

# **Glättungsverfahren für semidefinite Programme**

Dissertation zur Erlangung des  
naturwissenschaftlichen Doktorgrades  
der Bayerischen Julius-Maximilians-Universität Würzburg

vorgelegt von

Christian Nagel

aus

Hamburg

Würzburg 2003

Eingereicht am: 18.12.2003

bei der Fakultät für Mathematik und Informatik

1. Gutachter: Prof. Dr. Christian Kanzow
2. Gutachter: Prof. Dr. Florian Jarre

Tag der mündlichen Prüfung: 19.02.2004

# Inhaltsverzeichnis

<b>1. Einleitung</b>	<b>1</b>
<b>I. Theorie</b>	<b>3</b>
<b>2. Semidefinite Programme und NCP-Funktionen</b>	<b>5</b>
2.1. Semidefinite Programme . . . . .	5
2.2. NCP-Funktionen . . . . .	10
2.3. Lyapunov-Operatoren . . . . .	19
2.4. Eigenschaften von NCP-Funktionen . . . . .	25
<b>II. Verfahren</b>	<b>33</b>
<b>3. Ein Glättungsverfahren</b>	<b>35</b>
3.1. Algorithmus . . . . .	35
3.2. Globales und lokales Konvergenzverhalten . . . . .	43
3.3. Eigenschaften der Newton-Systeme . . . . .	52
3.4. Numerische Resultate . . . . .	59
<b>4. Trust-Region-Verfahren</b>	<b>67</b>
4.1. Trust-Region-Algorithmus . . . . .	71
4.2. Globale und lokale Konvergenzeigenschaften . . . . .	76
4.3. Numerische Resultate . . . . .	84
<b>5. Ein nichtglattes Newton-Verfahren</b>	<b>87</b>
5.1. Algorithmus . . . . .	87
5.2. Matrix-Vektor-Formulierung der Newton-Systeme . . . . .	90
5.3. Lokale Konvergenz . . . . .	95
5.4. Numerische Beispiele . . . . .	104
<b>6. Zusammenfassung</b>	<b>111</b>
<b>Literaturverzeichnis</b>	<b>115</b>



# 1. Einleitung

In dieser Arbeit beschreiben wir Algorithmen zur Lösung von so genannten semidefiniten Programmen. Unter Benutzung einiger Standardnotationen (welche formal am Anfang des nächsten Kapitels eingeführt werden) ist ein semidefinites Programm (kurz: SDP) ein restringiertes Optimierungsproblem, welches in primaler Form gegeben ist durch

$$\min C \bullet X \quad \text{u.d.N.} \quad A_i \bullet X = b_i, \quad i = 1, \dots, m, \quad X \succeq 0. \quad (1.1)$$

Hierbei sind der Vektor  $b \in \mathbb{R}^m$  ebenso wie die symmetrischen Matrizen  $C \in \mathbb{R}^{n \times n}$  und  $A_i \in \mathbb{R}^{n \times n}$ ,  $i = 1, \dots, m$ , die gegebenen Daten, während die symmetrische Matrix  $X \in \mathbb{R}^{n \times n}$  die Variable des semidefiniten Programms (1.1) bezeichnet.

Semidefinite Programme dieser Art haben zahlreiche Anwendungen, insbesondere im Bereich der kombinatorischen Optimierung, der Kontrolltheorie und den Ingenieurwissenschaften, vergleiche beispielsweise [4, 5, 26, 27, 28, 29, 39, 69].

Das semidefinite Programm (1.1) ist ein konvexes Optimierungsproblem und daher (unter einer geeigneten Regularitätsvoraussetzung) äquivalent zu seinen Optimalitätsbedingungen. Motiviert durch das Buch von Nesterov und Nemirovskii [55] haben viele Autoren versucht, diese Optimalitätsbedingungen durch (primal-duale) Innere-Punkte-Methoden (vergleiche zum Beispiel [3, 25, 39, 50, 56, 65, 66, 70]) zu lösen. Innere-Punkte-Methoden wenden üblicherweise ein Newton-ähnliches Verfahren auf eine Störung der Optimalitätsbedingungen, die so genannten Zentralen-Pfad-Bedingungen, an. Dabei wird die positive Definitheit der auftretenden Matrizen durch eine geeignete Schrittlängenbestimmung während der Iteration garantiert.

In dieser Arbeit wollen wir einen etwas anderen Ansatz verfolgen. Wir werden zunächst die Optimalitätsbedingungen bzw. die Zentralen-Pfad-Bedingungen durch matrixwertige NCP-Funktionen in ein nichtlineares Gleichungssystem überführen. Dieses nichtlineare und teilweise nicht differenzierbare Gleichungssystem lösen wir dann mit einem Newton-ähnlichem Verfahren. Durch die Umformulierung in ein nichtlineares Gleichungssystem müssen wir während der Iteration nicht mehr explizit die positive (Semi-)Definitheit der beteiligten Matrizen beachten. Weiter werden wir zeigen, dass dieser Ansatz im Gegensatz zu Inneren-Punkte-Methoden sofort symmetrische Suchrichtungen erzeugt. Um globale Konvergenz zu erhalten, werden wir verschie-

## 1. Einleitung

dene Globalisierungsstrategien (Schrittweitenbestimmung, Trust-Region-Ansatz) untersuchen.

In dieser Arbeit behandeln wir nur die oben eingeführten so genannten linearen semidefiniten Programme. Nichtlineare semidefinite Programme, d.h. Optimierungsprobleme mit nichtlinearer Zielfunktion oder nichtlinearen Nebenbedingungen und semidefiniten Nebenbedingungen der Form  $X \succeq 0$ , sind nicht Gegenstand dieser Dissertation. Für diese Problemklasse sei auf die Arbeiten [14, 34, 37, 38, 47, 23, 52, 61] und die darin enthaltenen Referenzen verwiesen.

Der Rest der Arbeit ist wie folgt organisiert: Im folgenden Kapitel 2 legen wir die nötigen theoretischen Grundlagen, bevor wir uns im zweiten Teil mit Verfahren zur Lösung der Optimalitätsbedingungen beschäftigen. In Kapitel 3 werden wir ein Prädiktor-Korrektor-Glättungsverfahren zur Lösung der Optimalitätsbedingungen angeben, theoretisch untersuchen und numerische Resultate präsentieren. In Kapitel 4 werden wir dann einen Trust-Region-Ansatz vorstellen und analysieren.

In Kapitel 5 untersuchen wir ein nichtglattes Newton-Verfahren auf lokal schnelle Konvergenz.

Wir schließen die Arbeit mit einer kurzen Zusammenfassung.

Die folgenden Kapitel beinhalten zum Teil Ergebnisse, die aus gemeinsamen Forschungsaktivitäten an den Universitäten Hamburg und Würzburg mit dem Betreuer dieser Dissertation, Herrn Prof. Dr. Christian Kanzow, hervorgegangen sind. Für die hervorragende Unterstützung und Förderung während dieser Zeit, ohne die diese Arbeit in dieser Form sicherlich nicht entstanden wäre, möchte ich mich bei meinem Doktorvater ausdrücklich bedanken.

Weiter möchte ich mich bei meinen Kollegen – in Würzburg sowie der ganzen Welt – und den anonymen Gutachtern der zur Veröffentlichung eingereichten Arbeiten bedanken, und zwar sowohl für manch konstruktive Diskussion und Anregung als auch für die unterhaltsamen Stunden auf den Konferenzen („Da ist ja Marzipan drin!“).

Ein besonderer Dank gebührt natürlich meiner Familie und meiner Freundin Stephanie für die moralische Unterstützung – auch über weite Entfernungen hinweg.

**Teil I.**  
**Theorie**





## 2. Semidefinite Programme und NCP-Funktionen

In diesem Kapitel betrachten wir die theoretischen Grundlagen für semidefinite Programme. Im Abschnitt 2.1 werden wir semidefinite Programme formal einführen und den Zusammenhang mit den Optimalitätsbedingungen und den Zentralen-Pfad-Bedingungen darstellen. Es folgt ein Abschnitt über matrixwertige NCP-Funktionen. Wir werden uns dabei auf die so genannte Minimum-Funktion und die Fischer-Burmeister-Funktion beschränken und deren Eigenschaften darstellen. Mit Hilfe dieser Funktionen werden wir sowohl die Optimalitätsbedingungen als auch die Zentralen-Pfad-Bedingungen in ein nichtlineares Gleichungssystem überführen. Die zur Lösung dieses Gleichungssystems notwendigen Verfahren werden wir dann im zweiten Teil dieser Arbeit darstellen und untersuchen.

### 2.1. Semidefinite Programme

Bevor wir ein semidefinites Programm formal angeben können, benötigen wir ein paar Notationen: Die Menge der natürlichen Zahlen bezeichnen wir mit  $\mathbb{N}$ . Es sei  $\mathbb{R}$  die Menge aller reellen Zahlen und  $\mathbb{R}_+$  bzw.  $\mathbb{R}_{++}$  die Teilmenge aller nichtnegativen bzw. positiven reellen Zahlen. Mit  $\mathcal{S}^{n \times n}$  bezeichnen wir die Menge aller symmetrischen Matrizen aus dem  $\mathbb{R}^{n \times n}$ , dem Raum aller Matrizen der Dimension  $n \times n$  mit reellwertigen Einträgen. Mit  $\mathcal{S}_+^{n \times n}$  bezeichnen wir den Raum der symmetrischen und positiv semidefiniten, mit  $\mathcal{S}_{++}^{n \times n}$  den Raum der symmetrischen und positiv definiten Matrizen. Ist  $A \in \mathcal{S}_+^{n \times n}$  (bzw.  $A \in \mathcal{S}_{++}^{n \times n}$ ), so schreiben wir kürzer  $A \succeq 0$  (bzw.  $A \succ 0$ ). Die Abkürzung  $A \succeq B$  bedeutet nichts anderes als  $A - B \succeq 0$ . Die Einträge einer Matrix  $A \in \mathbb{R}^{n \times n}$  bezeichnen wir im Allgemeinen mit  $a_{ij}$ , d.h. es gilt  $A = (a_{ij})_{i,j=1,\dots,n}$ . Mit  $\text{diag}(d_1, \dots, d_n)$  bezeichnen wir eine Diagonalmatrix, deren Diagonalelemente durch die Zahlen  $d_i$ ,  $i = 1, \dots, n$  gegeben sind. Weiter definieren wir für zwei Matrizen  $A, B \in \mathbb{R}^{n \times n}$

$$A \bullet B := \langle A, B \rangle := \text{tr}(AB^T) = \sum_{i,j=1}^n a_{ij}b_{ij}.$$

Dabei ist die Spur  $\text{tr}(C)$  für eine Matrix  $C \in \mathbb{R}^{n \times n}$  gegeben durch

$$\text{tr}(C) = \sum_{i=1}^n c_{ii}.$$

## 2. Semidefinite Programme und NCP-Funktionen

Die Abbildung  $\langle \cdot, \cdot \rangle$  ist ein Skalarprodukt auf dem Raum  $\mathbb{R}^{n \times n}$  und man verifiziert leicht, dass  $\text{tr}((AB)^T) = \text{tr}(AB) = \text{tr}(BA)$  gilt. Die durch dieses Skalarprodukt induzierte Norm

$$\|A\|_F := \sqrt{A \bullet A} = \left( \sum_{i,j=1}^n a_{ij}^2 \right)^{1/2} \quad (2.1)$$

ist die übliche *Frobenius-Norm*.

Das Skalarprodukt  $\langle \cdot, \cdot \rangle$  ist mit der Löwner'schen Halbordnung  $\succeq$  eng durch den folgenden Satz verbunden, dessen Beweis beispielsweise in Jarre und Stoer [39, Satz 8.5.2] gefunden werden kann.

**Satz 2.1 (Féjer)** *Eine symmetrische Matrix  $A \in \mathcal{S}^{n \times n}$  ist genau dann positiv semidefinit, wenn  $A \bullet B \geq 0$  für alle  $B \succeq 0$  gilt.*

Seien jetzt die Matrizen  $A_i \in \mathcal{S}^{n \times n}$ ,  $i = 1, \dots, m$  und  $C \in \mathcal{S}^{n \times n}$  sowie der Vektor  $b \in \mathbb{R}^m$  gegeben. Dann ist

$$\min C \bullet X \quad \text{u.d.N.} \quad A_i \bullet X = b_i, \quad i = 1, \dots, m, \quad X \succeq 0 \quad (2.2)$$

ein primales (lineares) semidefinites Programm mit der Variable  $X \in \mathcal{S}^{n \times n}$ . Bildet man hierzu das Lagrange-duale Problem (vergleiche Geiger und Kanzow [25]), so erhält man mit

$$\max b^T \lambda \quad \text{u.d.N.} \quad \sum_{i=1}^m \lambda_i A_i + S = C, \quad S \succeq 0 \quad (2.3)$$

ein duales (lineares) semidefinites Programm in den Variablen  $\lambda \in \mathbb{R}^m$  und  $S \in \mathcal{S}^{n \times n}$ . Da wir uns in dieser Arbeit nur mit linearen semidefiniten Programmen beschäftigen werden, werden wir den Zusatz „linear“ in Zukunft weglassen und einfach nur von semidefiniten Programmen sprechen.

Da eine symmetrische Matrix nur reelle Eigenwerte hat und genau dann positiv semidefinit ist, falls alle Eigenwerte nichtnegativ sind, lässt sich das primale semidefinite Programm als restringiertes Optimierungsproblem in der Form

$$\min C \bullet X \quad \text{u.d.N.} \quad A_i \bullet X = b_i, \quad i = 1, \dots, m, \quad \lambda_{\min}(X) \geq 0 \quad (2.4)$$

schreiben. Dabei bezeichnet  $\lambda_{\min}(X)$  den kleinsten Eigenwert der Matrix  $X$ . Da die Funktion  $X \mapsto \lambda_{\min}(X)$  bekanntermaßen konkav ist, ist das primale

SDP ein konvexes, restringiertes Optimierungsproblem. Wir betrachten daher die üblichen KKT-Bedingungen

$$\begin{aligned} \sum_{i=1}^m \lambda_i A_i + S &= C, \\ A_i \bullet X &= b_i \quad \forall i = 1, \dots, m, \\ X \succeq 0, S \succeq 0, X \bullet S &= 0 \end{aligned} \tag{2.5}$$

zu (2.2). Es ist klar, dass jede Lösung  $(X^*, \lambda^*, S^*)$  der KKT-Bedingungen (2.5) eine Lösung  $X^*$  des primalen semidefiniten Programms (2.2) und eine Lösung  $(\lambda^*, S^*)$  des dualen Programms (2.3) liefert. Die Umkehrung gilt nur unter einer geeigneten Regularitätsvoraussetzung. Da es sich um ein konvexes Optimierungsproblem handelt, liegt es nahe, die folgende Voraussetzung zu machen.

**Voraussetzung 2.2 (Slater-Bedingung)** *Es gibt einen Punkt  $(\hat{X}, \hat{\lambda}, \hat{S}) \in \mathcal{S}^{n \times n} \times \mathbb{R}^m \times \mathcal{S}^{n \times n}$  mit*

$$\begin{aligned} A_i \bullet \hat{X} &= b_i, \quad i = 1, \dots, m, \\ \sum_{i=1}^m \hat{\lambda}_i A_i + \hat{S} &= C, \\ \hat{X} \succ 0, \quad \hat{S} \succ 0. \end{aligned}$$

Die Slater-Bedingung 2.2 besagt, dass es einen bezüglich der linearen Restriktionen zulässigen Punkt gibt, welcher die Ungleichungen sogar strikt erfüllt. Wir werden in der gesamten Arbeit, auch wenn es nicht explizit erwähnt wird, davon ausgehen, dass die Slater-Bedingung erfüllt ist. Dann gilt der folgende Satz, vergleiche beispielsweise Jarre und Stoer [39, Satz 8.5.9], Geiger und Kanzow [25, Satz 4.23] oder Nesterov und Nemirovskii [55, Theorem 4.2.1].

