun erkennen, daß $\sum_{i=1}^{2s} (X_\tau)_{ii} = \sum_{i=1}^{2s} (\tilde{X}_\tau)_{ii}$ gilt, falls $\hat{V}_{0\tau}\check{V}_{0\tau}^T = 0$ ist. Zusätzlich muß aber, aufgrund der sonst nicht vorhandenen Stabilität von A_τ (s. Lemma 1.4), auch $\check{V}_{0\tau}\check{V}_{0\tau}^T \neq 0$ gelten.

Überträgt man dies wieder auf die nicht permutierte Schreibweise, so folgt die Aussage des Satzes.

Q.E.D.

Betrachten wir die Aussage von Lemma 4.2, so liegt die Vermutung nahe, daß, falls $\check{V}\check{V}^T$ gegen Null konvergiert (was gleichzeitig $\tilde{V}\tilde{V}^T \rightarrow 0$ zur Folge hat), der relative Fehler (4.3) ebenfalls gegen Null konvergiert. Diese Vermutung werden wir jedoch in diesem Kapitel widerlegen. Der relative Fehler der Spur verhält sich also nicht stetig bei Annäherung der Dämpfungsmatrix an die Form (4.4).

Problem 4.1 exakt zu untersuchen, hat sich daher als sehr schwer erwiesen. Das folgende Kapitel untersucht das Problem für $g = 1$, d.h. falls $\text{rang}(C) = 1$ ist und C damit aus einer Dyade cc^T besteht. In diesem Fall steht eine explizite Formel

für X zur Verfügung (s. [24]). Nach Bemerkung 1.6 müssen wir voraussetzen, daß die ω_i voneinander verschieden sind und c in allen Komponenten ungleich 0 ist. Im Anschluß hieran wird die entsprechende Fragestellung für die optimale Viskosität untersucht und anhand von numerischen Tests überprüft, ob die für einen Dämpfer gewonnenen Ergebnisse auch auf den Fall mehrerer Dämpfer übertragen werden können.

4.1 Abschneidung bei einem Dämpfer

In diesem Abschnitt wollen wir also Problem 4.1 für $C = \varepsilon cc^T$, $\varepsilon > 0$ und c ein Vektor der Länge n , untersuchen. Dazu formulieren wir zunächst ein Lemma.

Lemma 4.3

Sei

$$X = \begin{pmatrix} X_{11} & X_{12} \\ X_{12}^T & X_{22} \end{pmatrix}$$

die eindeutig bestimmte Lösung der Ljapunovgleichung

$$A^T X + X A = -I$$

mit

$$A = \begin{pmatrix} 0 & \Omega \\ -\Omega & -\varepsilon cc^T \end{pmatrix},$$

$\Omega = \text{diag}(\omega_1, \dots, \omega_n)$ mit $0 < \omega_1 < \dots < \omega_n$, $\varepsilon > 0$, $c = (c_1, \dots, c_n)^T$ mit $c_i \neq 0$ für alle $i = 1, \dots, n$.

Dann sind die Diagonalelemente von X gegeben durch

$$(X_{11})_{pp} = \frac{1}{\varepsilon c_p^2} + \frac{\varepsilon c_p^2}{2\omega_p^2} + \varepsilon \left[\sum_{\substack{j=1 \\ j \neq p}}^n \frac{2\omega_p^2 c_j^2 + c_p^2 \omega_j^2}{(\omega_p^2 - \omega_j^2)^2} + \frac{\omega_p^2}{c_p^2} \left(\sum_{\substack{k=1 \\ k \neq p}}^n \frac{c_k^2}{\omega_k^2 - \omega_p^2} \right)^2 \right],$$

$$(X_{22})_{pp} = \frac{1}{\varepsilon c_p^2} + \varepsilon \left[\sum_{\substack{j=1 \\ j \neq p}}^n \frac{\omega_p^2 c_j^2}{(\omega_p^2 - \omega_j^2)^2} + \sum_{\substack{j=1 \\ j \neq p}}^n \frac{\omega_j^2 (c_j^2 + c_p^2)}{(\omega_p^2 - \omega_j^2)^2} + \frac{\omega_p^2}{c_p^2} \left(\sum_{\substack{k=1 \\ k \neq p}}^n \frac{c_k^2}{\omega_k^2 - \omega_p^2} \right)^2 \right],$$

$p = 1, \dots, n$.

Beweis

Nach [24] Satz 3.1 können die Diagonalelemente von X auf folgende Weise berechnet werden (die in [24] vorhandenen Druckfehler wurden hier bereinigt):

Sei $X_{11} =: (\eta_{ij})$ und $X_{22} = (\xi_{ij})$. Dann gilt für $i \neq j$

$$\begin{aligned}\xi_{ij} &= \frac{\varepsilon}{\omega_j^2 - \omega_i^2} (z_i c_j \omega_i - z_j c_i \omega_j), \\ \eta_{ij} &= \frac{\varepsilon}{\omega_j^2 - \omega_i^2} (z_i c_j \omega_j - z_j c_i \omega_i)\end{aligned}$$

und für $i = j$

$$\begin{aligned}\xi_{ii} &= \frac{z_{ii}}{2c_i^2 \varepsilon} - \frac{1}{c_i} \sum_{\substack{j=1 \\ j \neq i}}^n \xi_{ij} c_j, \\ \eta_{ii} &= \xi_{ii} + \frac{\varepsilon z_i c_i}{\omega_i}\end{aligned}$$

mit

$$\begin{aligned}Z &= (z_{ij}) = 2I, \\ z &= (z_i) = X_{12}c, \\ X_{12} &= \frac{1}{2}\Omega^{-1} + \frac{1}{2}S\Omega^{-1}, \\ S &= (s_{ij}) \text{ mit} \\ s_{ij} &= \frac{\frac{c_i z_{jj} + c_j z_{ii}}{c_j} + \frac{c_i}{\omega_j - \omega_i}}{\frac{\omega_i}{\omega_j} - \frac{\omega_j}{\omega_i}} \text{ für } i \neq j \text{ und} \\ s_{ii} &= 0.\end{aligned}$$

Damit ergibt sich für $i \neq j$

$$s_{ij} = \frac{\frac{c_i z_{jj} + c_j z_{ii}}{c_j} + \frac{c_i}{\omega_j - \omega_i}}{\frac{\omega_i}{\omega_j} - \frac{\omega_j}{\omega_i}} = \frac{2c_i + \frac{2c_j}{c_i}}{\frac{\omega_i}{\omega_j} - \frac{\omega_j}{\omega_i}} = \frac{2\left(\frac{c_i}{c_j} + \frac{c_j}{c_i}\right)}{\frac{\omega_i^2 - \omega_j^2}{\omega_i \omega_j}} = \frac{2\omega_i \omega_j \left(\frac{c_i}{c_j} + \frac{c_j}{c_i}\right)}{\omega_i^2 - \omega_j^2},$$

und wir erhalten

$$z = (z_i) = X_{12}c = \left(\frac{1}{2}\Omega^{-1} + \frac{1}{2}S\Omega^{-1}\right)c,$$

also

$$\begin{aligned}z_i &= \frac{1}{2} \left(\frac{c_i}{\omega_i} + \sum_{\substack{k=1 \\ k \neq i}}^n s_{ik} \frac{c_k}{\omega_k} \right) = \frac{1}{2} \left(\frac{c_i}{\omega_i} + \sum_{\substack{k=1 \\ k \neq i}}^n \frac{2\omega_i \omega_k \left(\frac{c_i}{c_k} + \frac{c_k}{c_i}\right)}{\omega_i^2 - \omega_k^2} \frac{c_k}{\omega_k} \right) \\ &= \frac{c_i}{2\omega_i} + \sum_{\substack{k=1 \\ k \neq i}}^n \frac{\omega_i (c_i^2 + c_k^2)}{c_i (\omega_i^2 - \omega_k^2)},\end{aligned}$$

und für $p \neq j$ ist

$$\begin{aligned}
\xi_{pj} &= \frac{\varepsilon}{\omega_j^2 - \omega_p^2} (z_p c_j \omega_p - z_j c_p \omega_j) \\
&= \frac{\varepsilon}{\omega_j^2 - \omega_p^2} \left(\frac{c_p c_j}{2} + \sum_{\substack{k=1 \\ k \neq p}}^n \frac{\omega_p^2 (c_p^2 + c_k^2) c_j}{c_p (\omega_p^2 - \omega_k^2)} - \frac{c_p c_j}{2} - \sum_{\substack{k=1 \\ k \neq j}}^n \frac{\omega_j^2 (c_j^2 + c_k^2) c_p}{c_j (\omega_j^2 - \omega_k^2)} \right) \\
&= \frac{\varepsilon}{\omega_j^2 - \omega_p^2} \left(\sum_{\substack{k=1 \\ k \neq p}}^n \frac{\omega_p^2 (c_p^2 + c_k^2) c_j}{c_p (\omega_p^2 - \omega_k^2)} - \sum_{\substack{k=1 \\ k \neq j}}^n \frac{\omega_j^2 (c_j^2 + c_k^2) c_p}{c_j (\omega_j^2 - \omega_k^2)} \right).
\end{aligned}$$

Nun erhält man

$$\begin{aligned}
\xi_{pp} &= \frac{z_{pp}}{2\varepsilon c_p^2} - \frac{1}{c_p} \sum_{\substack{j=1 \\ j \neq p}}^n \xi_{pj} c_j \\
&= \frac{1}{\varepsilon c_p^2} - \frac{1}{c_p} \sum_{\substack{j=1 \\ j \neq p}}^n \xi_{pj} c_j \\
&= \frac{1}{\varepsilon c_p^2} - \sum_{\substack{j=1 \\ j \neq p}}^n \left[\frac{\varepsilon}{\omega_j^2 - \omega_p^2} \left(\sum_{\substack{k=1 \\ k \neq p}}^n \frac{\omega_p^2 (c_p^2 + c_k^2) c_j^2}{c_p^2 (\omega_p^2 - \omega_k^2)} - \sum_{\substack{k=1 \\ k \neq j}}^n \frac{\omega_j^2 (c_j^2 + c_k^2)}{(\omega_j^2 - \omega_k^2)} \right) \right] \\
&= \frac{1}{\varepsilon c_p^2} + \varepsilon \left[\sum_{\substack{j=1 \\ j \neq p}}^n \frac{\omega_p^2 (c_p^2 + c_j^2) c_j^2}{c_p^2 (\omega_p^2 - \omega_j^2)^2} + \sum_{\substack{j=1 \\ j \neq p}}^n \frac{1}{\omega_p^2 - \omega_j^2} \sum_{\substack{k=1 \\ k \neq p, j}}^n \frac{\omega_p^2 (c_p^2 + c_k^2) c_j^2}{c_p^2 (\omega_p^2 - \omega_k^2)} \right. \\
&\quad \left. + \sum_{\substack{j=1 \\ j \neq p}}^n \frac{\omega_j^2 (c_j^2 + c_p^2)}{(\omega_p^2 - \omega_j^2)^2} - \sum_{\substack{j=1 \\ j \neq p}}^n \frac{1}{\omega_p^2 - \omega_j^2} \sum_{\substack{k=1 \\ k \neq j, p}}^n \frac{\omega_j^2 (c_j^2 + c_k^2)}{\omega_j^2 - \omega_k^2} \right].
\end{aligned}$$

Für den letzten Term

$$h := \sum_{\substack{j=1 \\ j \neq p}}^n \frac{1}{\omega_p^2 - \omega_j^2} \sum_{\substack{k=1 \\ k \neq j, p}}^n \frac{\omega_j^2 (c_j^2 + c_k^2)}{\omega_j^2 - \omega_k^2}$$

gilt

$$h = \sum_{\substack{j=1 \\ j \neq p}}^n \sum_{\substack{k=1 \\ k \neq p}}^{j-1} \frac{\omega_j^2 (c_j^2 + c_k^2)}{(\omega_j^2 - \omega_k^2)(\omega_p^2 - \omega_j^2)} + \sum_{\substack{j=1 \\ j \neq p}}^n \sum_{\substack{k=j+1 \\ k \neq p}}^n \frac{\omega_j^2 (c_j^2 + c_k^2)}{(\omega_j^2 - \omega_k^2)(\omega_p^2 - \omega_j^2)}.$$

Für die erste Doppelsumme von h gilt weiter

$$\begin{aligned} \sum_{\substack{j=1 \\ j \neq p}}^n \sum_{\substack{k=1 \\ k \neq p}}^{j-1} \frac{\omega_j^2(c_j^2 + c_k^2)}{(\omega_j^2 - \omega_k^2)(\omega_p^2 - \omega_j^2)} &= \sum_{\substack{k=1 \\ k \neq p}}^n \sum_{\substack{j=k+1 \\ j \neq p}}^n \frac{\omega_j^2(c_j^2 + c_k^2)}{(\omega_j^2 - \omega_k^2)(\omega_p^2 - \omega_j^2)} \\ &= - \sum_{\substack{j=1 \\ j \neq p}}^n \sum_{\substack{k=j+1 \\ k \neq p}}^n \frac{\omega_k^2(c_k^2 + c_j^2)}{(\omega_j^2 - \omega_k^2)(\omega_p^2 - \omega_k^2)}, \end{aligned}$$

wobei im letzten Schritt nur j und k vertauscht worden sind.

Damit ergibt sich weiter für ξ_{pp}

$$\begin{aligned} \xi_{pp} &= \frac{1}{\varepsilon c_p^2} + \varepsilon \left[\sum_{\substack{j=1 \\ j \neq p}}^n \frac{\omega_p^2(c_p^2 + c_j^2)c_j^2}{c_p^2(\omega_p^2 - \omega_j^2)^2} + \sum_{\substack{j=1 \\ j \neq p}}^n \frac{\omega_j^2(c_j^2 + c_p^2)}{(\omega_p^2 - \omega_j^2)^2} + \sum_{\substack{j=1 \\ j \neq p}}^n \frac{1}{\omega_p^2 - \omega_j^2} \sum_{\substack{k=1 \\ k \neq p, j}}^n \frac{\omega_p^2(c_p^2 + c_k^2)c_j^2}{c_p^2(\omega_p^2 - \omega_k^2)} \right. \\ &\quad \left. - \sum_{\substack{j=1 \\ j \neq p}}^n \sum_{\substack{k=j+1 \\ k \neq p}}^n \left[\frac{1}{\omega_j^2 - \omega_k^2} \left(\frac{\omega_j^2(c_j^2 + c_k^2)}{\omega_p^2 - \omega_j^2} - \frac{\omega_k^2(c_k^2 + c_j^2)}{\omega_p^2 - \omega_k^2} \right) \right] \right] \\ &= \frac{1}{\varepsilon c_p^2} + \varepsilon \left[\sum_{\substack{j=1 \\ j \neq p}}^n \frac{\omega_p^2(c_p^2 + c_j^2)c_j^2}{c_p^2(\omega_p^2 - \omega_j^2)^2} + \sum_{\substack{j=1 \\ j \neq p}}^n \frac{\omega_j^2(c_j^2 + c_p^2)}{(\omega_p^2 - \omega_j^2)^2} + \sum_{\substack{j=1 \\ j \neq p}}^n \frac{1}{\omega_p^2 - \omega_j^2} \sum_{\substack{k=1 \\ k \neq p, j}}^n \frac{\omega_p^2(c_p^2 + c_k^2)c_j^2}{c_p^2(\omega_p^2 - \omega_k^2)} \right. \\ &\quad \left. - \sum_{\substack{j=1 \\ j \neq p}}^n \sum_{\substack{k=j+1 \\ k \neq p}}^n \frac{1}{\omega_j^2 - \omega_k^2} \left(\frac{\omega_p^2 c_j^2(\omega_j^2 - \omega_k^2) + \omega_p^2 c_k^2(\omega_j^2 - \omega_k^2)}{(\omega_p^2 - \omega_j^2)(\omega_p^2 - \omega_k^2)} \right) \right] \\ &= \frac{1}{\varepsilon c_p^2} + \varepsilon \left[\sum_{\substack{j=1 \\ j \neq p}}^n \frac{\omega_p^2(c_p^2 + c_j^2)c_j^2}{c_p^2(\omega_p^2 - \omega_j^2)^2} + \sum_{\substack{j=1 \\ j \neq p}}^n \frac{\omega_j^2(c_j^2 + c_p^2)}{(\omega_p^2 - \omega_j^2)^2} \right. \\ &\quad \left. + \sum_{\substack{j=1 \\ j \neq p}}^n \frac{1}{\omega_p^2 - \omega_j^2} \sum_{\substack{k=1 \\ k \neq p, j}}^n \frac{\omega_p^2(c_p^2 + c_k^2)c_j^2}{c_p^2(\omega_p^2 - \omega_k^2)} - \sum_{\substack{j=1 \\ j \neq p}}^n \sum_{\substack{k=j+1 \\ k \neq p}}^n \frac{\omega_p^2(c_j^2 + c_k^2)}{(\omega_p^2 - \omega_j^2)(\omega_p^2 - \omega_k^2)} \right] \\ &= \frac{1}{\varepsilon c_p^2} + \varepsilon \left[\sum_{\substack{j=1 \\ j \neq p}}^n \frac{\omega_p^2(c_p^2 + c_j^2)c_j^2}{c_p^2(\omega_p^2 - \omega_j^2)^2} + \sum_{\substack{j=1 \\ j \neq p}}^n \frac{\omega_j^2(c_j^2 + c_p^2)}{(\omega_p^2 - \omega_j^2)^2} \right. \\ &\quad \left. + \sum_{\substack{j=1 \\ j \neq p}}^n \sum_{\substack{k=j+1 \\ k \neq p}}^n \frac{\omega_p^2}{c_p^2(\omega_p^2 - \omega_j^2)(\omega_p^2 - \omega_k^2)} \left[(c_p^2 + c_k^2)c_j^2 + (c_p^2 + c_j^2)c_k^2 \right] \right. \\ &\quad \left. - \sum_{\substack{j=1 \\ j \neq p}}^n \sum_{\substack{k=j+1 \\ k \neq p}}^n \frac{\omega_p^2(c_j^2 + c_k^2)}{(\omega_p^2 - \omega_j^2)(\omega_p^2 - \omega_k^2)} \right] = \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{\varepsilon c_p^2} + \varepsilon \left[\sum_{\substack{j=1 \\ j \neq p}}^n \frac{\omega_p^2 (c_p^2 + c_j^2) c_j^2}{c_p^2 (\omega_p^2 - \omega_j^2)^2} + \sum_{\substack{j=1 \\ j \neq p}}^n \frac{\omega_j^2 (c_j^2 + c_p^2)}{(\omega_p^2 - \omega_j^2)^2} \right. \\
&\quad \left. + \sum_{\substack{j=1 \\ j \neq p}}^n \sum_{\substack{k=j+1 \\ k \neq p}}^n \frac{\omega_p^2}{c_p^2 (\omega_p^2 - \omega_j^2) (\omega_p^2 - \omega_k^2)} (c_p^2 c_j^2 + c_k^2 c_j^2 + c_p^2 c_k^2 + c_j^2 c_k^2 - c_p^2 c_j^2 - c_p^2 c_k^2) \right] \\
&= \frac{1}{\varepsilon c_p^2} + \varepsilon \left[\sum_{\substack{j=1 \\ j \neq p}}^n \frac{\omega_p^2 (c_p^2 + c_j^2) c_j^2}{c_p^2 (\omega_p^2 - \omega_j^2)^2} + \sum_{\substack{j=1 \\ j \neq p}}^n \frac{\omega_j^2 (c_j^2 + c_p^2)}{(\omega_p^2 - \omega_j^2)^2} + 2 \sum_{\substack{j=1 \\ j \neq p}}^n \sum_{\substack{k=j+1 \\ k \neq p}}^n \frac{c_k^2 c_j^2 \omega_p^2}{c_p^2 (\omega_p^2 - \omega_j^2) (\omega_p^2 - \omega_k^2)} \right].
\end{aligned}$$

Nun wollen wir ξ_{pp} in einer Form darstellen, die nur positive Terme enthält. Diese Darstellung wird später für die Abschätzungen benötigt. Wir erhalten

$$\begin{aligned}
\xi_{pp} &= \frac{1}{\varepsilon c_p^2} + \varepsilon \left[\sum_{\substack{j=1 \\ j \neq p}}^n \frac{\omega_p^2 c_j^2}{(\omega_p^2 - \omega_j^2)^2} + \frac{\omega_p^2}{c_p^2} \sum_{\substack{j=1 \\ j \neq p}}^n \frac{c_j^4}{(\omega_p^2 - \omega_j^2)^2} + \sum_{\substack{j=1 \\ j \neq p}}^n \frac{\omega_j^2 (c_j^2 + c_p^2)}{(\omega_p^2 - \omega_j^2)^2} \right. \\
&\quad \left. + \frac{\omega_p^2}{c_p^2} \sum_{\substack{j=1 \\ j \neq p}}^n \sum_{\substack{k=j+1 \\ k \neq p}}^n \frac{2c_k^2 c_j^2}{(\omega_p^2 - \omega_j^2) (\omega_p^2 - \omega_k^2)} \right]
\end{aligned}$$

und wegen

$$\sum_{\substack{j=1 \\ j \neq p}}^n \frac{c_j^4}{(\omega_p^2 - \omega_j^2)^2} + \sum_{\substack{j=1 \\ j \neq p}}^n \sum_{\substack{k=j+1 \\ k \neq p}}^n \frac{2c_k^2 c_j^2}{(\omega_p^2 - \omega_j^2) (\omega_p^2 - \omega_k^2)} = \left(\sum_{\substack{k=1 \\ k \neq p}}^n \frac{c_k^2}{\omega_k^2 - \omega_p^2} \right)^2$$

ergibt sich

$$\xi_{pp} = \frac{1}{\varepsilon c_p^2} + \varepsilon \left[\sum_{\substack{j=1 \\ j \neq p}}^n \frac{\omega_p^2 c_j^2}{(\omega_p^2 - \omega_j^2)^2} + \sum_{\substack{j=1 \\ j \neq p}}^n \frac{\omega_j^2 (c_j^2 + c_p^2)}{(\omega_p^2 - \omega_j^2)^2} + \frac{\omega_p^2}{c_p^2} \left(\sum_{\substack{k=1 \\ k \neq p}}^n \frac{c_k^2}{\omega_k^2 - \omega_p^2} \right)^2 \right].$$

Jetzt kann η_{pp} berechnet werden:

$$\begin{aligned}
\eta_{pp} &= \xi_{pp} + \frac{\varepsilon z_p c_p}{\omega_p} \\
&= \frac{1}{\varepsilon c_p^2} + \varepsilon \left[\sum_{\substack{j=1 \\ j \neq p}}^n \frac{\omega_p^2 c_j^2}{(\omega_p^2 - \omega_j^2)^2} + \sum_{\substack{j=1 \\ j \neq p}}^n \frac{\omega_j^2 (c_j^2 + c_p^2)}{(\omega_p^2 - \omega_j^2)^2} + \frac{\omega_p^2}{c_p^2} \left(\sum_{\substack{k=1 \\ k \neq p}}^n \frac{c_k^2}{\omega_k^2 - \omega_p^2} \right)^2 \right] \\
&\quad + \varepsilon \left[\frac{c_p^2}{2\omega_p^2} + \sum_{\substack{k=1 \\ k \neq p}}^n \frac{c_p^2 + c_k^2}{\omega_p^2 - \omega_k^2} \right] \\
&= \frac{1}{\varepsilon c_p^2} + \frac{\varepsilon c_p^2}{2\omega_p^2} + \varepsilon \left[\sum_{\substack{j=1 \\ j \neq p}}^n \frac{\omega_p^2 c_j^2 + \omega_j^2 (c_j^2 + c_p^2) + (c_p^2 + c_j^2)(\omega_p^2 - \omega_j^2)}{(\omega_p^2 - \omega_j^2)^2} \right. \\
&\quad \left. + \frac{\omega_p^2}{c_p^2} \left(\sum_{\substack{k=1 \\ k \neq p}}^n \frac{c_k^2}{\omega_k^2 - \omega_p^2} \right)^2 \right] \\
&= \frac{1}{\varepsilon c_p^2} + \frac{\varepsilon c_p^2}{2\omega_p^2} + \varepsilon \left[\sum_{\substack{j=1 \\ j \neq p}}^n \frac{2\omega_p^2 c_j^2 + c_p^2 \omega_p^2}{(\omega_p^2 - \omega_j^2)^2} + \frac{\omega_p^2}{c_p^2} \left(\sum_{\substack{k=1 \\ k \neq p}}^n \frac{c_k^2}{\omega_k^2 - \omega_p^2} \right)^2 \right].
\end{aligned}$$

Q.E.D.

Mit Lemma 4.3 wollen wir nun $\text{Tr}(XZ_s)$ berechnen mit Z_s aus (4.1). Sei wieder $X_{11} =: (\eta_{ij})$ und $X_{22} =: (\xi_{ij})$. Dann ist

$$\text{Tr}(XZ_s) = \sum_{p=1}^s (\eta_{pp} + \xi_{pp})$$

und mit Lemma 4.3 erhalten wir

$$\begin{aligned}
\eta_{pp} + \xi_{pp} &= \frac{1}{\varepsilon c_p^2} + \frac{\varepsilon c_p^2}{2\omega_p^2} + \varepsilon \left[\sum_{\substack{j=1 \\ j \neq p}}^n \frac{2\omega_p^2 c_j^2 + c_p^2 \omega_p^2}{(\omega_p^2 - \omega_j^2)^2} + \frac{\omega_p^2}{c_p^2} \left(\sum_{\substack{k=1 \\ k \neq p}}^n \frac{c_k^2}{\omega_k^2 - \omega_p^2} \right)^2 \right] \\
&\quad + \frac{1}{\varepsilon c_p^2} + \varepsilon \left[\sum_{\substack{j=1 \\ j \neq p}}^n \frac{\omega_p^2 c_j^2}{(\omega_p^2 - \omega_j^2)^2} + \sum_{\substack{j=1 \\ j \neq p}}^n \frac{\omega_j^2 (c_j^2 + c_p^2)}{(\omega_p^2 - \omega_j^2)^2} + \frac{\omega_p^2}{c_p^2} \left(\sum_{\substack{k=1 \\ k \neq p}}^n \frac{c_k^2}{\omega_k^2 - \omega_p^2} \right)^2 \right] \\
&= \frac{2}{\varepsilon c_p^2} + \frac{\varepsilon c_p^2}{2\omega_p^2} + \varepsilon \left[\sum_{\substack{j=1 \\ j \neq p}}^n \frac{3\omega_p^2 c_j^2 + \omega_p^2 c_p^2 + \omega_j^2 c_j^2 + \omega_j^2 c_p^2}{(\omega_p^2 - \omega_j^2)^2} + 2 \frac{\omega_p^2}{c_p^2} \left(\sum_{\substack{k=1 \\ k \neq p}}^n \frac{c_k^2}{\omega_k^2 - \omega_p^2} \right)^2 \right].
\end{aligned}$$

Damit erhalten wir den folgenden Ausdruck

$$\begin{aligned} \text{Tr}(XZ_s) = & \quad (4.6) \\ & \sum_{p=1}^s \left[\frac{2}{\varepsilon c_p^2} + \frac{\varepsilon c_p^2}{2\omega_p^2} + \varepsilon \left(\sum_{\substack{j=1 \\ j \neq p}}^n \frac{3\omega_p^2 c_j^2 + \omega_p^2 c_p^2 + \omega_j^2 c_j^2 + \omega_j^2 c_p^2}{(\omega_p^2 - \omega_j^2)^2} + 2 \frac{\omega_p^2}{c_p^2} \left(\sum_{\substack{k=1 \\ k \neq p}}^n \frac{c_k^2}{\omega_k^2 - \omega_p^2} \right)^2 \right) \right]. \end{aligned}$$

Später benötigen wir noch eine andere Darstellung für $\eta_{pp} + \xi_{pp}$. Daher formen wir diesen Ausdruck weiter um. Es ergibt sich

$$\begin{aligned} \eta_{pp} + \xi_{pp} &= \frac{2}{\varepsilon c_p^2} + \frac{\varepsilon c_p^2}{2\omega_p^2} + \varepsilon \left[\sum_{\substack{j=1 \\ j \neq p}}^n \frac{3\omega_p^2 c_j^2 + \omega_p^2 c_p^2 + \omega_j^2 c_j^2 + \omega_j^2 c_p^2}{(\omega_p^2 - \omega_j^2)^2} + \sum_{\substack{k=1 \\ k \neq p}}^n \frac{2\omega_p^2 c_k^4}{c_p^2 (\omega_p^2 - \omega_k^2)^2} \right. \\ &\quad \left. + 4 \sum_{\substack{j=1 \\ j \neq p}}^n \sum_{\substack{k=j+1 \\ k \neq p}}^n \frac{c_k^2 c_j^2 \omega_p^2}{c_p^2 (\omega_p^2 - \omega_j^2) (\omega_p^2 - \omega_k^2)} \right] \\ &= \frac{2}{\varepsilon c_p^2} + \frac{\varepsilon c_p^2}{2\omega_p^2} + \varepsilon \left[4 \sum_{\substack{j=1 \\ j \neq p}}^n \sum_{\substack{k=j+1 \\ k \neq p}}^n \frac{c_k^2 c_j^2 \omega_p^2}{c_p^2 (\omega_p^2 - \omega_j^2) (\omega_p^2 - \omega_k^2)} \right. \\ &\quad \left. + \sum_{\substack{j=1 \\ j \neq p}}^n \frac{(c_p^2 + c_j^2) [\omega_p^2 c_p^2 + 2\omega_p^2 c_j^2 + \omega_j^2 c_p^2]}{c_p^2 (\omega_p^2 - \omega_j^2)^2} \right]. \quad (4.7) \end{aligned}$$

Sei

$$\tilde{X} = \begin{pmatrix} \tilde{X}_{11} & \tilde{X}_{12} \\ \tilde{X}_{12} & \tilde{X}_{22} \end{pmatrix} \in M_{2r}, \quad r < n$$

die eindeutig bestimmte Lösung der Ljapunovgleichung

$$\tilde{A}^T \tilde{X} + \tilde{X} \tilde{A} = -I. \quad (4.8)$$

Hierbei ist

$$\tilde{A} = \begin{pmatrix} 0 & \tilde{\Omega} \\ -\tilde{\Omega} & -\varepsilon \tilde{c} \tilde{c}^T \end{pmatrix}$$

mit $\tilde{\Omega} = \text{diag}(\omega_1, \dots, \omega_r)$ und $\tilde{c} = (c_1, \dots, c_r)^T$. Sei außerdem \tilde{Z}_s eine Matrix der Dimension $2r$, $s \leq r < n$, mit

$$\tilde{Z}_s = \begin{pmatrix} I_s & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & I_s & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Im folgenden bezeichnen wir stets (4.8) als das bei r geschnittene System mit Bezug auf das Ursprungssystem $A^T X + X A = -I$. Die Indexmenge $\{1, \dots, s\}$ bezeichnen wir als *relevanten Bereich* und $\{r+1, \dots, n\}$ als den *Bereich hinter dem Schnitt*. Der Bereich $\{s+1, \dots, r\}$ wird als *Zwischenbereich* bezeichnet.

Nun wollen wir den relativen Fehler

$$\text{rF} [\text{Tr} (X Z_s)] := \left| \frac{\text{Tr} (X Z_s) - \text{Tr} (\tilde{X} \tilde{Z}_s)}{\text{Tr} (X Z_s)} \right|$$

für $s \leq r < n$ berechnen. Wir werden ihn im folgenden als *relativen Fehler der Spur* bezeichnen.

Sei wieder $X_{11} =: (\eta_{ij})$, $X_{22} =: (\xi_{ij})$ und entsprechend $\tilde{X}_{11} =: (\tilde{\eta}_{ij})$, $\tilde{X}_{22} =: (\tilde{\xi}_{ij})$. Dann ist

$$\text{Tr} (X Z_s) - \text{Tr} (\tilde{X} \tilde{Z}_s) = \sum_{p=1}^s [(\eta_{pp} + \xi_{pp}) - (\tilde{\eta}_{pp} + \tilde{\xi}_{pp})].$$

Zunächst berechnen wir $(\eta_{pp} + \xi_{pp}) - (\tilde{\eta}_{pp} + \tilde{\xi}_{pp})$ mit (4.7) .

$$\begin{aligned} & (\eta_{pp} + \xi_{pp}) - (\tilde{\eta}_{pp} + \tilde{\xi}_{pp}) \\ &= \frac{2}{\varepsilon c_p^2} + \frac{\varepsilon c_p^2}{2\omega_p^2} + \varepsilon \left[4 \sum_{\substack{j=1 \\ j \neq p}}^n \sum_{\substack{k=j+1 \\ k \neq p}}^n \frac{c_k^2 c_j^2 \omega_p^2}{c_p^2 (\omega_p^2 - \omega_j^2) (\omega_p^2 - \omega_k^2)} \right. \\ & \quad \left. + \sum_{\substack{j=1 \\ j \neq p}}^n \frac{(c_p^2 + c_j^2) [\omega_p^2 c_p^2 + 2\omega_p^2 c_j^2 + \omega_j^2 c_p^2]}{c_p^2 (\omega_p^2 - \omega_j^2)^2} \right] \\ & \quad - \frac{2}{\varepsilon c_p^2} - \frac{\varepsilon c_p^2}{2\omega_p^2} - \varepsilon \left[4 \sum_{\substack{j=1 \\ j \neq p}}^r \sum_{\substack{k=j+1 \\ k \neq p}}^r \frac{c_k^2 c_j^2 \omega_p^2}{c_p^2 (\omega_p^2 - \omega_j^2) (\omega_p^2 - \omega_k^2)} \right. \\ & \quad \left. + \sum_{\substack{j=1 \\ j \neq p}}^r \frac{(c_p^2 + c_j^2) [\omega_p^2 c_p^2 + 2\omega_p^2 c_j^2 + \omega_j^2 c_p^2]}{c_p^2 (\omega_p^2 - \omega_j^2)^2} \right] \\ &= \varepsilon \left[\sum_{j=r+1}^n \frac{(c_p^2 + c_j^2) [\omega_p^2 c_p^2 + 2\omega_p^2 c_j^2 + \omega_j^2 c_p^2]}{c_p^2 (\omega_j^2 - \omega_p^2)^2} + 4 \sum_{\substack{j=1 \\ j \neq p}}^r \sum_{k=r+1}^n \frac{c_k^2 c_j^2 \omega_p^2}{c_p^2 (\omega_p^2 - \omega_j^2) (\omega_p^2 - \omega_k^2)} \right. \\ & \quad \left. + 4 \sum_{j=r+1}^n \sum_{k=j+1}^n \frac{c_k^2 c_j^2 \omega_p^2}{c_p^2 (\omega_j^2 - \omega_p^2) (\omega_k^2 - \omega_p^2)} \right] = \end{aligned}$$

$$\begin{aligned}
&= \varepsilon \left[\sum_{j=r+1}^n \frac{(c_p^2 + c_j^2)[\omega_p^2 c_p^2 + 2\omega_p^2 c_j^2 + \omega_j^2 c_p^2]}{c_p^2(\omega_j^2 - \omega_p^2)^2} - 4 \sum_{j=1}^{p-1} \sum_{k=r+1}^n \frac{c_k^2 c_j^2 \omega_p^2}{c_p^2(\omega_p^2 - \omega_j^2)(\omega_k^2 - \omega_p^2)} \right. \\
&\quad \left. + 4 \sum_{j=p+1}^r \sum_{k=r+1}^n \frac{c_k^2 c_j^2 \omega_p^2}{c_p^2(\omega_j^2 - \omega_p^2)(\omega_k^2 - \omega_p^2)} + 4 \sum_{j=r+1}^n \sum_{k=j+1}^n \frac{c_k^2 c_j^2 \omega_p^2}{c_p^2(\omega_j^2 - \omega_p^2)(\omega_k^2 - \omega_p^2)} \right] \\
&= \varepsilon \left[\sum_{j=r+1}^n \frac{(c_p^2 + c_j^2)^2 \omega_p^2}{c_p^2(\omega_j^2 - \omega_p^2)^2} + \sum_{j=r+1}^n \frac{(c_p^2 + c_j^2)(\omega_p^2 c_j^2 + \omega_j^2 c_p^2)}{c_p^2(\omega_j^2 - \omega_p^2)^2} \right. \\
&\quad - 4 \sum_{j=1}^{p-1} \sum_{k=r+1}^n \frac{c_k^2 c_j^2 \omega_p^2}{c_p^2(\omega_p^2 - \omega_j^2)(\omega_k^2 - \omega_p^2)} + 4 \sum_{j=p+1}^r \sum_{k=r+1}^n \frac{c_k^2 c_j^2 \omega_p^2}{c_p^2(\omega_j^2 - \omega_p^2)(\omega_k^2 - \omega_p^2)} \\
&\quad \left. + 4 \sum_{j=r+1}^n \sum_{k=j+1}^n \frac{c_k^2 c_j^2 \omega_p^2}{c_p^2(\omega_j^2 - \omega_p^2)(\omega_k^2 - \omega_p^2)} \right].
\end{aligned}$$

Damit ist

$$\begin{aligned}
&\text{Tr}(XZ_s) - \text{Tr}(\tilde{X}\tilde{Z}_s) = \\
&\varepsilon \sum_{p=1}^s \left[\sum_{j=r+1}^n \frac{(c_p^2 + c_j^2)^2 \omega_p^2}{c_p^2(\omega_j^2 - \omega_p^2)^2} + \sum_{j=r+1}^n \frac{(c_p^2 + c_j^2)(\omega_p^2 c_j^2 + \omega_j^2 c_p^2)}{c_p^2(\omega_j^2 - \omega_p^2)^2} \right. \\
&\quad - 4 \sum_{j=1}^{p-1} \sum_{k=r+1}^n \frac{c_k^2 c_j^2 \omega_p^2}{c_p^2(\omega_p^2 - \omega_j^2)(\omega_k^2 - \omega_p^2)} + 4 \sum_{j=p+1}^r \sum_{k=r+1}^n \frac{c_k^2 c_j^2 \omega_p^2}{c_p^2(\omega_j^2 - \omega_p^2)(\omega_k^2 - \omega_p^2)} \\
&\quad \left. + 4 \sum_{j=r+1}^n \sum_{k=j+1}^n \frac{c_k^2 c_j^2 \omega_p^2}{c_p^2(\omega_j^2 - \omega_p^2)(\omega_k^2 - \omega_p^2)} \right]. \tag{4.9}
\end{aligned}$$

Im folgenden setzen wir o.B.d.A. $\varepsilon = 1$, da wir über alle c_k und ω_k , also alle Dämpfungen und Frequenzspektren variieren wollen. Eine Betrachtung der Viskosität ε , die (4.6) minimiert, folgt in Abschnitt 4.4, in dem wir viele der hier gewonnenen Ergebnisse übernehmen werden können.

Dann ist nach (4.6)

$$\begin{aligned}
&\text{Tr}(XZ_s) = \tag{4.10} \\
&\sum_{p=1}^s \left[\frac{2}{c_p^2} + \frac{c_p^2}{2\omega_p^2} + 2 \frac{\omega_p^2}{c_p^2} \left(\sum_{\substack{k=1 \\ k \neq p}}^n \frac{c_k^2}{\omega_k^2 - \omega_p^2} \right)^2 + \sum_{\substack{j=1 \\ j \neq p}}^n \frac{3\omega_p^2 c_j^2 + \omega_p^2 c_p^2 + \omega_j^2 c_j^2 + \omega_j^2 c_p^2}{(\omega_p^2 - \omega_j^2)^2} \right]
\end{aligned}$$

und für den relativen Fehler der Spur ergibt sich mit (4.9) und (4.10)

Satz 4.4

Sei X die eindeutig bestimmte Lösung der Ljapunovgleichung

$$A^T X + X A = -I$$

mit

$$A = \begin{pmatrix} 0 & \Omega \\ -\Omega & -c c^T \end{pmatrix},$$

$\Omega = \text{diag}(\omega_1, \dots, \omega_n)$ mit $0 < \omega_1 < \dots < \omega_n$, $c = (c_1, \dots, c_n)^T$ mit $c_i \neq 0$ für alle $i = 1, \dots, n$, und sei \tilde{X} die eindeutig bestimmte Lösung der Ljapunovgleichung

$$\tilde{A}^T \tilde{X} + \tilde{X} \tilde{A} = -I$$

mit

$$\tilde{A} = \begin{pmatrix} 0 & \tilde{\Omega} \\ -\tilde{\Omega} & -\tilde{c} \tilde{c}^T \end{pmatrix},$$

$\tilde{\Omega} = \text{diag}(\omega_1, \dots, \omega_r)$, $\tilde{c} = (c_1, \dots, c_r)^T$ mit $r < n$.

Sei

$$Z_s = \begin{pmatrix} I_s & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & I_s & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

eine Matrix der Dimension $2n$, wobei I_s die s -dimensionale Einheitsmatrix mit $s \leq r < n$ ist. \tilde{Z}_s sei die entsprechende $2r$ - dimensionale Matrix.

Dann gilt

$$\left| \frac{\text{Tr}(X Z_s) - \text{Tr}(\tilde{X} \tilde{Z}_s)}{\text{Tr}(X Z_s)} \right| = \frac{Z_{rF}}{N_{rF}} \quad (4.11)$$

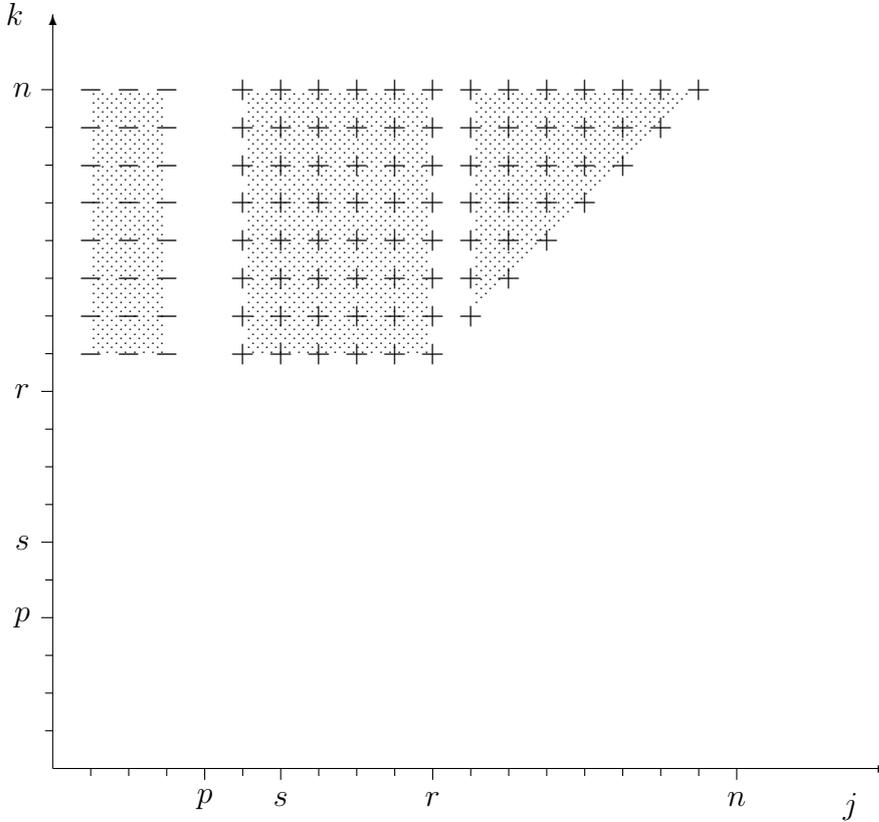
mit

$$\begin{aligned} Z_{rF} &= \left| \sum_{p=1}^s \left(\sum_{j=r+1}^n \frac{(c_p^2 + c_j^2)^2 \omega_p^2}{c_p^2 (\omega_j^2 - \omega_p^2)^2} + \sum_{j=r+1}^n \frac{(c_p^2 + c_j^2)(\omega_p^2 c_j^2 + \omega_j^2 c_p^2)}{c_p^2 (\omega_j^2 - \omega_p^2)^2} \right. \right. \\ &\quad \left. \left. + 4 \sum_{\substack{j=1 \\ j \neq p}}^r \sum_{k=r+1}^n S_{pjk} + 4 \sum_{j=r+1}^n \sum_{k=j+1}^n S_{pjk} \right) \right|, \\ N_{rF} &= \sum_{p=1}^s \left[\frac{2}{c_p^2} + \frac{c_p^2}{2\omega_p^2} + 2 \frac{\omega_p^2}{c_p^2} \left(\sum_{\substack{k=1 \\ k \neq p}}^n \frac{c_k^2}{\omega_k^2 - \omega_p^2} \right)^2 + \sum_{\substack{j=1 \\ j \neq p}}^n \frac{3\omega_p^2 c_j^2 + \omega_p^2 c_p^2 + \omega_j^2 c_j^2 + \omega_j^2 c_p^2}{(\omega_p^2 - \omega_j^2)^2} \right]. \end{aligned}$$

Dabei ist

$$S_{pjk} = \frac{c_k^2 c_j^2 \omega_p^2}{c_p^2 (\omega_j^2 - \omega_p^2) (\omega_k^2 - \omega_p^2)} \quad \text{mit } k > j \quad \text{und} \quad S_{pjk} \begin{cases} > 0 & \text{für } j > p \\ < 0 & \text{für } j < p \end{cases}.$$

Die folgende Zeichnung veranschaulicht die Summierung der S_{pjk} für ein festes p . Die Symbole $+$ und $-$ kennzeichnen das Vorzeichen der S_{pjk} .



Aus (4.11) folgt unmittelbar

Korollar 4.5

Es gelten die Voraussetzungen von Satz 4.4.

Sei $I = I_{rel} \cup I_{zw}$ mit $I_{rel} \subset \{1, \dots, s\}$, $I_{zw} \subset \{s+1, \dots, r\}$ und $c_q := \beta \cdot c_{oq}$, $\beta > 0$ für jedes $q \in I$.

(i) Für $\beta \rightarrow \infty$ gilt

$$\left| \frac{\text{Tr}(XZ_s) - \text{Tr}(\tilde{X}\tilde{Z}_s)}{\text{Tr}(XZ_s)} \right| = \frac{|P_1|}{Q_1} \beta^{-2} + O(\beta^{-4}),$$

falls $Q_1 \neq 0$, und

$$\left| \frac{\text{Tr}(XZ_s) - \text{Tr}(\tilde{X}\tilde{Z}_s)}{\text{Tr}(XZ_s)} \right| = \frac{|P_1|}{Q_2} + O(\beta^{-2}),$$

falls $Q_1 = 0$, wobei

$$Q_1 = 2 \sum_{\substack{p=1 \\ p \notin I_{rel}}}^s \frac{\omega_p^2}{c_p^2} \left(\sum_{q \in I} \frac{c_{oq}^2}{\omega_q^2 - \omega_p^2} \right)^2,$$

$$P_1 = \sum_{q \in I_{rel}} \sum_{j=r+1}^n \frac{c_{oq}^2 (\omega_q^2 + \omega_j^2)}{(\omega_j^2 - \omega_q^2)^2} + 4 \sum_{\substack{p=1 \\ p \notin I}}^s \sum_{q \in I} \sum_{\substack{k=r+1 \\ k > j}}^n \frac{c_{oq}^2 c_k^2 \omega_p^2}{c_p^2 (\omega_q^2 - \omega_p^2) (\omega_k^2 - \omega_p^2)},$$

$$Q_2 > 0.$$

(ii) Für $\beta \rightarrow 0$ gilt

$$\left| \frac{\text{Tr}(XZ_s) - \text{Tr}(\tilde{X}\tilde{Z}_s)}{\text{Tr}(XZ_s)} \right| = L + O(\beta^2),$$

wobei

$$L = \frac{|P_3|}{Q_4}$$

mit

$$P_3 = \sum_{q \in I_{rel}} \frac{1}{c_{oq}^2} \left[\sum_{j=r+1}^n \frac{2c_j^4 \omega_q^2}{(\omega_j^2 - \omega_q^2)^2} + 4 \sum_{\substack{j=1 \\ j \notin I}}^n \sum_{\substack{k=r+1 \\ k > j}}^n \frac{c_j^2 c_k^2 \omega_q^2}{(\omega_j^2 - \omega_q^2) (\omega_k^2 - \omega_q^2)} \right],$$

$$Q_4 = \sum_{q \in I_{rel}} \left[\frac{2}{c_{oq}^2} + \frac{2\omega_q^2}{c_{oq}^2} \left(\sum_{\substack{k=1 \\ k \notin I}}^n \frac{c_k^2}{\omega_k^2 - \omega_q^2} \right)^2 \right] > 0$$

für $I_{rel} \neq \emptyset$ und

$L \geq 0$ für $I_{rel} = \emptyset$.

Beweis

Es gilt mit den Bezeichnungen aus Satz 4.4

$$Z_{rF} = \left| P_1 \beta^2 + P_2 + \frac{P_3}{\beta^2} \right|,$$

$$N_{rF} = Q_1 \beta^4 + Q_2 \beta^2 + Q_3 + \frac{Q_4}{\beta^2}$$

mit P_1, Q_1, Q_2, P_3, Q_4 wie oben. Der Fall $Q_1 = 0, Q_2 \leq 0$ kann nicht auftreten, da dann N_{rF} für $\beta \rightarrow \infty$ beschränkt wäre oder negative Werte annehmen würde, was nicht der Fall ist (vgl. Satz 4.4). Im Fall $I_{rel} = \emptyset$ ist $L = \frac{|P_2|}{Q_3}$, wobei $Q_3 \geq \sum_{p=1}^s \frac{2}{c_p^2} > 0$ ist.

Q.E.D.

Die Aussagen von Korollar 4.5 und alle folgenden Aussagen sind in dem Sinne zu verstehen, daß alle Frequenzen und Dämpfungskomponenten, über die keine Aussagen gemacht werden, asymptotisch konstant bleiben.

Bemerkung 4.6

Korollar 4.5 beschreibt den Einfluß einer Skalierung einzelner Dämpfungskomponenten, die im relevanten Bereich und im Zwischenbereich liegen (genauer: deren Indices im relevanten Bereich bzw. Zwischenbereich liegen):

(i) Durch eine genügende simultane Vergrößerung der Dämpfungskomponenten c_q mit $q \in I$ kann der relative Fehler der Spur beliebig klein gemacht werden, wenn $Q_1 \neq 0$ ist. Das ist stets der Fall, wenn die Indexmenge I nur aus einem Element besteht (vgl. Beispiel 4.19) und dieses Element nicht den gesamten relevanten Bereich darstellt (vgl. Beispiele 4.19 und 4.21 (i)) oder wenn für die Indexmenge I nur $I = I_{zw}$ gilt, d.h. sie nur aus Elementen des Zwischenbereichs besteht (vgl. Beispiel 4.20 (i)). Hier ist es dann umso günstiger, je näher die Elemente aus I_{zw} am relevanten Bereich liegen, da dann Q_1 größer wird und gleichzeitig P_1 langsamer wächst als Q_1 (vgl. Beispiele 4.19 (iv),(v)).

Q_1 wird immer gleich Null, falls die Indexmenge I neben einer beliebigen Menge I_{zw} den gesamten relevanten Bereich enthält, d.h. es werden alle Dämpfungskomponenten c_q mit $q \in I_{rel} = \{1, \dots, s\}$ simultan vergrößert (vgl. Beispiel 4.21). In diesem Fall ist auch P_1 stets größer Null und damit führt eine genügende Vergrößerung der Dämpfungskomponenten c_q , $q \in I$ hier zu dem positiven Grenzwert $\frac{|P_1|}{Q_2}$. Ansonsten ist der Fall, daß $Q_1 = 0$ wird, nur in ganz singulären Fällen möglich. Eine notwendige Bedingung hierfür ist, daß jedes ω_p mit $1 \leq p \leq s$, $p \notin I$ isoliert zwischen zwei ω_k , ω_l mit $k, l \in I$ liegt. Auch dann hat der relative Fehler der Spur den Grenzwert $\frac{|P_1|}{Q_2}$ (s. o.), wobei jetzt $P_1 = 0$ gelten kann (vgl. Beispiel 4.20 (ii), das den Fall $Q_1 \neq 0$ zeigt und Beispiel 4.20 (iii), das den Fall $Q_1 = 0$ verdeutlicht).

(ii) Durch eine genügende simultane Verkleinerung der Dämpfungskomponenten c_q mit $q \in I$ konvergiert der relative Fehler der Spur gegen einen Grenzwert $\frac{|P_3|}{Q_4}$ bzw. $\frac{|P_2|}{Q_3}$. Dieser Grenzwert nimmt in nur ganz singulären Fällen den Wert Null an. Er ist stets größer Null, wenn $I = \{1, \dots, s\}$ ist, d.h. wenn alle Dämpfungskomponenten aus dem relevanten Bereich genügend verkleinert werden. (vgl. Beispiel 4.19 (i), (ii) und (iii)).

Korollar 4.7

Es gelten die Voraussetzungen von Satz 4.4.

Sei $I_{hs} \subset \{r+1, \dots, n\}$, $I_{hs} \neq \emptyset$, und $c_q := \beta \cdot c_{oq}$, $\beta > 0$ für jedes $q \in I_{hs}$.

Dann gilt

(i)

$$\left| \frac{\text{Tr}(XZ_s) - \text{Tr}(\tilde{X}\tilde{Z}_s)}{\text{Tr}(XZ_s)} \right| = 1 + O(\beta^{-2}) \quad \text{für } \beta \rightarrow \infty$$

und
(ii)

$$\left| \frac{\operatorname{Tr}(XZ_s) - \operatorname{Tr}(\tilde{X}\tilde{Z}_s)}{\operatorname{Tr}(XZ_s)} \right| = L' + O(\beta^2) \quad \text{für } \beta \rightarrow 0$$

mit $L' \geq 0$.

Beweis

Mit den Bezeichnungen aus Satz 4.4 gilt

$$\begin{aligned} Z_{rF} &= |P_1\beta^4 + P_2\beta^2 + P_3|, \\ N_{rF} &= Q_1\beta^4 + Q_2\beta^2 + Q_3 \end{aligned}$$

mit

$$\begin{aligned} P_1 &= Q_1 = 2 \sum_{p=1}^s \frac{\omega_p^2}{c_p^2} \left(\sum_{q \in I_{hs}} \frac{c_{o_q}^2}{\omega_q^2 - \omega_p^2} \right)^2, \\ Q_3 &> 0. \end{aligned}$$

Q.E.D.

Bemerkung 4.8

Korollar 4.7 beschreibt den Einfluß einer Skalierung einzelner Dämpfungskomponenten, die hinter dem Schnitt liegen:

(i) Der relative Fehler der Spur hat den Grenzwert 1 bei beliebiger simultaner Vergrößerung einer oder mehrerer Dämpfungskomponenten, die hinter dem Schnitt liegen (vgl. Beispiel 4.22).

(ii) Der relative Fehler der Spur hat den Grenzwert $L' = \frac{|P_3|}{Q_3}$ bei beliebiger simultaner Verkleinerung einer oder mehrerer Dämpfungskomponenten hinter dem Schnitt. Dieser Grenzwert nimmt in nur ganz singulären Fällen den Wert Null an. Er ist stets größer Null bei einer genügenden simultanen Verkleinerung aller Dämpfungskomponenten hinter dem Schnitt, da dann

$$P_3 = \sum_{p=1}^s \left(\sum_{j=r+1}^n \frac{c_p^4 \omega_p^2}{c_p^2 (\omega_j^2 - \omega_p^2)^2} + \sum_{j=r+1}^n \frac{c_p^4 \omega_j^2}{c_p^2 (\omega_j^2 - \omega_p^2)^2} \right)$$

gilt (vgl. Beispiel 4.22). Damit haben wir unsere anfängliche Vermutung, daß eine beliebige Verkleinerung aller Dämpfungskomponenten hinter dem Schnitt zu einer beliebigen Verkleinerung des relativen Fehlers der Spur führt, widerlegt.

Um weitere Aussagen über den relativen Fehler der Spur zu machen, wollen wir ihn in Lemma 4.9 und Satz 4.13 abschätzen:

Mit Satz 4.4 erhalten wir

$$\begin{aligned} & \left| \frac{\operatorname{Tr}(XZ_s) - \operatorname{Tr}(\tilde{X}\tilde{Z}_s)}{\operatorname{Tr}(XZ_s)} \right| \\ & \leq \frac{\sum_{p=1}^s \left(\sum_{j=r+1}^n \frac{(c_p^2 + c_j^2)^2 \omega_p^2}{c_p^2 (\omega_j^2 - \omega_p^2)^2} + \sum_{j=r+1}^n \frac{(c_p^2 + c_j^2)(\omega_j^2 c_j^2 + \omega_j^2 c_p^2)}{c_p^2 (\omega_j^2 - \omega_p^2)^2} \right)}{N_{rF}} \\ & \quad + \frac{4 \sum_{p=1}^s \left(\sum_{j=p+1}^r \sum_{k=r+1}^n S_{pjk} + \sum_{j=r+1}^n \sum_{k=j+1}^n S_{pjk} + \sum_{j=1}^{p-1} \sum_{k=r+1}^n |S_{pjk}| \right)}{N_{rF}}, \end{aligned}$$

wobei wir zum einen die negativen S_{pjk} durch ihren Betrag und zum anderen in der zweiten Summe im Zähler ω_p durch ω_j abgeschätzt haben. Also ergibt sich

$$\begin{aligned} & \left| \frac{\operatorname{Tr}(XZ_s) - \operatorname{Tr}(\tilde{X}\tilde{Z}_s)}{\operatorname{Tr}(XZ_s)} \right| \\ & \leq \frac{\sum_{p=1}^s \left(\sum_{j=r+1}^n \frac{(c_p^2 + c_j^2)^2 (\omega_j^2 + \omega_p^2)}{c_p^2 (\omega_j^2 - \omega_p^2)^2} + 4 \sum_{j=p+1}^r \sum_{k=r+1}^n \frac{c_k^2 c_j^2 \omega_p^2}{c_p^2 (\omega_j^2 - \omega_p^2)(\omega_k^2 - \omega_p^2)} \right)}{N_{rF}} \quad (4.12) \\ & \quad + \frac{\sum_{p=1}^s \left(4 \sum_{j=r+1}^n \sum_{k=j+1}^n \frac{c_k^2 c_j^2 \omega_p^2}{c_p^2 (\omega_j^2 - \omega_p^2)(\omega_k^2 - \omega_p^2)} + 4 \sum_{j=1}^{p-1} \sum_{k=r+1}^n \frac{c_j^2 c_k^2 \omega_p^2}{c_p^2 (\omega_p^2 - \omega_j^2)(\omega_k^2 - \omega_p^2)} \right)}{N_{rF}}. \end{aligned}$$

Lemma 4.9

Es gelten die Voraussetzungen von Satz 4.4.