**Satz 2.3** *Sei die Slater-Bedingung 2.2 erfüllt. Dann gelten die folgenden Aussagen:*

- (a) *Das primale semidefinite Programm (2.2) hat eine Lösung  $X^*$ .*
- (b) *Das duale semidefinite Programm (2.3) hat eine Lösung  $(\lambda^*, S^*)$ .*
- (c) *Die KKT-Bedingungen (2.5) haben eine Lösung  $(X^*, \lambda^*, S^*)$ .*

In der gesamten Arbeit werden wir daher versuchen, die KKT-Bedingungen (2.5) zu lösen, um so Lösungen der zugrunde liegenden semidefiniten Programme zu erhalten.

## 2. Semidefinite Programme und NCP-Funktionen

Der Term  $X \bullet S = 0$  in den KKT-Bedingungen bedeutet nichts anderes, als dass die Dualitätslücke

$$C \bullet X - b^T \lambda = S \bullet X - \sum_{i=1}^m \lambda_i A_i \bullet X - b^T \lambda = X \bullet S$$

verschwinden soll. Allgemein gilt für ein primal zulässiges  $X$  und dual zulässiges  $(\lambda, S)$  wegen Satz 2.1

$$C \bullet X - b^T \lambda = X \bullet S \geq 0.$$

Die in (2.5) angegebene Formulierung der Optimalitätsbedingungen ist für unsere Zwecke nicht ganz geeignet, da dort mehr Unbekannte als Gleichungen auftreten. Bevor wir jedoch das zur Umformulierung notwendige Lemma formulieren können, bringen wir die zwei folgenden Sätze über Spektralzerlegungen von symmetrischen Matrizen.

**Satz 2.4 (Spektralsatz)** Sei  $A \in \mathcal{S}^{n \times n}$  eine symmetrische Matrix. Dann existieren eine orthogonale Matrix  $Q \in \mathbb{R}^{n \times n}$  und eine Diagonalmatrix  $D \in \mathbb{R}^{n \times n}$  mit  $A = QDQ^T$ . Die Diagonalelemente von  $D$  sind dabei die Eigenwerte der Matrix  $A$ .

**Satz 2.5** Seien  $A, B \in \mathcal{S}^{n \times n}$  zwei symmetrische Matrizen. Dann kommutieren  $A$  und  $B$  genau dann, wenn  $A$  und  $B$  eine gemeinsame Spektralzerlegung haben, d.h. wenn eine orthogonale Matrix  $Q \in \mathbb{R}^{n \times n}$  und Diagonalmatrizen  $D_A$  und  $D_B$  existieren mit  $A = QD_AQ^T$  und  $B = QD_BQ^T$ .

Die Beweise für diese Sätze kann man beispielsweise in dem Buch von Horn und Johnson [35] finden. Weiter benötigen wir noch das folgende Resultat.

**Lemma 2.6** Sei  $A = (a_{ij}) \in \mathcal{S}_+^{n \times n}$  eine symmetrische und positiv semidefinite Matrix. Ist  $a_{ii} = 0$  für ein  $i \in \{1, \dots, n\}$ , so gilt  $a_{ij} = a_{ji} = 0$  für alle  $j = 1, \dots, n$ .

**Beweis:** Sei  $e_i$  der  $i$ -te Einheitsvektor im  $\mathbb{R}^n$ . Angenommen, es ist  $a_{ij} \neq 0$  für ein  $j \in \{1, \dots, n\}$ . Setzen wir  $x := \xi e_i + e_j$  mit  $\xi \in \mathbb{R}$  beliebig, so gilt wegen  $A \succeq 0$

$$0 \leq x^T A x = \sum_{l,k=1}^n a_{lk} x_l x_k = 2a_{ij}\xi + a_{jj} \quad \forall \xi \in \mathbb{R}.$$

Je nachdem, ob  $a_{ij}$  positiv oder negativ ist, bedeutet dies

$$\xi \geq \frac{-a_{jj}}{2a_{ij}} \quad \text{bzw.} \quad \xi \leq \frac{-a_{jj}}{2a_{ij}}$$

für alle  $\xi \in \mathbb{R}$ , ein Widerspruch.  $\square$

Damit können wir das folgende Lemma beweisen.

**Lemma 2.7 (Alizadeh [1])** *Seien  $X, S \in \mathcal{S}_+^{n \times n}$  zwei symmetrische positiv semidefinite Matrizen. Dann ist  $XS = 0$  genau dann, wenn  $X \bullet S = 0$  gilt.*

**Beweis:** Sei zunächst  $XS = 0$ . Dann folgt sofort

$$X \bullet S = \text{tr}(XS^T) = \text{tr}(XS) = \text{tr}(0) = 0.$$

Sei daher umgekehrt  $X \bullet S = 0$  und  $S = QDQ^T$  eine nach Satz 2.4 existierende Spektralzerlegung der symmetrischen Matrix  $S$  mit einer orthogonalen Matrix  $Q$  und einer Diagonalmatrix  $D = \text{diag}(\lambda_1, \dots, \lambda_n)$ . Da  $S$  positiv semidefinit ist, gilt  $\lambda_i \geq 0$  für alle  $i = 1, \dots, n$ . Setzen wir  $A := Q^T X Q$ , so ist  $A$  symmetrisch und positiv semidefinit. Insbesondere sind alle Diagonalelemente  $a_{ii}$  von  $A$  nichtnegativ. Da  $Q$  regulär ist, reicht es zu zeigen, dass  $AD = Q^T X S Q = 0$  gilt. Wegen  $X \bullet S = 0$  ist

$$0 = X \bullet S = \text{tr}(XS) = \text{tr}(Q^T X S Q) = \text{tr}(AD) = \sum_{i=1}^n a_{ii} \lambda_i.$$

Da alle Summanden nichtnegativ sind, folgt, dass jeder einzelne Summand verschwindet. Ist jetzt  $\lambda_i > 0$ , so folgt  $a_{ii} = 0$  und wegen  $A \succeq 0$  sind dann nach Lemma 2.6 sowohl die  $i$ -te Spalte als auch die  $i$ -Zeile von  $A$  gleich Null.

Angenommen, es ist  $(AD)_{ij} = a_{ij} \lambda_j \neq 0$  für ein  $i, j$ . Dann ist  $\lambda_j > 0$  und nach obiger Überlegung somit sowohl die  $j$ -Spalte als auch die  $j$ -te Zeile von  $A$  gleich Null. Insbesondere ist also  $a_{ij} = 0$  und somit  $a_{ij} \lambda_j = 0$ , ein Widerspruch zur Annahme.  $\square$

Unter Benutzung von Lemma 2.7 schreiben sich die KKT-Bedingungen (2.5) als *Optimalitätsbedingungen* wie folgt:

$$\begin{aligned} \sum_{i=1}^m \lambda_i A_i + S &= C, \\ A_i \bullet X &= b_i \quad \forall i = 1, \dots, m, \\ X \succeq 0, S \succeq 0, XS &= 0. \end{aligned} \tag{2.6}$$

Wir werden in Zukunft diese Formulierung benutzen und versuchen, diese Optimalitätsbedingungen zu lösen.

## 2. Semidefinite Programme und NCP-Funktionen

Motiviert durch das Buch von Nesterov und Nemirovskii [55] haben viele Autoren (primal-duale) Innere-Punkte-Methoden zur Lösung der Optimalitätsbedingungen betrachtet. Diese Inneren-Punkte-Methoden betrachten typischer Weise die folgende Störung der Optimalitätsbedingungen, die so genannten *Zentralen-Pfad-Bedingungen*:

$$\begin{aligned} \sum_{i=1}^m \lambda_i A_i + S &= C, \\ A_i \bullet X &= b_i \quad \forall i = 1, \dots, m, \\ X \succ 0, S \succ 0, XS &= \tau^2 I. \end{aligned} \tag{2.7}$$

Hierbei ist  $I$  die Einheitsmatrix passender Dimension. Wir haben hier die Zentralen-Pfad-Bedingungen mit dem Parameter  $\tau^2$  (anstelle von  $\tau$ ) parametrisiert. Dies wird die folgenden Formeln vereinfachen.

Innere-Punkte-Methoden wenden jetzt typischer Weise ein Newton-ähnliches Verfahren auf eine Symmetrisierung der Gleichungen innerhalb der Zentralen-Pfad-Bedingungen (2.7) an und garantieren die positive Definitheit der Iterierten  $X \succ 0$  und  $S \succ 0$  durch eine geeignete Schrittweitenstrategie. Dabei wird während der Iteration der Parameter  $\tau$  geeignet reduziert, so dass man im Grenzwert eine Lösung der Optimalitätsbedingungen erhält. Für genauere Ausführungen zu den Inneren-Punkte-Methoden vergleiche man beispielsweise die Arbeiten [1, 3, 6, 39, 50, 51, 65, 66, 70].

Wir wollen jetzt jedoch ein wenig anders vorgehen und die Optimalitätsbedingungen (2.6) bzw. die Zentralen-Pfad-Bedingungen (2.7) weiter umformulieren. Ziel ist es, diese Systeme als nichtlineare (und nichtglatte) Gleichungssysteme zu schreiben. D.h. es sollen explizit keine Ungleichungen der Form  $X \succ 0$ ,  $X \succeq 0$  und  $S \succ 0$ ,  $S \succeq 0$  mehr vorliegen. Die dafür notwendigen matrixwertigen NCP-Funktionen werden wir in dem folgenden Abschnitt einführen.

### 2.2. NCP-Funktionen

Ziel dieses Abschnittes ist es, das in den Optimalitätsbedingungen auftretende Komplementaritätsproblem

$$X \succeq 0, \quad S \succeq 0, \quad XS = 0$$

durch eine einzige nichtlineare Gleichung der Form  $\phi(X, S) = 0$  zu ersetzen. Für den reellwertigen Fall (d.h. für  $n = 1$ ) sind dafür die so genannten *NCP-Funktionen* bekannt. Die Abkürzung NCP steht dabei für den englischen Ausdruck *nonlinear complementarity problem*. Eine NCP-Funktion ist

eine Funktion  $\varphi : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$  mit

$$\varphi(a, b) = 0 \iff a \geq 0, b \geq 0, ab = 0.$$

Zwei bekanntere Vertreter von reellwertigen NCP-Funktionen sind die von Fischer [22] eingeführte *Fischer-Burmeister-Funktion*

$$\varphi(a, b) := a + b - \sqrt{a^2 + b^2} \quad (2.8)$$

und die *Minimum-Funktion*

$$\varphi(a, b) := a + b - \sqrt{(a - b)^2} = 2 \min\{a, b\}. \quad (2.9)$$

Dass diese Funktionen NCP-Funktionen sind, lässt sich leicht beweisen (vergleiche auch [41]). Wir wollen jetzt den Begriff der NCP-Funktionen auf den Raum der symmetrischen Matrizen verallgemeinern.

**Definition 2.8** *Ein Funktion  $\phi : \mathcal{S}^{n \times n} \times \mathcal{S}^{n \times n} \rightarrow \mathcal{S}^{n \times n}$  heißt matrixwertige NCP-Funktion, falls die Äquivalenz*

$$\phi(X, S) = 0 \iff X \succeq 0, S \succeq 0, XS = 0$$

*gilt.*

Da der reellwertige Fall ( $n = 1$ ) durch die Definition bereits mit abgedeckt ist und wir uns im Folgenden nur mit matrixwertigen NCP-Funktionen beschäftigen, werden wir den Zusatz „matrixwertig“ meist weglassen und nur von NCP-Funktionen sprechen. Man beachte, dass wir fordern, dass eine NCP-Funktion eine Abbildung in den Raum der symmetrischen Matrizen ist. Diese Forderung wird uns nachher automatisch symmetrische Suchrichtungen liefern.

Um die oben angegebenen reellwertigen NCP-Funktionen auf den Raum der symmetrischen Matrizen zu verallgemeinern, benötigen wir die Quadratwurzel einer symmetrischen, positiv semidefiniten Matrix. Wir werden uns daher zunächst der Existenz und Eindeutigkeit der Quadratwurzel im  $\mathcal{S}_+^{n \times n}$  widmen. Diese liefert der folgende Satz aus Horn und Johnson [35, Theorem 7.2.6].

**Satz 2.9** *Sei  $A \in \mathbb{R}^{n \times n}$  eine symmetrische, positiv semidefinite Matrix. Dann existiert eine eindeutig bestimmte symmetrische und positiv semidefinite Matrix  $B$  mit  $B^2 = A$  sowie  $\text{Rang } A = \text{Rang } B$ . Insbesondere ist  $B$  genau dann regulär (und somit positiv definit), wenn  $A$  dies ist.*

## 2. Semidefinite Programme und NCP-Funktionen

**Beweis:** Nach dem Spektralsatz 2.4 besitzt die Matrix  $A \succeq 0$  eine Spektralzerlegung  $A = QDQ^T$  mit einer orthogonalen Matrix  $Q \in \mathbb{R}^{n \times n}$  und einer Diagonalmatrix  $D = \text{diag}(\lambda_1, \dots, \lambda_n)$  mit  $\lambda_i \geq 0$  für alle  $i = 1 \dots, n$ . Wir setzen jetzt  $B := QD^{1/2}Q^T$ , wobei  $D^{1/2} := \text{diag}(\lambda_1^{1/2}, \dots, \lambda_n^{1/2})$  und in allen Fällen die positive Wurzel genommen wird. Dann ist  $B^2 = A$  und  $B \succeq 0$  klar. Der Rang von  $B$  ist die Anzahl der positiven  $\lambda_i$ , und dies ist gerade der Rang von  $A$ .

Den (eher technischen) Beweis der Eindeutigkeit findet man beispielsweise in dem Buch von Bunse und Bunse-Gerstner [9, Lemma 1.6.9].  $\square$

Wir werden im Folgenden die eindeutig bestimmte Quadratwurzel  $B \succeq 0$  aus Satz 2.9 mit  $A^{1/2}$  bezeichnen. Weiter sei bemerkt, dass der Satz 2.9 auch ein praktisches Verfahren zur Berechnung der Quadratwurzel von  $A \succeq 0$  liefert: Man muss „nur“ eine Spektralzerlegung  $A = QDQ^T$  von  $A$  berechnen und  $A^{1/2} = QD^{1/2}Q^T$  setzen. Die Berechnung der Spektralzerlegung erfordert jedoch in jedem Fall  $O(n^3)$  Rechenoperationen, da auch für die schnellsten bisher bekannten Verfahren, das *QR-Verfahren* für  $n \leq 25$  und *Divide-and-Conquer* für  $n > 25$ , die Matrix zunächst auf Tridiagonalgestalt gebracht werden muss, vergleiche Demmel [16]. Dafür kann man dann die Quadratwurzel stabil berechnen.