Dann ist

$$\begin{aligned} \left| \frac{\operatorname{Tr}(XZ_s) - \operatorname{Tr}(\tilde{X}\tilde{Z}_s)}{\operatorname{Tr}(XZ_s)} \right| & \leq \max_{1 \leq i \leq n} (c_i^4) \left[\sum_{p=1}^s \left(\sum_{j=r+1}^n \frac{2}{(\omega_j - \omega_p)^2} + \sum_{j=p+1}^r \frac{(n-r)}{2(\omega_j - \omega_p)^2} \right. \right. \\ & \quad \left. \left. + \sum_{j=r+1}^n \frac{(n-j)}{2(\omega_j - \omega_p)^2} + 2 \sum_{j=1}^{p-1} \sum_{k=r+1}^n \frac{\omega_p^2}{(\omega_p^2 - \omega_j^2)(\omega_k^2 - \omega_p^2)} \right) \right]. \end{aligned}$$

Beweis

Es gilt mit (4.11)

$$N_{rF} \geq \sum_{p=1}^s \frac{2}{c_p^2},$$

und damit erhalten wir nach (4.12) für den relativen Fehler der Spur

$$\begin{aligned}
& \left| \frac{\text{Tr}(XZ_s) - \text{Tr}(\tilde{X}\tilde{Z}_s)}{\text{Tr}(XZ_s)} \right| \\
& \leq \frac{\sum_{p=1}^s \left(\sum_{j=r+1}^n \frac{(c_p^2 + c_j^2)^2 (\omega_j^2 + \omega_p^2)}{c_p^2 (\omega_j^2 - \omega_p^2)^2} + 4 \sum_{j=p+1}^r \sum_{k=r+1}^n \frac{c_k^2 c_j^2 \omega_p^2}{c_p^2 (\omega_j^2 - \omega_p^2) (\omega_k^2 - \omega_p^2)} \right)}{\sum_{p=1}^s \frac{2}{c_p^2}} \\
& \quad + \frac{\sum_{p=1}^s \left(4 \sum_{j=r+1}^n \sum_{k=j+1}^n \frac{c_k^2 c_j^2 \omega_p^2}{c_p^2 (\omega_j^2 - \omega_p^2) (\omega_k^2 - \omega_p^2)} + 4 \sum_{j=1}^{p-1} \sum_{k=r+1}^n \frac{c_j^2 c_k^2 \omega_p^2}{c_p^2 (\omega_p^2 - \omega_j^2) (\omega_k^2 - \omega_p^2)} \right)}{\sum_{p=1}^s \frac{2}{c_p^2}} \\
& \leq \sum_{p=1}^s \left[\sum_{j=r+1}^n \frac{(c_p^2 + c_j^2)^2 (\omega_j^2 + \omega_p^2)}{2(\omega_j^2 - \omega_p^2)^2} + 2 \sum_{j=p+1}^r \sum_{k=r+1}^n \frac{c_k^2 c_j^2 \omega_p^2}{(\omega_j^2 - \omega_p^2) (\omega_k^2 - \omega_p^2)} \right. \\
& \quad \left. + 2 \sum_{j=r+1}^n \sum_{k=j+1}^n \frac{c_k^2 c_j^2 \omega_p^2}{(\omega_j^2 - \omega_p^2) (\omega_k^2 - \omega_p^2)} + 2 \sum_{j=1}^{p-1} \sum_{k=r+1}^n \frac{c_j^2 c_k^2 \omega_p^2}{(\omega_p^2 - \omega_j^2) (\omega_k^2 - \omega_p^2)} \right]. \tag{4.13}
\end{aligned}$$

Also erhalten wir

$$\begin{aligned}
& \left| \frac{\text{Tr}(XZ_s) - \text{Tr}(\tilde{X}\tilde{Z}_s)}{\text{Tr}(XZ_s)} \right| \\
& \leq \max_{1 \leq i \leq n} (c_i^4) \left[\sum_{p=1}^s \left(\sum_{j=r+1}^n \frac{2(\omega_j^2 + \omega_p^2)}{(\omega_j^2 - \omega_p^2)^2} + 2 \sum_{j=p+1}^r \sum_{k=r+1}^n \frac{\omega_p^2}{(\omega_j^2 - \omega_p^2) (\omega_k^2 - \omega_p^2)} \right. \right. \\
& \quad \left. \left. + 2 \sum_{j=r+1}^n \sum_{k=j+1}^n \frac{\omega_p^2}{(\omega_j^2 - \omega_p^2) (\omega_k^2 - \omega_p^2)} + 2 \sum_{j=1}^{p-1} \sum_{k=r+1}^n \frac{\omega_p^2}{(\omega_p^2 - \omega_j^2) (\omega_k^2 - \omega_p^2)} \right) \right] \\
& \leq \max_{1 \leq i \leq n} (c_i^4) \left[\sum_{p=1}^s \left(2 \sum_{j=r+1}^n \frac{(\omega_j^2 + \omega_p^2)}{(\omega_j - \omega_p)^2 (\omega_j + \omega_p)^2} + 2 \sum_{j=p+1}^r \frac{\omega_p^2 (n-r)}{(\omega_j^2 - \omega_p^2)^2} \right. \right. \\
& \quad \left. \left. + 2 \sum_{j=r+1}^n \frac{\omega_p^2 (n-j)}{(\omega_j^2 - \omega_p^2)^2} + 2 \sum_{j=1}^{p-1} \sum_{k=r+1}^n \frac{\omega_p^2}{(\omega_p^2 - \omega_j^2) (\omega_k^2 - \omega_p^2)} \right) \right] \\
& \leq \max_{1 \leq i \leq n} (c_i^4) \left[\sum_{p=1}^s \left(\sum_{j=r+1}^n \frac{2}{(\omega_j - \omega_p)^2} + \sum_{j=p+1}^r \frac{(n-r)}{2(\omega_j - \omega_p)^2} \right. \right. \\
& \quad \left. \left. + \sum_{j=r+1}^n \frac{(n-j)}{2(\omega_j - \omega_p)^2} + 2 \sum_{j=1}^{p-1} \sum_{k=r+1}^n \frac{\omega_p^2}{(\omega_p^2 - \omega_j^2) (\omega_k^2 - \omega_p^2)} \right) \right]. \tag{4.14}
\end{aligned}$$

Q.E.D.

Gilt weiter $|\omega_s^2 - \omega_1^2| \leq |\omega_{r+1}^2 - \omega_s^2|$, dann gilt für $1 \leq p \leq s$ auch $|\omega_p^2 - \omega_j^2| \leq |\omega_k^2 - \omega_p^2|$ für jedes $j = 1, \dots, p-1$ und $k = r+1, \dots, n$.

Damit vereinfacht sich (4.14) zu

$$\begin{aligned}
& \left| \frac{\text{Tr}(XZ_s) - \text{Tr}(\tilde{X}\tilde{Z}_s)}{\text{Tr}(XZ_s)} \right| \\
& \leq \max_{1 \leq i \leq n} (c_i^4) \left[\sum_{p=1}^s \left(\sum_{j=r+1}^n \frac{2}{(\omega_j - \omega_p)^2} + \sum_{j=p+1}^r \frac{(n-r)}{2(\omega_j - \omega_p)^2} + \sum_{j=r+1}^n \frac{(n-j)}{2(\omega_j - \omega_p)^2} \right. \right. \\
& \quad \left. \left. + \sum_{j=1}^{p-1} \frac{2(n-r)}{(\omega_p - \omega_j)^2} \right) \right] \\
& = \max_{1 \leq i \leq n} (c_i^4) \left[\sum_{p=1}^s \left(\sum_{j=r+1}^n \frac{2}{(\omega_j - \omega_p)^2} + \sum_{\substack{j=1 \\ j \neq p}}^r \frac{2(n-r)}{(\omega_j - \omega_p)^2} + \sum_{j=r+1}^n \frac{(n-j)}{2(\omega_j - \omega_p)^2} \right) \right] \\
& = \max_{1 \leq i \leq n} (c_i^4) \left[\sum_{p=1}^s \left(2(n-r) \sum_{\substack{j=1 \\ j \neq p}}^r \frac{1}{(\omega_j - \omega_p)^2} + \sum_{j=r+1}^n \frac{4+n-j}{2(\omega_j - \omega_p)^2} \right) \right].
\end{aligned}$$

Lemma 4.9 stellt den Einfluß einer Skalierung von allen Dämpfungskomponenten und allen Frequenzen auf den relativen Fehler der Spur dar:

Korollar 4.10

Wenn alle Dämpfungskomponenten mit einem Faktor β und alle Frequenzen mit einem Faktor γ skaliert werden, gilt für $\beta \rightarrow 0$ bzw. $\gamma \rightarrow \infty$

$$\left| \frac{\text{Tr}(XZ_s) - \text{Tr}(\tilde{X}\tilde{Z}_s)}{\text{Tr}(XZ_s)} \right| = O(\beta^4) \quad \text{bzw.} \quad \left| \frac{\text{Tr}(XZ_s) - \text{Tr}(\tilde{X}\tilde{Z}_s)}{\text{Tr}(XZ_s)} \right| = O(\gamma^{-2}).$$

Dagegen bleibt der relative Fehler der Spur mit $\beta \rightarrow \infty$ bzw. $\gamma \rightarrow 0$ asymptotisch konstant.

Beweis

Die ersten beiden Aussagen folgen sofort aus Lemma 4.9, die letzten beiden aus (4.11).

Q.E.D.

Die Beispiele 4.23 und 4.24 verdeutlichen die Aussagen des Korollars 4.10.

Abgeschwächt gilt bereits

Korollar 4.11

Sei $I \supset \{1, \dots, s\} \cup \{r+1, \dots, n\}$ und $c_q := \beta \cdot c_{o_q}$, $\beta > 0$ für jedes $q \in I$. Dann gilt mit den Voraussetzungen von Satz 4.4

$$\left| \frac{\text{Tr}(XZ_s) - \text{Tr}(\tilde{X}\tilde{Z}_s)}{\text{Tr}(XZ_s)} \right| = O(\beta^2) \quad \text{für } \beta \rightarrow 0.$$

Beweis

Folgt unmittelbar aus (4.13).

Q.E.D.

Bemerkung 4.12

In Korollar 4.5 (ii) bzw. Bemerkung 4.6 (ii) sowie in Korollar 4.7 (ii) bzw. Bemerkung 4.8 (ii) wurde gezeigt, daß eine beliebige Verkleinerung aller Dämpfungskomponenten im relevanten Bereich oder aller Dämpfungskomponenten hinter dem Schnitt allein nicht zu einer beliebigen Verkleinerung des relativen Fehlers der Spur führt. Dagegen zeigt nun Korollar 4.11, daß durch eine beliebige Verkleinerung aller Dämpfungskomponenten im relevanten Bereich und gleichzeitig aller hinter dem Schnitt der relative Fehler der Spur gegen Null konvergiert.

Satz 4.13

Es gelten die Voraussetzungen von Satz 4.4.

Für $1 \leq s \leq r$ gilt dann

$$\begin{aligned} \left| \frac{\text{Tr}(XZ_s) - \text{Tr}(\tilde{X}\tilde{Z}_s)}{\text{Tr}(XZ_s)} \right| &\leq \sum_{p=1}^s \sum_{\substack{j=1 \\ j \neq p}}^r \sum_{k=r+1}^n \frac{c_k^2 |\omega_j^2 - \omega_p^2|}{c_p^2 (\omega_k^2 - \omega_p^2)} \\ &+ \alpha \left(\sum_{p=1}^s \sum_{k=r+1}^n \left[\frac{(c_p^2 + c_k^2)^2}{c_p^2 (\omega_k - \omega_p)^2} + \sum_{l=k+1}^n \frac{c_l^2 c_k^2}{c_p^2 (\omega_k - \omega_p)(\omega_l - \omega_p)} \right] \right) \end{aligned} \quad (4.15)$$

mit

$$\alpha = \begin{cases} \min \left\{ \left\{ 2 \frac{(\omega_m - \omega_s)^2}{c_m^2 + c_s^2}, s < m \leq r \right\} \cup \left\{ \frac{(\omega_m - \omega_{m-1})^2}{(c_m^2 + c_{m-1}^2)}, 1 < m \leq s \right\} \right\}, & r > 1 \\ 2 \frac{\omega_1^2}{c_1^2}, & r = 1 \end{cases}$$

Beweis

Sei zunächst $s = r = 1$.

Dann kann der relative Fehler der Spur nach (4.12) abgeschätzt werden durch

$$\begin{aligned}
& \left| \frac{\text{Tr}(XZ_s) - \text{Tr}(\tilde{X}\tilde{Z}_s)}{\text{Tr}(XZ_s)} \right| \\
& \leq \frac{\sum_{j=2}^n \frac{(c_1^2 + c_j^2)^2 (\omega_j^2 + \omega_1^2)}{c_1^2 (\omega_j^2 - \omega_1^2)^2} + 4 \sum_{j=2}^n \sum_{k=j+1}^n \frac{c_k^2 c_j^2 \omega_1^2}{c_1^2 (\omega_j^2 - \omega_1^2) (\omega_k^2 - \omega_1^2)}}{\frac{2}{c_1^2} + \frac{c_1^2}{2\omega_1^2} + 4 \sum_{j=2}^n \sum_{k=j+1}^n \frac{c_k^2 c_j^2 \omega_1^2}{c_1^2 (\omega_j^2 - \omega_1^2) (\omega_k^2 - \omega_1^2)} + \sum_{j=2}^n \frac{(c_1^2 + c_j^2) (\omega_1^2 c_1^2 + 2\omega_1^2 c_j^2 + \omega_j^2 c_1^2)}{c_1^2 (\omega_j^2 - \omega_1^2)^2}} \\
& \leq \frac{\sum_{j=2}^n \frac{(c_1^2 + c_j^2)^2 (\omega_j^2 + \omega_1^2)}{c_1^2 (\omega_j^2 - \omega_1^2)^2} + 4 \sum_{j=2}^n \sum_{k=j+1}^n \frac{c_k^2 c_j^2 \omega_1^2}{c_1^2 (\omega_j^2 - \omega_1^2) (\omega_k^2 - \omega_1^2)}}{\frac{c_1^2}{2\omega_1^2}} \\
& = 2 \sum_{j=2}^n \frac{(c_1^2 + c_j^2)^2}{c_1^4} \cdot \frac{(\omega_j^2 + \omega_1^2)}{(\omega_j + \omega_1)^2} \cdot \frac{\omega_1^2}{(\omega_j - \omega_1)^2} \\
& \quad + 8 \sum_{j=2}^n \sum_{k=j+1}^n \frac{c_k^2 c_j^2}{c_1^4} \cdot \frac{\omega_1^4}{(\omega_j + \omega_1)(\omega_k + \omega_1)} \cdot \frac{1}{(\omega_j - \omega_1)(\omega_k - \omega_1)} \\
& \leq 2 \sum_{j=2}^n \frac{(c_1^2 + c_j^2)^2 \omega_1^2}{c_1^4 (\omega_j - \omega_1)^2} + 2 \sum_{j=2}^n \sum_{k=j+1}^n \frac{\omega_1^2 c_k^2 c_j^2}{c_1^4 (\omega_j - \omega_1)(\omega_k - \omega_1)} \\
& = 2 \frac{\omega_1^2}{c_1^2} \left(\sum_{j=2}^n \left[\frac{(c_1^2 + c_j^2)^2}{c_1^2 (\omega_j - \omega_1)^2} + \sum_{k=j+1}^n \frac{c_k^2 c_j^2}{c_1^2 (\omega_j - \omega_1)(\omega_k - \omega_1)} \right] \right).
\end{aligned}$$

Damit gilt die Behauptung des Satzes für diesen Fall.

Sei im folgenden nun $r > 1$.

Damit ist $s > 1$ oder $s < r$, und mit (4.11) gilt

$$\begin{aligned}
N_{rF} & \geq \sum_{p=1}^s \sum_{\substack{j=1 \\ j \neq p}}^n \frac{3\omega_p^2 c_j^2 + \omega_p^2 c_p^2 + \omega_j^2 c_j^2 + \omega_j^2 c_p^2}{(\omega_p^2 - \omega_j^2)^2} \\
& \geq \sum_{p=1}^s \sum_{\substack{j=1 \\ j \neq p}}^r \frac{3\omega_p^2 c_j^2 + \omega_p^2 c_p^2 + \omega_j^2 c_j^2 + \omega_j^2 c_p^2}{(\omega_p^2 - \omega_j^2)^2} \\
& = \sum_{p=1}^s \sum_{j=p+1}^s \left(\frac{3\omega_p^2 c_j^2 + \omega_p^2 c_p^2 + \omega_j^2 c_j^2 + \omega_j^2 c_p^2}{(\omega_j^2 - \omega_p^2)^2} + \frac{3\omega_j^2 c_p^2 + \omega_j^2 c_j^2 + \omega_p^2 c_p^2 + \omega_p^2 c_j^2}{(\omega_j^2 - \omega_p^2)^2} \right) \\
& \quad + \sum_{p=1}^s \sum_{j=s+1}^r \frac{3\omega_p^2 c_j^2 + \omega_p^2 c_p^2 + \omega_j^2 c_j^2 + \omega_j^2 c_p^2}{(\omega_j^2 - \omega_p^2)^2} =
\end{aligned}$$

$$\begin{aligned}
&= \sum_{p=1}^s \sum_{j=p+1}^s \frac{4\omega_p^2 c_j^2 + 4\omega_j^2 c_p^2 + 2\omega_p^2 c_p^2 + 2\omega_j^2 c_j^2}{(\omega_j^2 - \omega_p^2)^2} \\
&\quad + \sum_{p=1}^s \sum_{j=s+1}^r \frac{3\omega_p^2 c_j^2 + \omega_p^2 c_p^2 + \omega_j^2 c_j^2 + \omega_j^2 c_p^2}{(\omega_j^2 - \omega_p^2)^2} \\
&=: \hat{N}_{rF} > 0.
\end{aligned} \tag{4.16}$$

Damit kann der relative Fehler der Spur abgeschätzt werden.

Mit (4.12) ist

$$\begin{aligned}
&\left| \frac{\text{Tr}(XZ_s) - \text{Tr}(\tilde{X}\tilde{Z}_s)}{\text{Tr}(XZ_s)} \right| \\
&\leq \frac{\sum_{p=1}^s \left(\sum_{j=r+1}^n \frac{(c_p^2 + c_j^2)^2 (\omega_j^2 + \omega_p^2)}{c_p^2 (\omega_j^2 - \omega_p^2)^2} + 4 \sum_{j=p+1}^r \sum_{k=r+1}^n \frac{c_k^2 c_j^2 \omega_p^2}{c_p^2 (\omega_j^2 - \omega_p^2) (\omega_k^2 - \omega_p^2)} \right)}{\hat{N}_{rF}} \\
&\quad + 4 \frac{\sum_{p=1}^s \left(\sum_{j=r+1}^n \sum_{k=j+1}^n \frac{c_k^2 c_j^2 \omega_p^2}{c_p^2 (\omega_j^2 - \omega_p^2) (\omega_k^2 - \omega_p^2)} + \sum_{j=1}^{p-1} \sum_{k=r+1}^n \frac{c_j^2 c_k^2 \omega_p^2}{c_p^2 (\omega_p^2 - \omega_j^2) (\omega_k^2 - \omega_p^2)} \right)}{\hat{N}_{rF}} \\
&= \frac{\sum_{p=1}^s \left(\sum_{j=r+1}^n \frac{(c_p^2 + c_j^2)^2 (\omega_j^2 + \omega_p^2)}{c_p^2 (\omega_j^2 - \omega_p^2)^2} + 4 \sum_{j=r+1}^n \sum_{k=j+1}^n \frac{c_k^2 c_j^2 \omega_p^2}{c_p^2 (\omega_j^2 - \omega_p^2) (\omega_k^2 - \omega_p^2)} \right)}{\hat{N}_{rF}} \\
&\quad + 4 \frac{\sum_{p=1}^s \left(\sum_{j=p+1}^s \sum_{k=r+1}^n \frac{c_k^2 c_j^2 \omega_p^2}{c_p^2 (\omega_j^2 - \omega_p^2) (\omega_k^2 - \omega_p^2)} + \sum_{j=1}^{p-1} \sum_{k=r+1}^n \frac{c_j^2 c_k^2 \omega_p^2}{c_p^2 (\omega_p^2 - \omega_j^2) (\omega_k^2 - \omega_p^2)} \right)}{\hat{N}_{rF}} \tag{4.17} \\
&\quad + 4 \frac{\sum_{p=1}^s \left(\sum_{j=s+1}^r \sum_{k=r+1}^n \frac{c_k^2 c_j^2 \omega_p^2}{c_p^2 (\omega_j^2 - \omega_p^2) (\omega_k^2 - \omega_p^2)} \right)}{\hat{N}_{rF}}.
\end{aligned}$$

Wir unterscheiden die folgenden Fälle:

Fall 1: $1 < s < r$

Mit (4.17) ist

$$\left| \frac{\text{Tr}(XZ_s) - \text{Tr}(\tilde{X}\tilde{Z}_s)}{\text{Tr}(XZ_s)} \right| \leq$$

$$\begin{aligned}
&\leq \frac{\sum_{p=1}^s \left(\sum_{j=r+1}^n \frac{(c_p^2+c_j^2)^2(\omega_j^2+\omega_p^2)}{c_p^2(\omega_j^2-\omega_p^2)^2} + 4 \sum_{j=r+1}^n \sum_{k=j+1}^n \frac{c_k^2 c_j^2 \omega_p^2}{c_p^2(\omega_j^2-\omega_p^2)(\omega_k^2-\omega_p^2)} \right)}{\hat{N}_{rF}} \\
&+ 4 \frac{\sum_{p=1}^s \left(\sum_{j=p+1}^s \sum_{k=r+1}^n \frac{c_k^2 c_j^2 \omega_p^2}{c_p^2(\omega_j^2-\omega_p^2)(\omega_k^2-\omega_p^2)} + \sum_{j=1}^{p-1} \sum_{k=r+1}^n \frac{c_j^2 c_k^2 \omega_p^2}{c_p^2(\omega_p^2-\omega_j^2)(\omega_k^2-\omega_p^2)} \right)}{\hat{N}_{rF}} \\
&+ 4 \frac{\sum_{p=1}^s \left(\sum_{j=s+1}^r \sum_{k=r+1}^n \frac{c_k^2 c_j^2 \omega_p^2}{c_p^2(\omega_j^2-\omega_p^2)(\omega_k^2-\omega_p^2)} \right)}{\sum_{p=1}^s \sum_{j=s+1}^r \frac{(3\omega_p^2+\omega_j^2)c_j^2}{(\omega_j^2-\omega_p^2)^2}} \\
&\leq \frac{\sum_{p=1}^s \left(\sum_{j=r+1}^n \frac{(c_p^2+c_j^2)^2(\omega_j^2+\omega_p^2)}{c_p^2(\omega_j^2-\omega_p^2)^2} + 4 \sum_{j=r+1}^n \sum_{k=j+1}^n \frac{c_k^2 c_j^2 \omega_p^2}{c_p^2(\omega_j^2-\omega_p^2)(\omega_k^2-\omega_p^2)} \right)}{\hat{N}_{rF}} \\
&+ 4 \frac{\sum_{p=1}^s \left(\sum_{j=p+1}^s \sum_{k=r+1}^n \frac{c_k^2 c_j^2 \omega_p^2}{c_p^2(\omega_j^2-\omega_p^2)(\omega_k^2-\omega_p^2)} + \sum_{j=1}^{p-1} \sum_{k=r+1}^n \frac{c_j^2 c_k^2 \omega_p^2}{c_p^2(\omega_p^2-\omega_j^2)(\omega_k^2-\omega_p^2)} \right)}{\hat{N}_{rF}} \\
&+ 4 \frac{\sum_{p=1}^s \left(\sum_{j=s+1}^r \sum_{k=r+1}^n \frac{c_k^2 c_j^2 \omega_p^2}{c_p^2(\omega_j^2-\omega_p^2)(\omega_k^2-\omega_p^2)} \right)}{\sum_{p=1}^s \sum_{j=s+1}^r \frac{4\omega_p^2 c_j^2}{(\omega_j^2-\omega_p^2)^2}} \\
&\leq \frac{\sum_{p=1}^s \left(\sum_{j=r+1}^n \frac{(c_p^2+c_j^2)^2(\omega_j^2+\omega_p^2)}{c_p^2(\omega_j^2-\omega_p^2)^2} + 4 \sum_{j=r+1}^n \sum_{k=j+1}^n \frac{c_k^2 c_j^2 \omega_p^2}{c_p^2(\omega_j^2-\omega_p^2)(\omega_k^2-\omega_p^2)} \right)}{\hat{N}_{rF}} \\
&+ 4 \frac{\sum_{p=1}^s \left(\sum_{j=p+1}^s \sum_{k=r+1}^n \frac{c_k^2 c_j^2 \omega_p^2}{c_p^2(\omega_j^2-\omega_p^2)(\omega_k^2-\omega_p^2)} + \sum_{j=1}^{p-1} \sum_{k=r+1}^n \frac{c_j^2 c_k^2 \omega_p^2}{c_p^2(\omega_p^2-\omega_j^2)(\omega_k^2-\omega_p^2)} \right)}{\hat{N}_{rF}} \quad (4.18) \\
&+ \sum_{p=1}^s \sum_{j=s+1}^r \sum_{k=r+1}^n \frac{c_k^2(\omega_j^2 - \omega_p^2)}{c_p^2(\omega_k^2 - \omega_p^2)}.
\end{aligned}$$

Mit (4.18) läßt sich der relative Fehler der Spur weiter abschätzen durch

$$\begin{aligned}
& \frac{\sum_{p=1}^s \sum_{j=r+1}^n \frac{(c_p^2+c_j^2)^2(\omega_j^2+\omega_p^2)}{c_p^2(\omega_j^2-\omega_p^2)^2} + 4 \sum_{p=1}^s \sum_{j=r+1}^n \sum_{k=j+1}^n \frac{c_k^2 c_j^2 \omega_p^2}{c_p^2(\omega_j^2-\omega_p^2)(\omega_k^2-\omega_p^2)}}{\sum_{p=1}^s \sum_{j=s+1}^r \frac{3\omega_p^2 c_j^2 + \omega_p^2 c_p^2 + \omega_j^2 c_p^2 + \omega_j^2 c_j^2}{(\omega_j^2-\omega_p^2)^2} + \sum_{p=1}^s \sum_{j=p+1}^s \frac{4\omega_p^2 c_j^2 + 4\omega_j^2 c_p^2 + 2\omega_p^2 c_p^2 + 2\omega_j^2 c_j^2}{(\omega_j^2-\omega_p^2)^2}} \\
& + 4 \frac{\sum_{p=1}^s \sum_{j=p+1}^s \sum_{k=r+1}^n \frac{c_k^2 c_j^2 \omega_p^2}{c_p^2(\omega_j^2-\omega_p^2)(\omega_k^2-\omega_p^2)}}{\sum_{p=1}^s \sum_{j=p+1}^s \frac{4\omega_p^2 c_j^2}{(\omega_j^2-\omega_p^2)^2}} + 4 \frac{\sum_{p=1}^s \sum_{j=1}^{p-1} \sum_{k=r+1}^n \frac{c_j^2 c_k^2 \omega_p^2}{c_p^2(\omega_p^2-\omega_j^2)(\omega_k^2-\omega_p^2)}}{\sum_{p=1}^s \sum_{l=p+1}^s \frac{4\omega_l^2 c_p^2}{(\omega_l^2-\omega_p^2)^2}} \\
& + \sum_{p=1}^s \sum_{j=s+1}^r \sum_{k=r+1}^n \frac{c_k^2(\omega_j^2 - \omega_p^2)}{c_p^2(\omega_k^2 - \omega_p^2)} ;
\end{aligned}$$

hierin gilt weiter

$$\begin{aligned}
4 \frac{\sum_{p=1}^s \sum_{j=1}^{p-1} \sum_{k=r+1}^n \frac{c_j^2 c_k^2 \omega_p^2}{c_p^2(\omega_p^2-\omega_j^2)(\omega_k^2-\omega_p^2)}}{\sum_{p=1}^s \sum_{l=p+1}^s \frac{4\omega_l^2 c_p^2}{(\omega_l^2-\omega_p^2)^2}} &= \frac{4 \sum_{p=1}^s \sum_{j=1}^{p-1} \sum_{k=r+1}^n \frac{c_j^2 c_k^2 \omega_p^2}{c_p^2(\omega_p^2-\omega_j^2)(\omega_k^2-\omega_p^2)}}{\sum_{p=1}^s \sum_{j=1}^{p-1} \frac{4\omega_p^2 c_j^2}{(\omega_p^2-\omega_j^2)^2}} \\
&\leq \sum_{p=1}^s \sum_{j=1}^{p-1} \sum_{k=r+1}^n \frac{c_k^2(\omega_p^2 - \omega_j^2)}{c_p^2(\omega_k^2 - \omega_p^2)} , \\
4 \frac{\sum_{p=1}^s \sum_{j=p+1}^s \sum_{k=r+1}^n \frac{c_k^2 c_j^2 \omega_p^2}{c_p^2(\omega_j^2-\omega_p^2)(\omega_k^2-\omega_p^2)}}{\sum_{p=1}^s \sum_{j=p+1}^s \frac{4\omega_p^2 c_j^2}{(\omega_j^2-\omega_p^2)^2}} &\leq \sum_{p=1}^s \sum_{j=p+1}^s \sum_{k=r+1}^n \frac{c_k^2(\omega_j^2 - \omega_p^2)}{c_p^2(\omega_k^2 - \omega_p^2)} ,
\end{aligned}$$

und insgesamt ergibt sich damit für den relativen Fehler im Fall $1 < s < r$

$$\begin{aligned}
& \left| \frac{\text{Tr}(XZ_s) - \text{Tr}(\tilde{X}\tilde{Z}_s)}{\text{Tr}(XZ_s)} \right| \\
& \leq \sum_{p=1}^s \sum_{j=1}^{p-1} \sum_{k=r+1}^n \frac{c_k^2(\omega_p^2 - \omega_j^2)}{c_p^2(\omega_k^2 - \omega_p^2)} + \sum_{p=1}^s \sum_{j=p+1}^s \sum_{k=r+1}^n \frac{c_k^2(\omega_j^2 - \omega_p^2)}{c_p^2(\omega_k^2 - \omega_p^2)} \\
& + \sum_{p=1}^s \sum_{j=s+1}^r \sum_{k=r+1}^n \frac{c_k^2(\omega_j^2 - \omega_p^2)}{c_p^2(\omega_k^2 - \omega_p^2)} + E + F \\
& = \sum_{p=1}^s \sum_{\substack{j=1 \\ j \neq p}}^r \sum_{k=r+1}^n \frac{c_k^2 |\omega_j^2 - \omega_p^2|}{c_p^2(\omega_k^2 - \omega_p^2)} + E + F , \tag{4.19}
\end{aligned}$$

mit

$$E := \frac{\sum_{p=1}^s \sum_{j=r+1}^n \frac{(c_p^2 + c_j^2)^2 (\omega_j^2 + \omega_p^2)}{c_p^2 (\omega_j^2 - \omega_p^2)^2}}{\sum_{p=1}^s \sum_{l=s+1}^r \frac{3\omega_p^2 c_l^2 + \omega_p^2 c_p^2 + \omega_l^2 c_p^2 + \omega_l^2 c_l^2}{(\omega_l^2 - \omega_p^2)^2} + \sum_{p=1}^s \sum_{j=p+1}^s \frac{4\omega_p^2 c_j^2 + 4\omega_j^2 c_p^2 + 2\omega_p^2 c_p^2 + 2\omega_j^2 c_j^2}{(\omega_j^2 - \omega_p^2)^2}},$$

$$F := 4 \frac{\sum_{p=1}^s \sum_{j=r+1}^n \sum_{k=j+1}^n \frac{c_k^2 c_j^2 \omega_p^2}{c_p^2 (\omega_j^2 - \omega_p^2) (\omega_k^2 - \omega_p^2)}}{\sum_{p=1}^s \sum_{j=s+1}^r \frac{3\omega_p^2 c_j^2 + \omega_p^2 c_p^2 + \omega_j^2 c_p^2 + \omega_j^2 c_j^2}{(\omega_j^2 - \omega_p^2)^2} + \sum_{p=1}^s \sum_{j=p+1}^s \frac{4\omega_p^2 c_j^2 + 4\omega_j^2 c_p^2 + 2\omega_p^2 c_p^2 + 2\omega_j^2 c_j^2}{(\omega_j^2 - \omega_p^2)^2}}.$$

Als nächstes werden wir E und F abschätzen. Zunächst sei m beliebig mit $s < m \leq r$. Dann ist

$$\begin{aligned} E &\leq \frac{\sum_{p=1}^s \sum_{j=r+1}^n \frac{(c_p^2 + c_j^2)^2 (\omega_j^2 + \omega_p^2)}{c_p^2 (\omega_j^2 - \omega_p^2)^2}}{\sum_{p=1}^s \sum_{l=s+1}^r \frac{3\omega_p^2 c_l^2 + \omega_p^2 c_p^2 + \omega_l^2 c_p^2 + \omega_l^2 c_l^2}{(\omega_l^2 - \omega_p^2)^2}} \\ &\leq \frac{\sum_{p=1}^s \sum_{j=r+1}^n \frac{(c_p^2 + c_j^2)^2 (\omega_j^2 + \omega_p^2)}{c_p^2 (\omega_j^2 - \omega_p^2)^2}}{\sum_{p=1}^s \sum_{l=s+1}^r \frac{(\omega_l^2 + \omega_p^2)(c_p^2 + c_l^2)}{(\omega_l^2 - \omega_p^2)^2}} \\ &\leq \frac{\sum_{p=1}^s \sum_{j=r+1}^n \frac{(c_p^2 + c_j^2)^2 (\omega_j^2 + \omega_p^2)}{c_p^2 (\omega_j^2 - \omega_p^2)^2}}{\frac{(\omega_m^2 + \omega_s^2)(c_m^2 + c_s^2)}{(\omega_m^2 - \omega_s^2)^2}} \\ &= \sum_{p=1}^s \sum_{j=r+1}^n \frac{(c_p^2 + c_j^2)^2}{c_p^2 (c_m^2 + c_s^2)} \cdot \frac{(\omega_j^2 + \omega_p^2)}{(\omega_j + \omega_p)^2} \cdot \frac{(\omega_m + \omega_s)^2}{(\omega_m^2 + \omega_s^2)} \cdot \frac{(\omega_m - \omega_s)^2}{(\omega_j - \omega_p)^2} \\ &\leq 2 \sum_{p=1}^s \sum_{j=r+1}^n \frac{(c_p^2 + c_j^2)^2}{c_p^2 (c_m^2 + c_s^2)} \cdot \frac{(\omega_m - \omega_s)^2}{(\omega_j - \omega_p)^2}, \\ F &\leq \frac{4 \sum_{p=1}^s \sum_{j=r+1}^n \sum_{k=j+1}^n \frac{c_k^2 c_j^2 \omega_p^2}{c_p^2 (\omega_j^2 - \omega_p^2) (\omega_k^2 - \omega_p^2)}}{\frac{(\omega_m^2 + \omega_s^2)(c_m^2 + c_s^2)}{(\omega_m^2 - \omega_s^2)^2}} \\ &= 4 \sum_{p=1}^s \sum_{j=r+1}^n \sum_{k=j+1}^n \frac{c_k^2 c_j^2 \omega_p^2 (\omega_m^2 - \omega_s^2)^2}{c_p^2 (c_m^2 + c_s^2) (\omega_m^2 + \omega_s^2) (\omega_j^2 - \omega_p^2) (\omega_k^2 - \omega_p^2)} \\ &= 4 \sum_{p=1}^s \sum_{j=r+1}^n \sum_{k=j+1}^n \frac{c_k^2 c_j^2}{c_p^2 (c_m^2 + c_s^2)} \cdot \frac{\omega_p^2}{(\omega_j + \omega_p)(\omega_k + \omega_p)} \cdot \\ &\quad \frac{(\omega_m + \omega_s)^2}{(\omega_m^2 + \omega_s^2)} \cdot \frac{(\omega_m - \omega_s)^2}{(\omega_j - \omega_p)(\omega_k - \omega_p)} \leq \end{aligned}$$

$$\leq 2 \sum_{p=1}^s \sum_{j=r+1}^n \sum_{k=j+1}^n \frac{c_k^2 c_j^2}{c_p^2 (c_m^2 + c_s^2)} \cdot \frac{(\omega_m - \omega_s)^2}{(\omega_j - \omega_p)(\omega_k - \omega_p)}.$$

Sei nun m beliebig mit $1 < m \leq s$. Dann ist

$$\begin{aligned} E &\leq \frac{\sum_{p=1}^s \sum_{j=r+1}^n \frac{(c_p^2 + c_j^2)^2 (\omega_j^2 + \omega_p^2)}{c_p^2 (\omega_j^2 - \omega_p^2)^2}}{\sum_{p=1}^s \sum_{l=p+1}^s \frac{4\omega_p^2 c_l^2 + 4\omega_l^2 c_p^2 + 2\omega_p^2 c_p^2 + 2\omega_l^2 c_l^2}{(\omega_l^2 - \omega_p^2)^2}} \\ &\leq \frac{\sum_{p=1}^s \sum_{j=r+1}^n \frac{(c_p^2 + c_j^2)^2 (\omega_j^2 + \omega_p^2)}{c_p^2 (\omega_j^2 - \omega_p^2)^2}}{\sum_{p=1}^s \sum_{l=p+1}^s \frac{2(\omega_l^2 + \omega_p^2)(c_l^2 + c_p^2)}{(\omega_l^2 - \omega_p^2)^2}} \\ &\leq \frac{\sum_{p=1}^s \sum_{j=r+1}^n \frac{(c_p^2 + c_j^2)^2 (\omega_j^2 + \omega_p^2)}{c_p^2 (\omega_j^2 - \omega_p^2)^2}}{\frac{2(c_m^2 + c_{m-1}^2)(\omega_m^2 + \omega_{m-1}^2)}{(\omega_m^2 - \omega_{m-1}^2)^2}} \\ &= \sum_{p=1}^s \sum_{j=r+1}^n \frac{(c_p^2 + c_j^2)^2}{2c_p^2 (c_{m-1}^2 + c_m^2)} \cdot \frac{(\omega_j^2 + \omega_p^2)}{(\omega_j + \omega_p)^2} \cdot \frac{(\omega_m + \omega_{m-1})^2}{(\omega_m^2 + \omega_{m-1}^2)} \cdot \frac{(\omega_m - \omega_{m-1})^2}{(\omega_j - \omega_p)^2} \\ &\leq \sum_{p=1}^s \sum_{j=r+1}^n \frac{(c_p^2 + c_j^2)^2}{c_p^2 (c_m^2 + c_{m-1}^2)} \cdot \frac{(\omega_m - \omega_{m-1})^2}{(\omega_j - \omega_p)^2}, \end{aligned}$$

$$\begin{aligned} F &\leq \frac{4 \sum_{p=1}^s \sum_{j=r+1}^n \sum_{k=j+1}^n \frac{c_k^2 c_j^2 \omega_p^2}{c_p^2 (\omega_j^2 - \omega_p^2)(\omega_k^2 - \omega_p^2)}}{\frac{2(c_m^2 + c_{m-1}^2)(\omega_m^2 + \omega_{m-1}^2)}{(\omega_m^2 - \omega_{m-1}^2)^2}} \\ &\leq 4 \sum_{p=1}^s \sum_{j=r+1}^n \sum_{k=j+1}^n \frac{c_k^2 c_j^2}{2c_p^2 (c_m^2 + c_{m-1}^2)} \cdot \frac{\omega_p^2}{(\omega_j + \omega_p)(\omega_k + \omega_p)} \\ &\quad \cdot \frac{(\omega_m + \omega_{m-1})^2}{(\omega_m^2 + \omega_{m-1}^2)} \cdot \frac{(\omega_m - \omega_{m-1})^2}{(\omega_j - \omega_p)(\omega_k - \omega_p)} \\ &\leq \sum_{p=1}^s \sum_{j=r+1}^n \sum_{k=j+1}^n \frac{c_k^2 c_j^2}{c_p^2 (c_m^2 + c_{m-1}^2)} \cdot \frac{(\omega_m - \omega_{m-1})^2}{(\omega_j - \omega_p)(\omega_k - \omega_p)}. \end{aligned}$$

Insgesamt ergibt sich also mit (4.19) für $1 < s < r$ für den relativen Fehler der Spur

a) für $s < m \leq r$

$$\begin{aligned} & \left| \frac{\text{Tr}(XZ_s) - \text{Tr}(\tilde{X}\tilde{Z}_s)}{\text{Tr}(XZ_s)} \right| \\ & \leq \sum_{p=1}^s \sum_{\substack{j=1 \\ j \neq p}}^r \sum_{k=r+1}^n \frac{c_k^2 |\omega_j^2 - \omega_p^2|}{c_p^2 (\omega_k^2 - \omega_p^2)} \\ & \quad + \frac{(\omega_m - \omega_s)^2}{c_m^2 + c_s^2} \left(2 \sum_{p=1}^s \sum_{j=r+1}^n \left[\frac{(c_p^2 + c_j^2)^2}{c_p^2 (\omega_j - \omega_p)^2} + \sum_{k=j+1}^n \frac{c_k^2 c_j^2}{c_p^2 (\omega_j - \omega_p)(\omega_k - \omega_p)} \right] \right), \end{aligned}$$

b) für $1 < m \leq s$

$$\begin{aligned} & \left| \frac{\text{Tr}(XZ_s) - \text{Tr}(\tilde{X}\tilde{Z}_s)}{\text{Tr}(XZ_s)} \right| \\ & \leq \sum_{p=1}^s \sum_{\substack{j=1 \\ j \neq p}}^r \sum_{k=r+1}^n \frac{c_k^2 |\omega_j^2 - \omega_p^2|}{c_p^2 (\omega_k^2 - \omega_p^2)} \\ & \quad + \frac{(\omega_m - \omega_{m-1})^2}{c_m^2 + c_{m-1}^2} \left(\sum_{p=1}^s \sum_{j=r+1}^n \left[\frac{(c_p^2 + c_j^2)^2}{c_p^2 (\omega_j - \omega_p)^2} + \sum_{k=j+1}^n \frac{c_k^2 c_j^2}{c_p^2 (\omega_j - \omega_p)(\omega_k - \omega_p)} \right] \right), \end{aligned}$$

und damit gilt die Behauptung des Satzes.

Es fehlen noch die Fälle $s = 1 < r$ und $s = r > 1$.

Fall 2: $s = 1 < r$

In diesem Fall ist (s. (4.16))

$$\hat{N}_{rF} = \sum_{j=2}^r \frac{3\omega_1^2 c_j^2 + \omega_1^2 c_1^2 + \omega_j^2 c_1^2 + \omega_j^2 c_j^2}{(\omega_j^2 - \omega_1^2)^2} \geq \sum_{j=2}^r \frac{(\omega_j^2 + \omega_1^2)(c_1^2 + c_j^2)}{(\omega_j^2 - \omega_1^2)^2}$$

und der relative Fehler der Spur läßt sich mit (4.18) (das auch jetzt gilt) abschätzen durch

$$\frac{\sum_{j=r+1}^n \frac{(c_1^2 + c_j^2)^2 (\omega_j^2 + \omega_1^2)}{c_1^2 (\omega_j^2 - \omega_1^2)^2} + 4 \sum_{j=r+1}^n \sum_{k=j+1}^n \frac{c_k^2 c_j^2 \omega_1^2}{c_1^2 (\omega_j^2 - \omega_1^2)(\omega_k^2 - \omega_1^2)}}{\sum_{j=2}^r \frac{(\omega_j^2 + \omega_1^2)(c_1^2 + c_j^2)}{(\omega_j^2 - \omega_1^2)^2}} + \sum_{j=2}^r \sum_{k=r+1}^n \frac{c_k^2 (\omega_j^2 - \omega_1^2)}{c_1^2 (\omega_k^2 - \omega_1^2)}.$$

Sei m mit $1 < m \leq r$ beliebig. Dann ergibt sich für den relativen Fehler der Spur

$$\begin{aligned}
& \left| \frac{\text{Tr}(XZ_s) - \text{Tr}(\tilde{X}\tilde{Z}_s)}{\text{Tr}(XZ_s)} \right| \\
& \leq \frac{\sum_{j=r+1}^n \frac{(c_1^2 + c_j^2)^2 (\omega_j^2 + \omega_1^2)}{c_1^2 (\omega_j^2 - \omega_1^2)^2} + 4 \sum_{j=r+1}^n \sum_{k=j+1}^n \frac{c_k^2 c_j^2 \omega_1^2}{c_1^2 (\omega_j^2 - \omega_1^2) (\omega_k^2 - \omega_1^2)}}{\frac{(\omega_m^2 + \omega_1^2)(c_1^2 + c_m^2)}{(\omega_m^2 - \omega_1^2)^2}} + \sum_{j=2}^r \sum_{k=r+1}^n \frac{c_k^2 (\omega_j^2 - \omega_1^2)}{c_1^2 (\omega_k^2 - \omega_1^2)} \\
& = \sum_{j=r+1}^n \frac{(c_1^2 + c_j^2)^2}{c_1^2 (c_1^2 + c_m^2)} \cdot \frac{(\omega_j^2 + \omega_1^2)}{(\omega_j + \omega_1)^2} \cdot \frac{(\omega_m + \omega_1)^2}{(\omega_m^2 + \omega_1^2)} \cdot \frac{(\omega_m - \omega_1)^2}{(\omega_j - \omega_1)^2} \\
& \quad + 4 \sum_{j=r+1}^n \sum_{k=j+1}^n \frac{c_k^2 c_j^2}{c_1^2 (c_1^2 + c_m^2)} \cdot \frac{(\omega_m + \omega_1)^2}{(\omega_m^2 + \omega_1^2)} \cdot \frac{\omega_1^2}{(\omega_j + \omega_1)(\omega_k + \omega_1)} \cdot \frac{(\omega_m - \omega_1)^2}{(\omega_j - \omega_1)(\omega_k - \omega_1)} \\
& \quad + \sum_{j=2}^r \sum_{k=r+1}^n \frac{c_k^2 (\omega_j^2 - \omega_1^2)}{c_1^2 (\omega_k^2 - \omega_1^2)} \\
& \leq 2 \sum_{j=r+1}^n \frac{(c_1^2 + c_j^2)^2 (\omega_m - \omega_1)^2}{c_1^2 (c_1^2 + c_m^2) (\omega_j - \omega_1)^2} + 2 \sum_{j=r+1}^n \sum_{k=j+1}^n \frac{c_k^2 c_j^2 (\omega_m - \omega_1)^2}{c_1^2 (c_1^2 + c_m^2) (\omega_j - \omega_1) (\omega_k - \omega_1)} \\
& \quad + \sum_{j=2}^r \sum_{k=r+1}^n \frac{c_k^2 (\omega_j^2 - \omega_1^2)}{c_1^2 (\omega_k^2 - \omega_1^2)} ;
\end{aligned}$$

damit gilt für diesen Fall für den relativen Fehler der Spur

$$\begin{aligned}
& \left| \frac{\text{Tr}(XZ_1) - \text{Tr}(\tilde{X}\tilde{Z}_1)}{\text{Tr}(XZ_1)} \right| \\
& \leq \sum_{j=2}^r \sum_{k=r+1}^n \frac{c_k^2 (\omega_j^2 - \omega_1^2)}{c_1^2 (\omega_k^2 - \omega_1^2)} \\
& \quad + 2 \frac{(\omega_m - \omega_1)^2}{c_m^2 + c_1^2} \left(\sum_{j=r+1}^n \left[\frac{(c_1^2 + c_j^2)^2}{c_1^2 (\omega_j - \omega_1)^2} + \sum_{k=j+1}^n \frac{c_k^2 c_j^2}{c_1^2 (\omega_j - \omega_1) (\omega_k - \omega_1)} \right] \right),
\end{aligned}$$

also die Behauptung des Satzes.

Fall 3: $s = r > 1$

In diesem Fall ist \hat{N}_{rF} ist gegeben durch

$$\hat{N}_{rF} = \sum_{p=1}^r \sum_{j=p+1}^r \frac{4\omega_p^2 c_j^2 + 4\omega_j^2 c_p^2 + 2\omega_p^2 c_p^2 + 2\omega_j^2 c_j^2}{(\omega_j^2 - \omega_p^2)^2},$$

und damit erhalten wir mit (4.17)

$$\begin{aligned}
& \left| \frac{\text{Tr}(XZ_s) - \text{Tr}(\tilde{X}\tilde{Z}_s)}{\text{Tr}(XZ_s)} \right| \\
& \leq \frac{\sum_{p=1}^r \sum_{j=r+1}^n \frac{(c_p^2+c_j^2)^2(\omega_j^2+\omega_p^2)}{c_p^2(\omega_j^2-\omega_p^2)^2} + 4 \sum_{p=1}^r \sum_{j=r+1}^n \sum_{k=j+1}^n \frac{c_k^2 c_j^2 \omega_p^2}{c_p^2(\omega_j^2-\omega_p^2)(\omega_k^2-\omega_p^2)}}{\sum_{p=1}^r \sum_{j=p+1}^r \frac{4\omega_p^2 c_j^2 + 4\omega_j^2 c_p^2 + 2\omega_p^2 c_p^2 + 2\omega_j^2 c_j^2}{(\omega_j^2-\omega_p^2)^2}} \\
& + 4 \frac{\sum_{p=1}^r \left(\sum_{j=p+1}^r \sum_{k=r+1}^n \frac{c_k^2 c_j^2 \omega_p^2}{c_p^2(\omega_j^2-\omega_p^2)(\omega_k^2-\omega_p^2)} + \sum_{j=1}^{p-1} \sum_{k=r+1}^n \frac{c_j^2 c_k^2 \omega_p^2}{c_p^2(\omega_p^2-\omega_j^2)(\omega_k^2-\omega_p^2)} \right)}{\sum_{p=1}^r \sum_{j=p+1}^r \frac{4\omega_p^2 c_j^2 + 4\omega_j^2 c_p^2 + 2\omega_p^2 c_p^2 + 2\omega_j^2 c_j^2}{(\omega_j^2-\omega_p^2)^2}} \\
& \leq \frac{\sum_{p=1}^r \sum_{j=r+1}^n \frac{(c_p^2+c_j^2)^2(\omega_j^2+\omega_p^2)}{c_p^2(\omega_j^2-\omega_p^2)^2} + 4 \sum_{p=1}^r \sum_{j=r+1}^n \sum_{k=j+1}^n \frac{c_k^2 c_j^2 \omega_p^2}{c_p^2(\omega_j^2-\omega_p^2)(\omega_k^2-\omega_p^2)}}{\sum_{p=1}^r \sum_{j=p+1}^r \frac{4\omega_p^2 c_j^2 + 4\omega_j^2 c_p^2 + 2\omega_p^2 c_p^2 + 2\omega_j^2 c_j^2}{(\omega_j^2-\omega_p^2)^2}} \\
& + 4 \frac{\sum_{p=1}^r \sum_{j=p+1}^r \sum_{k=r+1}^n \frac{c_k^2 c_j^2 \omega_p^2}{c_p^2(\omega_j^2-\omega_p^2)(\omega_k^2-\omega_p^2)}}{\sum_{p=1}^r \sum_{j=p+1}^r \frac{4\omega_p^2 c_j^2}{(\omega_p^2-\omega_j^2)^2}} + 4 \frac{\sum_{p=1}^r \sum_{j=1}^{p-1} \sum_{k=r+1}^n \frac{c_j^2 c_k^2 \omega_p^2}{c_p^2(\omega_p^2-\omega_j^2)(\omega_k^2-\omega_p^2)}}{\sum_{p=1}^r \sum_{l=p+1}^r \frac{4\omega_l^2 c_p^2}{(\omega_l^2-\omega_p^2)^2}} \\
& \leq \frac{\sum_{p=1}^r \sum_{j=r+1}^n \frac{(c_p^2+c_j^2)^2(\omega_j^2+\omega_p^2)}{c_p^2(\omega_j^2-\omega_p^2)^2} + 4 \sum_{p=1}^r \sum_{j=r+1}^n \sum_{k=j+1}^n \frac{c_k^2 c_j^2 \omega_p^2}{c_p^2(\omega_j^2-\omega_p^2)(\omega_k^2-\omega_p^2)}}{\sum_{p=1}^r \sum_{j=p+1}^r \frac{4\omega_p^2 c_j^2 + 4\omega_j^2 c_p^2 + 2\omega_p^2 c_p^2 + 2\omega_j^2 c_j^2}{(\omega_j^2-\omega_p^2)^2}} \\
& + \sum_{p=1}^r \sum_{j=p+1}^r \sum_{k=r+1}^n \frac{c_k^2(\omega_j^2 - \omega_p^2)}{c_p^2(\omega_k^2 - \omega_p^2)} + \sum_{p=1}^r \sum_{j=1}^{p-1} \sum_{k=r+1}^n \frac{c_k^2(\omega_p^2 - \omega_j^2)}{c_p^2(\omega_k^2 - \omega_p^2)} \\
& \leq \frac{\sum_{p=1}^r \sum_{j=r+1}^n \frac{(c_p^2+c_j^2)^2(\omega_j^2+\omega_p^2)}{c_p^2(\omega_j^2-\omega_p^2)^2} + 4 \sum_{p=1}^r \sum_{j=r+1}^n \sum_{k=j+1}^n \frac{c_k^2 c_j^2 \omega_p^2}{c_p^2(\omega_j^2-\omega_p^2)(\omega_k^2-\omega_p^2)}}{\sum_{p=1}^r \sum_{j=p+1}^r \frac{4\omega_p^2 c_j^2 + 4\omega_j^2 c_p^2 + 2\omega_p^2 c_p^2 + 2\omega_j^2 c_j^2}{(\omega_j^2-\omega_p^2)^2}} \\
& + \sum_{p=1}^r \sum_{j=1}^r \sum_{k=r+1}^n \frac{c_k^2 |\omega_j^2 - \omega_p^2|}{c_p^2(\omega_k^2 - \omega_p^2)}. \tag{4.20}
\end{aligned}$$

Ähnlich wie oben gilt dann für m beliebig mit $1 < m \leq r$

$$\begin{aligned}
& \frac{\sum_{p=1}^r \sum_{j=r+1}^n \frac{(c_p^2 + c_j^2)^2 (\omega_j^2 + \omega_p^2)}{c_p^2 (\omega_j^2 - \omega_p^2)^2}}{\sum_{p=1}^r \sum_{j=p+1}^r \frac{4\omega_p^2 c_j^2 + 4\omega_j^2 c_p^2 + 2\omega_p^2 c_p^2 + 2\omega_j^2 c_j^2}{(\omega_j^2 - \omega_p^2)^2}} \leq \sum_{p=1}^r \sum_{j=r+1}^n \frac{(c_p^2 + c_j^2)^2 (\omega_m - \omega_{m-1})^2}{c_p^2 (c_m^2 + c_{m-1}^2) (\omega_j - \omega_p)^2}, \\
& \frac{4 \sum_{p=1}^r \sum_{j=r+1}^n \sum_{k=j+1}^n \frac{c_k^2 c_j^2 \omega_p^2}{c_p^2 (\omega_j^2 - \omega_p^2) (\omega_k^2 - \omega_p^2)}}{\sum_{p=1}^r \sum_{j=p+1}^r \frac{4\omega_p^2 c_j^2 + 4\omega_j^2 c_p^2 + 2\omega_p^2 c_p^2 + 2\omega_j^2 c_j^2}{(\omega_j^2 - \omega_p^2)^2}} \leq \sum_{p=1}^r \sum_{j=r+1}^n \sum_{k=j+1}^n \frac{c_k^2 c_j^2 (\omega_m - \omega_{m-1})^2}{c_p^2 (c_m^2 + c_{m-1}^2) (\omega_j - \omega_p) (\omega_k - \omega_p)},
\end{aligned}$$

und damit gilt nach (4.20) für den relativen Fehler für den Fall $s = r > 1$

$$\begin{aligned}
& \left| \frac{\text{Tr}(XZ_s) - \text{Tr}(\tilde{X})}{\text{Tr}(XZ_s)} \right| \\
& \leq \sum_{p=1}^r \sum_{\substack{j=1 \\ j \neq p}}^r \sum_{k=r+1}^n \frac{c_k^2 |\omega_j^2 - \omega_p^2|}{c_p^2 (\omega_k^2 - \omega_p^2)} \\
& \quad + \frac{(\omega_m - \omega_{m-1})^2}{c_m^2 + c_{m-1}^2} \left(\sum_{p=1}^r \sum_{j=r+1}^n \left[\frac{(c_p^2 + c_j^2)^2}{c_p^2 (\omega_j - \omega_p)^2} + \sum_{k=j+1}^n \frac{c_k^2 c_j^2}{c_p^2 (\omega_j - \omega_p) (\omega_k - \omega_p)} \right] \right).
\end{aligned}$$

Damit gilt auch für diesen Fall die Behauptung des Satzes.

Q.E.D.

Korollar 4.14

Mit den Voraussetzungen von Satz 4.4 gilt

$$\left| \frac{\text{Tr}(XZ_s) - \text{Tr}(\tilde{X}\tilde{Z}_s)}{\text{Tr}(XZ_s)} \right| = O(|\omega_{r+1} - \omega_s|^{-2}) \quad \text{für } |\omega_{r+1} - \omega_s| \rightarrow \infty.$$

Beweis

Folgt unmittelbar aus Satz 4.13.

Q.E.D.

Bemerkung 4.15

Im Gegensatz zu den Korollaren 4.5, 4.7 und 4.10, die den Einfluß einer Skalierung auf die Frequenzen bzw. Dämpfungskomponenten untersuchen, ist die Aussage des Satzes 4.13 unabhängig von einer Skalierung und zeigt Auswirkungen auf, die die Relationen der Frequenzen bzw. Dämpfungskomponenten untereinander haben. Dies

soll im folgenden erläutert werden.

Als erstes betrachten wir den für die Praxis unrealistischen Fall $r = 1$, d.h. $s = 1$. Dann ist nach Satz 4.13

$$\begin{aligned} & \left| \frac{\operatorname{Tr}(XZ_s) - \operatorname{Tr}(\tilde{X}\tilde{Z}_s)}{\operatorname{Tr}(XZ_s)} \right| \leq 2 \sum_{k=2}^n \left[\frac{(c_1^2 + c_k^2)^2 \omega_1^2}{c_1^4 (\omega_k - \omega_1)^2} + \sum_{l=k+1}^n \frac{c_l^2 c_k^2 \omega_1^2}{c_1^4 (\omega_k - \omega_1)(\omega_l - \omega_1)} \right] \\ & = 2 \sum_{k=2}^n \frac{\omega_1^2}{(\omega_k - \omega_1)^2} + 2 \sum_{k=2}^n \left(\frac{2c_k^2 \omega_1^2}{c_1^2 (\omega_k - \omega_1)^2} + \frac{c_k^4 \omega_1^2}{c_1^4 (\omega_k - \omega_1)^2} + \sum_{l=k+1}^n \frac{c_l^2 c_k^2 \omega_1^2}{c_1^4 (\omega_k - \omega_1)(\omega_l - \omega_1)} \right). \end{aligned}$$

Diese Summanden werden klein, wenn

- ω_1 klein wird und ω_2 nicht gleichzeitig sehr dicht bei ω_1 liegt und c_1 außergewöhnlich klein wird;
- der Abstand der Frequenzen hinterm Schnitt zu ω_1 möglichst groß ist;
- $|c_1|$ gegenüber den anderen Dämpfungskomponenten $|c_k|, k > 1$ groß ist (dies verkleinert nur die zweite Summe, die erste Summe bleibt hiervon unberührt).