Es stellt sich daher die Frage, ob die Quadratwurzel durch andere Verfahren schnell berechnet werden kann. Eine nahe liegende Idee ist, die Quadratwurzel durch Anwendung des Newton-Verfahrens auf die Gleichung

$$X^2 - A = 0$$

zu berechnen. Wählen wir speziell den Startvektor  $X_0 := A$ , so ergibt sich nach einigen Umformulierungen die Iterationsvorschrift

$$X_{k+1} := \frac{1}{2}(X_k + X_k^{-1}A), \quad X_0 := A.$$

Diese Iteration zur Berechnung von  $A^{1/2}$  hat einige Nachteile: Zunächst einmal muss in jedem Iterationsschritt eine Matrix invertiert werden, was einen Aufwand von  $O(n^3)$  Rechenoperationen (je Iteration!) bedeutet. Weiter ist dieses Verfahren numerisch instabil. Man vergleiche hierzu die Ausführungen von Higham [32] und die darin enthaltenen Referenzen. Und zu guter Letzt haben wir den Startvektor bei dieser Iteration fest gewählt. Wählt man hier nicht speziell  $X_0 = A$ , so ist die Iterationsvorschrift für das Newton-Verfahren komplizierter und man muss eine Lyapunov-Gleichung (siehe Abschnitt 2.3) in jeder Iteration lösen. Dafür wird aber wiederum eine Spektralzerlegung



benötigt, so dass sich das Newton-Verfahren insgesamt als zu rechenintensiv und instabil herausstellt.

In den Arbeiten [7, 31, 32, 53] werden noch weitere Verfahren zur Berechnung der Quadratwurzel beschrieben. Jedoch sind auch diese Verfahren entweder numerisch instabil oder nicht weniger aufwendig als der Zugang über die Spektralzerlegung. Dabei ist zu beachten, dass bei den meisten Verfahren der Startvektor der Iteration nicht beliebig gewählt werden kann. Die Konvergenz gegen die Lösung  $A^{1/2}$  lässt sich nur für bestimmte Startwerte zeigen. Aus diesen Gründen, und dies ergab auch eine Diskussion mit Higham [33], ist in diesem Fall die Berechnung über die Spektralzerlegung die beste Alternative. Insbesondere wird sich herausstellen, dass wir die benötigte Spektralzerlegung teilweise auch an anderen Stellen der Algorithmen benötigen.

Nachdem jetzt die Existenz und Eindeutigkeit der Quadratwurzel von  $A \succeq 0$  geklärt ist, zeigen wir jetzt, dass die Wurzelfunktion stetig ist.

**Lemma 2.10** *Die Abbildung  $A \mapsto A^{1/2}$  ist stetig auf  $\mathcal{S}_+^{n \times n}$ .*

**Beweis:** Sei  $A \in \mathcal{S}_+^{n \times n}$  fest gegeben und  $\{A_k\} \subseteq \mathcal{S}_+^{n \times n}$  eine beliebige Folge mit  $A_k \rightarrow A$ . Wir müssen jetzt zeigen, dass  $A_k^{1/2} \rightarrow A^{1/2}$  gilt. Zu diesem Zweck verifizieren wir zunächst, dass die Folge  $\{A_k^{1/2}\}$  beschränkt ist: Wäre nämlich  $\{\|A_k^{1/2}\|\}_K \rightarrow \infty$  auf einer Teilfolge, so wäre wegen der Beschränktheit von  $\{A_k\}$  dann

$$\left\{ \left( \frac{A_k^{1/2}}{\|A_k^{1/2}\|} \right)^2 \right\}_K = \left\{ \frac{A_k}{\|A_k^{1/2}\|^2} \right\}_K \rightarrow 0.$$

Dies kann jedoch nicht sein, da jedes Element der Folge

$$\left\{ \frac{A_k^{1/2}}{\|A_k^{1/2}\|} \right\}$$

die Norm Eins hat und somit jeder Häufungspunkt dieser Folge von Null verschieden ist. Also ist die Folge  $\{A_k^{1/2}\}$  beschränkt.

Sei nun  $\tilde{A}$  ein Häufungspunkt von  $\{A_k^{1/2}\}$ . Offensichtlich ist  $\tilde{A}$  dann symmetrisch und positiv semidefinit und es gilt

$$A_k = A_k^{1/2} A_k^{1/2} \rightarrow \tilde{A} \tilde{A} = \tilde{A}^2.$$

Andererseits gilt  $A_k \rightarrow A$  nach Voraussetzung, woraus sich  $\tilde{A}^2 = A$  und aus der Eindeutigkeit der Quadratwurzel daher  $\tilde{A} = A^{1/2}$  ergibt. Jeder Häufungspunkt von  $\{A_k^{1/2}\}$  ist somit  $A^{1/2}$ . Da  $A$  aber nach Satz 2.9 genau eine symmetrische und positiv semidefinite Quadratwurzel besitzt und die Folge  $\{A_k^{1/2}\}$  beschränkt ist, konvergiert die gesamte Folge  $\{A_k^{1/2}\}$  gegen  $A^{1/2}$ , was

## 2. Semidefinite Programme und NCP-Funktionen

zu zeigen war.  $\square$

Wir können jetzt die in (2.8) und (2.9) definierten Funktionen verallgemeinern. Wir tun dies zunächst für die Fischer-Burmeister-Funktion und definieren analog zum reellwertigen Fall die (matrixwertige) Fischer-Burmeister-Funktion  $\phi : \mathcal{S}^{n \times n} \times \mathcal{S}^{n \times n} \rightarrow \mathcal{S}^{n \times n}$  durch

$$\phi(X, S) := X + S - (X^2 + S^2)^{1/2}. \quad (2.10)$$

Aus Lemma 2.10 folgt sofort, dass die Fischer-Burmeister-Funktion stetig auf dem gesamten  $\mathcal{S}^{n \times n} \times \mathcal{S}^{n \times n}$  ist. Wir müssen jetzt natürlich nachweisen, dass dies tatsächlich eine matrixwertige NCP-Funktion ist. Dies geschieht in dem folgenden Lemma, welches auf Tseng [67] zurück geht. Wir benutzen hier jedoch eine etwas andere Beweistechnik, welche sich später noch bei einer entsprechenden Charakterisierung für die Zentralen-Pfad-Bedingungen als nützlich herausstellen wird.

**Lemma 2.11 (Tseng [67])** *Sei  $\phi : \mathcal{S}^{n \times n} \times \mathcal{S}^{n \times n} \rightarrow \mathcal{S}^{n \times n}$  die in (2.10) definierte Fischer-Burmeister-Funktion. Dann gilt*

$$\phi(X, S) = 0 \iff X \succeq 0, S \succeq 0, XS = 0,$$

*d.h. die Fischer-Burmeister-Funktion ist eine NCP-Funktion.*

**Beweis:** Sei zunächst  $X \succeq 0, S \succeq 0$  mit  $XS = 0$ . Dies impliziert  $XS + SX = 0$  und daher gilt

$$(X + S)^2 = X^2 + S^2.$$

Da  $X$  und  $S$  symmetrisch und positiv semidefinit sind, folgt mit Satz 2.9

$$X + S = (X^2 + S^2)^{1/2},$$

da die Quadratwurzel einer symmetrischen und positiv semidefiniten Matrix im Raum der symmetrischen und positiv semidefiniten Matrizen eindeutig bestimmt ist. Dies impliziert offensichtlich  $\phi(X, S) = 0$ .

Sei nun umgekehrt  $\phi(X, S) = 0$  für zwei symmetrische Matrizen  $X, S \in \mathcal{S}^{n \times n}$ , also  $X + S = (X^2 + S^2)^{1/2}$ . Quadrieren beider Seiten dieser Gleichung liefert

$$(X + S)^2 = X^2 + S^2 \quad \text{und} \quad X + S \in \mathcal{S}_+^{n \times n}.$$

Dies ist äquivalent zu

$$XS + SX = 0 \quad \text{und} \quad X + S \in \mathcal{S}_+^{n \times n}. \quad (2.11)$$

Sei nun  $X = QDQ^T$  mit einer orthogonalen Matrix  $Q \in \mathbb{R}^{n \times n}$  und einer Diagonalmatrix  $D = \text{diag}(\lambda_1, \dots, \lambda_n)$  eine nach Satz 2.4 existierende Spektralzerlegung der symmetrischen Matrix  $X$ . Damit kann (2.11) geschrieben werden als

$$QDQ^T S + SQDQ^T = 0 \quad \text{und} \quad QDQ^T + S \in \mathcal{S}_+^{n \times n}.$$

Multiplikation dieser Gleichung mit  $Q^T$  von links und  $Q$  von rechts ergibt

$$DQ^T S Q + Q^T S Q D = 0 \quad \text{und} \quad D + Q^T S Q \in \mathcal{S}_+^{n \times n}.$$

Mit der Abkürzung  $A := Q^T S Q$  schreibt sich dies als

$$DA + AD = 0 \quad \text{und} \quad D + A \in \mathcal{S}_+^{n \times n}. \quad (2.12)$$

Komponentenweise bedeutet dies

$$(\lambda_i + \lambda_j)a_{ij} = 0 \quad \text{und} \quad D + A \in \mathcal{S}_+^{n \times n} \quad (2.13)$$

für alle  $i, j = 1, \dots, n$ . Wählen wir speziell  $i = j$ , so ergibt sich  $2\lambda_i a_{ii} = 0$  und  $\lambda_i + a_{ii} \geq 0$  für alle  $i = 1, \dots, n$ . Daraus folgt sofort  $\lambda_i \geq 0$  für alle  $i = 1, \dots, n$ , was bedeutet, dass  $X$  positiv semidefinit ist. Ein analoges Argument (basierend auf einer Spektralzerlegung von  $S$ ) zeigt, dass auch  $S$  positiv semidefinit ist.

Weiter folgt nun aus (2.11)

$$X \bullet S = \text{tr}(XS) = \frac{1}{2} \text{tr}(XS + SX) = 0.$$

Mit Lemma 2.7 folgt daher  $XS = 0$ . Damit ist alles gezeigt.  $\square$

Als Nächstes definieren wir die Minimum-Funktion  $\phi : \mathcal{S}^{n \times n} \times \mathcal{S}^{n \times n} \rightarrow \mathcal{S}^{n \times n}$  durch

$$\phi(X, S) := X + S - ((X - S)^2)^{1/2}. \quad (2.14)$$

Die Minimum-Funktion ist wegen Lemma 2.10 ebenfalls stetig. Sie ist auch eine NCP-Funktion, wie das folgende Lemma zeigt.

**Lemma 2.12 (Tseng [67])** *Sei  $\phi : \mathcal{S}^{n \times n} \times \mathcal{S}^{n \times n} \rightarrow \mathcal{S}^{n \times n}$  die in (2.14) definierte Minimum-Funktion. Dann gilt*

$$\phi(X, S) = 0 \iff X \succeq 0, S \succeq 0, XS = 0,$$

*d.h. die Minimum-Funktion ist eine NCP-Funktion.*

## 2. Semidefinite Programme und NCP-Funktionen

Der Beweis erfolgt völlig analog zu dem von Lemma 2.11 und wird daher hier weggelassen.

Wir sind nun in der Lage, eine Charakterisierung der Optimalitätsbedingungen anzugeben. Sei dazu  $\phi$  entweder die in (2.10) definierte Fischer-Burmeister-Funktion oder die Minimum-Funktion aus (2.14). Damit definieren wir die Abbildung  $\Phi : \mathcal{S}^{n \times n} \times \mathbb{R}^m \times \mathcal{S}^{n \times n} \rightarrow \mathcal{S}^{n \times n} \times \mathbb{R}^m \times \mathcal{S}^{n \times n}$  durch

$$\Phi(X, \lambda, S) := \begin{pmatrix} \sum_{i=1}^m \lambda_i A_i + S - C \\ A_i \bullet X - b_i \quad (i = 1, \dots, m) \\ \phi(X, S) \end{pmatrix}. \quad (2.15)$$

Dann ergibt sich aus den Lemmata 2.11 und 2.12 sofort das folgende Resultat.

**Satz 2.13** *Sei  $\Phi$  definiert durch (2.15), wobei  $\phi$  entweder durch (2.10) oder (2.14) gegeben ist. Dann sind die folgenden Aussagen äquivalent:*

- (a)  $(X^*, \lambda^*, S^*)$  ist eine Lösung der Optimalitätsbedingungen (2.6).
- (b)  $(X^*, \lambda^*, S^*)$  ist eine Lösung des nichtlinearen Gleichungssystems

$$\Phi(X, \lambda, S) = 0.$$

Wie man leicht sieht, ist die reellwertige Fischer-Burmeister-Funktion im Punkt  $(a, b) = (0, 0)$  nicht differenzierbar, da die Wurzelfunktion im Ursprung nicht differenzierbar ist. Entsprechendes gilt für die Minimum-Funktion in allen Punkten  $(a, b)$  mit  $a = b$ . Ein entsprechendes Verhalten überträgt sich natürlich auch auf die verallgemeinerten, matrixwertigen NCP-Funktionen. Wir werden die Differenzierbarkeit der beiden NCP-Funktionen später noch formal untersuchen, wollen dies jedoch jetzt als Motivation nehmen, für einen Parameter  $\tau > 0$  die Approximation (vergleiche auch Lemma 2.23)

$$\phi_\tau(X, S) := X + S - (X^2 + S^2 + 2\tau^2 I)^{1/2} \quad (2.16)$$

der Fischer-Burmeister-Funktion bzw.

$$\phi_\tau(X, S) := X + S - ((X - S)^2 + 4\tau^2 I)^{1/2} \quad (2.17)$$

als Approximation der Minimum-Funktion einzuführen. Es wird sich später herausstellen, dass diese Funktionen für  $\tau > 0$  stetig differenzierbar, d.h. tatsächlich glatte Approximationen an die entsprechenden NCP-Funktionen sind. Diese Funktionen heißen daher auch *geglättete Fischer-Burmeister-Funktion* bzw. *geglättete Minimum-Funktion*. Wir haben zwar die Nicht-Differenzierbarkeit der NCP-Funktionen (2.10) und (2.14) als Motivation zur Einführung dieser geglätteten Varianten angegeben, es stellt sich jedoch heraus, dass diese Funktionen eine neue Charakterisierung der Zentralen-Pfad-Bedingungen liefern. Dazu dient das folgende Lemma, welches auf [42] zurückgeht.

**Lemma 2.14** Sei  $\tau > 0$  eine beliebige positive reelle Zahl und  $\phi_\tau : \mathcal{S}^{n \times n} \times \mathcal{S}^{n \times n} \rightarrow \mathcal{S}^{n \times n}$  die in (2.16) definierte geglättete Fischer-Burmeister-Funktion. Dann gilt

$$\phi_\tau(X, S) = 0 \iff X \succ 0, S \succ 0, XS = \tau^2 I.$$

**Beweis:** Sei zunächst  $X \succ 0, S \succ 0$  und  $XS = \tau^2 I$ . Dies impliziert  $XS + SX = 2\tau^2 I$  und daher  $(X + S)^2 = X^2 + S^2 + 2\tau^2 I$ . Da  $X$  und  $S$  beide symmetrisch und positiv definit sind, folgt  $X + S = (X^2 + S^2 + 2\tau^2 I)^{1/2}$  und somit  $\phi_\tau(X, S) = 0$ .