Nun betrachten wir den Fall, daß $r > 1$ ist.

Zuerst untersuchen wir die erste Dreifachsumme aus (4.15),

$$s_1 := \sum_{p=1}^s \sum_{\substack{j=1 \\ j \neq p}}^r \sum_{\substack{k=r+1 \\ k \neq p}}^n \frac{c_k^2 |\omega_j^2 - \omega_p^2|}{c_p^2 (\omega_k^2 - \omega_p^2)}. \quad (4.21)$$

Hierbei ist stets $\frac{|\omega_j^2 - \omega_p^2|}{(\omega_k^2 - \omega_p^2)} < \frac{(\omega_p^2 - \omega_1^2)}{(\omega_{r+1}^2 - \omega_s^2)}$ für alle $1 \leq p \leq s, 1 \leq j \leq r, j \neq p, r+1 \leq k \leq n$.

Die Dreifachsumme wird kleiner,

- je weiter die Frequenzen, die hinter dem Schnitt liegen, vom relevanten Frequenzbereich entfernt sind (vgl. Korollar 4.14 und Beispiel 4.25) und
- je näher die Frequenzen im gesamten Bereich vor dem Schnitt beieinander liegen.

Außerdem wird s_1 kleiner,

- je kleiner die Dämpfungskomponenten, die hinter dem Schnitt liegen, gegenüber den Dämpfungskomponenten im relevanten Bereich sind. (Besonders wichtig ist dies für die c_k , deren k direkt hinter dem Schnitt liegt.) (Vgl. Beispiel 4.22.)

Durch jeden einzelnen dieser Effekte kann s_1 beliebig klein gemacht werden.

Als nächstes betrachten wir den zweiten Teil von (4.15), d.h. die Summe

$$\alpha \sum_{p=1}^s \sum_{k=r+1}^n \left(\frac{(c_p^2 + c_k^2)^2}{c_p^2(\omega_k - \omega_p)^2} + \sum_{l=k+1}^n \frac{c_l^2 c_k^2}{c_p^2(\omega_k - \omega_p)(\omega_l - \omega_p)} \right),$$

wobei α das Minimum der Menge

$$\left\{ \left\{ \frac{2(\omega_m - \omega_s)^2}{c_m^2 + c_s^2}, s < m \leq r \right\} \cup \left\{ \frac{(\omega_m - \omega_{m-1})^2}{c_m^2 + c_{m-1}^2}, 1 < m \leq s \right\} \right\}$$

ist.

Zuerst untersuchen wir den Term

$$s_3 := \alpha \sum_{p=1}^s \sum_{k=r+1}^n \sum_{l=k+1}^n \frac{c_l^2 c_k^2}{c_p^2(\omega_k - \omega_p)(\omega_l - \omega_p)}. \quad (4.22)$$

Hierfür gilt das für s_1 Gesagte im verstärkten Maße, denn wir haben z.B.

$$\begin{aligned} & \alpha \sum_{p=1}^s \sum_{k=r+1}^n \sum_{l=k+1}^n \frac{c_l^2 c_k^2}{c_p^2(\omega_k - \omega_p)(\omega_l - \omega_p)} \\ & \leq \sum_{p=1}^s \sum_{k=r+1}^n \sum_{l=k+1}^n \frac{c_l^2 c_k^2 (\omega_m - \omega_{m-1})^2}{c_p^2 (c_m^2 + c_{m-1}^2) (\omega_k - \omega_p)(\omega_l - \omega_p)} \end{aligned} \quad (4.23)$$

für jedes m mit $1 < m \leq s$. Hier sind die von den Frequenzen gelieferten Faktoren

$$\frac{(\omega_m - \omega_{m-1})^2}{(\omega_k - \omega_p)(\omega_l - \omega_p)}$$

wesentlich kleiner als bei s_1 und die von den Dämpfungskomponenten gelieferten Faktoren sind Produkte von denen aus s_1 . Zusätzliche Effekte werden noch in der folgenden Betrachtung von (4.24) aufgeführt.

Zuletzt betrachten wir die Doppelsumme

$$s_2 := \alpha \sum_{p=1}^s \sum_{k=r+1}^n \frac{(c_p^2 + c_k^2)^2}{c_p^2(\omega_k - \omega_p)^2}. \quad (4.24)$$

Wie bei s_1 wirken sich große Dämpfungskomponenten hinter dem Schnitt negativ auf s_2 aus (vgl. Korollar 4.7 (i), wo der entsprechende Effekt schon genauer formuliert wurde und Beispiel 4.22); dagegen verbessern kleine c_k , $k > r$, s_2 nur bis zu einem gewissen Grenzwert (vgl. Korollar 4.7 (ii) und Beispiel 4.22). Dagegen kann durch Vergrößerung des Abstandes der Frequenzen hinter dem Schnitt vom relevanten

Frequenzbereich s_2 beliebig klein gemacht werden (vgl. Korollar 4.14 und Beispiel 4.25).

Wie man aus (vgl. (4.23))

$$s_2 \leq \sum_{p=1}^s \sum_{k=r+1}^n \frac{(c_p^2 + c_k^2)^2 (\omega_m - \omega_{m-1})^2}{c_p^2 (c_m^2 + c_{m-1}^2) (\omega_k - \omega_p)^2}, \quad 1 < m \leq s \quad (4.25)$$

bzw.

$$s_2 \leq 2 \sum_{p=1}^s \sum_{k=r+1}^n \frac{(c_p^2 + c_k^2)^2 (\omega_m - \omega_s)^2}{c_p^2 (c_m^2 + c_s^2) (\omega_k - \omega_p)^2}, \quad s < m \leq r \quad (4.26)$$

ersieht, wirken sich zusätzlich folgende Konstellationen positiv auf s_2 (und auch auf s_3) aus:

- zwei dicht beieinander liegende Frequenzen im relevanten Bereich, wenn nicht gleichzeitig beide zugehörigen Dämpfungskomponenten außergewöhnlich klein sind (s. (4.25)); dieser Effekt wirkt sich auf s_1 nicht signifikant aus;
- eine einzelne Dämpfungskomponente c_{p_0} im Bereich vor dem Schnitt, die besonders groß gegenüber den Komponenten im relevanten Bereich ist (vgl. Beispiel 4.19).

Dies kann Einfluß auf alle drei Summen haben:

Liegt die Dämpfungskomponente c_{p_0} im relevanten Bereich, dann wirkt sich eine Vergrößerung von c_{p_0}

- auf s_1 positiv aus für alle Summanden, für die $p = p_0$ gilt, die anderen Summanden in (4.21) verändern sich nicht,
- auf s_3 gleichmäßig positiv aus für alle Summanden (s. (4.23)) und
- auf s_2 positiv aus für die Summanden, für die $p \neq p_0$ gilt, die Summanden, für die $p = p_0$ gilt, verhalten sich asymptotisch konstant.

Liegt die Dämpfungskomponente c_{p_0} im Zwischenbereich, dann hat eine Vergrößerung von c_{p_0} auf (4.21) keinen Einfluß. Hierbei ist es dann umso günstiger, je näher c_{p_0} am relevanten Bereich liegt. Dies deckt sich weitgehend mit der Aussage von Korollar 4.5 (i) (vgl. Bemerkung 4.6 (i)); die dortige Aussage war insofern schärfer, als der relative Fehler durch eine einzelne beliebig große Dämpfungskomponente beliebig klein gemacht werden kann.

Zum Schluß zeigen wir, daß unter sehr allgemeinen Bedingungen die Aussage von Satz 4.4 einen Grenzwert für $n \rightarrow \infty$ liefert. Außerdem zeigen wir, daß der relative Fehler der Spur für diesen Fall gegen Null konvergiert, falls r gegen ∞ strebt, wobei die Relation zwischen r und n keine Rolle spielt.

Satz 4.16

Für $k \in \mathbb{N}$ seien ω_k, c_k aus \mathbb{R} mit $0 < \omega_1 < \omega_2 < \dots, c_k \neq 0$ für alle k . Es gelte weiter

$$\sum_{k=1}^{\infty} \frac{1}{\omega_k^2} < \infty, \quad |c_k| \leq M < \infty \quad \text{für alle } k. \quad (4.27)$$

Seien $s, r \in \mathbb{N}$ fest, $s \leq r$.

Mit den Bezeichnungen aus Satz 4.4 existieren dann

$$\lim_{n \rightarrow \infty} \text{Tr}(XZ_s) \quad \text{und} \quad \lim_{n \rightarrow \infty} \left| \frac{\text{Tr}(XZ_s) - \text{Tr}(\tilde{X}\tilde{Z}_s)}{\text{Tr}(XZ_s)} \right|$$

und sind endlich.

Beweis

Der Einfachheit halber betrachten wir zunächst $c_k = c = \text{const}$.

Dann ist (vgl. 4.11)

$$\begin{aligned} Z_{rF} = & \left| \sum_{p=1}^s \left[6c^2\omega_p^2 \sum_{j=r+1}^n \frac{1}{(\omega_j^2 - \omega_p^2)^2} + 2c^2 \sum_{j=r+1}^n \frac{\omega_j^2}{(\omega_j^2 - \omega_p^2)^2} \right. \right. \\ & + 4c^2\omega_p^2 \sum_{\substack{j=1 \\ j \neq p}}^r \frac{1}{\omega_j^2 - \omega_p^2} \left(\sum_{k=r+1}^n \frac{1}{\omega_k^2 - \omega_p^2} \right) \\ & \left. \left. + 4c^2\omega_p^2 \sum_{j=r+1}^n \frac{1}{\omega_j^2 - \omega_p^2} \left(\sum_{k=j+1}^n \frac{1}{\omega_k^2 - \omega_p^2} \right) \right] \right|, \end{aligned} \quad (4.28)$$

$$\begin{aligned} \text{Tr}(XZ_s) = N_{rF} = & \sum_{p=1}^s \left[\frac{2}{c^2} + \frac{c}{2\omega_p^2} + 2\omega_p^2 c^2 \left(\sum_{\substack{k=1 \\ k \neq p}}^n \frac{1}{\omega_k^2 - \omega_p^2} \right)^2 \right. \\ & \left. + c^2 \left(4\omega_p^2 \sum_{\substack{j=1 \\ j \neq p}}^n \frac{1}{(\omega_j^2 - \omega_p^2)^2} + 2 \sum_{\substack{j=1 \\ j \neq p}}^n \frac{\omega_j^2}{(\omega_j^2 - \omega_p^2)^2} \right) \right], \end{aligned}$$

und alle Summen mit n als Obergrenze konvergieren offensichtlich für $n \rightarrow \infty$, falls $\sum_{k=1}^{\infty} \frac{1}{\omega_k^2}$ konvergiert. Für $|c_k| \leq M < \infty$ verläuft der Beweis analog, da im Nenner der betrachteten Terme jeweils nur endlich viele $c_p \neq 0$ auftreten.

Q.E.D.

Satz 4.17

Seien ω_k, c_k wie in Satz 4.16. Wir betrachten den relativen Fehler der Spur für festes $s \in \mathbb{N}$ und beliebige $r, n \in \mathbb{N}$ mit $s \leq r \leq n$; dann gilt

$$\left| \frac{\text{Tr}(XZ_s) - \text{Tr}(\tilde{X}\tilde{Z}_s)}{\text{Tr}(XZ_s)} \right| \longrightarrow 0, \quad \text{wenn } r \rightarrow \infty.$$

Beweis

Aus $\sum_{k=1}^{\infty} \frac{1}{\omega_k^2} < \infty$ folgt $\sum_{k>r} \frac{1}{\omega_k^2} \rightarrow 0$ für $r \rightarrow \infty$; damit folgt in (4.28) $Z_{rF} \rightarrow 0$ für $r \rightarrow \infty$. Wegen $N_{rF} > \sum_{p=1}^s \frac{2}{c^2}$ folgt daraus die Behauptung, wenn $c_k = c = \text{const.}$ Für $|c_k| \leq M < \infty$ verläuft der Beweis wieder analog.

Q.E.D.

Bemerkung 4.18

Die Voraussetzung an die Frequenzen in (4.27) ist z.B. schon für $\omega_k = \alpha \cdot k^\beta$ für $\beta > \frac{1}{2}$ erfüllt.

4.2 Fazit bei einem Dämpfer

Zusammenfassend kann also gesagt werden, daß eine beliebige Verkleinerung des relativen Fehlers der Spur durch einen der folgenden Punkte erreicht wird:

- genügende simultane Verkleinerung aller Dämpfungskomponenten im relevanten Bereich und gleichzeitig im Bereich hinter dem Schnitt, oder genügende simultane Verkleinerung aller Dämpfungskomponenten (s. Korollare 4.10, 4.11 und Beispiel 4.23);
- genügende Vergrößerung einer einzelnen Dämpfungskomponente c_{p_0} , die vor dem Schnitt liegt (eine Ausnahme bildet hier der Fall, daß der relevante Bereich nur aus p_0 besteht (s. u.)). Liegt p_0 im Zwischenbereich, dann ist es günstiger, je näher p_0 am relevanten Bereich liegt (s. Korollar 4.5 (i), Bemerkung 4.6 (i) und Beispiele 4.19, 4.21 (i));
- genügende simultane Vergrößerung einer oder mehrerer Dämpfungskomponenten im Zwischenbereich (s. Korollar 4.5 (i), Bemerkung 4.6 (i) und Beispiele 4.19 (iv), (v), 4.20 (i); die Auswirkungen einer simultanen Vergrößerung mehrerer Dämpfungskomponenten im relevanten Bereich werden ebenfalls in Korollar 4.5 (i), Bemerkung 4.6 (i) erläutert);
- genügende Vergrößerung des Abstandes der Frequenzen hinter dem Schnitt zu den Frequenzen im relevanten Bereich (s. Korollar 4.14 und Beispiel 4.25);
- genügende simultane Vergrößerung aller Frequenzen (s. Korollar 4.10 und Beispiel 4.24).

Negativ auf den relativen Fehler wirken sich die folgenden Faktoren aus:

- beliebig simultane Vergrößerung einer oder mehrerer Dämpfungskomponenten hinter dem Schnitt (s. Korollar 4.7 (i), Bemerkung 4.15 und Beispiel 4.22);
- Vergrößerung der Abstände der Frequenzen vor dem Schnitt (s. Bemerkung 4.15);

Bis zu bestimmten Grenzwerten wirken sich eine

- simultane Verkleinerung einer oder mehrerer Dämpfungskomponenten hinter dem Schnitt tendenziell positiv (s. Korollar 4.7 (ii), Bemerkung 4.15 und Beispiel 4.22),
- simultane Verkleinerung einzelner Dämpfungskomponenten vor dem Schnitt tendenziell negativ (s. Korollar 4.5 (ii), Bemerkung 4.15 und Beispiel 4.19 (i)-(iii)),
- simultane Vergrößerung aller Dämpfungskomponenten tendenziell negativ (s. Korollar 4.10 und Beispiel 4.23),
- simultane Verkleinerung aller Frequenzen tendenziell negativ (s. Korollar 4.10 und Beispiel 4.24)

auf den relativen Fehler der Spur aus. „Tendenziell negativ“ heißt hier, daß eine hinreichende Veränderung in umgekehrter Richtung eine eindeutige Verbesserung mit sich bringt, und umgekehrt.

- Die simultane Verkleinerung selbst aller Dämpfungskomponenten hinter dem Schnitt führt aber entgegen früheren Vermutungen nicht zu einer **beliebigen** Verkleinerung des relativen Fehlers!

Keine eindeutig bewertbare Aussage kann dagegen über eine beliebige simultane Vergrößerung oder Verkleinerung aller Dämpfungskomponenten im relevanten Bereich gemacht werden, da in beide Richtungen positive Grenzwerte angenommen werden (s. Korollare 4.5 (i), (ii) bzw. Bemerkungen 4.6 (i), (ii) und Beispiel 4.21).

(Bei allen Punkten des Fazits wird davon ausgegangen, daß die nicht angesprochenen Dämpfungskomponenten und Frequenzen nicht wesentlich verändert werden.) Diese Aussagen wollen wir nun anhand mehrerer Beispiele weiter verdeutlichen.

4.3 Beispiele bei einem Dämpfer

In den folgenden Beispielen wird in den Tabellen jeweils der relative Fehler der Spur $\left| \frac{\text{Tr}(XZ_s) - \text{Tr}(\tilde{X}\tilde{Z}_s)}{\text{Tr}(XZ_s)} \right|$ für s mit $1 \leq s \leq r$ abgetragen.

Dabei wurden alle Werte nach (4.11) mit Matlab berechnet.

Bei allen Beispielen ist $n = 7$ und $r = 5$.

Die Tabellen der Beispiele 4.19 und 4.20 (i) und (ii) beinhalten die relativen Fehler der Spuren für den Fall $s = 3$ für zwei verschiedene Frequenzspektren Ω_i . Hier ist also der relevante Bereich durch die Indexmenge $\{1, 2, 3\}$, der Zwischenbereich durch $\{4, 5\}$ und der Bereich hinter dem Schnitt durch $\{6, 7\}$ gegeben.

Beispiel 4.19

Sei $c = (\alpha, \beta, \gamma, 1, 1, 1, 1)^T$, $s = 3$ und

$\Omega_1 = \text{diag}(1, 2, 3, 4, 5, 6, 7)$, $\Omega_2 = \text{diag}(1, 4, 9, 16, 25, 36, 49)$.

(i)

α	β	γ	Ω	Ω_1	Ω_2
0.01	1	1		4.2366E-02	2.0152E-04
0.1	1	1		4.3254E-02	2.2246E-04
1	1	1		2.3539E-02	4.4043E-04
10	1	1		4.2715E-03	5.0952E-04
100	1	1		4.3783E-05	5.3757E-06
1000	1	1		4.3794E-07	5.3785E-08

(ii)

α	β	γ	Ω	Ω_1	Ω_2
1	0.01	1		1.0340E-02	1.6708E-03
1	0.1	1		1.0340E-02	1.6090E-03
1	1	1		2.3539E-02	4.4043E-04
1	10	1		3.1735E-03	9.1727E-04
1	100	1		3.3473E-05	9.9611E-06
1	1000	1		3.3491E-07	9.9685E-08

(iii)

α	β	$\gamma \backslash \Omega$	Ω_1	Ω_2
1	1	0.01	9.6017E-02	3.8876E-03
1	1	0.1	8.8815E-02	3.7810E-03
1	1	1	2.3539E-02	4.4043E-04
1	1	10	7.9149E-03	2.9937E-03
1	1	100	8.3597E-05	3.3400E-05
1	1	1000	8.3641E-07	3.3417E-07

Diese Beispiele zeigen die Auswirkungen einer hinreichenden Vergrößerung einer einzigen Dämpfungskomponenten im relevanten Bereich. Unabhängig davon, welche Komponente hinreichend vergrößert wird, wird der relative Fehler quadratisch kleiner (hier für α , β bzw. $\gamma \geq 10$ zu sehen) (vgl. Korollar 4.5 (i) bzw. Bemerkung 4.6 (i) und Bemerkung 4.15). Dagegen konvergiert der relative Fehler der Spur gegen einen positiven Grenzwert, wenn die einzelnen Dämpfungskomponenten beliebig verkleinert werden (s. Korollar 4.5 (ii), Bemerkung 4.15).

Die nächsten beiden Tabellen zeigen, daß auch eine hinreichende Vergrößerung einer einzelnen Dämpfungskomponenten im Zwischenbereich zu einer beliebigen Verkleinerung des relativen Fehlers führt (zur besseren Übersicht der Übersichtlichkeit der Tabellen werden die Fälle, daß α , $\beta < 1$ sind, hier nicht mehr aufgeführt. Wie in den obigen Beispielen konvergiert der relative Fehler für diese Fälle auch gegen einen positiven Grenzwert).

Sei $c = (1, 1, 1, \alpha, \beta, 1, 1)^T$.

(iv)

α	$\beta \backslash \Omega$	Ω_1	Ω_2
1	1	2.3539E-02	4.4043E-04
10	1	9.2631E-03	4.1295E-03
100	1	9.3448E-05	4.5293E-05
1000	1	9.3452E-07	4.5279E-07

(v)

α	$\beta \backslash \Omega$	Ω_1	Ω_2
1	1	2.3539E-02	4.4043E-04
1	10	2.0024E-02	6.5739E-03
1	100	2.0511E-04	1.3800E-04
1	1000	2.0512E-06	1.3786E-06

Hier sieht man, daß es günstiger ist, α statt β zu erhöhen, da α näher am relevanten Bereich liegt (vgl. Korollar 4.5 (i), Bemerkung 4.6 (i) und Bemerkung 4.15).

Beispiel 4.20

Sei $s = 3$ und $\Omega_1 = \text{diag}(1, 2, 3, 4, 5, 6, 7)$, $\Omega_2 = \text{diag}(1, 4, 9, 16, 25, 26, 49)$.

(i) $c = (1, 1, 1, \alpha, \beta, 1, 1)^T$.

α	β	Ω	Ω_1	Ω_2
1	1		2.3539E-02	4.4043E-04
10	10		6.4345E-03	3.2690E-03
100	100		6.4375E-05	3.4122E-05
1000	1000		6.4374E-07	3.4113E-07

(ii) $c = (\alpha, \beta, \gamma, 1, 1, 1, 1)^T$.

α	β	γ	Ω	Ω_1	Ω_2
1	1	1		2.3539E-02	4.4043E-04
10	10	1		3.0881E-03	6.4204E-04
100	100	1		3.2004E-05	7.0696E-06
1000	1000	1		3.2015E-07	7.0766E-08

In Beispiel 4.20 werden mehrere Dämpfungskomponenten simultan vergrößert. Zuerst alle Komponenten im Zwischenbereich, in der zweiten Tabelle dann zwei Komponenten aus dem relevanten Bereich. Durch beides kann der relative Fehler beliebig verkleinert werden. Beides sind Beispiele zu Korollar 4.5 (i); die Indexmenge I besteht in Beispiel 4.20 (i) nur aus I_{zw} , in Beispiel 2 (ii) ist I gegeben durch $\{1, 2\}$. Damit ist in diesem Fall immer $N_1 \neq 0$ (s. Korollar 4.5 (i), Bemerkung 4.6 (i)), und der relative Fehler konvergiert quadratisch gegen Null bei genügender Vergrößerung der Dämpfungskomponenten c_1 und c_2 .

Die nächste Tabelle ist ein Beispiel für Korollar 4.5 (i), bzw. Bemerkung 4.6 (i) für den Fall $N_1 = 0$ für $s = 1$, $s = 2$ und $s = 3$. Man sieht, daß für diese s der relative Fehler der Spur für $\alpha \rightarrow \infty$ gegen einen von Null verschiedenen Grenzwert konvergiert, während er für alle anderen s gegen Null konvergiert.

(iii) Sei $\Omega = \text{diag}(1, 5, 7, 10, 12, 15, 20)$ und $c = (\alpha, 1, \alpha, 1, 1, 1, 1)^T$.

α^s	1	2	3	4	5
1	6.0746E-03	1.0426E-02	4.6486E-03	3.6003E-03	1.1618E-02
10	1.1531E-02	7.6015E-03	1.4675E-02	5.9914E-03	9.9114E-03
100	1.1426E-02	7.3041E-03	1.4779E-02	6.5187E-05	1.0532E-04
1000	1.1425E-02	7.3010E-03	1.4780E-02	6.5241E-07	1.0538E-06

Beispiel 4.21

Sei $\Omega = \text{diag}(1, 4, 9, 16, 25, 36, 49)$.

(i) $c = (\alpha, 1, 1, 1, 1, 1, 1)^T$

$\alpha \backslash s$	1	2	3	4	5
1	1.0272E-03	3.5132E-04	4.4043E-04	1.4477E-03	2.9986E-03
10	2.0162E-03	2.5920E-04	5.0952E-04	8.0002E-04	1.2293E-03
100	2.0005E-03	2.7628E-06	5.3757E-06	8.4017E-06	1.2855E-05
1000	2.0003E-03	2.7646E-08	5.3785E-08	8.4056E-08	1.2860E-07

(ii) Sei $c = (\alpha, \alpha, 1, 1, 1, 1, 1)^T$

$\alpha \backslash s$	1	2	3	4	5
1	1.0272E-03	3.5132E-04	4.4043E-04	1.4477E-03	2.9986E-03
10	1.7499E-03	2.0175E-03	6.4204E-04	1.2005E-03	2.0103E-03
100	1.7293E-03	2.0372E-03	7.0696E-06	1.2995E-05	2.1647E-05
1000	1.7291E-03	2.0374E-03	7.0766E-08	1.3006E-07	2.1663E-07

(iii) Sei $c = (\alpha, \alpha, \alpha, 1, 1, 1, 1)^T$

$\alpha \backslash s$	1	2	3	4	5
1	1.0272E-03	3.5132E-04	4.4043E-04	1.4477E-03	2.9986E-03
10	1.7095E-03	2.0459E-03	2.5400E-03	1.4592E-03	2.9256E-03
100	1.6882E-03	2.0559E-03	2.6138E-03	1.6757E-05	3.2667E-05
1000	1.6880E-03	2.0560E-03	2.6145E-03	1.6782E-07	3.2704E-07

Anhand dieses Beispiels werden die Auswirkungen einer simultanen Vergrößerung aller Dämpfungskomponenten im relevanten Bereich verdeutlicht. Durch eine Vergrößerung kann der relative Fehler der Spur nicht beliebig verkleinert werden, er konvergiert gegen einen positiven Grenzwert (s. Korollar 4.5 (i) bzw. Bemerkung 4.6 (i)). In Beispiel 4.21 (i) wird nur die Dämpfungskomponente c_1 vergrößert. Damit können alle relativen Fehler für $s > 1$ beliebig verkleinert werden, während für $s = 1$ dies nicht möglich ist. Entsprechendes ergibt sich für (ii) und (iii); in (ii) werden die ersten beiden Dämpfungskomponenten vergrößert, so daß sowohl für $s = 1$ als auch für $s = 2$ eine beliebige Verkleinerung des relativen Fehlers nicht möglich ist, in (iii) werden die ersten drei Komponenten vergrößert, so daß hier für $s = 1$, $s = 2$ und $s = 3$ keine beliebige Verkleinerung des relativen Fehlers erreicht werden kann.

Beispiel 4.22

Sei $\Omega = \text{diag}(1, 2, 3, 4, 5, 6, 7)$.

(i) Sei $c = (1, 1, 1, 1, 1, \alpha, 1)^T$.

$\alpha \backslash s$	1	2	3	4	5
0.01	2.3592E-02	1.9627E-02	1.5258E-02	9.1686E-03	2.0003E-03
0.1	2.3776E-02	1.9761E-02	1.5322E-02	9.1058E-03	1.6204E-03
1	4.2139E-02	3.3800E-02	2.3539E-02	7.3760E-03	2.0807E-02
10	8.3854E-01	9.0897E-01	9.5297E-01	9.7879E-01	9.9375E-01
100	9.9997E-01	9.9999E-01	9.9999E-01	1.0000E-00	1.0000E-00
1000	1.0000E-00	1.0000E-00	1.0000E-00	1.0000E-00	1.0000E-00

(ii) Sei $c = (1, 1, 1, 1, 1, 1, \alpha)^T$.

$\alpha \backslash s$	1	2	3	4	5
0.01	2.8746E-02	2.3774E-02	1.7940E-02	8.7981E-03	6.5370E-03
0.1	2.8878E-02	2.3870E-02	1.7987E-02	8.7661E-03	6.7169E-03
1	4.2139E-02	3.3800E-02	2.3539E-02	7.3760E-03	2.0807E-02
10	7.5981E-01	8.4395E-01	9.0679E-01	9.4822E-01	9.7502E-01
100	9.9994E-01	9.9998E-01	9.9999E-01	9.9999E-01	1.0000E-00
1000	1.0000E-00	1.0000E-00	1.0000E-00	1.0000E-00	1.0000E-00

(iii) Sei $c = (1, 1, 1, 1, 1, \alpha, \alpha)^T$.

$\alpha \backslash s$	1	2	3	4	5
0.01	1.0131E-02	1.0500E-02	1.2191E-02	1.6321E-02	3.1033E-02
0.1	1.0449E-02	1.0712E-02	1.2251E-02	1.6116E-02	3.0210E-02
1	4.2139E-02	3.3800E-02	2.3539E-02	7.3760E-03	2.0807E-02
10	9.2812E-01	9.6471E-01	9.8255E-01	9.9188E-01	9.9720E-01
100	9.9999E-01	1.0000E-00	1.0000E-00	1.0000E-00	1.0000E-00
1000	1.0000E-00	1.0000E-00	1.0000E-00	1.0000E-00	1.0000E-00

Beispiel 4.22 zeigt, wie sich der relative Fehler der Spur für alle s verschlechtert, wenn Dämpfungskomponenten hinter dem Schnitt vergrößert werden (vgl. Korollar 4.7 (i) und Bemerkung 4.15). In allen Fällen nähert sich der relative Fehler sehr schnell dem Grenzwert 1. Man sieht, daß eine hinreichende Vergrößerung der Dämpfungskomponente, die direkt hinter dem Schnitt liegt, zu einer schnelleren Konvergenz führt als eine hinreichende Vergrößerung einer Komponente, die weiter

vom relevanten Bereich entfernt ist (s. Bemerkung 4.15). Dagegen führt eine Verkleinerung der Dämpfungskomponenten hinter dem Schnitt (selbst aller!) für sich allein entgegen früheren Vermutungen nicht zu einer beliebigen Verkleinerung des relativen Fehlers der Spur (s. Korollar 4.7 (ii) bzw. Bemerkung 4.8 (ii))!

Beispiel 4.23

Sei $\Omega = \text{diag}(1, 2, 3, 4, 5, 6, 7)$ und $c = \beta \cdot (1, 1, 1, 1, 1, 1, 1)^T$.

$\beta \backslash s$	1	2	3	4	5
100	6.7707E-02	5.2004E-02	3.5456E-02	1.0936E-02	3.0386E-02
10	6.7703E-02	5.2001E-02	3.5454E-02	1.0935E-02	3.0384E-02
1	4.2139E-02	3.3800E-02	2.3539E-02	7.3760E-03	2.0807E-02
0.1	1.1157E-05	9.6544E-06	7.0023E-06	2.2654E-06	6.5991E-06
0.01	1.1159E-09	9.6562E-10	7.0036E-10	2.2659E-10	6.6005E-10
0.001	1.1159E-13	9.6562E-14	7.0036E-14	2.2659E-14	6.6005E-14

Beispiel 4.23 zeigt die Auswirkungen einer Skalierung von allen Dämpfungskomponenten und ist damit ein Beispiel für Korollar 4.10. Man sieht, daß der relative Fehler der Spur für alle s mit $s \leq 5$ mit einer Geschwindigkeit von der vierten Potenz des Skalierungsfaktors β gegen Null läuft, falls β gegen Null geht (dies ist hier für $\beta < 0.1$ der Fall). Dagegen bleibt der relative Fehler der Spur für alle s mit $s \leq 5$ asymptotisch konstant für $\beta \rightarrow \infty$ (hier für $\beta \geq 1$ zu sehen).

Beispiel 4.24

Sei $\Omega = \gamma \cdot \text{diag}(1, 2, 3, 4, 5, 6, 7)$, $c = (1, 1, 1, 1, 1, 1, 1)^T$.

$\gamma \backslash s$	1	2	3	4	5
0.01	6.7703E-02	5.2001E-02	3.5454E-02	1.0935E-02	3.0384E-02
0.1	6.7298E-02	5.1725E-02	3.5277E-02	1.0883E-02	3.0246E-02
1	4.2139E-02	3.3800E-02	2.3539E-02	7.3760E-03	2.0807E-02
10	1.0978E-03	9.4802E-04	6.8680E-04	2.2199E-04	6.4602E-04
100	1.1157E-05	9.6544E-06	7.0023E-06	2.2654E-06	6.5991E-06
1000	1.1159E-07	9.6562E-08	7.0036E-08	2.2659E-08	6.6005E-08

Hier wird der Einfluß einer Skalierung aller Frequenzen auf den relativen Fehler

der Spur verdeutlicht. Lauft der Skalierungsfaktor γ gegen ∞ , dann verkleinert sich der relative Fehler der Spur mit der zweiten Potenz von γ (hier fur $\gamma \geq 10$). Dagegen bleibt er asymptotisch konstant fur kleine Skalierungen (hier fur $\gamma \leq 1$) (vgl. Korollar 4.10).

Beispiel 4.25

Sei $c = (1, 1, 1, 1, 1, 1, 1)^T$.

(i) Sei $\Omega = \text{diag}(1, 2, 3, 4, 5, \alpha \cdot 6, \alpha \cdot 7)$.

$\alpha \backslash s$	1	2	3	4	5
10^0	4.2139E-02	3.3800E-02	2.3539E-02	7.3760E-03	2.0807E-02
$10^{0.5}$	4.0694E-03	2.7779E-03	1.1744E-03	1.0579E-03	4.5656E-03
10^1	4.0543E-04	2.7242E-04	1.0928E-04	1.1215E-04	4.4743E-04
$10^{1.5}$	4.0528E-05	2.7189E-05	1.0849E-05	1.1273E-05	4.4651E-05
10^2	4.0527E-06	2.7184E-06	1.0841E-06	1.1279E-06	4.4642E-06

(ii) Sei $\Omega = \text{diag}(1, 4, 9, 16, 25, \alpha \cdot 36, \alpha \cdot 49)$.

$\alpha \backslash s$	1	2	3	4	5
10^0	1.0272E-03	3.5132E-04	4.4043E-04	1.4477E-03	2.9986E-03
$10^{0.5}$	1.0250E-04	3.3058E-05	5.0061E-05	1.5496E-04	3.0783E-04
10^1	1.0248E-05	3.2855E-06	5.0593E-06	1.5544E-05	3.0530E-05
$10^{1.5}$	1.0248E-06	3.2834E-07	5.0645E-07	1.5549E-06	3.0504E-06
10^2	1.0248E-07	3.2832E-08	5.0651E-08	1.5549E-07	3.0501E-07

In diesem Beispiel werden fur zwei Frequenzspektren die Auswirkungen einer Vergroerung des Abstandes der Frequenzen, die hinter dem Schnitt liegen, zu den Frequenzen im relevanten Bereich verdeutlicht. Durch eine Vergroerung des Abstandes kann der relative Fehler der Spur beliebig verkleinert werden (s. Korollar 4.14).

Dies kann auch bei den vorherigen Beispielen, bei denen die zwei verschiedene Frequenzspektren $\Omega_1 = \text{diag}(1, 2, 3, 4, 5, 6, 7)$ und $\Omega_2 = \text{diag}(1, 4, 9, 16, 25, 36, 49)$ betrachtet wurden, gesehen werden. Hier ist immer der relative Fehler der Spur fur Ω_2 besser als fur Ω_1 , da der Abstand der Frequenzen bei der Frequenzverteilung Ω_2 groer als bei Ω_1 ist.

4.4 Zum relativen Fehler der optimalen Viskositäten bei einem Dämpfer

In den vorangegangenen Abschnitten haben wir die Spur der Lösung des geschnittenen Systems mit der Spur der Lösung des ungeschnittenen Systems verglichen. Eine andere Fragestellung ist der Vergleich der Minimumstelle des geschnittenen bzw. ungeschnittenen Systems. Hiermit wollen wir uns in diesem Abschnitt beschäftigen. Wir wollen untersuchen, welche Bedingungen die Frequenzen und Dämpfungskomponenten erfüllen müssen, damit der relative Fehler der Minimumstelle möglichst gering wird. Wir werden sehen, daß viele der in Abschnitt 4.1 bewiesenen Aussagen übertragbar sind. Zunächst wollen wir die Minimumstelle von $\text{Tr}(XZ_s)$ bestimmen, die sich direkt aus der Minimierung von (4.6) ergibt.

Lemma 4.26

Es gelten die Voraussetzungen von Lemma 4.3. Dann nimmt die Funktion $\text{Tr}(XZ_s)$ für

$$\varepsilon = \left(\frac{\sum_{p=1}^s \frac{2}{c_p^2}}{\sum_{p=1}^s \left[\frac{c_p^2}{2\omega_p^2} + \left(\sum_{\substack{j=1 \\ j \neq p}}^n \frac{3\omega_p^2 c_j^2 + \omega_p^2 c_p^2 + \omega_j^2 c_j^2 + \omega_j^2 c_p^2}{(\omega_p^2 - \omega_j^2)^2} + \frac{2\omega_p^2}{c_p^2} \left(\sum_{\substack{k=1 \\ k \neq p}}^n \frac{c_k^2}{\omega_k^2 - \omega_p^2} \right)^2 \right) \right]} \right)^{\frac{1}{2}} \quad (4.29)$$

ihr Minimum an.

Entsprechend nimmt die Spur des bei r geschnittenen Systems $\text{Tr}(\tilde{X}\tilde{Z}_s)$ bei

$$\tilde{\varepsilon} = \left(\frac{\sum_{p=1}^s \frac{2}{c_p^2}}{\sum_{p=1}^s \left[\frac{c_p^2}{2\omega_p^2} + \left(\sum_{\substack{j=1 \\ j \neq p}}^r \frac{3\omega_p^2 c_j^2 + \omega_p^2 c_p^2 + \omega_j^2 c_j^2 + \omega_j^2 c_p^2}{(\omega_p^2 - \omega_j^2)^2} + \frac{2\omega_p^2}{c_p^2} \left(\sum_{\substack{k=1 \\ k \neq p}}^r \frac{c_k^2}{\omega_k^2 - \omega_p^2} \right)^2 \right) \right]} \right)^{\frac{1}{2}} \quad (4.30)$$

ihr Minimum an.

Im folgenden betrachten wir der Einfachheit halber den relativen Fehler der Quadrate der Minimumstelle

$$\left| \frac{\varepsilon^2 - \tilde{\varepsilon}^2}{\varepsilon^2} \right|, \quad (4.31)$$

der wegen

$$\left| \frac{\varepsilon - \tilde{\varepsilon}}{\varepsilon} \right| \leq \left| \frac{\varepsilon^2 - \tilde{\varepsilon}^2}{\varepsilon^2} \right| = \left| \frac{\varepsilon - \tilde{\varepsilon}}{\varepsilon} \left(2 - \frac{\varepsilon - \tilde{\varepsilon}}{\varepsilon} \right) \right|$$

(was direkt aus $\frac{\varepsilon^2 - \tilde{\varepsilon}^2}{\varepsilon^2} = \frac{\varepsilon - \tilde{\varepsilon}}{\varepsilon} \cdot \frac{\varepsilon + \tilde{\varepsilon}}{\varepsilon}$ folgt) auch ein Maß für $\left| \frac{\varepsilon - \tilde{\varepsilon}}{\varepsilon} \right|$ darstellt. Analog zu den entsprechenden Berechnungen zu Satz 4.4 erhalten wir dafür

Satz 4.27

Es gelten die Voraussetzungen von Satz 4.4. Dann gilt

$$\left| \frac{\varepsilon^2 - \tilde{\varepsilon}^2}{\varepsilon^2} \right| = \frac{Z_{rF\varepsilon^2}}{N_{rF\varepsilon^2}}$$

mit

$$\begin{aligned} Z_{rF\varepsilon^2} &= \left| \sum_{p=1}^s \left(\sum_{\substack{j=r+1 \\ j \neq p}}^n \frac{(c_p^2 + c_j^2)^2 \omega_p^2}{c_p^2 (\omega_j^2 - \omega_p^2)^2} + \sum_{j=r+1}^n \frac{(c_p^2 + c_j^2)(\omega_p^2 c_j^2 + \omega_j^2 c_p^2)}{c_p^2 (\omega_j^2 - \omega_p^2)^2} \right. \right. \\ &\quad \left. \left. + 4 \sum_{\substack{j=1 \\ j \neq p}}^r \sum_{k=r+1}^n S_{pjk} + 4 \sum_{j=r+1}^n \sum_{k=j+1}^n S_{pjk} \right) \right|, \\ N_{rF\varepsilon^2} &= \sum_{p=1}^s \left[\frac{c_p^2}{2\omega_p^2} + 2 \frac{\omega_p^2}{c_p^2} \left(\sum_{\substack{k=1 \\ k \neq p}}^r \frac{c_k^2}{\omega_k^2 - \omega_p^2} \right)^2 + \sum_{\substack{j=1 \\ j \neq p}}^r \frac{3\omega_p^2 c_j^2 + \omega_p^2 c_p^2 + \omega_j^2 c_j^2 + \omega_j^2 c_p^2}{(\omega_p^2 - \omega_j^2)^2} \right]. \end{aligned}$$

Dabei ist

$$S_{pjk} = \frac{c_k^2 c_j^2 \omega_p^2}{c_p^2 (\omega_j^2 - \omega_p^2) (\omega_k^2 - \omega_p^2)} \quad \text{mit } k > j \text{ und } S_{pjk} \begin{cases} > 0 & \text{für } j > p \\ < 0 & \text{für } j < p \end{cases}.$$

Verglichen mit (4.11) gilt also $Z_{rF\varepsilon^2} = Z_{rF}$, und $N_{rF\varepsilon^2}$ stimmt mit N_{rF} bis auf die fehlenden Terme $\frac{2}{c_p^2}$ und den obersten Summenindex r (statt n) überein. Nun übertragen sich die Ergebnisse aus Korollar 4.5 (i), Korollar 4.7 (ii), Satz 4.13, Korollar 4.14 unverändert auf den relativen Fehler der Quadrate der Minimumstelle. In Korollar 4.5 (ii) kann Q_4 jetzt in singulären Fällen zu 0 werden, in Korollar 4.7 (i) ist Q_1 stets gleich 0, was jeweils dazu führt, daß der relative Fehler der Quadrate der Minimumstelle quadratisch divergiert; dies hat wiederum zur Folge, daß der relative Fehler der Minimumstelle mindestens linear divergiert. Statt Korollar 4.10 erhält man wie zu erwarten die Unabhängigkeit des relativen Fehlers der Minimumstelle von einer simultanen Skalierung aller c_k und/oder ω_k . Außerdem gelten – mit kleinen Anpassungen in den Beweisen – Aussagen analog zu den Sätzen 4.16 und 4.17.

4.5 Abschneidung bei mehreren Dämpfern

In den vorherigen Abschnitten des Kapitels 4 haben wir Problem 4.1 explizit für $C = cc^T$, $c \in \mathbb{R}^n$, d.h. für einen Dämpfer untersucht und die in Abschnitt 4.2 zusammengefaßten Ergebnisse bewiesen; anschließend wurde in Abschnitt 4.4 untersucht, welche dieser Ergebnisse auf den relativen Fehler der Minimumstellen übertragbar sind. Dieser Abschnitt beinhaltet nun die Untersuchung von Problem 4.1 für $C = VV^T$, $V \in M_{n,g}$, $g > 1$, wobei die Spalten von V linear unabhängig sind, d.h. für mehrere Dämpfer, und anschließend eine entsprechende Untersuchung des relativen Fehlers der Minimumstellen. C hat damit die Gestalt $C = V \cdot \text{diag}(\varepsilon_1, \dots, \varepsilon_g) \cdot V^T$. Im Gegensatz zu den vorherigen Abschnitten beruhen die Untersuchungen aus diesem Abschnitt 4.5 auf numerischer Basis. Dabei wurden für einige wesentliche der für einen Dämpfer herausgearbeiteten Effekte Testreihen durchgeführt, die untersuchen sollten, ob die Effekte auch auf den Fall mehrerer Dämpfer übertragbar sind. Wir werden sehen, daß dies in einigen Fällen problemlos, in anderen aber nicht ohne weiteres möglich ist. Wir benutzen im folgenden wieder die in 4.1 eingeführten Bezeichnungen *relevanter Bereich*, *Zwischenbereich* und *Bereich hinter dem Schnitt*. Für $j = 1, \dots, g$ bezeichnen die Komponenten $(V)_{ij}$, $i = 1, \dots, s$ die Dämpfungskomponenten im relevanten Bereich, die Komponenten $(V)_{ij}$, $i = s + 1, \dots, r$ werden mit Dämpfungskomponenten im Zwischenbereich und $(V)_{ij}$, $i = r + 1, \dots, n$ mit Dämpfungskomponenten im Bereich hinter dem Schnitt bezeichnet. Sprechen wir von der k -ten Dämpfungskomponente von V , so meinen wir $(V)_{kj}$ für alle $j = 1, \dots, g$. Sprechen wir von der k -ten Dämpfungskomponente des l -ten Dämpfers, dann meinen wir $(V)_{kl}$.

Zunächst betrachten wir den relativen Fehler der Spur.

Es wurden die folgenden Untersuchungen bzgl. des Frequenzspektrums durchgeführt:

- (a) Simultane Variation aller Frequenzen,
- (b) Simultane Vergrößerung aller Frequenzen hinter dem Schnitt.

Es wurden die folgenden Untersuchungen bzgl. der Dämpfungskomponenten durchgeführt:

- (c) Simultane Variation aller Dämpfungskomponenten,
- (d) Simultane Variation aller Dämpfungskomponenten hinter dem Schnitt,
- (e) Variation einer Dämpfungskomponente von V im relevanten Bereich,
- (f) Variation einer Dämpfungskomponente von V im Zwischenbereich.

Mit Variation einer Dämpfungskomponente von V ist gemeint, daß die i -te Zeile von V mit i fest und für alle $j = 1, \dots, g$ durch $\beta \cdot (V)_{ij}$ variiert wird. Für alle

oben genannten Untersuchungen wurden jeweils 3 Testreihen durchgeführt, wobei eine Testreihe den Fall $n = 7, r = 5$, die zweite den Fall $n = 10, r = 7$ und die dritte Testreihe den Fall $n = 15, r = 10$ untersuchte. In jeder Testreihe durchlief s alle Werte zwischen 1 und r , die Anzahl der Dämpfer variierte.

Alle Tests, die den relativen Fehler der Spur betrachten, wurden auf einem PC Pentium 100 mit 32 MB RAM unter Windows 95 mit Matlab 5.1 durchgeführt.

Zu (a): Für die Untersuchung der Auswirkung einer simultanen Variation aller Frequenzen wurden die Tests zum einen für das Frequenzspektrum $\Omega = \Omega_l = \gamma \cdot \text{diag}(1, 2, \dots, n)$ und zum anderen für $\Omega = \Omega_q = \gamma \cdot \text{diag}(1, 4, 9, \dots, n^2)$ durchgeführt. γ wurde innerhalb eines Intervalls $\Gamma = [10^a, 10^b]$ variiert. V wurde dabei mit Rang g gewählt.

In allen durchgeführten Tests konnten disjunkte Teilintervalle $\Gamma_1, \Gamma_2 \subset \Gamma$ mit $\gamma_1 < \gamma_2$ für alle $\gamma_1 \in \Gamma_1, \gamma_2 \in \Gamma_2$ ermittelt werden, so daß eine Verkleinerung von $\gamma \in \Gamma_1$ nahezu keine Auswirkungen auf den relativen Fehler der Spur hatte, während eine Vergrößerung von $\gamma \in \Gamma_2$ zu einer Verkleinerung des relativen Fehlers der Spur führte. Die Tests lassen vermuten, daß hier wie auch bei einem Dämpfer

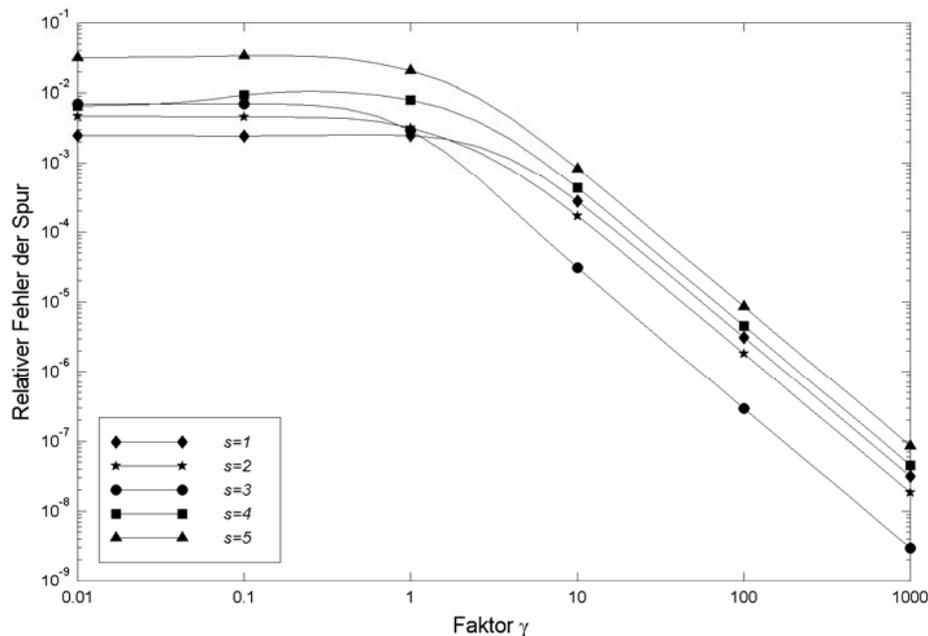


Abb. 4.1: Simultane Variation aller Frequenzen

(vgl. Korollar 4.10) für $\gamma \rightarrow \infty$ gilt

$$\left| \frac{\text{Tr}(XZ_s) - \text{Tr}(\tilde{X}\tilde{Z}_s)}{\text{Tr}(XZ_s)} \right| = O(\gamma^{-2}),$$

während für $\gamma \rightarrow 0$ der relative Fehler der Spur asymptotisch konstant bleibt. (In den Tests konnte γ nur bis zu einer bestimmten Grenze vergrößert bzw. verkleinert werden, da sowohl sehr große als auch sehr kleine Werte aufgrund numerischer Instabilität zu unbrauchbaren Ergebnissen führten.)

Abbildung 4.1 auf der vorherigen Seite verdeutlicht die Ergebnisse des Tests mit $n = 7$, $r = 5$ und $\Omega = \Omega_q$ für $g = 3$, also für drei Dämpfer. Dabei wurde γ zwischen 0,01 und 1000 variiert, V hatte die Gestalt $V = [v_1, v_2, v_3]$ mit $v_1 = (2, 1, 1, 1, 1, 1, 1)^T$, $v_2 = (1, 2, 1, 1, 1, 1, 1)^T$, $v_3 = (1, 1, 2, 1, 1, 1, 1)^T$. Man sieht, daß für γ zwischen 0,01 und 1 der relative Fehler der Spur nahezu konstant bleibt, erst für $\gamma > 1$ fällt der Fehler ab.

Zu (b): Für die Untersuchungen einer simultanen Vergrößerung der Frequenzen hinter dem Schnitt wurden Tests wie oben beschrieben durchgeführt. Dabei wurden die zwei Frequenzspektren $\Omega = \Omega_l = \text{diag}(1, 2, \dots, r, \gamma \cdot (r + 1), \dots, \gamma \cdot n)$ und $\Omega = \Omega_q = \text{diag}(1, 4, 9, \dots, r^2, \gamma \cdot (r + 1)^2, \dots, \gamma \cdot n^2)$ für γ zwischen 1 und 1000 untersucht. Alle durchgeführten Tests ergaben, daß eine Vergrößerung von γ zu einer

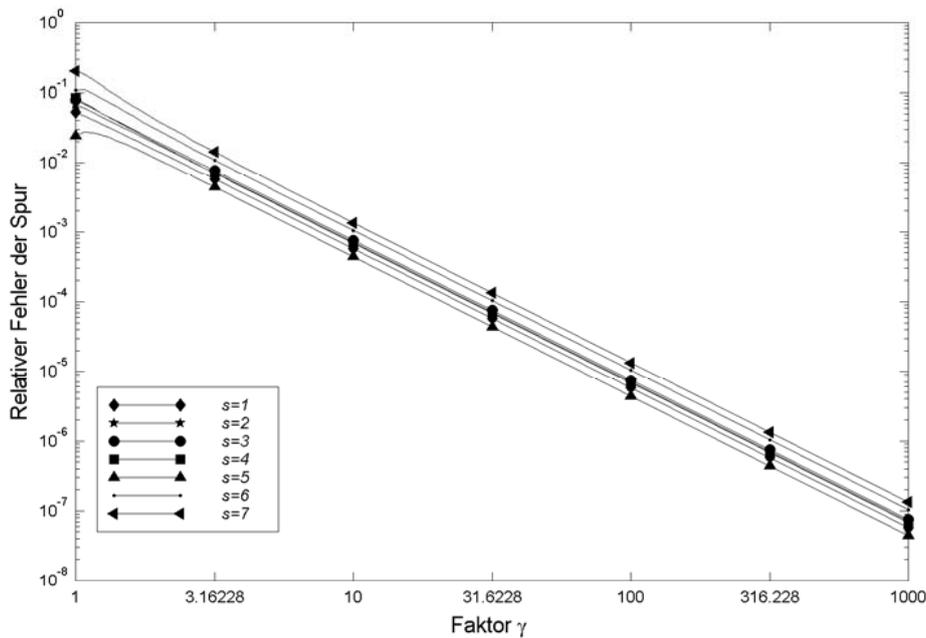


Abb. 4.2: Vergrößerung aller Frequenzen im Bereich hinter dem Schnitt

Verkleinerung des relativen Fehlers der Spur führte. Die Tests lassen vermuten, daß für $\gamma \rightarrow \infty$ die Beziehung (vgl. Korollar 4.14)

$$\left| \frac{\text{Tr}(XZ_s) - \text{Tr}(\tilde{X}\tilde{Z}_s)}{\text{Tr}(XZ_s)} \right| = O(\gamma^{-2})$$

gilt. Der Fall $\gamma \rightarrow 0$ ist hier aufgrund der aufsteigend sortierten Frequenzen nicht möglich.

Abbildung 4.2 auf der vorherigen Seite verdeutlicht den Fall $n = 10$, $r = 7$, $g = 4$ und $\Omega = \Omega_l$. Alle anderen Tests führten zu ähnlichen Graphiken.

Zu (c): Für die Untersuchung der Auswirkung einer simultanen Variation aller Dämpfungskomponenten wurden wieder Tests wie oben beschrieben für die Frequenzspektren $\Omega = \Omega_l = \text{diag}(1, 2, \dots, n)$ und für $\Omega = \Omega_q = \text{diag}(1, 4, \dots, n^2)$ durchgeführt. Für verschiedene g wurde V mit Rang g bestimmt und mittels β skaliert, wobei β innerhalb eines Intervalls $\Gamma = [10^a, 10^b]$ variierte. Für ein festes V wurde also die Dämpfungsmatrix C innerhalb eines Tests durch $C = (\beta V)(\beta V^T)$ variiert. In allen durchgeführten Tests konnten disjunkte Teilintervalle $\Gamma_1, \Gamma_2 \subset \Gamma$ mit $\beta_1 > \beta_2$ für alle $\beta_1 \in \Gamma_1, \beta_2 \in \Gamma_2$ ermittelt werden, so daß eine Vergrößerung von $\beta \in \Gamma_1$ nahezu keine Auswirkungen auf den relativen Fehler der Spur hatte,

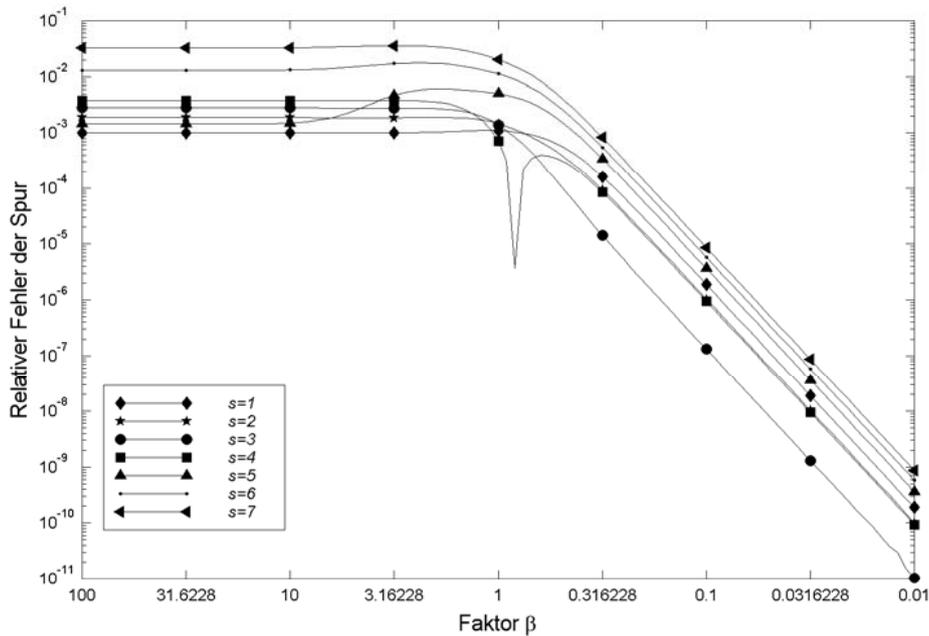


Abb. 4.3: Simultane Verkleinerung aller Dämpfungskomponenten

während eine Verkleinerung von $\beta \in \Gamma_2$ zu einer Verkleinerung des relativen Fehlers der Spur führte. Die Tests lassen vermuten, daß für $\beta \rightarrow 0$

$$\left| \frac{\text{Tr}(XZ_s) - \text{Tr}(\tilde{X}\tilde{Z}_s)}{\text{Tr}(XZ_s)} \right| = O(\beta^4)$$

gilt, während für $\beta \rightarrow \infty$ der relative Fehler asymptotisch konstant bleibt. Dies war auch das Ergebnis für einen Dämpfer (s. Korollar 4.10).

Abbildung 4.3 stellt den Testfall $n = 10$, $\Omega = \Omega_q$, $g = 4$ und $V = \beta \cdot (v_1, v_2, v_3, v_4)$ mit $v_1 = (2, 1, 1, 1, 1, 1, 1, 1, 1, 1)^T$, $v_2 = (1, 2, 1, 1, 1, 1, 1, 1, 1, 1)^T$, $v_3 = (1, 1, 2, 1, 1, 1, 1, 1, 1, 1)^T$, $v_4 = (1, 1, 1, 2, 1, 1, 1, 1, 1, 1)^T$ dar. β wurde zwischen 0,01 und 100 gewählt. Man sieht, daß eine Verkleinerung von β von 100 auf 1 auf den relativen Fehler kaum Auswirkung zeigt, während er für eine Verkleinerung von β zwischen 1 und 0,01 wie $O(\beta^4)$ fällt.