Umgekehrt sei  $\phi_\tau(X, S) = 0$  für zwei symmetrische Matrizen  $X, S \in \mathcal{S}^{n \times n}$ . Das bedeutet  $X + S = (X^2 + S^2 + 2\tau^2 I)^{1/2}$ . Quadrieren dieser Gleichung auf beiden Seiten ergibt

$$X^2 + S^2 + 2\tau^2 I = (X + S)^2 \quad \text{und} \quad X + S \in \mathcal{S}_{++}^{n \times n}.$$

Dies ist äquivalent zu

$$XS + SX = 2\tau^2 I \quad \text{und} \quad X + S \in \mathcal{S}_{++}^{n \times n}. \quad (2.18)$$

Sei nun  $X = QDQ^T$  mit einer orthogonalen Matrix  $Q \in \mathbb{R}^{n \times n}$  und einer Diagonalmatrix  $D = \text{diag}(\lambda_1, \dots, \lambda_n)$  eine Spektralzerlegung der Matrix  $X$ . Analog zu der Vorgehensweise im Beweis zu Lemma 2.11 und mit der Definition  $A := Q^T S Q$  kann Gleichung (2.18) geschrieben werden als

$$DA + AD = 2\tau^2 I \quad \text{und} \quad D + A \in \mathcal{S}_{++}^{n \times n}. \quad (2.19)$$

Komponentenweise schreibt sich dies als

$$(\lambda_i + \lambda_j)a_{ij} = 2\tau^2 \delta_{ij} \quad \text{und} \quad D + A \in \mathcal{S}_{++}^{n \times n} \quad (2.20)$$

für alle  $i, j = 1, \dots, n$ , wobei  $\delta_{ij}$  das übliche Kronecker-Symbol bezeichnet:

$$\delta_{ij} := \begin{cases} 1, & \text{falls } i = j, \\ 0, & \text{sonst.} \end{cases}$$

Speziell für  $i = j$  erhalten wir  $2\lambda_i a_{ii} = 2\tau^2$  und  $\lambda_i + a_{ii} > 0$  für alle  $i = 1, \dots, n$ . Offensichtlich impliziert dies  $\lambda_i > 0$  für alle  $i = 1, \dots, n$ . Also ist die symmetrische Matrix  $X$  positiv definit. Auf gleiche Weise (unter Benutzung einer Spektralzerlegung von  $S$  anstelle von  $X$ ) kann man zeigen, dass auch  $S$  positiv definit ist.

Um zu zeigen, dass  $XS = \tau^2 I$  gilt, bemerken wir zunächst, dass aus (2.20) sofort  $a_{ij} = 0$  für alle  $i \neq j$  folgt, da ja  $\lambda_i + \lambda_j > 0$  gilt. Also ist  $A$  eine

## 2. Semidefinite Programme und NCP-Funktionen

Diagonalmatrix und somit  $DA = AD$ . Daher erhalten wir aus (2.19) die Identität  $DA = \tau^2 I$ . Multiplikation von links mit  $Q$  und von rechts mit  $Q^T$  ergibt  $XS = QDQ^T S = QDAQ^T = \tau^2 I$ .  $\square$

Ein entsprechendes Resultat gilt auch für die geglättete Minimum-Funktion.

**Lemma 2.15** *Sei  $\tau > 0$  eine positive reelle Zahl und  $\phi_\tau : \mathcal{S}^{n \times n} \times \mathcal{S}^{n \times n} \rightarrow \mathcal{S}^{n \times n}$  die in (2.17) definierte geglättete Minimum-Funktion. Dann gilt*

$$\phi_\tau(X, S) = 0 \iff X \succ 0, S \succ 0, XS = \tau^2 I.$$

**Beweis:** Sei zunächst  $X \succ 0, S \succ 0$  und  $XS = \tau^2 I$ . Dann folgt

$$(X - S)^2 = (X + S)^2 - 2(XS + SX) = (X + S)^2 - 4\tau^2 I$$

und das bilden der Quadratwurzel auf beiden Seiten liefert, da  $X$  und  $S$  symmetrisch und positiv definit sind,  $\phi_\tau(X, S) = 0$ .

Sei nun umgekehrt  $\phi_\tau(X, S) = 0$  für zwei symmetrische Matrizen  $X, S \in \mathcal{S}^{n \times n}$ . Dann gilt

$$X + S = ((X - S)^2 + 4\tau^2 I)^{1/2}$$

und somit  $(X - S)^2 + 4\tau^2 I = (X + S)^2$  sowie  $X + S \in \mathcal{S}_{++}^{n \times n}$ . Dies ist äquivalent zu  $XS + SX = 2\tau^2 I$  und  $X + S \in \mathcal{S}_{++}^{n \times n}$ . Analog zum Beweis von Lemma 2.14 folgt dann, dass  $X \succ 0, S \succ 0$  und  $XS = \tau^2 I$  gilt.  $\square$

Ist jetzt  $\phi_\tau$  entweder die geglättete Fischer-Burmeister-Funktion aus (2.16) oder die geglättete Minimum-Funktion aus (2.17), so definieren wir die Abbildung  $\Phi_\tau : \mathcal{S}^{n \times n} \times \mathbb{R}^m \times \mathcal{S}^{n \times n} \rightarrow \mathcal{S}^{n \times n} \times \mathbb{R}^m \times \mathcal{S}^{n \times n}$  durch

$$\Phi_\tau(X, \lambda, S) := \begin{pmatrix} \sum_{i=1}^m \lambda_i A_i + S - C \\ A_i \bullet X - b_i \quad (i = 1, \dots, m) \\ \phi_\tau(X, S) \end{pmatrix}. \quad (2.21)$$

Dann erhalten wir aus den Lemmata 2.14 und 2.15 sofort die folgende Charakterisierung der Zentralen-Pfad-Bedingungen (2.7) für semidefinite Programme.

**Satz 2.16** *Sei  $\Phi_\tau$  definiert durch (2.21), wobei  $\phi_\tau$  entweder durch (2.16) oder (2.17) gegeben ist. Dann sind die folgenden Aussagen für  $\tau > 0$  äquivalent:*

- (a)  $(X^*, \lambda^*, S^*)$  ist eine Lösung der Zentralen-Pfad-Bedingungen (2.7).
- (b)  $(X^*, \lambda^*, S^*)$  ist eine Lösung des nichtlinearen Gleichungssystems

$$\Phi_\tau(X, \lambda, S) = 0.$$

## 2.3. Lyapunov-Operatoren

Bevor wir uns im nächsten Abschnitt mit weiteren Eigenschaften der NCP-Funktionen beschäftigen wollen, werden wir in diesem Abschnitt so genannte Lyapunov-Operatoren betrachten. Sei dazu  $A \succeq 0$  eine symmetrische und positiv semidefinite Matrix. Dann ist der *Lyapunov-Operator*  $L_A : \mathcal{S}^{n \times n} \rightarrow \mathcal{S}^{n \times n}$  definiert durch

$$L_A[X] := AX + XA. \quad (2.22)$$

Es stellt sich zunächst die Frage, ob dieser Operator invertierbar ist. Für den Fall, dass  $A$  positiv definit ist, beantwortet dies der folgende Satz, dessen Beweis beispielsweise in dem Buch von Geiger und Kanzow [25, Satz 4.29] gefunden werden kann.

**Satz 2.17** *Sei  $A \in \mathcal{S}_{++}^{n \times n}$  eine symmetrische und positiv definite Matrix und  $B \in \mathcal{S}^{n \times n}$  symmetrisch. Dann besitzt die Lyapunov-Gleichung*

$$AX + XA = B \quad (2.23)$$

*genau eine Lösung  $X \in \mathcal{S}^{n \times n}$ .*

Wir verzichten zwar an dieser Stelle auf den formalen Beweis, wollen jedoch erwähnen, dass sich aus dem Beweis ein konstruktives Verfahren zur Berechnung von  $X \in \mathcal{S}^{n \times n}$  ergibt. Ist  $A = QDQ^T$  eine Spektralzerlegung von  $A$  mit  $D = \text{diag}(d_1, \dots, d_n)$ , so gilt

$$X = Q(\Lambda \star Q^T B Q)Q^T.$$

Dabei ist  $\Lambda \in \mathcal{S}^{n \times n}$  die Matrix mit den Elementen

$$\lambda_{ij} = \frac{1}{d_i + d_j}$$

und

$$A \star B = (a_{ij} b_{ij})_{i,j=1,\dots,n}$$

bezeichnet das so genannte *Hadamard-Produkt* von  $A, B \in \mathbb{R}^{n \times n}$ .

Im Fall  $A \succ 0$  bezeichnen wir mit  $L_A^{-1}[B]$  die nach Satz 2.17 eindeutig bestimmte Lösung  $X \in \mathcal{S}^{n \times n}$  von (2.23).

Zwar wird in den meisten Fällen die Matrix  $A$  positiv definit sein, an einigen Stellen werden wir es jedoch lediglich mit einer positiv semidefiniten Matrix  $A$  zu tun haben. In diesem Fall existiert die Inverse des Lyapunov  $L_A$  nicht mehr auf dem ganzen Raum  $\mathcal{S}^{n \times n}$ . Wir definieren daher den Unterraum  $\mathcal{S}_A^{n \times n}$

## 2. Semidefinite Programme und NCP-Funktionen

von  $\mathcal{S}^{n \times n}$  als die Menge aller Matrizen  $B \in \mathcal{S}^{n \times n}$ , in deren Kern der Kern von  $A$  enthalten ist, d.h.

$$\mathcal{S}_A^{n \times n} := \{B \in \mathcal{S}^{n \times n} \mid \text{Kern}(A) \subseteq \text{Kern}(B)\}.$$

Es gilt  $\mathcal{S}_A^{n \times n} = \mathcal{S}^{n \times n}$ , falls  $A$  positiv definit und somit regulär ist. Allgemein liefert der folgende Satz eine Darstellung von  $\mathcal{S}_A^{n \times n}$ , welche man (ohne Beweis) auch in Tseng [67] findet.

**Satz 2.18** Sei  $A \in \mathcal{S}_+^{n \times n}$  eine symmetrische und positiv semidefinite Matrix und

$$A = Q \begin{pmatrix} \tilde{A}_{II} & 0 \\ 0 & 0 \end{pmatrix} Q^T \quad (2.24)$$

mit einer orthogonalen Matrix  $Q \in \mathbb{R}^{n \times n}$ , einer geeigneten Teilmenge  $I \subset \{1, \dots, n\}$  und einer positiv definiten Untermatrix  $\tilde{A}_{II}$ . Dann gilt

$$\mathcal{S}_A^{n \times n} = \left\{ B \in \mathcal{S}^{n \times n} \mid B = Q \begin{pmatrix} \tilde{B}_{II} & 0 \\ 0 & 0 \end{pmatrix} Q^T \text{ mit einer Untermatrix } \tilde{B}_{II} \right\}.$$

**Beweis:** Wir setzen

$$\tilde{\mathcal{S}}_A^{n \times n} := \left\{ B \in \mathcal{S}^{n \times n} \mid B = Q \begin{pmatrix} \tilde{B}_{II} & 0 \\ 0 & 0 \end{pmatrix} Q^T \text{ mit einer Untermatrix } \tilde{B}_{II} \right\}$$

und müssen  $\mathcal{S}_A^{n \times n} = \tilde{\mathcal{S}}_A^{n \times n}$ , also  $\mathcal{S}_A^{n \times n} \subseteq \tilde{\mathcal{S}}_A^{n \times n}$  und  $\tilde{\mathcal{S}}_A^{n \times n} \subseteq \mathcal{S}_A^{n \times n}$  zeigen.

Zu  $\tilde{\mathcal{S}}_A^{n \times n} \subseteq \mathcal{S}_A^{n \times n}$ : Sei  $B \in \tilde{\mathcal{S}}_A^{n \times n}$ . Für  $x \in \text{Kern}(A)$  beliebig gilt dann mit  $J := \{1, \dots, n\} \setminus I$  und  $\tilde{x} := Q^T x = (\tilde{x}_I^T, \tilde{x}_J^T)^T$

$$0 = Ax = Q \begin{pmatrix} \tilde{A}_{II} & 0 \\ 0 & 0 \end{pmatrix} Q^T x = Q \begin{pmatrix} \tilde{A}_{II} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \tilde{x}_I \\ \tilde{x}_J \end{pmatrix} = Q \begin{pmatrix} \tilde{A}_{II} \tilde{x}_I \\ 0 \end{pmatrix}.$$

Da  $Q$  orthogonal und  $\tilde{A}_{II}$  positiv definit ist, folgt daraus  $\tilde{x}_I = 0$  und somit

$$Bx = Q \begin{pmatrix} \tilde{B}_{II} & 0 \\ 0 & 0 \end{pmatrix} Q^T x = Q \begin{pmatrix} \tilde{B}_{II} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \tilde{x}_I \\ \tilde{x}_J \end{pmatrix} = Q \begin{pmatrix} \tilde{B}_{II} \tilde{x}_I \\ 0 \end{pmatrix} = 0.$$

Also ist  $x \in \text{Kern}(B)$  und somit  $B \in \mathcal{S}_A^{n \times n}$ .

Zu  $\mathcal{S}_A^{n \times n} \subseteq \tilde{\mathcal{S}}_A^{n \times n}$ : Sei  $B \in \mathcal{S}_A^{n \times n}$ , also  $\text{Kern}(A) \subseteq \text{Kern}(B)$ . Da die Matrix  $B$  symmetrisch ist, lässt sich  $B$  mit  $J := \{1, \dots, n\} \setminus I$  schreiben als

$$B = Q \begin{pmatrix} \tilde{B}_{II} & \tilde{B}_{IJ} \\ \tilde{B}_{IJ}^T & \tilde{B}_{JJ} \end{pmatrix} Q^T.$$



Sei jetzt  $\tilde{x}_J$  beliebig. Dann gilt wegen (2.24)

$$x := Q \begin{pmatrix} 0 \\ \tilde{x}_J \end{pmatrix} \in \text{Kern}(A) \subseteq \text{Kern}(B)$$

und somit

$$0 = Bx = Q \begin{pmatrix} \tilde{B}_{II} & \tilde{B}_{IJ} \\ \tilde{B}_{IJ}^T & \tilde{B}_{JJ} \end{pmatrix} Q^T x = Q \begin{pmatrix} \tilde{B}_{II} & \tilde{B}_{IJ} \\ \tilde{B}_{IJ}^T & \tilde{B}_{JJ} \end{pmatrix} \begin{pmatrix} 0 \\ \tilde{x}_J \end{pmatrix} = Q \begin{pmatrix} \tilde{B}_{IJ}\tilde{x}_J \\ \tilde{B}_{JJ}\tilde{x}_J \end{pmatrix}.$$

Multiplikation von links mit  $Q^T$  liefert  $\tilde{B}_{IJ}\tilde{x}_J = 0$  und  $\tilde{B}_{JJ}\tilde{x}_J = 0$ . Da  $\tilde{x}_J$  beliebig war, folgt daraus sofort  $\tilde{B}_{IJ} = 0$  und  $\tilde{B}_{JJ} = 0$ , also  $B \in \tilde{\mathcal{S}}_A^{n \times n}$ . Damit ist alles gezeigt.  $\square$

Ohne Beschränkung der Allgemeinheit kann in der Zerlegung (2.24) die Matrix  $\tilde{A}_{II}$  als positiv definite Diagonalmatrix gewählt werden. Dann lässt sich leicht zeigen, dass der zugehörige Lyapunov-Operator  $L_A : \mathcal{S}_A^{n \times n} \rightarrow \mathcal{S}_A^{n \times n}$  auf dem Raum  $\mathcal{S}_A^{n \times n}$  strikt monoton ist. Für alle  $X \in \mathcal{S}_A^{n \times n}$ ,  $X \neq 0$  gilt mit den in Satz 2.18 eingeführten Notationen nämlich

$$X \bullet L_A[X] = 2 \text{tr}[\tilde{A}_{II}\tilde{X}_{II}^2] = 2 \sum_{i \in I} \tilde{a}_{ii} \left( \sum_{j \in I} \tilde{x}_{ij}^2 \right).$$

Dabei gilt  $\tilde{a}_{ii} > 0$  und wegen  $X \neq 0$  und somit  $\tilde{X}_{II} \neq 0$  ist auch die zweite Summe für mindestens ein  $i \in I$  positiv. Daher gilt  $X \bullet L_A[X] > 0$ , so dass  $L_A$  strikt monoton auf  $\mathcal{S}_A^{n \times n}$  ist. Somit existiert die Inverse  $L_A^{-1}$ . Für jedes  $B \in \mathcal{S}_A^{n \times n}$  bezeichnet daher  $L_A^{-1}[B]$  die eindeutig bestimmte Matrix  $X \in \mathcal{S}_A^{n \times n}$  mit  $AX + XA = B$ . Analog zum Fall  $A \succ 0$  kann auch hier eine explizite Form angegeben werden: Wird die Matrix  $\tilde{A}_{II}$  wieder speziell als Diagonalmatrix gewählt, so gilt

$$L_A^{-1}[B] = Q \begin{pmatrix} \tilde{\Lambda}_{II} \star \tilde{B}_{II} & 0 \\ 0 & 0 \end{pmatrix} Q^T. \quad (2.25)$$

Dabei bezeichnet  $\tilde{\Lambda}_{II}$  die Matrix mit den Elementen

$$\tilde{\lambda}_{ij} = \frac{1}{\tilde{a}_{ii} + \tilde{a}_{jj}}, \quad i, j \in I.$$

In den folgenden Resultaten sind einige Eigenschaften des Lyapunov-Operators zusammengefasst, welche teilweise (jedoch ohne Beweis) auch in Tseng [67] gefunden werden können.