Zu (d): Für die Untersuchung der Auswirkung einer simultanen Skalierung aller Dämpfungskomponenten hinter dem Schnitt wurden Tests wie in (c) beschrieben durchgeführt, wobei V nur im Bereich hinter dem Schnitt durch β skaliert wurde. In allen Tests ergab sich, daß eine beliebige Verkleinerung aller Dämpfungskomponenten hinter dem Schnitt nicht zu einer beliebigen Verkleinerung des relativen Fehlers der Spur führte. Dies kann man deutlich in Abbildung 4.4 sehen, wo für $\beta < 0,1$

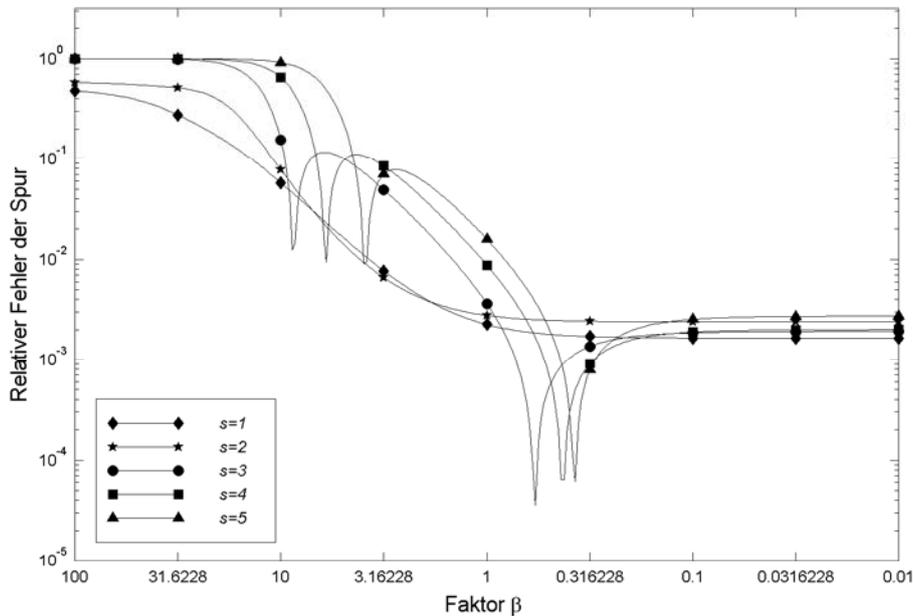


Abb. 4.4: Skalierung aller Dämpfungskomponenten im Bereich hinter dem Schnitt

der relative Fehler der Spur für alle s nahezu konstant bleibt. Dieses Phänomen hatten wir auch für einen Dämpfer in Korollar 4.7 bzw. Bemerkung 4.8 gezeigt. Für eine genügende Vergrößerung von β zeigen die Tests, daß der relative Fehler der Spur für die Fälle $s > g$ gegen 1 konvergiert, während er für die Fälle $s \leq g$ gegen einen möglicherweise kleineren Wert als 1 konvergiert. Auch hier konnten die Tests aufgrund der Rechnerungenauigkeit nur für begrenzt große β durchgeführt werden. Dies hat zur Folge, daß nicht überprüft werden konnte, ob für $s \leq g$ und β genügend groß der relative Fehler doch noch gegen 1 konvergiert. Im wesentlichen decken sich die Beobachtungen in diesen Tests mit den Ergebnissen für einen Dämpfer. Für einen Dämpfer führte allerdings, wie in Korollar 4.7 bewiesen, für alle s (d.h. auch für $s = 1$) eine genügende Vergrößerung der Dämpfungskomponenten hinter dem Schnitt zu dem relativen Fehler 1.

Abbildung 4.4 zeigt die Auswirkung einer simultanen Skalierung aller Dämpfungskomponenten hinter dem Schnitt für den Fall $n = 7$ und $g = 2$ mit $\Omega = \Omega_q$. Man erkennt hier, daß für $\beta > 10$ der relative Fehler der Spur für $s = 3, 4, 5$ bereits den Wert 1 angenommen hat, während er sich für $s = 1$ und $s = 2$ jeweils einem Wert annähert, der möglicherweise unter 1 liegt.

Zu (e): Um die Auswirkungen der Variation einer Dämpfungskomponente von V im relevanten Bereich zu untersuchen, wurden Tests wie in (d) durchgeführt, wobei V diesmal nur in der i -ten Zeile mit $i \in \{1, \dots, s\}$ durch β skaliert wurde, d.h. eine Variation erfolgte durch $\beta \cdot (V)_{ij}$ für alle $j = 1, \dots, g$. Für einen Dämpfer (s. Korollar 4.5 bzw. Bemerkung 4.6) führt eine genügende Vergrößerung einer Dämpfungskomponente im relevanten Bereich zu einer beliebigen Verkleinerung des relativen Fehlers der Spur. Die einzige Ausnahme bildet für einen Dämpfer der Fall, daß der relevante Bereich nur aus einem Element besteht (vgl. Bemerkung 4.6). In diesem Fall kann der relative Fehler nicht durch eine genügende Vergrößerung von β beliebig verkleinert werden. Die Tests für mehrere Dämpfer führten zu ähnlichen, aber nicht ganz auf einen Dämpfer übertragbaren, Ergebnissen. Es zeigte sich, daß prinzipiell für jede Anzahl von Dämpfern eine genügende Vergrößerung einer Dämpfungskomponente von V im relevanten Bereich durch die Skalierung mit β zu einer Verkleinerung des relativen Fehlers der Spur führte. Jedoch gab es auch hier Ausnahmen. Bei Variation einer beliebigen Dämpfungskomponente von V im relevanten Bereich lassen alle durchgeführten Tests vermuten, daß für $g \geq 2$ und $s \neq g + 1$ stets für $\beta \rightarrow \infty$

$$\left| \frac{\text{Tr}(XZ_s) - \text{Tr}(\tilde{X}\tilde{Z}_s)}{\text{Tr}(XZ_s)} \right| = O(\beta^{-2})$$

gilt. Für $s = g + 1$ ergaben die Tests, daß der relative Fehler der Spur durch eine genügende Vergrößerung einer Dämpfungskomponente im relevanten Bereich ebenfalls wie $O(\beta^{-2})$ fällt, wenn $i \neq s$. Eine Ausnahme bildete in allen Tests der Fall, daß für $s = g + 1$ die $(g + 1)$ -te Dämpfungskomponente von V vergrößert wurde.

In diesem Fall blieb der relative Fehler der Spur asymptotisch konstant. Die Tests lassen außerdem vermuten, daß der relative Fehler der Spur für eine genügende Verkleinerung von β gegen einen positiven Grenzwert konvergiert.

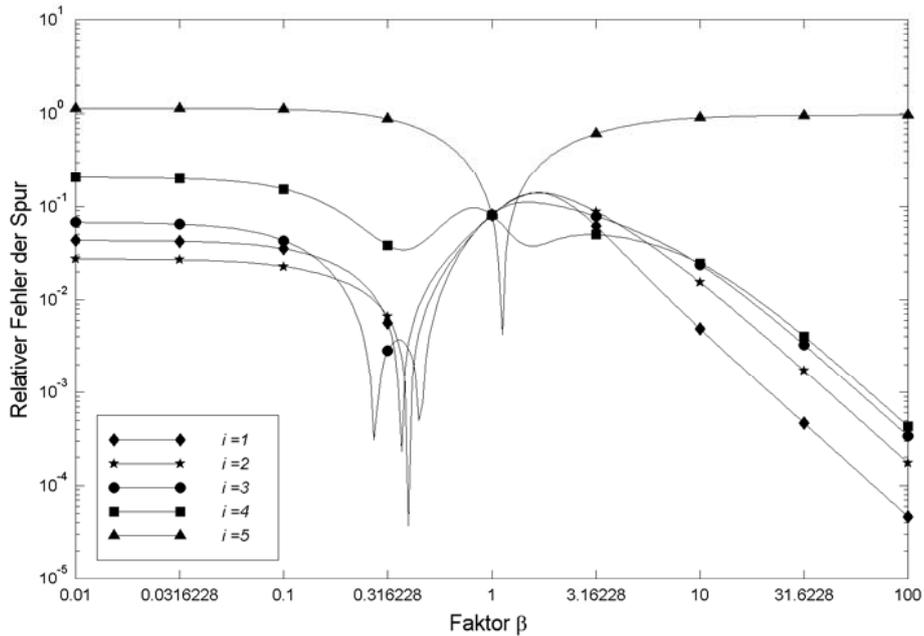


Abb. 4.5: Vergrößerung einer Dämpfungskomponente im relevanten Bereich

Abbildung 4.5 veranschaulicht die Auswirkungen einer Variation einer Dämpfungskomponente von V im relevanten Bereich auf den relativen Fehler der Spur. Die Graphik zeigt die Ergebnisse für den Fall $n = 7$, $r = 5$, $\Omega = \Omega_l$, $g = 4$ und $s = 5$. Hier wurde jede der 5 Dämpfungskomponenten $(V)_{ij}$, $j = 1, \dots, g$, $i \in \{1, \dots, 5\}$ im relevanten Bereich durch β skaliert. Jeder Graph repräsentiert die Skalierung einer Dämpfungskomponente. Man erkennt sehr deutlich, daß für $\beta < 0,1$ der relative Fehler der Spur für alle i nahezu konstant bleibt. Für $\beta > 3$ dagegen fällt der relative Fehler für eine Skalierung der Dämpfungskomponenten $i = 1, 2, 3, 4$ wie $O(\beta^{-2})$ ab. Im Gegensatz hierzu konvergiert der relative Fehler der Spur für $i = 5$, d.h. für die $(g + 1)$ -te Dämpfungskomponente, gegen einen Grenzwert, der bei 1 liegt.

Zu (f): Um die Auswirkungen der Variation einer Dämpfungskomponente von V im Zwischenbereich zu untersuchen, wurden Tests wie in (d) durchgeführt, wobei V diesmal nur in der i -ten Zeile mit $i \in \{s + 1, \dots, r\}$ durch β skaliert wurde, d.h. eine Variation erfolgte durch $\beta \cdot (V)_{ij}$ für alle $j = 1, \dots, g$. In allen durchgeführten Tests konnten die folgenden Phänomene beobachtet werden. Für $s \leq g$ blieb der relative

Fehler der Spur für eine Verkleinerung und für eine Vergrößerung einer Dämpfungskomponente im Zwischenbereich asymptotisch konstant. Für $s > g$ dagegen führte eine Vergrößerung einer Dämpfungskomponente im Zwischenbereich zu einer Verkleinerung des relativen Fehlers. Die Tests lassen vermuten, daß für den Fall $s > g$ für $\beta \rightarrow \infty$

$$\left| \frac{\text{Tr}(XZ_s) - \text{Tr}(\tilde{X}\tilde{Z}_s)}{\text{Tr}(XZ_s)} \right| = O(\beta^{-2})$$

gilt. Dies deckt sich im wesentlichen mit den Ergebnissen für einen Dämpfer, die in Korollar 4.5 bzw. Bemerkung 4.6 bewiesen wurden. Für einen Dämpfer führt allerdings für alle s , insbesondere auch für $s = 1$, eine genügende Vergrößerung einer Dämpfungskomponente im Zwischenbereich zu einer beliebigen Verkleinerung des relativen Fehlers.

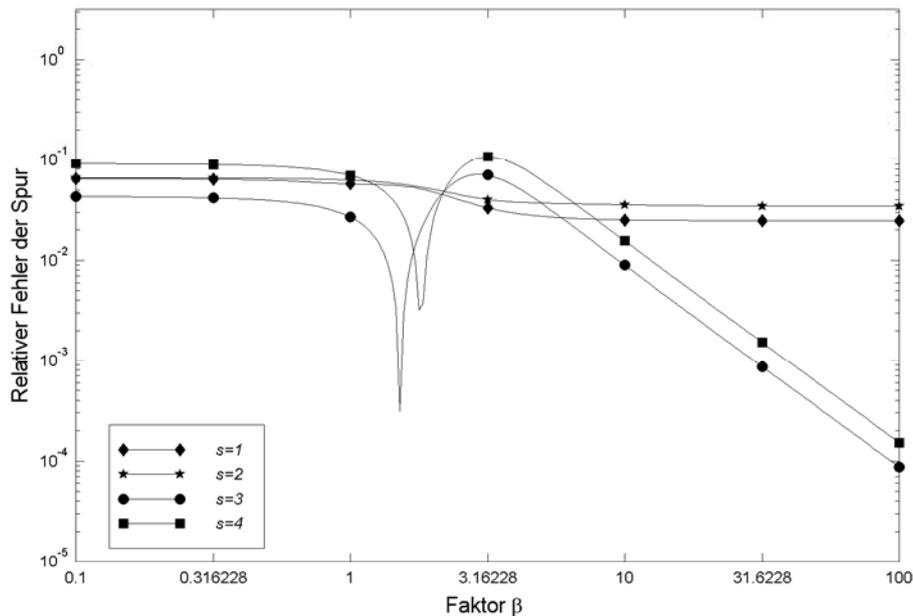


Abb. 4.6: Vergrößerung der ersten Dämpfungskomponente im Zwischenbereich

Abbildung 4.6 veranschaulicht eine Variation der ersten Dämpfungskomponente von V hinter dem Schnitt, d.h. eine Skalierung erfolgt durch $\beta \cdot (V)_{(s+1)j}$ für $j = 1, \dots, g$ für den Fall $n = 7$, $g = 2$ und $\Omega = \Omega_l$, $r = 5$. s wurde nur bis 4 gewählt, da für $s = 5$ kein Zwischenbereich vorhanden wäre. Man erkennt deutlich, daß im Bereich zwischen 0,1 und 1 eine Skalierung durch β auf den relativen Fehler kaum Einfluß hat. Dies bleibt für $s = 1$ und $s = 2$ auch im weiteren Verlauf von β erhalten. Dagegen fällt der relative Fehler für $s = 3$ und $s = 4$ für $\beta > 3$ wie $O(\beta^{-2})$ ab.

Anschließend sollen noch einige Testreihen beschrieben werden, die durchgeführt wurden, um die optimalen Viskositäten des ungeschnittenen Systems mit denen des geschnittenen Systems zu vergleichen. Dabei wurde das Gütekriterium

$$\max_{i=1,\dots,g} \left\{ \frac{|\varepsilon_i^{(\min)} - \tilde{\varepsilon}_i^{(\min)}|}{\varepsilon_i^{(\min)}} \right\} \quad (4.32)$$

gewählt, wobei $\varepsilon^{(\min)} = (\varepsilon_1^{(\min)}, \dots, \varepsilon_g^{(\min)})$ der Viskositätenvektor ist, der $\text{Tr}(XZ_s)$ minimiert und $\tilde{\varepsilon}^{(\min)} = (\tilde{\varepsilon}_1^{(\min)}, \dots, \tilde{\varepsilon}_g^{(\min)})$ der Vektor, der $\text{Tr}(\tilde{X}\tilde{Z}_s)$ minimiert. Wir werden (4.32) im folgenden als *relativen Fehler der Minimumstelle* bezeichnen. Entsprechend zu den Testreihen (b), (d), (e) und (f) des relativen Fehlers der Spur wurden für die Untersuchung von (4.32) die folgenden Tests durchgeführt:

- (b') Simultane Vergrößerung aller Frequenzen hinter dem Schnitt,
- (d') Simultane Variation aller Dämpfungskomponenten hinter dem Schnitt,
- (e') Variation einer Dämpfungskomponente von V im relevanten Bereich,
- (f') Variation einer Dämpfungskomponente von V im Zwischenbereich.

Eine simultane Variation aller Dämpfungskomponenten bzw. aller Frequenzen hat auf die Minimierung von $\text{Tr}(XZ_s)$ keinen Einfluß (siehe auch Abschnitt 4.4), so daß entsprechende Tests zu (a) und (c) hier nicht durchgeführt wurden. In allen Tests wurden die optimalen Viskositäten $\varepsilon^{(\min)}$ und $\tilde{\varepsilon}^{(\min)}$ mittels des im Anhang (Kapitel 5) beschriebenen Newtonverfahren ermittelt. Ansonsten wurden die Testreihen analog zu den Testreihen für den relativen Fehler der Spur durchgeführt, so daß sie im folgenden nicht extra beschrieben werden. Wie schon in Abschnitt 2.2 angeführt, trat der Fehler $\varepsilon_i^{(\min)} = 0$ in den durchgeführten Tests niemals auf, so daß die Betrachtung von (4.32) ohne Komplikationen möglich war.

Alle Tests, die den relativen Fehler der Minimumstelle untersuchen, wurden auf einer Sun Ultra 2/1300 mit 128 MB RAM unter Solaris mit Matlab 5.0 durchgeführt.

Zu (b'): Für die Untersuchung der Auswirkung einer simultanen Vergrößerung der Frequenzen hinter dem Schnitt auf den Fehler (4.32) wurden Tests wie in (b) beschrieben durchgeführt. Es ergaben sich tendenziell die selben Ergebnisse wie in (b), insbesondere lassen auch hier die Tests vermuten, daß

$$\max_{i=1,\dots,g} \left\{ \frac{|\varepsilon_i^{(\min)} - \tilde{\varepsilon}_i^{(\min)}|}{\varepsilon_i^{(\min)}} \right\} = O(\gamma^{-2})$$

für $\gamma \rightarrow \infty$ gilt.

Abbildung 4.7 verdeutlicht den Fall $n = 12$, $g = 2$, $r = 9$ und $\Omega = \Omega_l$. Man erkennt, daß die Abbildungen 4.7 und 4.2 ähnlich zueinander sind.

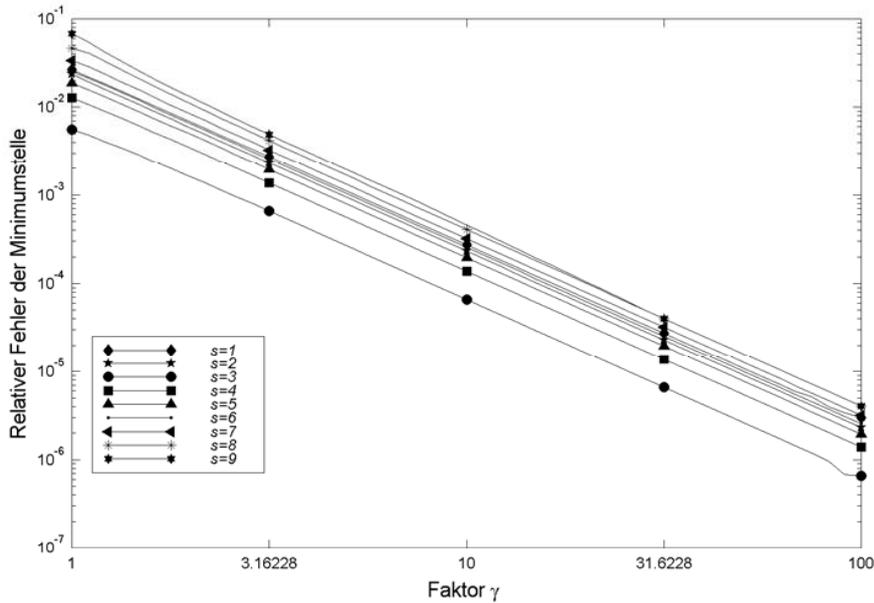


Abb. 4.7: Vergrößerung aller Frequenzen im Bereich hinter dem Schnitt

Zu (d'): Auch die durchgeführten Tests zur Untersuchung der Auswirkung einer simultanen Variation aller Dämpfungskomponenten hinter dem Schnitt auf den relativen Fehler der Minimumstelle (4.32) ergaben ähnliche Ergebnisse wie die entsprechenden Tests für den relativen Fehler der Spur.

Abbildung 4.8 auf der nächsten Seite zeigt exemplarisch das Ergebnis des Tests für $n = 7$, $g = 2$, $\Omega = \Omega_q$. Man kann erkennen, daß eine simultane Verkleinerung der Dämpfungskomponenten (hier etwa für $\beta < 0.2$) hinter dem Schnitt dazu führt, daß der Fehler (4.32) asymptotisch konstant bleibt. Eine Verkleinerung der Dämpfungskomponenten hinter dem Schnitt um den Skalierungsfaktor β mit $\beta \rightarrow 0$ führt also nicht zu einer beliebigen Verkleinerung des relativen Fehlers der Minimumstelle (dies galt auch für den relativen Fehler der Spur). Ein Unterschied zu dem Ergebnis beim relativen Fehler der Spur ergibt sich bei einer beliebigen Vergrößerung des Skalierungsfaktors β . Hier läuft der Fehler (4.32) gegen ∞ , während er in Abbildung 4.4 gegen 1 läuft. Dieser Unterschied trat auch schon bei einem Dämpfer auf (s. Kap. 4.4).

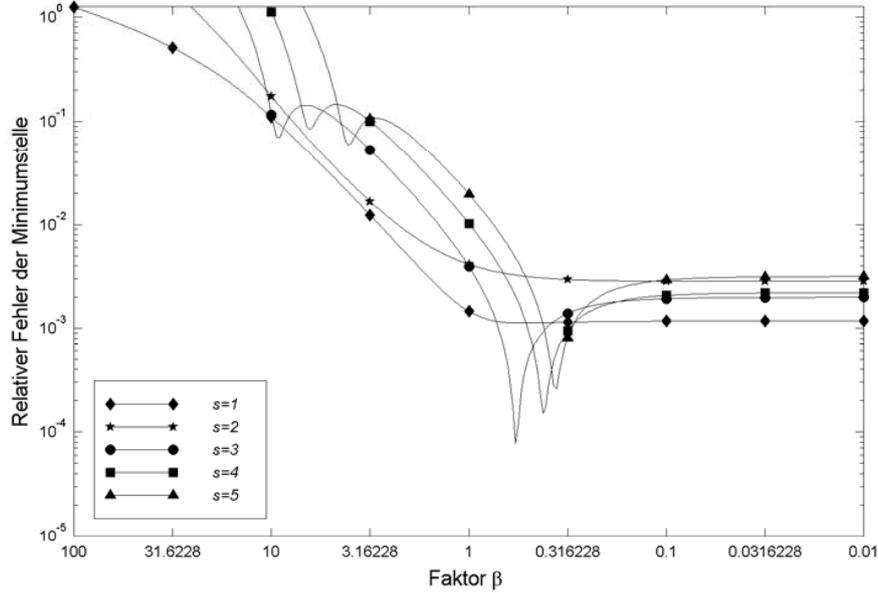


Abb. 4.8: Simultane Variation aller Dämpfungskomponenten hinter dem Schnitt

Zu (e'): Um die Auswirkung der Variation einer Dämpfungskomponente von V im relevanten Bereich zu untersuchen, wurden Tests wie in (e) beschrieben durchgeführt. V wurde also in der i -ten Zeile mit $i \in \{1, \dots, s\}$ durch β skaliert. Für einen Dämpfer ergaben sich hierbei ähnliche Ergebnisse wie bei der Betrachtung des relativen Fehlers der Spur (vgl. Abschnitt 4.4). Auch die durchgeführten Tests für mehrere Dämpfer zeigen eine Übertragbarkeit der Ergebnisse. Auch hier lassen die Tests (vgl. (e)) vermuten, daß für $g \geq 2$ und $s \neq g + 1$ stets für $\beta \rightarrow \infty$

$$\max_{i=1, \dots, g} \left\{ \frac{|\varepsilon_i^{(\min)} - \tilde{\varepsilon}_i^{(\min)}|}{\varepsilon_i^{(\min)}} \right\} = O(\beta^{-2}) \quad (4.33)$$

gilt. Für $s = g + 1$ ergaben die Tests ebenfalls, daß der relative Fehler der Minimumstelle durch eine genügende Vergrößerung von β wie $O(\beta^{-2})$ fällt, falls $i \neq s$ gilt. Eine Ausnahme bildet dabei der Fall (wie auch schon beim relativen Fehler der Spur), daß für $s = g + 1$ die $(g + 1)$ -te Dämpfungskomponente von V vergrößert wurde. In diesem Fall blieb der relative Fehler der Minimumstelle asymptotisch konstant. Außerdem lassen die Tests vermuten, daß der relative Fehler der Minimumstelle für eine genügende Verkleinerung von β gegen einen positiven Grenzwert konvergiert.

Abbildung 4.9 veranschaulicht den Fall $n = 7$, $r = 5$, $\Omega = \Omega_l$, $g = 4$ und $s = 5$. Es wurde jede der 5 Dämpfungskomponenten $(V)_{ij}$, $j = 1, \dots, g$, $i \in \{1, \dots, 5\}$ im

relevanten Bereich durch β skaliert. Wie auch in Abbildung 4.5 repräsentiert jeder Graph die Skalierung einer Dämpfungskomponente. Man erkennt, daß für $\beta < 0,1$ der relative Fehler der Minimumstelle für alle i nahezu konstant bleibt. Für $\beta > 3$ dagegen fällt er für eine Skalierung der i -ten Dämpfungskomponente mit $i = 1, 2, 3, 4$ wie $O(\beta^{-2})$ ab, wobei er für $i = 5$ gegen einen Grenzwert konvergiert, der bei 1 liegt. Im Bereich zwischen $\beta = 0,1$ und $\beta = 3$ sind die Graphen sehr unruhig (vgl. den entsprechenden Bereich in Abbildung 4.5). Dies mag daran liegen, daß bei allen Tests, die zur Untersuchung des relativen Fehlers der Minimumstelle durchgeführt wurden, stets mehrfach die Ljapunovgleichung gelöst werden mußte, was zu Ungenauigkeiten führen kann.

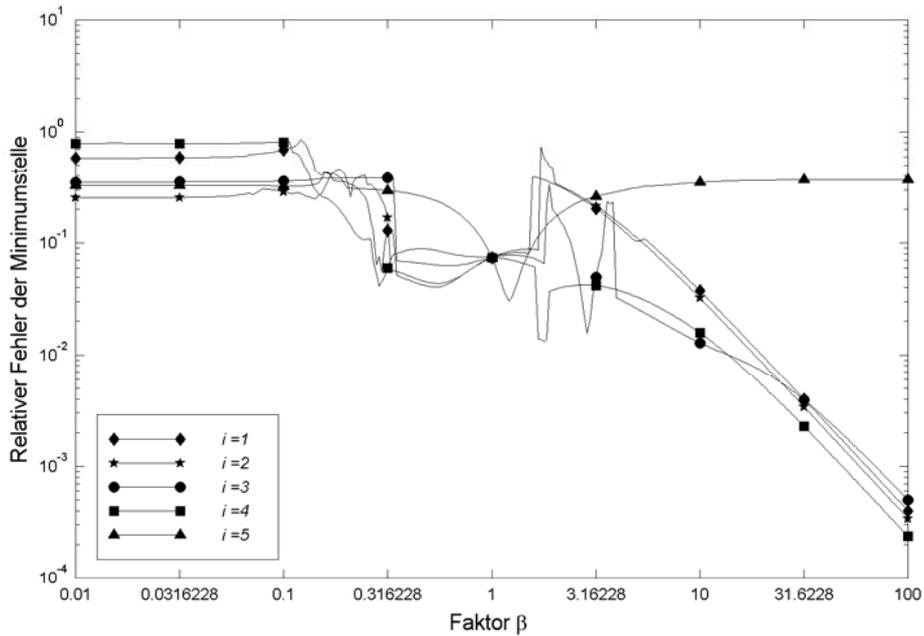


Abb. 4.9: Vergrößerung einer Dämpfungskomponente im relevanten Bereich

Zu (f'): Bei den Tests zur Untersuchung der Auswirkung einer Variation einer Dämpfungskomponente von V im Zwischenbereich auf den Fehler (4.32) wurde wieder Tests wie in (f) beschrieben durchgeführt, d.h. V wurde in der i -ten Zeile mit $i \in \{s + 1, \dots, r\}$ durch β skaliert. Die Ergebnisse des Tests lassen vermuten, daß für den Fall $s \neq g$ für $\beta \rightarrow \infty$

$$\max_{i=1, \dots, g} \left\{ \frac{|\varepsilon_i^{(\min)} - \tilde{\varepsilon}_i^{(\min)}|}{\varepsilon_i^{(\min)}} \right\} = O(\beta^{-2}) \quad (4.34)$$

gilt. Für $s = g$ dagegen bleibt der Fehler (4.32) für $\beta \rightarrow \infty$ asymptotisch konstant. Dies deckt sich im wesentlichen mit den Ergebnissen aus der Testreihe (f). Nur der

Fall $s < g$ unterscheidet sich (der relative Fehler der Spur blieb auch für $s < g$ asymptotisch konstant).

Abbildung 4.10 zeigt den Fall $n = 7$, $g = 2$, $\Omega = \Omega_l$, $i = s + 1$.