## 2. Semidefinite Programme und NCP-Funktionen

**Lemma 2.19** Seien  $A, B \in \mathcal{S}_+^{n \times n}$  zwei symmetrische, positiv semidefinite Matrizen und  $L_A, L_B$  die durch (2.22) definierten zugehörigen Lyapunov-Operatoren, deren Inversen (auf  $\mathcal{S}_A^{n \times n}$  bzw.  $\mathcal{S}_B^{n \times n}$ ) wir mit  $L_A^{-1}, L_B^{-1}$  bezeichnen. Dann gelten die folgenden Aussagen:

(a)  $L_A$  und  $L_B$  sind selbst-adjungiert.

(b)  $L_A^{-1}$  und  $L_B^{-1}$  sind selbst-adjungiert.

Ist zusätzlich  $A, B \succ 0$  und  $AB + BA \succeq 0$ , so gilt:

(c)  $L_A \circ L_B$  und  $L_B \circ L_A$  sind gleichmäßig monoton.

(d)  $L_A^{-1} \circ L_B$  und  $L_B^{-1} \circ L_A$  sind gleichmäßig monoton.

**Beweis:** (a) Wir müssen nur zeigen, dass  $L_A$  selbst-adjungiert ist. Dies folgt sofort aus der Gültigkeit von

$$\begin{aligned}
 L_A[X] \bullet Y &= \text{tr}(L_A[X]Y) \\
 &= \text{tr}((AX + XA)Y) \\
 &= \text{tr}(AXY) + \text{tr}(XAY) \\
 &= \text{tr}(XYA) + \text{tr}(XAY) \\
 &= \text{tr}(X(A Y + Y A)) \\
 &= \text{tr}(X L_A[Y]) \\
 &= X \bullet L_A[Y]
 \end{aligned}$$

für alle  $X, Y \in \mathcal{S}^{n \times n}$ .

(b) Wir zeigen, dass  $L_A^{-1}$  selbst-adjungiert ist. Da  $L_A^{-1}$  der inverse Operator zu  $L_A$  ist, ergibt sich aus Teil (a)

$$L_A^{-1}[X] \bullet Y = L_A^{-1}[X] \bullet L_A[L_A^{-1}[Y]] = L_A[L_A^{-1}[X]] \bullet L_A^{-1}[Y] = X \bullet L_A^{-1}[Y]$$

für alle  $X, Y \in \mathcal{S}_A^{n \times n}$ .

(c) Unter Verwendung von Teil (a) ergibt sich

$$\begin{aligned}
 (L_A \circ L_B[X]) \bullet X &= L_B[X] \bullet L_A[X] \\
 &= \text{tr}(L_B[X]L_A[X]) \\
 &= \text{tr}((BX + XB)(AX + XA)) \\
 &= \text{tr}(BXAX + XBAX + BXXA + XBXA) \quad (2.26) \\
 &= \text{tr}(2BXAX + X^2(BA + AB)) \\
 &= 2 \text{tr}(BXAX) + \text{tr}(X^2(BA + AB)) \\
 &= 2 \|B^{1/2} X A^{1/2}\|_F^2 + \text{tr}(X(BA + AB)X)
 \end{aligned}$$

für alle  $X \in \mathcal{S}^{n \times n}$ . Wegen  $BA + AB \succeq 0$  ist auch die Matrix  $X(BA + AB)X$  symmetrisch und positiv semidefinit und daher gilt

$$\operatorname{tr}(X(BA + AB)X) \geq 0 \quad (2.27)$$

für alle  $X \in \mathcal{S}^{n \times n}$ . Da die Abbildung

$$X \mapsto \|B^{1/2}XA^{1/2}\|_F$$

eine Norm definiert und in endlich-dimensionalen Räumen alle Normen äquivalent sind, existiert eine Konstante  $\mu > 0$ , so dass

$$\|B^{1/2}XA^{1/2}\|_F \geq \mu \|X\|_F \quad (2.28)$$

für alle  $X \in \mathcal{S}^{n \times n}$  gilt. Unter Benutzung von (2.26)–(2.28) ergibt sich

$$(L_A \circ L_B[X]) \bullet X \geq 2 \|B^{1/2}XA^{1/2}\|_F^2 \geq 2\mu^2 \|X\|_F^2,$$

d.h.  $L_A \circ L_B$  ist gleichmäßig monoton auf dem Raum  $\mathcal{S}^{n \times n}$ . Um zu zeigen, dass auch  $L_B \circ L_A$  gleichmäßig monoton ist, müssen wir nur die Rollen von  $A$  und  $B$  vertauschen.

(d) Da  $L_A$  nach Teil (a) selbst-adjungiert ist, erhalten wir für jedes  $X \in \mathcal{S}^{n \times n}$  (mit  $Y := L_A^{-1}[X]$ )

$$(L_A^{-1} \circ L_B[X]) \bullet X = (L_A^{-1} \circ L_B \circ L_A[Y]) \bullet L_A[Y] = (L_B \circ L_A[Y]) \bullet Y.$$

Nun ist aber  $L_B \circ L_A$  gleichmäßig monoton nach Teil (c). Daher folgt (d) aus (c).  $\square$

In Lemma 2.19 haben wir in Teil (c) und (d) neben  $A, B \succ 0$  auch noch  $AB + BA \succeq 0$  vorausgesetzt. Es stellt sich nun die Frage, ob diese Ungleichung nicht bereits aus der positiven Definitheit von  $A$  und  $B$  folgt. Wenn  $A$  und  $B$  kommutieren, so folgt dies sofort aus der gemeinsamen Spektralzerlegung von  $A$  und  $B$ . Im Allgemeinen gilt dies jedoch nicht: Setzen wir

$$A := \begin{pmatrix} 10 & 5 \\ 5 & 6 \end{pmatrix} \succ 0, \quad B := \begin{pmatrix} 1 & 2 \\ 2 & 6 \end{pmatrix} \succ 0,$$

so zeigt eine einfache Rechnung

$$AB + BA = \begin{pmatrix} 40 & 67 \\ 67 & 92 \end{pmatrix}.$$

Diese Matrix ist nicht positiv semidefinit, da  $\det(AB + BA) = -809 < 0$  gilt.

Die folgenden zwei Lemmata werden sich im späteren Verlauf als nützlich erweisen.

## 2. Semidefinite Programme und NCP-Funktionen

**Lemma 2.20** Sei  $A \succeq 0$  und  $X \in \mathcal{S}_A^{n \times n}$ . Dann gilt die Äquivalenz

$$XL_A^{-1}[X] = 0 \iff X = 0.$$

**Beweis:** Sei  $XL_A^{-1}[X] = 0$  und

$$A = Q \begin{pmatrix} D_I & 0 \\ 0 & 0 \end{pmatrix} Q^T$$

mit einer regulären Diagonalmatrix  $D_I$  und einer orthogonalen Matrix  $Q \in \mathbb{R}^{n \times n}$  eine Spektralzerlegung von  $A$ . Wegen  $X \in \mathcal{S}_A^{n \times n}$  lässt sich  $X$  schreiben als

$$X = Q \begin{pmatrix} \tilde{X}_{II} & 0 \\ 0 & 0 \end{pmatrix} Q^T$$

mit einer symmetrischen Untermatrix  $\tilde{X}_{II}$ . Dann folgt aus  $XL_A^{-1}[X] = 0$  sofort  $\tilde{X}_{II}^2 = 0$  (vergleiche (2.25)) und aus der Symmetrie von  $\tilde{X}_{II}$  folgt somit  $\tilde{X}_{II} = 0$ . Also gilt  $X = 0$ . Die andere Richtung ist offensichtlich.  $\square$

**Lemma 2.21** Seien  $A, B \in \mathcal{S}^{n \times n}$  zwei kommutierende Matrizen und  $L_A, L_B$  die zugehörigen Lyapunov-Operatoren. Dann gilt  $L_A \circ L_B = L_B \circ L_A$ .

**Beweis:** Seien  $A$  und  $B$  zwei vertauschbare Matrizen, d.h. es gilt  $AB = BA$ . Dann folgt

$$\begin{aligned} (L_A \circ L_B)[X] &= L_A[L_B[X]] = L_A[BX + XB] \\ &= A(BX + XB) + (BX + XB)A \\ &= B(AX + XA) + (AX + XA)B \\ &= (L_B \circ L_A)[X] \end{aligned}$$

für jede Matrix  $X \in \mathcal{S}^{n \times n}$ . Dies impliziert  $L_A \circ L_B = L_B \circ L_A$ .  $\square$

Wir beenden den Abschnitt mit einem Resultat über die Stetigkeit des inversen Lyapunov-Operators.

**Lemma 2.22 (Chen, Tseng [15])** Die Abbildung  $(X, A) \mapsto L_A^{-1}[X]$  ist stetig im Punkt  $(X, A) \in \mathcal{S}^{n \times n} \times \mathcal{S}_{++}^{n \times n}$ .

**Beweis:** Sei  $\{(X^k, A^k)\}$  eine gegen  $(X, A) \in \mathcal{S}^{n \times n} \times \mathcal{S}_{++}^{n \times n}$  konvergente Folge. Wir setzen  $Y^k := L_{A^k}^{-1}[X^k]$ , d.h. es gilt

$$A^k Y^k + Y^k A^k = X^k$$

für alle  $k \in \mathbb{N}$ . Angenommen, es gilt  $\|Y^k\|_2 \rightarrow \infty$ . Dann gilt

$$A^k \left( \frac{Y^k}{\|Y^k\|_2} \right) + \left( \frac{Y^k}{\|Y^k\|_2} \right) A^k = \frac{X^k}{\|Y^k\|_2} \rightarrow 0.$$

Sei nun  $\tilde{Y}$  ein Häufungspunkt von  $Y^k / \|Y^k\|_2$ . Dann ist  $\|\tilde{Y}\|_2 = 1$  und Grenzübergang auf einer gegen  $\tilde{Y}$  konvergenten Teilfolge liefert

$$A\tilde{Y} + \tilde{Y}A = 0.$$

Wegen  $A \succ 0$  ist dann  $\tilde{Y} = 0$ , ein Widerspruch zu  $\|\tilde{Y}\|_2 = 1$ . Daher ist die Folge  $\{Y^k\}$  beschränkt. Dann erfüllt aber jeder Häufungspunkt  $Y$  von  $\{Y^k\}$  die Gleichung

$$AY + YA = X,$$

also ist  $Y = L_A^{-1}[X]$  nach Satz 2.17 und die gesamte Folge  $\{Y^k\}$  konvergiert gegen  $Y$ . Damit ist alles gezeigt.  $\square$

## 2.4. Eigenschaften von NCP-Funktionen

In Abschnitt 2.2 haben wir die Fischer-Burmeister-Funktion und die Minimum-Funktion eingeführt und gezeigt, dass dies NCP-Funktionen sind. In diesem Abschnitt wollen wir weitere Eigenschaften, insbesondere die Differenzierbarkeit im Sinne von Fréchet, dieser Funktionen und ihrer geglätteten Varianten untersuchen. Da sich die ungeglätteten Funktionen jeweils als Spezialfall  $\tau = 0$  aus den geglätteten Varianten ergeben, werden wir uns bei den folgenden Untersuchungen auf die geglättete Fischer-Burmeister- bzw. Minimum-Funktion beschränken.

In den vorhergehenden Betrachtungen haben wir die nichtnegative Zahl  $\tau$  als einen Parameter betrachtet. Im Gegensatz dazu werden wir jetzt  $\tau$  als eine eigenständige Variable ansehen. Um dies auch in der Notation klar zu machen, setzen wir  $\phi(X, S, \tau) := \phi_\tau(X, S)$ , d.h. wir schreiben

$$\phi(X, S, \tau) := X + S - (X^2 + S^2 + 2\tau^2 I)^{1/2} \quad (2.29)$$

für die geglättete Fischer-Burmeister-Funktion aus (2.16) und

$$\phi(X, S, \tau) := X + S - ((X - S)^2 + 4\tau^2 I)^{1/2} \quad (2.30)$$

für die geglättete Minimum-Funktion aus (2.17).

Wir beginnen unsere Analyse der Funktionen  $\phi$  mit den folgenden beiden Resultaten aus Chen und Tseng [15].

## 2. Semidefinite Programme und NCP-Funktionen

**Lemma 2.23 (Chen, Tseng [15])** *Sei  $\phi$  entweder die geglättete Fischer-Burmeister-Funktion aus (2.29) oder die geglättete Minimum-Funktion aus (2.30). Dann gilt für alle  $X, S \in \mathcal{S}^{n \times n}$  und alle  $\tau > \nu \geq 0$  die Ungleichung*

$$\kappa(\tau - \nu)I \succeq \phi(X, S, \nu) - \phi(X, S, \tau) \succeq 0,$$

wobei  $\kappa$  eine von  $\phi$ , nicht jedoch von  $X, S, \tau$  und  $\nu$  abhängige (bekannte) Konstante ist.

**Beweis:** Seien  $X, S \in \mathcal{S}^{n \times n}$  und  $\tau > \nu > 0$  fest gegeben sowie  $\phi$  die in (2.30) definierte geglättete Minimum-Funktion. Setzen wir  $A := (X - S)^2$  und sei  $A = QDQ^T$  mit einer orthogonalen Matrix  $Q \in \mathbb{R}^{n \times n}$  und einer Diagonalmatrix  $D = \text{diag}(\lambda_1, \dots, \lambda_n) \in \mathbb{R}^{n \times n}$  eine Spektralzerlegung von  $A$ , so erhalten wir

$$\begin{aligned} \phi(X, S, \tau) &= X + S - (A + 4\tau^2 I)^{1/2} \\ &= X + S - Q \text{diag}[(\lambda_1 + 4\tau^2)^{1/2}, \dots, (\lambda_n + 4\tau^2)^{1/2}] Q^T \end{aligned}$$

und analog

$$\phi(X, S, \nu) = X + S - Q \text{diag}[(\lambda_1 + 4\nu^2)^{1/2}, \dots, (\lambda_n + 4\nu^2)^{1/2}] Q^T.$$

Daraus folgt

$$\phi(X, S, \nu) - \phi(X, S, \tau) = Q \text{diag}[(\lambda_i + 4\tau^2)^{1/2} - (\lambda_i + 4\nu^2)^{1/2}]_{i=1}^n Q^T.$$

Da  $A$  positiv semidefinit ist, gilt  $\lambda_i \geq 0$  für alle  $i = 1, \dots, n$ . Daher folgt

$$0 < (\lambda_i + 4\tau^2)^{1/2} - (\lambda_i + 4\nu^2)^{1/2} \leq 2(\tau - \nu),$$

wobei wir in der zweiten Ungleichung benutzt haben, dass für jedes  $\lambda \in \mathbb{R}_+$  die Funktion  $h(\tau) := (\lambda + 4\tau^2)^{1/2}$  konvex und differenzierbar auf  $\mathbb{R}_{++}$  mit  $h'(\tau) = 4(\lambda + 4\tau^2)^{-1/2}\tau \leq 2$  ist, so dass  $h(\tau) - h(\nu) \leq h'(\tau)(\tau - \nu) \leq 2(\tau - \nu)$  gilt. Also folgt

$$2(\tau - \nu)I \succeq \phi(X, S, \nu) - \phi(X, S, \tau) \succ 0.$$

Setzen wir jetzt  $\kappa := 2$ , so folgt die Behauptung für den Fall  $\nu > 0$ . Da die Behauptung für alle  $\nu \in (0, \tau)$  gilt, folgt durch Grenzübergang  $\nu \rightarrow 0$  die Aussage auch für  $\nu = 0$ , da  $\phi(X, S, \nu)$  stetig in  $\nu$  ist.

Der Beweis für die geglättete Fischer-Burmeister-Funktion ergibt sich analog mit  $\kappa := \sqrt{2}$  und unter Benutzung einer Spektralzerlegung von  $X^2 + S^2$  anstelle von  $(X - S)^2$ . Man vergleiche [15] für weitere Details.  $\square$

Daraus folgt sofort das folgende Resultat.

**Korollar 2.24 (Chen, Tseng [15])** Sei  $\phi$  gegeben durch (2.29) oder (2.30) und sei  $\kappa$  die Konstante aus Lemma 2.23. Dann gilt die Ungleichung

$$\|\phi(X, S, \nu) - \phi(X, S, \tau)\|_F \leq \kappa\sqrt{n}(\tau - \nu)$$

für alle  $X, S \in \mathcal{S}^{n \times n}$  und alle  $\tau > \nu \geq 0$ .

**Beweis:** Da  $\|I\|_F = \sqrt{n}$  gilt, folgt die Behauptung direkt aus den Definitionen der Abbildungen  $\phi$  und der Konstante  $\kappa$  in Lemma 2.23: Seien  $\lambda_1, \dots, \lambda_n$  die Eigenwerte der symmetrischen Matrix  $\phi(X, S, \nu) - \phi(X, S, \tau)$ . Nach Lemma 2.23 gilt  $\kappa(\tau - \nu) \geq \lambda_i \geq 0$ . Daraus folgt

$$\|\phi(X, S, \nu) - \phi(X, S, \tau)\|_F = \sqrt{\lambda_1^2 + \dots + \lambda_n^2} \leq \kappa\sqrt{n}(\tau - \nu)$$

und somit die Behauptung.  $\square$

Wir wollen als Nächstes zeigen, dass die beiden Funktionen  $\phi$  aus (2.29) und (2.30) (zumindest unter geeigneten Voraussetzungen) stetig differenzierbar in den Argumenten  $X, S$  und  $\tau$  sind. Diese Eigenschaften (mit  $\tau$  als Parameter und nicht als Variable) wurden als erstes von Chen und Tseng [15] gezeigt. Die Aussagen ergeben sich auch aus dem Artikel von Sun und Sun [62].

Wir wollen hier jedoch einen etwas anderen Beweis angeben, da (zumindest aus Sicht des Autors) der Beweis von Chen und Tseng in [15] nicht sehr konstruktiv ist, da nicht klar wird, woher die eher komplizierten Formeln für die Ableitung der Funktionen  $\phi$  kommen. Der hier angegebene Beweis basiert auf den beiden folgenden Lemmata, welche in Horn und Johnson [35, Abschnitt 7.2] als Aufgaben angegeben sind. Dabei bezeichnet  $\|A\|_2 = \sqrt{\lambda_{\max}(A^T A)}$  die Spektralnorm der Matrix  $A$ , wobei  $\lambda_{\max}(C)$  der größte Eigenwert der symmetrischen Matrix  $C$  ist. Dies ist die durch die übliche euklidische Vektornorm  $\|\cdot\|_2$  induzierte Matrixnorm, d.h. es gilt

$$\|A\|_2 = \max_{\|x\|_2=1} \|Ax\|_2.$$

Weiter definieren wir für die folgenden Betrachtungen die Normen

$$\|(X, \lambda, S)\| := \sqrt{\|X\|_F^2 + \|\lambda\|_2^2 + \|S\|_F^2}$$

auf dem  $\mathbb{R}^{n \times n} \times \mathbb{R}^m \times \mathbb{R}^{n \times n}$ ,

$$\|(X, S, \tau)\| := \sqrt{\|X\|_F^2 + \|S\|_F^2 + \tau^2}$$

## 2. Semidefinite Programme und NCP-Funktionen

auf dem  $\mathbb{R}^{n \times n} \times \mathbb{R}^{n \times n} \times \mathbb{R}$  und

$$\|(X, \lambda, S, \tau)\| := \sqrt{\|X\|_F^2 + \|\lambda\|_2^2 + \|S\|_F^2 + \tau^2}$$

auf dem  $\mathbb{R}^{n \times n} \times \mathbb{R}^m \times \mathbb{R}^{n \times n} \times \mathbb{R}$ . Wir haben zwar das gleiche Symbol für alle drei Normen benutzt, da es sich jedoch um unterschiedliche Räume handelt, auf denen die Normen definiert sind, ist eine Verwechslung auszuschließen. Man beachte, dass sich diese Normen auf die übliche euklidische Vektornorm reduzieren, wenn man die auftretenden Matrizen als lange Vektoren auffasst. Die durch diese Normen induzierte Operatornorm werden wir einfach mit  $\|\cdot\|$  bezeichnen.

**Lemma 2.25** *Seien  $A, B \in \mathcal{S}_+^{n \times n}$  zwei gegebene symmetrische, positiv semidefinite Matrizen, wobei mindestens eine der beiden Matrizen  $A$  oder  $B$  regulär sei. Dann gilt die Ungleichung*

$$\|A - B\|_2 \leq \|A^2 - B^2\|_2 / (\lambda_{\min}(A) + \lambda_{\min}(B)).$$

**Beweis:** Sei  $E := A - B$  und  $\lambda$  ein Eigenwert der symmetrischen Matrix  $E$  mit  $|\lambda| = \|E\|_2$ , d.h.  $\lambda$  ist ein betragsgrößerer Eigenwert von  $E$ . Sei  $x \neq 0$  ein zugehöriger Eigenvektor mit  $\|x\|_2 = 1$ . Unter Benutzung von  $Ex = \lambda x$  sowie der Gleichung

$$A^2 - B^2 = AE + EB$$

erhalten wir

$$\begin{aligned} \|A^2 - B^2\|_2 &= \|AE + EB\|_2 \\ &= \max_{\|y\|=1} |y^T (AE + EB)y| \\ &\geq |x^T (AE + EB)x| \\ &= |x^T AEx + x^T EBx| \\ &= |\lambda x^T Ax + \lambda x^T Bx| \\ &= |\lambda| \cdot |x^T Ax + x^T Bx| \\ &= |\lambda| \cdot (x^T Ax + x^T Bx) \\ &\geq \|A - B\|_2 \cdot (\lambda_{\min}(A) + \lambda_{\min}(B)). \end{aligned}$$

Da  $A$  und  $B$  beide positiv semidefinit sind und eine der beiden Matrizen sogar als regulär vorausgesetzt war, gilt  $\lambda_{\min}(A) + \lambda_{\min}(B) > 0$ . Division hierdurch liefert die behauptete Ungleichung.  $\square$



**Lemma 2.26** Seien  $A \in \mathcal{S}_{++}^{n \times n}$  und  $B \in \mathcal{S}_{+}^{n \times n}$  zwei gegebene Matrizen. Dann gilt

$$\|A^{1/2} - B^{1/2}\|_2 \leq \|A^{-1/2}\|_2 \cdot \|A - B\|_2.$$

**Beweis:** Unter Benutzung der Identität

$$\|A^{-1/2}\|_2 = \lambda_{\max}(A^{-1/2}) = \frac{1}{\lambda_{\min}(A^{1/2})}$$

und durch Anwendung von Lemma 2.25 erhalten wir

$$\begin{aligned} \|A^{1/2} - B^{1/2}\|_2 &\leq \|A - B\|_2 / (\lambda_{\min}(A^{1/2}) + \lambda_{\min}(B^{1/2})) \\ &\leq \|A - B\|_2 / \lambda_{\min}(A^{1/2}) \\ &= \|A - B\|_2 \cdot \|A^{-1/2}\|_2, \end{aligned}$$

wobei die zweite Ungleichung aus  $\lambda_{\min}(B^{1/2}) \geq 0$  folgt.  $\square$

Wir sind nun in der Position, um die Formeln für die Ableitungen der Abbildungen  $\phi$  herzuleiten. Wir werden dabei Gebrauch von den Landau-Symbolen (vergleiche [39, Abschnitt 4.1.2]) machen und schreiben für zwei nichtnegative Zahlenfolgen  $\{r_k\}, \{t_k\}$

$$r_k = O(t_k) \quad :\iff \quad \limsup \frac{r_k}{t_k} < \infty.$$

In diesem Fall existiert also eine Konstante  $M > 0$  mit  $r_k \leq Mt_k$  für alle  $k \in \mathbb{N}$ . Gilt  $t_k \rightarrow 0$ , so konvergiert die Folge  $r_k$  mindestens so schnell gegen Null wie  $t_k$  (unter Vernachlässigung der Konstanten  $M$ ). Weiter schreiben wir für  $t_k \rightarrow 0$

$$r_k = o(t_k) \quad :\iff \quad \lim_{k \rightarrow \infty} \frac{r_k}{t_k} = 0.$$

Gilt  $t_k \rightarrow 0$ , so konvergiert die Folge  $r_k$  schneller gegen Null als  $t_k$ . In vielen Fällen hängen die Größen  $r$  und  $t$  voneinander ab, ohne dass ein Iterationsindex  $k$  auftritt. Bei der Betrachtung von Grenzwerten  $t \rightarrow 0$  verallgemeinert man dann die O-Notation entsprechend: So bedeutet beispielsweise  $r = O(t)$ , dass es ein  $\varepsilon > 0$  und ein  $M < \infty$  mit

$$t < \varepsilon \implies r < Mt$$

gibt.

Sei jetzt zunächst  $\phi$  die geglättete Fischer-Burmeister-Funktion aus (2.29). Wir müssen zeigen, dass

## 2. Semidefinite Programme und NCP-Funktionen

$$\begin{aligned} \|\phi(X + U, S + V, \tau + \mu) - \phi(X, S, \tau) - \nabla\phi(X, S, \tau)(U, V, \mu)\|_2 \\ = o(\|(U, V, \mu)\|) \end{aligned}$$

für alle gegen  $(0, 0, 0)$  strebenden  $(U, V, \mu) \in \mathcal{S}^{n \times n} \times \mathcal{S}^{n \times n} \times \mathbb{R}$  gilt, wobei  $\nabla\phi(X, S, \tau)$  einen geeigneten linearen Operator bezeichnet, der für die Ableitung von  $\phi$  im Punkt  $(X, S, \tau)$  steht. Hierfür zerlegen wir die Abbildung  $\phi$  in

$$\phi(X, S, \tau) = \phi_1(X, S, \tau) - \phi_2(X, S, \tau)$$

mit

$$\phi_1(X, S, \tau) := X + S \quad \text{und} \quad \phi_2(X, S, \tau) := (X^2 + S^2 + 2\tau^2 I)^{1/2}. \quad (2.31)$$

Dann sieht man sofort ein, dass  $\phi_1$  differenzierbar ist mit

$$\nabla\phi_1(X, S, \tau)(U, V, \mu) = U + V.$$

Die Situation für  $\phi_2$  ist komplizierter. Wir setzen

$$E := (X^2 + S^2 + 2\tau^2 I)^{1/2}$$

und nehmen an, dass  $E$  positiv definit ist, was insbesondere für  $\tau > 0$  erfüllt ist. Definieren wir weiter die Matrix

$$D := ((X + U)^2 + (S + V)^2 + 2(\tau + \mu)^2 I)^{1/2},$$

so zeigt eine einfache Rechnung (die für beliebige  $D, E$  gültige) Gleichung

$$D^2 - E^2 = L_E[D - E] + (D - E)^2, \quad (2.32)$$

wobei  $L_E$  den in Abschnitt 2.3 eingeführten Lyapunov-Operator bezeichnet. Wegen  $E \succ 0$  ist  $L_E$  invertierbar und Anwendung des inversen Lyapunov-Operator  $L_E^{-1}$  auf (2.32) und anschließendes Umordnen der Terme ergibt

$$\begin{aligned} E - D \\ = L_E^{-1}[(D - E)^2 - (D^2 - E^2)] \\ = L_E^{-1}[(E - D)^2 - (XU + UX + SV + VS + 4\tau\mu I + U^2 + V^2 + 2\mu^2 I)]. \end{aligned}$$

Die Linearität von  $L_E^{-1}$  liefert dann

$$\begin{aligned} \phi_2(X + U, S + V, \tau + \mu) - \phi_2(X, S, \tau) - \nabla\phi_2(X, S, \tau)(U, V, \mu) \\ = -\nabla\phi_2(X, S, \tau)(U, V, \mu) - (E - D) \\ = -\nabla\phi_2(X, S, \tau)(U, V, \mu) + L_E^{-1}[XU + UX + SV + VS + 4\tau\mu I] \\ + L_E^{-1}[U^2 + V^2 + 2\mu^2 I] - L_E^{-1}[(E - D)^2]. \end{aligned} \quad (2.33)$$

## 2.4. Eigenschaften von NCP-Funktionen

Offensichtlich gilt dann  $\|L_E^{-1}[U^2 + V^2 + 2\mu^2 I]\|_F = O(\|(U, V, \mu)\|^2)$ . Mit Blick auf Lemma 2.26 gilt auch

$$\begin{aligned} \|(E - D)^2\|_F &\leq \|E - D\|_F^2 \leq \gamma_1 \|E^2 - D^2\|_F^2 \\ &= \gamma_1 \|XU + UX + SV + VS + 4\tau\mu I + U^2 + V^2 + 2\mu^2 I\|_F^2 \\ &= O(\|(U, V, \mu)\|^2) \end{aligned}$$

mit einer von  $U$ ,  $V$  und  $\mu$  unabhängigen Konstanten  $\gamma_1 > 0$ . Dies impliziert

$$\|L_E^{-1}[(E - D)^2]\|_F = O(\|(U, V, \mu)\|^2).$$

Setzen wir daher

$$\nabla\phi_2(X, S, \tau)(U, V, \mu) := L_E^{-1}[XU + UX + SV + VS + 4\tau\mu I],$$

so folgt sofort aus (2.33), dass  $\phi_2$  differenzierbar ist in  $(X, S, \tau)$ . Dies impliziert, dass die Funktion  $\phi$  selbst in diesem Punkt differenzierbar ist. Dies zeigt den Hauptteil der ersten Aussage in dem folgenden Resultat.