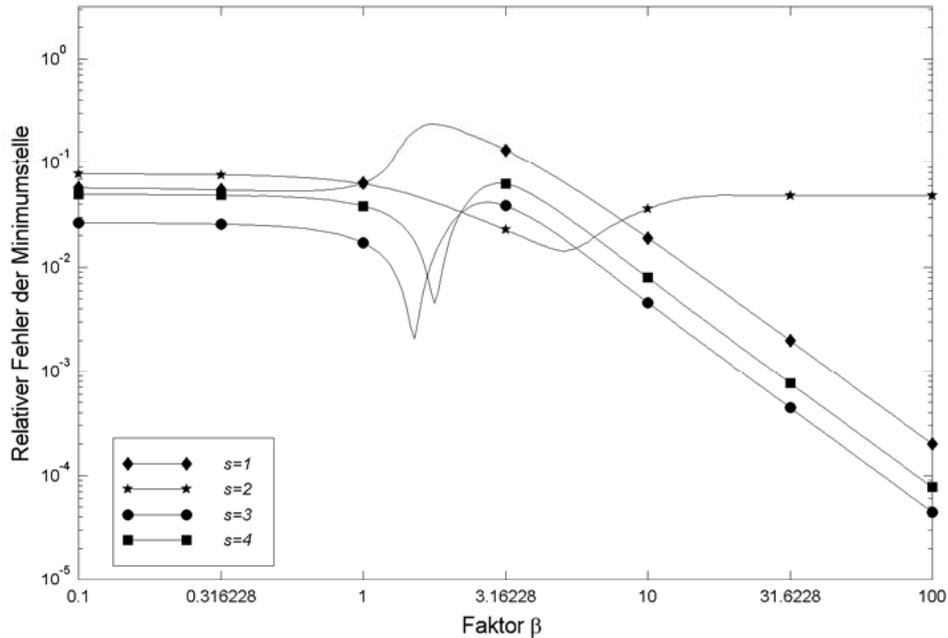


Abb. 4.10: Variation einer Dämpfungskomponente im Zwischenbereich

Die durchgeführten Testreihen zum relativen Fehler der Spur und zum relativen Fehler der Minimumstelle (4.32) haben ergeben, daß auch im Falle mehrerer Dämpfer die Ergebnisse im wesentlichen übertragbar sind. Die wenigen Unterschiede zwischen dem relativen Fehler der Spur und dem relativen Fehler der Minimumstelle (siehe Kap. 4.4) scheinen auch im den Fall mehrerer Dämpfer aufzutreten. Vergleicht man die Ergebnisse der durchgeführten Testreihen auf den in Abschnitt 4.1 und 4.4 beschriebenen Fall für einen Dämpfer, so zeigt sich, daß von den Effekten aus Abschnitt 4.1 diejenigen, die von den Frequenzen abhängen, auch im Fall mehrerer Dämpfer auftreten. Bei denen, die mit der Variation von Dämpfungskomponenten zusammenhängen, ist das nicht unbedingt der Fall, insbesondere dann nicht, wenn es um die Variation einzelner Dämpfungskomponenten geht. Bereits die relativ kleine Auswahl von Effekten, zu denen Testreihen durchgeführt wurden, zeigt bei der Betrachtung des relativen Fehlers der Spur, daß insbesondere für $g > 1$, $g \geq s$ andere, teils weniger günstige Effekte auftreten (siehe (f)), während man vermuten kann, daß die Fälle $s > g > 1$ sich eher bei den Fällen $s \geq g = 1$ einreihen.

Aber auch für $s = g + 1$ tauchten Besonderheiten auf, die bei einem Dämpfer nicht vorhanden sind (siehe (e)). Entsprechendes scheint für den relativen Fehler der Minimumstelle zu gelten (abgesehen davon, daß der Fall $s = g$ wieder eher günstiger einzuschätzen ist). Eine systematische Untersuchung müßte derartig komplexe Fallunterscheidungen treffen (Variation einer bzw. aller Dämpfungskomponenten eines bzw. aller Dämpfer in einem oder mehreren Bereichen), daß sie im Rahmen dieser Arbeit nicht durchführbar ist. Dabei ist noch zu beachten, daß schon im Fall $g = 1$ singuläre Fälle auftreten (vgl. Korollar 4.5 und Korollar 4.7 sowie die entsprechenden Ergebnisse aus Abschnitt 4.4), die sich einer Untersuchung durch numerische Testreihen entziehen. Die exakte Untersuchung der Abschneidung für den Fall mehrerer Dämpfer bleibt ein offenes Problem.

5 Anhang

Programmpaket zur Optimierung einer Dämpferkonfiguration

Der Anhang beschreibt das in Zusammenhang mit Kapitel 3 erstellte Programmpaket zur Optimierung einer Dämpferkonfiguration über das Gradienten- oder Newtonverfahren.

Für die praktische Berechnungen betrachtet man dabei in dem gedämpften linearen Schwingungssystem (1.1)

$$M\ddot{x} + D\dot{x} + Kx = 0 \quad (5.1)$$

Dämpfungsmatrizen der Form

$$D = D_\alpha + P \cdot \text{diag}(\varepsilon_1, \dots, \varepsilon_g) \cdot P^T, \quad (5.2)$$

wobei, wie in Kapitel 3 angegeben, die Informationen über Anzahl, Positionen und Viskositäten der Dämpfer in der Matrix $P \cdot \text{diag}(\varepsilon_1, \dots, \varepsilon_g) \cdot P^T$ enthalten sind. Die Matrix P der Dimension (n, g) enthält die euklidischen Einheitsvektoren, die den Dämpfungspositionen entsprechen, und $\varepsilon_1, \dots, \varepsilon_g$ sind die zugehörigen (nichtnegativen) Viskositäten der Dämpfer. Die Viskositäten können auch durch die typischen Einschränkungen

$$\varepsilon_{i+1} = \gamma \cdot \varepsilon_i, \quad \varepsilon_{i+2} = \beta \cdot \varepsilon_i \quad (5.3)$$

mit $i = 1, 4, \dots, (g-2)$ mit einander verbunden sein. In diesem Fall muß g durch 3 teilbar sein. Für die Anzahl g der Dämpfer gilt theoretisch nur $g \leq n$; in der Praxis ist g meist viel kleiner als n .

Die Matrix D_α enthält Informationen über die sogenannte Materialdämpfung. Im allgemeinen setzt man $D_\alpha > 0$ klein. Dadurch wird die Stabilität des Systems gesichert. Damit erhalten wir für die Matrix A aus (1.16)

$$\begin{aligned} A &= \begin{pmatrix} 0 & \Omega \\ -\Omega & -C \end{pmatrix} \text{ mit} \\ C &= \Phi^T D_\alpha \Phi + V \cdot \text{diag}(\varepsilon_1, \dots, \varepsilon_g) \cdot V^T \\ &= 2\alpha\Omega + V \cdot \text{diag}(\varepsilon_1, \dots, \varepsilon_g) \cdot V^T, \end{aligned} \quad (5.4)$$

wobei die letzte Zeile gleichzeitig als Definition von D_α dient; gewöhnlich wird $\alpha = 0.02$ gesetzt. Problem 3.1 ist nun mit diesem A zu lösen, wobei gleichzeitig davon ausgegangen wird, daß die ω_i aufsteigend sortiert sind; s wird durch $s = \max\{i : \omega_i < \omega_{\max}\}$ definiert, wobei ω_{\max} vorgegeben ist.

Mit dem System Matlab steht eine Programmiersprache zur Verfügung, welche eine effiziente Umsetzung des beschriebenen Optimierungsproblems 3.1 zuläßt. Die nachfolgend beschriebenen Matlab-Programme `transfer`, `calcvisc`, `minnew` und `mingrad` mit ihren Unterprogrammen sind eine Implementierung des Optimierungsproblems auf numerischer Ebene. Mit ihnen lassen sich, ausgehend von der Massenmatrix M , der Steifigkeitsmatrix K , dem Wert für die Materialdämpfung α und der Anzahl und den Positionen der Dämpfer, Viskositäten der Dämpfer bestimmen, so daß $Tr(XZ_s)$ minimal wird.

Analog zu den Programmen `calcvisc` und `mingrad` existieren die Programme `calcvis3` und `mingrad3`, die im Falle der durch (5.3) gegebenen Abhängigkeit der Viskositäten einzusetzen sind.

Die Programme `transfer`, `calcvisc`, `minnew`, `mingrad`

Ausgehend von einer Massenmatrix M und einer Steifigkeitsmatrix K , sowie von den Positionen und Anzahl der vorhandenen Dämpfer erzeugt die Funktion `transfer` den zum System gehörigen aufsteigend sortierten positiven Frequenzvektor ω sowie die Matrix V , die die transformierten Dämpferpositionen enthält. Dies sind die Komponenten der $2n$ -dimensionalen Phasenraummatrix A aus (5.4).

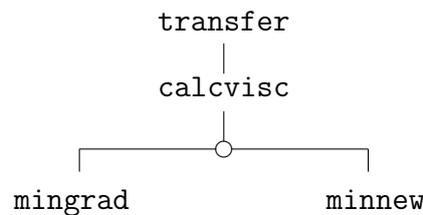


Abb. 5.1: Zeitliche Reihenfolge der Programmabläufe

Die Funktion `calcvisc` bestimmt mit Hilfe der Startwertberechnung aus Kapitel 3.1.1 einen positiven Viskositätenvektor $(\varepsilon_1, \dots, \varepsilon_g)$ aus dem Frequenzvektor ω und der Matrix V unter Berücksichtigung von ω_{\max} und der Materialdämpfungsmatrix $2\alpha \text{diag}(\omega) = 2\alpha\Omega$. Dieser Viskositätenvektor kann in den Programmen `mingrad` und `minnew` als Startwert für die Minimierung von $Tr(XZ_s)$ durch das Gradienten- bzw. Newtonverfahren verwendet werden. Diese Programme liefern dann einen neuen Viskositätenvektor, für den die durch ω_{\max} geschnittene Spur der Ljapunovgleichung

$A^T X + X A = -I$ bis zu einer gegebenen Genauigkeit numerisch minimiert ist.

In Abbildung 5.1 wie auch in den folgenden Abbildungen bedeutet \circ die Ausführung eines der angegebenen Äste, \bullet die Ausführung beider Äste.

Programmbeschreibungen

Die folgende Darstellung geht kurz auf die Besonderheiten der Programme (und ihrer Unterprogramme) ein. Des weiteren werden die Ein- und Ausgabeparameter der Programme (und ihrer Unterprogramme) dargelegt.

Sei n die Dimension der gegebenen Massenmatrix und g die Anzahl der Dämpfer, dann gibt es im wesentlichen folgende Übergabeparameter:

<i>mass</i>	=	reelle positiv definite Matrix der Ordnung n (entspricht Massenmatrix M);
<i>stiff</i>	=	reelle positiv definite Matrix der Ordnung n (entspricht Steifigkeitsmatrix K);
<i>posdamp</i>	=	Zeilenvektor der Länge g , der die Dämpferpositionen enthält (entspricht Matrix P , falls jede Position mit ihrem entsprechenden Einheitsvektor der Länge n identifiziert wird);
<i>omega</i>	=	Spaltenvektor der Länge n , der die aufsteigend sortierten Frequenzen des Systems enthält (entspricht ω);
<i>damp0</i>	=	Matrix der Dimension (n, g) . <i>damp0</i> enthält die transformierten Dämpfungspositionen (entspricht V);
<i>damp1</i>	=	Wert für Materialdämpfung (entspricht α);
<i>omax</i>	=	das betrachtete Frequenzspektrum wird durch <i>omax</i> begrenzt (entspricht ω_{max});
<i>visc</i>	=	Spaltenvektor der Länge g , der die Viskositäten der Dämpfer enthält (entspricht $(\varepsilon_1, \dots, \varepsilon_g)$).

Sämtliche Programme benutzen zur Lösung der Ljapunovgleichungen die Programme *ljap*, *tljap*, *tljapy*, *ljaputy*, die aus dem Programmpaket, das in [25] beschrieben wird, übernommen worden sind. Sie unterscheiden sich nur in den verschiedenen Rückgabewerten. Auf sie wird hier nicht näher eingegangen.

• **transfer**

Eingabe

mass, *stiff*, *posdamp*

Ausgabe

omega, *damp0*

Funktionsweise

(siehe oben).

- `calcvisc`

Eingabe

omega, damp0, damp1, omac

Ausgabe

visc, otv

otv = Matrix, deren Spalten jeweils Ergebnisse einer Lösung einer Ljapunovgleichung (1.17) enthalten. Jede Spalte enthält an dritter bis $(g + 2)$ -ter Stelle die Viskositäten der Dämpfer (variiert nach der Anmerkung zum Unterprogramm `olsearch`, s.u.), an erster Stelle steht die entsprechende durch ω_{max} geschnittene Spur von X , die sich mit den jeweiligen Viskositäten ergibt, an 2. Stelle steht das Maximum der Realteile aller Eigenwerte der zugehörigen Matrix A , deren Imaginärteile innerhalb des durch ω_{max} begrenzten Frequenzstreifens liegen.

Funktionsweise

Das Programm `calcvisc` benutzt das Programm `minguess`, um den Viskositätenvektor $\tilde{\varepsilon} = (\tilde{\varepsilon}_1, \dots, \tilde{\varepsilon}_g)^T$ aus Kapitel 3.1.1 zu ermitteln. Die eigentliche Berechnung von $\tilde{\varepsilon}$ findet dabei in den weiteren Unterprogrammen `vguess` oder `cguess` statt.

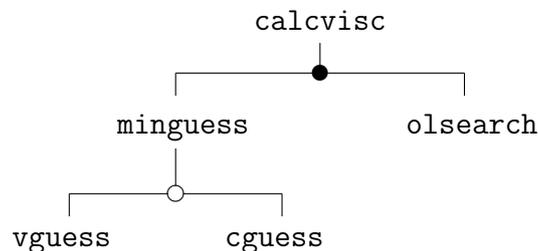


Abb. 5.2: Unterprogramme von `calcvisc`

Beide Programme berechnen $\tilde{\varepsilon}$, indem sie (3.6) mit Hilfe der Methode der kleinsten Quadrate lösen. Aus Optimierungsgründen wurden in der Umsetzung der Methode der kleinsten Quadrate jedoch die Fälle $s = n$ und $s < n$ unterschieden. Der Fall $s = n$ ist in `vguess` und der Fall $s < n$ ist in `cguess` implementiert.

Das Unterprogramm `olsearch` wurde aus dem Programmpaket, das in [25] beschrieben wird, übernommen. Es überprüft die Güte des berechneten Viskositätenvektors $\tilde{\varepsilon}$, indem es einige Spuren der Lösungen der Ljapunovgleichungen für verschiedene andere Viskositätenvektoren der Form $(1 + \delta) \cdot \tilde{\varepsilon}$,

$|\delta| < k$ zusammen mit $\tilde{\varepsilon}$ liefert. Diese Ergebnisse werden über die Matrix otv zurückgegeben. Das Programm `calcvisc` gibt die berechnete Matrix otv sowie den Viskositätenvektor $visc$ zurück. $visc$ ist dabei der Viskositätenvektor, der in der Matrix otv die kleinste Spur liefert.

- **mingrad**

Eingabe

ω , $damp0$, $visc$, $damp1$, $omax$, $StepSize$, $MinStepSize$, $ExaktGrd$

- $StepSize$ = beginnende Schrittweite für Annäherung in Richtung des Gradienten;
- $MinStepSize$ = Schrittweite, die bei Unterschreitung zum Ende des Verfahrens führt;
- $ExaktGrd$ = relativer Exaktheitsgrad für Abbruch des Verfahrens, d.h. mit $ExaktGrd$ wird das Verfahren abgebrochen, falls $Spur_{j-1} - Spur_j < \frac{Spur_j}{ExaktGrd}$

Ausgabe

Erg , $ViscNeu$, $AnzIt$

- Erg = Ergebnismatrix; jede Zeile beinhaltet das Ergebnis eines Iterationsschrittes; die Matrix enthält in
 - Spalte 1 $Tr(XZ_s)$ am Ende des jeweiligen Iterationsschrittes,
 - Spalte 2 den maximalen Realteil alle Eigenwerte von A , deren Imaginärteile innerhalb des durch $\omega_{max}=omax$ begrenzten Frequenzstreifens liegen;
 - die restlichen Spalten bilden den Viskositätenvektor des Iterationsschrittes;
- $ViscNeu$ = der durch das Verfahren bestimmte optimierte Viskositätenvektor; (entspricht dem Viskositätenvektor der letzten Zeile der Matrix Erg);
- $AnzIt$ = Anzahl der durchgeführten Iterationen.

Funktionsweise

Das Programm `mingrad` dient der Minimierung von $Tr(XZ_s)$ über der Menge der Viskositäten $\{\Sigma\}$ mittels des Gradientenverfahrens. Die eigentliche Minimierung findet in den weiteren Unterprogrammen `vollgrad` oder `cutgrad` statt.

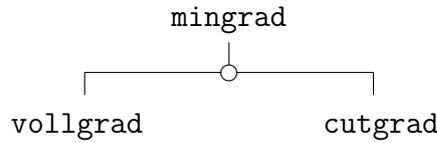


Abb. 5.3: Unterprogramme von mingrad

Das Programm `mingrad` unterscheidet die Fälle $s < n$ und $s = n$, in denen eine unterschiedliche Gradientenberechnung erfolgt. Der Fall $s = n$ ist in `vollgrad` und der Fall $s < n$ ist in `cutgrad` implementiert. Neben der einmaligen Lösung der Ljapunovgleichung $A^T X + X A = -I$ pro Iterationsschritt zur Berechnung von $Tr(X Z_s)$ in beiden Fällen, wird für den Fall $s < n$ für die Berechnung des gesamten Gradientenvektors zusätzlich eine Ljapunovgleichung pro Iterationsschritt gelöst. Im Fall $s = n$ erfolgt die gesamte Berechnung des Gradientenvektors ohne weitere Lösung einer Ljapunovgleichung. Damit wird im Fall $s < n$ pro Iterationsschritt zweimal die Lösung einer Ljapunovgleichung berechnet, im Fall $s = n$ pro Iterationsschritt nur einmal. (Im allgemeinen wäre für jede Komponente des Gradientenvektors die Lösung einer Ljapunovgleichung notwendig, was zur Lösung von $g + 1$ verschiedenen Ljapunovgleichungen pro Iterationsschritt führen würde.)

Die Unterprogramme `cutgrad` und `vollgrad` besitzen die gleichen Parameter wie das Programm `mingrad`, außer daß `vollgrad` nicht den Parameter `omax` übergeben bekommt.

Werden die Programme `cutgrad` und `vollgrad` separat (ohne `mingrad`) benutzt, so ist zu beachten, daß in Matlab-Versionen vor 5.0 `cutgrad` nur ausgeführt werden kann, wenn $s < n$ gilt. Für den Fall $s = n$ ist dann `vollgrad` zu benutzen. In Matlab ab Version 5.0 kann `cutgrad` auch ausgeführt werden, wenn $s = n$ gilt. Es empfiehlt sich aber in diesem Fall dennoch, die Funktion `vollgrad` zu benutzen, da sie (wie oben beschrieben) einen effizienteren Algorithmus benutzt.

- `minnew`

Eingabe

omega, damp0, visc, damp1, omax, MinStepSize, ExaktGrd

MinStepSize = Schrittweite, die bei Unterschreitung zum Ende des Verfahrens führt;
ExaktGrd = relativer Exaktheitsgrad für Abbruch des Verfahrens, d.h. mit *ExaktGrd* wird das Verfahren abgebrochen,
 falls $\text{Spur}_{j-1} - \text{Spur}_j < \frac{\text{Spur}_j}{\text{ExaktGrd}}$

Ausgabe

Erg, *ViscNeu*, *AnzIt*

Erg = Ergebnismatrix; jede Zeile beinhaltet das Ergebnis eines Iterationsschrittes; die Matrix enthält in

- Spalte 1 $\text{Tr}(XZ_s)$ am Ende des jeweiligen Iterationsschrittes,
- Spalte 2 den maximalen Realteil alle Eigenwerte von A , deren Imaginärteile innerhalb des durch $\omega_{max}=omax$ begrenzten Frequenzstreifens liegen;
- die restlichen Spalten bilden den Viskositätenvektor des Iterationsschrittes;

ViscNeu = der durch das Verfahren bestimmte optimierte Viskositätenvektor; (entspricht dem Viskositätenvektor der letzten Zeile der Matrix *Erg*);

AnzIt = Anzahl der durchgeführten Iterationen.

Funktionsweise

Das Programm `minnew` dient der Minimierung von $\text{Tr}(XZ_s)$ über der Menge der Viscositäten $\{\Sigma\}$ mittels des Newtonverfahrens. Die eigentliche Minimierung findet in den weiteren Unterprogrammen `vollnew` oder `cutnew` statt. Das Programm `minnew` unterscheidet die Fälle $s < n$ und $s = n$, in denen eine unterschiedliche Berechnung der Hesseschen Matrix erfolgt. Der Fall $s = n$ ist in `vollnew` und der Fall $s < n$ ist in `cutnew` implementiert.

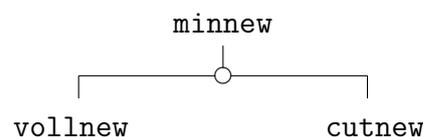


Abb. 5.4: Unterprogramme von `minnew`

In beiden Programmen wird die Ljapunovgleichung zur Lösung von $Tr(XZ_s)$ einmal pro Iterationsschritt berechnet. Außerdem werden für jede Komponente des Gradientenvektors (d.h. g -mal) verschiedene Ljapunovgleichungen pro Iterationsschritt berechnet. Für die gesamte Berechnung der Hesseschen Matrix dagegen ist für den Fall $s < n$ nur die Lösung einer Ljapunovgleichung erforderlich, der Fall $s = n$ kommt sogar ganz ohne Berechnung einer zusätzlichen Lösung einer Ljapunovgleichung aus. (Im allgemeinen wäre für die Berechnung jeder Komponente der Hesseschen Matrix eine Ljapunovgleichung zu lösen, was – wegen der Symmetrie der Hesseschen Matrix – zur zusätzlichen Lösung von $\frac{(g^2+g)}{2}$ Ljapunovgleichungen führen würde.)

Die Unterprogramme `cutnew` und `vollnew` besitzen die gleichen Parameter wie das Programm `minnew`, außer daß `vollnew` nicht den Parameter `omax` übergeben bekommt.

Werden die Programme `cutnew` und `vollnew` separat (ohne `minnew`) benutzt, so ist zu beachten, daß in Matlab-Versionen vor 5.0 `cutnew` nur ausgeführt werden kann, wenn $s < n$ gilt. Für den Fall $s = n$ ist dann `vollnew` zu benutzen. In Matlab ab Version 5.0 kann `cutnew` auch ausgeführt werden, wenn $s = n$ gilt. Es empfiehlt sich aber in diesem Fall dennoch, die Funktion `vollnew` zu benutzen, da sie (wie oben beschrieben) einen effizienteren Algorithmus benutzt.

Zu den Programmen `calcvisc`, `minguess`, `cguess`, `vguess`, `mingrad`, `vollgrad`, `cutgrad` existieren analoge Programme `calcvis3`, `mingues3`, `cguess3`, `vguess3`, `mingrad3`, `vollgra3`, `cutgrad3`, welche die gleiche Funktionalität unter Berücksichtigung der in (5.3) beschriebenen Abhängigkeit aufweisen. Sie unterscheiden sich jeweils nur um die weiteren Eingabeparameter β und γ (siehe (5.3)).

Genauere Erläuterungen zu jedem einzelnen Programm erhält man direkt unter Matlab durch den Aufruf `help <Programmname>`.

Literaturverzeichnis

- [1] Barnett, S. *Introduction to mathematical control theory*. Clarendon Press, Oxford, 1975.
- [2] Bartels, R.H. and Stewart, G. W. *Solution of the Matrix Equation $AX + XB = C$* . Comm. ACM, 15 (1972), pp. 820–826.
- [3] Bendsøe, M.P. and Olhoff, N. and Taylor, J.E. *On the Design of Structure and Controls for Optimal Performance of Actively Controlled Flexible Structures*. Mech. Structures and Mach., 15 (3) (1987) pp. 265–295.
- [4] Cox, S. J. *Designing for optimal energy absorption, III: numerical minimization of the spectral abscissa*. Structural Optimization, 13 (1997), Springer-Verlag pp. 17-22.
- [5] Cox, S. J. *Designing for optimal energy absorption, I: lumped parameter systems*. J. Vib. and Acoustics, (1996) (to appear).
- [6] Cox, S. J. and Zuazua, E. *The rate at which energy decays in a damped string*. Commun. in Partial Differential Equations, 19 (1994) pp. 213-243.
- [7] Gantmacher, F. R. *Applications of the Theory of Matrices*. Interscience Publishers, 1959.
- [8] Golub, G. H. and Van Loan, C. F. *Matrix Computations*. The Johns Hopkins University Press, 2nd edition, 1989.
- [9] Golub, G. H., Nash, S. and Van Loan, C. *A Hessenberg-Schur Method for the Problem $AX + XB = C$* . IEEE Trans. Autom. Control, AC-24 (1979), pp. 909–913.
- [10] Gürgöze, M. and Müller, P.C. *Optimal Positioning of Dampers in Multi-Body Systems*. Journal of Sound and Vibration, 158 (1992) pp. 517–530.
- [11] Hagander, P. *Numerical Solution of $A^T S + SA + Q = 0$* . Inform. Sci., 4 (1972), pp. 35–40.
- [12] Hammarling, S. J. *Numerical Solution of the Stable, Non-negative Definite Lyapunov Equation*. IMA Journal of Numerical Analysis, 2 (1982), pp. 303–323.

- [13] Horn, R. A. and Johnson, C. R. *Matrix Analysis*. Cambridge University Press, 1990.
- [14] Horn, R. A. and Johnson, C. R. *Topics in Matrix Analysis*. Cambridge University Press, 1991.
- [15] Kato, T. *Perturbation Theory for Linear Operators*. Springer-Verlag, Berlin, 2nd edition, 1976.
- [16] Lancaster, P. and Tismenetsky, M. *The Theory of Matrices with Applications*. Academic Press, 2nd edition, 1985.
- [17] Lawson, C. L. and Hanson, R. J. *Solving Least Squares Problems*. Prentice-Hall, Inc, 1974.
- [18] Müller, P.C. and Gürgöze, M. *Optimale Dämpfungsstärke eines viskosen Dämpfers bei einem mehrläufigen Schwingungssystem*. ZAMM Z. angew. Math. Mech., 71 (1991) pp. T 60–T 63.
- [19] Müller, P. C. and Schiehlen, W. O. *Lineare Schwingungen*. Akademische Verlagsgesellschaft, Wiesbaden, 1976.
- [20] Pace, I. C. and Barnett, S. *Comparison of numerical methods for solving Liapunov matrix equations*. Int. J. Control, 15 (1972), pp. 907–915.
- [21] Saad, Y. *Numerical Solution of Large Lyapunov Equations*. Signal processing scattering and operator theory, and numerical methods (Amsterdam, 1989), Progr. Systems Control Theory 5, Birkhäuser Boston, 1990, pp. 503–511.
- [22] Snyders, J. and Zakai, M. *On nonnegative solutions of the equation $AD + DA' = -C$* . SIAM J. Appl. Math., 18 (1970), pp. 704–714.
- [23] Stoer, J. *Einführung in die Numerische Mathematik 1*. Springer-Verlag, Berlin, 3. edition, 1979.
- [24] Veselić, K. *On linear vibrational systems with one dimensional damping II*. Integral Equations and Operator Theory, 13 (1990) pp. 883–897.
- [25] Veselić, K. *Optimierung einer Dämpfungskonfiguration*. Lehrgebiet Mathematische Physik, Fernuniversität Hagen, Postfach 940, D-58084 Hagen, 1991.
- [26] Veselić, K. *On linear vibrational systems with one dimensional damping*. Appl. Anal., 29 (1988) pp. 1–18.
- [27] Veselić, K. *Exponential decay of semigroups in Hilbert space*. Semigroup Forum, 55 (1997) pp. 325–331.
- [28] Veselić, K. *Estimating the operator exponential*. Fernuniversität Hagen, preprint 1997, to appear in Linear Algebra Appl.

- [29] Veselić, K., Delinić, K. *Passive control of linear systems*. Preprint, Lehrgebiet Mathematische Physik, Fernuniversität Hagen, Postfach 940, D-58084 Hagen, Januar 1996.
- [30] Wachspress, E. L. *Iterative Solution of the Lyapunov Matrix Equation*. Appl. Math. Lett., 1 (1988), pp. 87–90.

Abbildungsverzeichnis

2.1	Tr($X(\varepsilon_1, \varepsilon_2)$) im Vergleich zu Tr($X(C_{\min})$)	48
3.1	Vergleich der Startwertberechnung	59
3.2	Test (G1a), Vergleich von ($P_{Grad I}$) und ($P_{Grad II}$) bei voller Spur . . .	64
3.3	Test (G2b), Vergleich von ($P_{Grad I}$) und ($P_{Grad II}$) bei nichtvoller Spur	65
3.4	Test (G2a), Vergleich von ($P_{Grad I}$) und ($P_{Grad II}$) bei voller Spur . . .	66
3.5	Test (N2a), Vergleich von ($P_{Newton I}$) und ($P_{Newton II}$) bei voller Spur	69
4.1	Simultane Variation aller Frequenzen	119
4.2	Vergrößerung aller Frequenzen im Bereich hinter dem Schnitt	120
4.3	Simultane Verkleinerung aller Dämpfungskomponenten	121
4.4	Skalierung aller Dämpfungskomponenten im Bereich hinter dem Schnitt	122
4.5	Vergrößerung einer Dämpfungskomponente im relevanten Bereich . . .	124
4.6	Vergrößerung der ersten Dämpfungskomponente im Zwischenbereich .	125
4.7	Vergrößerung aller Frequenzen im Bereich hinter dem Schnitt	127
4.8	Simultane Variation aller Dämpfungskomponenten hinter dem Schnitt	128
4.9	Vergrößerung einer Dämpfungskomponente im relevanten Bereich . . .	129
4.10	Variation einer Dämpfungskomponente im Zwischenbereich	130
5.1	Zeitliche Reihenfolge der Programmabläufe	134
5.2	Unterprogramme von calcvisc	136
5.3	Unterprogramme von mingrad	138
5.4	Unterprogramme von minnew	139

Lebenslauf

Name		Katrin Brabender
Geburt		14.06.1968 in Wuppertal
Schulbildung	08/74 – 06/78	Grundschule in Wuppertal
	08/78 – 06/88	Gymnasium in Velbert–Langenberg
Schulabschluß	06/88	Abitur
Studium	10/88 – 01/94	Studium der Mathematik an der Ruhr-Universität Bochum
Studienabschluß	01/94	Diplom in Mathematik
Tätigkeit	04/91 – 01/94	Studentische Hilfskraft am mathematischen Institut der Ruhr-Universität Bochum
	seit 04/94	Wissenschaftliche Mitarbeiterin am Lehrgebiet Mathematische Physik der FernUniversität Hagen

Hagen, 6. Mai 1998

[1]