**Satz 2.27** *Seien  $X, S \in \mathcal{S}^{n \times n}$  zwei gegebene Matrizen und  $\tau \geq 0$ .*

- (a) *Ist  $\phi$  gegeben durch (2.29) und die Matrix  $X^2 + S^2 + 2\tau^2 I$  positiv definit, so ist  $\phi$  stetig differenzierbar in  $(X, S, \tau)$  mit*

$$\nabla\phi(X, S, \tau)(U, V, \mu) = U + V - L_E^{-1}[XU + UX + SV + VS + 4\tau\mu I]. \quad (2.34)$$

*Dabei ist  $E := (X^2 + S^2 + 2\tau^2 I)^{1/2}$ .*

- (b) *Ist  $\phi$  gegeben durch (2.30) und  $(X - S)^2 + 4\tau^2 I \succ 0$ , dann ist  $\phi$  stetig differenzierbar in  $(X, S, \tau)$  mit*

$$\begin{aligned} \nabla\phi(X, S, \tau)(U, V, \mu) &= U + V \\ &\quad - L_E^{-1}[(X - S)(U - V) + (U - V)(X - S) + 8\tau\mu I]. \end{aligned} \quad (2.35)$$

*Dabei ist  $E := ((X - S)^2 + 4\tau^2 I)^{1/2}$ .*

**Beweis:** (a) Die Differenzierbarkeit der geglätteten Fischer-Burmeister-Funktion folgt aus der vorhergehenden Diskussion. Da  $E = (X^2 + S^2 + 2\tau^2 I)^{1/2} \succ 0$  gilt und die Wurzelabbildung nach Lemma 2.10 stetig ist, folgt die Stetigkeit der Ableitung aus Lemma 2.22.

- (b) Dieser Teil kann völlig analog zu Teil (a) gezeigt werden.  $\square$

Satz 2.27 impliziert insbesondere, dass die Funktionen  $\phi$  stetig differenzierbar sind, falls  $\tau > 0$  gilt. Wir werden dies in den Algorithmen entsprechend ausnutzen.

## 2. Semidefinite Programme und NCP-Funktionen

# **Teil II.**

## **Verfahren**



### 3. Ein Glättungsverfahren

Wir wollen jetzt die theoretischen Resultate des vorhergehenden Teils nutzen, um Algorithmen zur Lösung der Optimalitätsbedingungen (2.6) und somit der zugehörigen semidefiniten Programme zu entwickeln. Dazu betrachten wir die schon in Kapitel 2 eingeführte Abbildung  $\Phi : \mathcal{S}^{n \times n} \times \mathbb{R}^m \times \mathcal{S}^{n \times n} \rightarrow \mathcal{S}^{n \times n} \times \mathbb{R}^m \times \mathcal{S}^{n \times n}$ , welche definiert ist durch

$$\Phi(X, \lambda, S) := \begin{pmatrix} \sum_{i=1}^m \lambda_i A_i + S - C \\ A_i \bullet X - b_i \quad (i = 1, \dots, m) \\ \phi(X, S) \end{pmatrix}. \quad (3.1)$$

Dabei ist  $\phi$  entweder die in (2.10) definierte Fischer-Burmeister-Funktion oder die Minimum-Funktion aus (2.14). Dann besagt Satz 2.13, dass  $(X^*, \lambda^*, S^*)$  genau dann eine Lösung der Optimalitätsbedingungen ist, wenn  $(X^*, \lambda^*, S^*)$  das nichtlineare Gleichungssystem  $\Phi(X, \lambda, S) = 0$  löst.

Zur Lösung des nichtlinearen Gleichungssystems  $\Phi(X, \lambda, S) = 0$  bietet sich ein Newton-Verfahren an. Leider ist die Funktion  $\phi$  (und somit auch  $\Phi$ ) nicht überall differenzierbar (vergleiche Satz 2.27), so dass hierfür nur ein nichtglattes Newton-Verfahren in Frage kommt. Wir werden das lokale Konvergenzverhalten eines solchen nichtglatten Newton-Verfahrens in Kapitel 5 untersuchen.

Um bei der numerischen Implementation die Berechnung der Subgradienten zu vermeiden, werden wir den nichtdifferenzierbaren Anteil  $\phi$  von  $\Phi$  durch die glatte Approximation  $\phi_\tau$  ersetzen. Dies führt auf die so genannten Glättungsverfahren, welche in den Kapiteln 3 und 4 behandelt werden. Der Unterschied dieser beiden Glättungsverfahren besteht in der Globalisierung. Während in diesem Kapitel ein Verfahren behandelt wird, welches zur Erreichung von globaler Konvergenz eine geeignete Schrittweitenstrategie entlang einer Suchrichtung einführt, wird in Kapitel 4 ein Trust-Region-Ansatz zur Globalisierung verwendet.

#### 3.1. Algorithmus

Zur Herleitung eines Glättungsverfahrens betrachten wir das nichtlineare Gleichungssystem  $\Phi(X, \lambda, S) = 0$  mit der durch (3.1) gegebenen Funktion  $\Phi$ . Es handelt sich hierbei um ein nichtglattes Problem, da die Funktion  $\phi$  und somit die Funktion  $\Phi$  nicht überall differenzierbar ist. Die Idee ist daher, die in

### 3. Ein Glättungsverfahren

(3.1) auftretende Funktion  $\phi$  durch die geglättete Approximation (2.16) bzw. (2.17) zu ersetzen und für gegebenen Parameter  $\tau > 0$  das Gleichungssystem  $\Phi_\tau(X, \lambda, S) = 0$  mit der differenzierbaren Funktion

$$\Phi_\tau(X, \lambda, S) := \begin{pmatrix} \sum_{i=1}^m \lambda_i A_i + S - C \\ A_i \bullet X - b_i \quad (i = 1, \dots, m) \\ \phi_\tau(X, S) \end{pmatrix}$$

zu betrachten. Hier bezeichnet natürlich  $\phi_\tau$  entweder die geglättete Fischer-Burmeister-Funktion (2.16) oder die geglättete Minimum-Funktion (2.17). In einem entsprechenden Algorithmus muss dann  $\tau \rightarrow 0$  durch eine geeignete Aufdatierung von  $\tau$  garantiert werden. Dieser Ansatz wurde von Chen und Tseng [15] untersucht, auch wenn sie nicht die Äquivalenz des nichtlinearen Gleichungssystems  $\Phi_\tau(X, \lambda, S) = 0$  mit den Zentralen-Pfad-Bedingungen erkannt haben (vergleiche Satz 2.16).

Wir wollen in dieser Arbeit einen etwas anderen Ansatz verfolgen und eine von Jiang [40] in Zusammenhang mit nichtlinearen Komplementaritätsproblemen gemachte Idee aufgreifen und  $\tau$  als eine unabhängige Variable betrachten. Wir definieren daher die Abbildung  $\Theta : \mathcal{S}^{n \times n} \times \mathbb{R}^m \times \mathcal{S}^{n \times n} \times \mathbb{R} \rightarrow \mathcal{S}^{n \times n} \times \mathbb{R}^m \times \mathcal{S}^{n \times n} \times \mathbb{R}$  durch

$$\Theta(X, \lambda, S, \tau) := \begin{pmatrix} \sum_{i=1}^m \lambda_i A_i + S - C \\ A_i \bullet X - b_i \quad (i = 1, \dots, m) \\ \phi(X, S, \tau) \\ \tau \end{pmatrix}, \quad (3.2)$$

wobei  $\phi$  eine der durch (2.29) oder (2.30) definierten Funktionen bezeichnet. Abgesehen von der Tatsache, dass  $\tau$  jetzt als eine Variable und nicht als Parameter betrachtet wird, unterscheidet sich die Funktion  $\Theta$  von  $\Phi_\tau$  durch eine zusätzlich eingeführte Komponente. Das System

$$\Theta(X, \lambda, S, \tau) = 0 \quad (3.3)$$

impliziert somit  $\tau = 0$ , so dass (3.3) äquivalent zu den Optimalitätsbedingungen (2.6) selbst und nicht zu den Zentralen-Pfad-Bedingungen (2.7) ist. Dies ist ein Vorteil gegenüber der Formulierung  $\Phi_\tau(X, \lambda, S) = 0$ , da wir ja die Optimalitätsbedingungen und nicht die Zentralen-Pfad-Bedingungen lösen wollen.

Aus Satz 2.27 folgt weiterhin, dass  $\Theta$  in allen Punkten  $(X, \lambda, S, \tau)$  mit  $\tau > 0$  differenzierbar ist. Wir werden die Aufdatierungsregeln in dem Algorithmus daher so konstruieren, dass die Positivität der Glättungsvariablen  $\tau$  zu jeder Zeit garantiert ist. Wie numerische Experimente von Engelke und Kanzow [18, 19, 20] in Zusammenhang mit linearen Programmen zeigen, hat diese



Umformulierung Vorteile gegenüber der Betrachtung des Gleichungssystems  $\Phi_\tau(X, \lambda, S) = 0$ . Es hat im Kontext der linearen Komplementaritätsprobleme auch theoretische Vorteile (vergleiche Burke und Xu [11, 10]). Es ist allerdings nicht klar, ob diese Resultate auf semidefinite Programme übertragen werden können.

Die Hauptidee des hier zu behandelnden Algorithmus ist, das Gleichungssystem (3.3) mit dem Newton-Verfahren zu lösen. Globale Konvergenz wird dadurch erreicht, dass sämtliche Iterierten in einer gewissen Umgebung des zentralen Pfades bleiben sollen. Diese Umgebung ist definiert durch

$$\mathcal{N}(\beta) = \left\{ (X, \lambda, S, \tau) \mid \begin{aligned} &A_i \bullet X = b_i, \quad i = 1, \dots, m, \\ &\sum_{i=1}^m \lambda_i A_i + S = C, \quad \|\phi(X, S, \tau)\|_F \leq \beta\tau \end{aligned} \right\},$$

wobei  $\beta$  eine gegebene, positive Zahl ist. Lokal schnelle Konvergenz wird durch einen geeigneten Prädiktor-Schritt garantiert.

Um die Formulierung des Algorithmus zu vereinfachen, führen wir noch die Abkürzung  $W^k := (X^k, \lambda^k, S^k)$  ein, wobei  $k$  den Iterationsindex bezeichnet.

Wir geben jetzt unser Glättungsverfahren zur Lösung der Optimalitätsbedingungen formal an.

### Algorithmus 3.1

(S.0) (Initialisierung)

Wähle  $W^0 = (X^0, \lambda^0, S^0) \in \mathcal{S}^{n \times n} \times \mathbb{R}^m \times \mathcal{S}^{n \times n}$  mit

$$\sum_{i=1}^m \lambda_i^0 A_i + S^0 = C \quad \text{und} \quad A_i \bullet X^0 = b_i \quad (i = 1, \dots, m).$$

Wähle  $\tau_0 > 0$ ,  $\beta > 0$  mit  $\|\phi(X^0, S^0, \tau_0)\|_F \leq \beta\tau_0$  sowie  $\hat{\sigma}, \alpha_1, \alpha_2 \in (0, 1)$  und setze  $k := 0$ .

(S.1) (Prädiktor-Schritt)

Sei  $(\Delta W^k, \Delta \tau_k) = (\Delta X^k, \Delta \lambda^k, \Delta S^k, \Delta \tau_k) \in \mathcal{S}^{n \times n} \times \mathbb{R}^m \times \mathcal{S}^{n \times n} \times \mathbb{R}$  eine Lösung des linearen Systems

$$\nabla \Theta(W^k, \tau_k)(\Delta W, \Delta \tau) = -\Theta(W^k, \tau_k). \quad (3.4)$$

Ist  $\|\phi(X^k + \Delta X^k, S^k + \Delta S^k, 0)\|_F = 0$ : STOPP.

Ist ansonsten  $\|\phi(X^k + \Delta X^k, S^k + \Delta S^k, \tau_k)\|_F > \beta\tau_k$ , so setze  $\hat{W}^k := W^k$ ,  $\hat{\tau}_k := \tau_k$  und  $\eta_k := 1$ , andernfalls setze  $\eta_k = \alpha_1^s$ , wobei  $s$  diejenige natürliche Zahl mit

$$\|\phi(X^k + \Delta X^k, S^k + \Delta S^k, \alpha_1^r \tau_k)\|_F \leq \beta\tau_k \alpha_1^r, \quad \forall r = 0, 1, 2, \dots, s,$$

### 3. Ein Glättungsverfahren

$$\left\| \phi(X^k + \Delta X^k, S^k + \Delta S^k, \alpha_1^{s+1} \tau_k) \right\|_F > \beta \tau_k \alpha_1^{s+1},$$

ist und setze

$$\hat{\tau}_k := \eta_k \tau_k \quad \text{und} \quad \hat{W}^k := \begin{cases} W^k, & \text{falls } s = 0, \\ W^k + \Delta W^k, & \text{sonst.} \end{cases}$$

(S.2) (Korrektor-Schritt)

Sei  $(\Delta \hat{W}^k, \Delta \hat{\tau}_k) = (\Delta \hat{X}^k, \Delta \hat{\lambda}^k, \Delta \hat{S}^k, \Delta \hat{\tau}_k)$  eine Lösung des linearen Systems

$$\nabla \Theta(\hat{W}^k, \hat{\tau}_k)(\Delta \hat{W}, \Delta \hat{\tau}) = -\Theta(\hat{W}^k, \hat{\tau}_k) + \begin{pmatrix} 0 \\ (1 - \hat{\sigma}) \hat{\tau}_k \end{pmatrix} \quad (3.5)$$

und  $\hat{\eta}_k$  die GröÙte der Zahlen  $1, \alpha_2, \alpha_2^2, \dots$  mit

$$\left\| \phi(\hat{X}^k + \hat{\eta}_k \Delta \hat{X}^k, \hat{S}^k + \hat{\eta}_k \Delta \hat{S}^k, \hat{\tau}_k + \hat{\eta}_k \Delta \hat{\tau}_k) \right\|_F \leq (1 - \hat{\sigma} \hat{\eta}_k) \beta \hat{\tau}_k. \quad (3.6)$$

Setze  $W^{k+1} := \hat{W}^k + \hat{\eta}_k \Delta \hat{W}^k$ ,  $\tau_{k+1} := (1 - \hat{\sigma} \hat{\eta}_k) \hat{\tau}_k$ ,  $k \leftarrow k + 1$ , und gehe zu (S.1).

Wir machen zunächst ein paar allgemeine Bemerkungen: Da der Startvektor in Schritt (S.0) primal und dual zulässig gewählt wird, folgt aus den Newton-Gleichungen (3.4) und (3.5) sofort, dass sämtliche durch Algorithmus 3.1 erzeugte Iterierte  $(X^k, \lambda^k, S^k)$  und  $(\hat{X}^k, \hat{\lambda}^k, \hat{S}^k)$  zulässig in dem Sinne sind, dass

$$\sum_{i=1}^m \lambda_i^k A_i + S^k = C, \quad A_i \bullet X^k = b_i \quad (i = 1, \dots, m) \quad (3.7)$$

und

$$\sum_{i=1}^m \hat{\lambda}_i^k A_i + \hat{S}^k = C, \quad A_i \bullet \hat{X}^k = b_i \quad (i = 1, \dots, m)$$

für alle  $k \in \mathbb{N}$  gilt. Weiter werden wir im Folgenden zeigen, dass die Matrizen  $X^k$ ,  $S^k$  und  $\hat{X}^k$ ,  $\hat{S}^k$  stets symmetrisch sind. Dies ist ein Unterschied zu den Inneren-Punkte-Methoden, wo dies durch eine gesonderte Symmetrisierung garantiert werden muss. Im Gegensatz zu Inneren-Punkte-Methoden müssen die durch Algorithmus 3.1 erzeugten Iterierten  $X^k$ ,  $S^k$  und  $\hat{X}^k$ ,  $\hat{S}^k$  jedoch nicht positiv definit oder positiv semidefinit sein.

Die Abbruchbedingung in Schritt (S.1) ist durch die Lemmata 2.11 und 2.12 gerechtfertigt: Zusammen mit der Zulässigkeit der Iterierten implizieren diese Resultate, dass

$$\|\phi(X^k + \Delta X^k, S^k + \Delta S^k, 0)\|_F = 0 \iff W^k + \Delta W^k \text{ löst (2.6)}$$

gilt. Für die folgenden theoretischen Untersuchungen des Konvergenzverhaltens werden wir immer annehmen, dass das Abbruchkriterium niemals erfüllt ist und somit Algorithmus 3.1 eine unendliche Folge von Iterierten erzeugt.

Weiter wollen wir bemerken, dass die Aufdatierungstechnik für  $\tau_{k+1}$  in Schritt (S.2) äquivalent ist zu der Formulierung  $\tau_{k+1} = \hat{\tau}_k + \hat{\eta}_k \Delta \hat{\tau}_k$ . Dies folgt sofort aus der letzten Blockzeile  $\Delta \hat{\tau}_k = -\hat{\sigma} \hat{\tau}_k$  des linearen Systems (3.5) im Korrektor-Schritt.

In Algorithmus 3.1 muss sowohl im Prädiktor- als auch im Korrektor-Schritt ein lineares System gelöst werden. Die dabei auftretenden Operatoren  $\nabla \Theta(W, \tau)$  können verschieden sein, so dass dieses Vorgehen mehr Berechnungsaufwand erfordert, als dies bei Inneren-Punkte-Methoden üblicherweise der Fall ist. Es wird sich jedoch zeigen, dass alle Resultate weiterhin ihre Gültigkeit behalten, wenn wir die folgende Modifikation von Algorithmus 3.1 betrachten: Wird der Prädiktor-Schritt mit  $\eta_k < 1$  akzeptiert, dann überspringen wir den Korrektor-Schritt, d.h. wir setzen  $W^{k+1} := W^k + \Delta W^k$ ,  $\tau_{k+1} := \eta_k \tau_k$ ,  $k \leftarrow k + 1$  und gehen zu Schritt (S.1). Der so modifizierte Algorithmus muss in jeder Iteration entweder ein lineares System im Prädiktor-Schritt lösen, oder aber zwei lineare Systeme mit dem gleichen Operator. Diese Modifikation wurde auch implementiert, um die numerischen Resultate aus Abschnitt 3.4 zu erhalten. Die Struktur der linearen Systeme und ihre numerische Lösbarkeit werden wir zuvor in Abschnitt 3.3 genauer untersuchen.

Den Rest dieses Abschnittes wollen wir verwenden, um zu zeigen, dass Algorithmus 3.1 wohldefiniert ist. Dafür müssen wir zunächst einmal zeigen, dass die linearen Systeme (3.4) und (3.5) eine eindeutig bestimmte Lösung haben. Wir werden dabei Lemma 2.19 (c), (d) speziell mit  $A = E - S$ ,  $B = E - X$  und  $E = (X^2 + S^2 + 2\tau^2 I)^{1/2}$  im Falle der Fischer-Burmeister-Funktion bzw.  $A = E - (X - S)$ ,  $B = E + (X - S)$  und  $E = ((X - S)^2 + 4\tau^2 I)^{1/2}$  im Falle der Minimum-Funktion anwenden. Daher müssen wir zeigen, dass in diesen Fällen die Ungleichung  $AB + BA \succeq 0$  erfüllt ist.

**Lemma 3.2** *Seien  $X, S \in \mathcal{S}^{n \times n}$  und  $\tau \in \mathbb{R}$  gegeben.*

- (a) *Mit  $E := (X^2 + S^2 + 2\tau^2 I)^{1/2}$ ,  $A := E - S$  und  $B := E - X$  gilt  $BA + AB \succeq 0$ .*
- (b) *Mit  $E := ((X - S)^2 + 4\tau^2 I)^{1/2}$ ,  $A := E - (X - S)$  und  $B := E + (X - S)$  gilt  $BA + AB \succeq 0$ .*

### 3. Ein Glättungsverfahren

**Beweis:** (a) Es gilt

$$\begin{aligned}
BA + AB &= (E - X)(E - S) + (E - S)(E - X) \\
&= E^2 - XE - ES + XS + E^2 - SE - EX + SX \\
&= X^2 + S^2 + 2\tau^2 I - E(X + S) - (X + S)E + E^2 + XS + SX \\
&= (X + S)^2 - E(X + S) - (X + S)E + E^2 + 2\tau^2 I \\
&= (X + S - E)^2 + 2\tau^2 I \succeq 0.
\end{aligned}$$

(b) Analog zu (a) ergibt sich

$$\begin{aligned}
BA + AB &= (E - (X - S))(E + (X - S)) + (E + (X - S))(E - (X - S)) \\
&= E^2 - (X - S)E + E(X - S) - (X - S)^2 \\
&\quad + E^2 + (X - S)E - E(X - S) - (X - S)^2 \\
&= 2E^2 - 2(X - S)^2 \\
&= 8\tau^2 I \succeq 0.
\end{aligned}$$

Damit ist alles gezeigt.  $\square$

Weiter benötigen wir noch das folgende Hilfsresultat, dessen Beweis analog zu dem von Lemma 6.1 (c) in Tseng [67] erfolgt.

**Lemma 3.3** Für alle  $A \succ 0$ ,  $B \in \mathcal{S}^{n \times n}$  mit  $A^2 - B^2 \succ 0$  gilt  $A - B \succ 0$ .

**Beweis:** Seien  $A \succ 0$ ,  $B \in \mathcal{S}^{n \times n}$  mit  $A^2 - B^2 \succ 0$  gegeben. Wir zeigen nur  $A - |B| \succ 0$  (mit  $|B| := (B^2)^{1/2}$ ), da dann mit  $|B| - B \succeq 0$  die Behauptung folgt.

Angenommen, es gilt nicht  $A - |B| \succ 0$ . Dann existiert ein  $v \in \mathbb{R}^n$ ,  $v \neq 0$  und ein  $\lambda \leq 0$  mit  $(A - |B|)v = \lambda v$ . Daraus folgt mit  $A + |B| \succ 0$

$$\begin{aligned}
0 &< 2v^T(A^2 - B^2)v \\
&= 2v^T(A^2 - |B|^2)v \\
&= v^T((A + |B|)(A - |B|) + (A - |B|)(A + |B|))v \\
&= v^T(A + |B|)\lambda v + \lambda v^T(A + |B|)v \\
&= 2\lambda v^T(A + |B|)v \\
&\leq 0,
\end{aligned}$$

ein Widerspruch.  $\square$

Um zu zeigen, dass die linearen Systeme (3.4) und (3.5) eine eindeutig bestimmte Lösung haben, zeigen wir, dass die lineare Abbildung  $\nabla\Theta(X, \lambda, S, \tau)$

invertierbar ist. Dafür benötigen wir die folgende Voraussetzung, welche zumindest theoretisch keine Einschränkung darstellt.

**Voraussetzung 3.4** Die Matrizen  $A_i$ ,  $i = 1, \dots, m$  sind linear unabhängig.

Mit Lemma 2.19 und Voraussetzung 3.4 sind wir nun in der Lage, zu zeigen, dass  $\nabla\Theta(X, \lambda, S, \tau)$  eine Bijektion darstellt. Dies impliziert dann, dass die Suchrichtung  $(\Delta X^k, \Delta \lambda^k, \Delta S^k, \Delta \tau_k)$  im Prädiktor-Schritt sowie die Suchrichtung  $(\Delta \hat{X}^k, \Delta \hat{\lambda}^k, \Delta \hat{S}^k, \Delta \hat{\tau}_k)$  im Korrektor-Schritt wohldefiniert sind.

**Lemma 3.5** Sei Voraussetzung 3.4 erfüllt. Dann ist die lineare Abbildung  $\nabla\Theta(X, \lambda, S, \tau)$  mit  $\phi$  gegeben durch (2.29) oder (2.30) für alle  $(X, \lambda, S, \tau) \in \mathcal{S}^{n \times n} \times \mathbb{R}^m \times \mathcal{S}^{n \times n} \times \mathbb{R}_{++}$  bijektiv.

**Beweis:** Wir betrachten nur den Fall der geglätteten Fischer-Burmeister-Funktion (2.29). Der Fall der geglätteten Minimum-Funktion folgt analog.

Sei  $(X, \lambda, S, \tau) \in \mathcal{S}^{n \times n} \times \mathbb{R}^m \times \mathcal{S}^{n \times n} \times \mathbb{R}_{++}$  fest gegeben. Da  $\nabla\Theta(X, \lambda, S, \tau)$  eine lineare Abbildung des endlich-dimensionalen Vektorraums  $\mathcal{S}^{n \times n} \times \mathbb{R}^m \times \mathcal{S}^{n \times n} \times \mathbb{R}$  in sich selbst darstellt, müssen wir nur zeigen, dass die Abbildung injektiv ist. Dafür reicht es zu zeigen, dass das System

$$\nabla\Theta(X, \lambda, S, \tau)(\Delta X, \Delta \lambda, \Delta S, \Delta \tau) = (0, 0, 0, 0)$$

oder äquivalent

$$\sum_{i=1}^m \Delta \lambda_i A_i + \Delta S = 0, \quad (3.8)$$

$$A_i \bullet \Delta X = 0, \quad i = 1, \dots, m, \quad (3.9)$$

$$\nabla\phi(X, S, \tau)(\Delta X, \Delta S, \Delta \tau) = 0, \quad (3.10)$$

$$\Delta \tau = 0 \quad (3.11)$$

nur den Vektor  $(\Delta X, \Delta \lambda, \Delta S, \Delta \tau) = (0, 0, 0, 0)$  als Lösung besitzt.

Aus (3.11) folgt sofort  $\Delta \tau = 0$ . Setzen wir jetzt  $E := (X^2 + S^2 + 2\tau^2 I)^{1/2}$ , so ist  $E \succ 0$  wegen  $\tau > 0$  und aus (3.10) und Satz 2.27 folgt, dass

$$\Delta X + \Delta S - L_E^{-1} [X \Delta X + \Delta X X + S \Delta S + \Delta S S] = 0$$

gilt. Anwendung des Lyapunov-Operators  $L_E$  auf beide Seiten und Umordnung der Terme liefert

$$L_{E-X}[\Delta X] + L_{E-S}[\Delta S] = 0.$$

### 3. Ein Glättungsverfahren

Wegen  $E \succ 0$  und  $E^2 - S^2 = X^2 + 2\tau^2 I \succ 0$  (beachte  $\tau > 0$ ) gilt  $E - S \succ 0$  nach Lemma 3.3. Daher existiert der inverse Lyapunov-Operator  $L_{E-S}^{-1}$  und wir erhalten

$$(L_{E-S}^{-1} \circ L_{E-X})[\Delta X] + \Delta S = 0. \quad (3.12)$$

Unter Benutzung von (3.8) und (3.9) liefert skalare Multiplikation mit  $\Delta X$

$$\begin{aligned} 0 &= (L_{E-S}^{-1} \circ L_{E-X})[\Delta X] \bullet \Delta X - \sum_{i=1}^m \Delta \lambda_i \underbrace{A_i \bullet \Delta X}_{=0} \\ &= (L_{E-S}^{-1} \circ L_{E-X})[\Delta X] \bullet \Delta X. \end{aligned} \quad (3.13)$$

Mit  $E - X \succ 0$  (dies folgt wieder aus Lemma 3.3) und  $E - S \succ 0$  folgt aus Lemma 2.19 (d) und Lemma 3.2 (a), dass der Operator  $L_{E-S}^{-1} \circ L_{E-X}$  gleichmäßig monoton ist. Daher folgt aus Gleichung (3.13) sofort  $\Delta X = 0$ . Dies impliziert  $\Delta S = 0$  wegen (3.12). Die angenommene lineare Unabhängigkeit der Matrizen  $A_i$ ,  $i = 1, \dots, m$  und (3.8) zeigt dann  $\Delta \lambda = 0$ .  $\square$

Mit den vorherigen Resultaten können wir jetzt zeigen, dass Algorithmus 3.1 unter der Voraussetzung 3.4 wohldefiniert ist.

**Satz 3.6** *Algorithmus 3.1 ist wohldefiniert unter der Voraussetzung 3.4. Weiter liegen alle Iterierten  $W^k = (X^k, \lambda^k, S^k, \tau_k)$  sowie  $\hat{W}^k = (\hat{X}^k, \hat{\lambda}^k, \hat{S}^k, \hat{\tau}_k)$  in der Umgebung  $\mathcal{N}(\beta)$  des zentralen Pfades.*

**Beweis:** Mit Blick auf Lemma 3.5 bleibt nur noch zu zeigen, dass die beiden Schrittweitenstrategien in (S.1) und (S.2) wohldefiniert sind.

Betrachten wir zunächst die Schrittweitenstrategie in Schritt (S.1). Da wir annehmen, dass Algorithmus 3.1 eine unendliche Folge von Iterierten erzeugt, ist das Abbruchkriterium in Schritt (S.1) nie erfüllt. Daher gilt

$$\|\phi(X^k + \Delta X^k, S^k + \Delta S^k, 0)\|_F > 0$$

für alle  $k \in \mathbb{N}$ . Da die Abbildung  $\tau \mapsto \|\phi(X, S, \tau)\|_F$  stetig ist, impliziert dies, dass die Schrittweite in Schritt (S.1) in endlich vielen Schritten berechnet wird. Daher sind  $\hat{W}^k$  und  $\hat{\tau}_k$  wohldefiniert und es gilt  $(\hat{W}^k, \hat{\tau}_k) \in \mathcal{N}(\beta)$ .

Zur Schrittweitenstrategie in Schritt (S.2): Definieren wir  $\psi(X, S, \tau) := \|\phi(X, S, \tau)\|_F$ , so ergibt sich

$$\begin{aligned} &\psi'(\hat{X}^k, \hat{S}^k, \hat{\tau}_k)(\Delta \hat{X}^k, \Delta \hat{S}^k, \Delta \hat{\tau}_k) \\ &= \frac{\phi(\hat{X}^k, \hat{S}^k, \hat{\tau}_k) \bullet \nabla \phi(\hat{X}^k, \hat{S}^k, \hat{\tau}_k)(\Delta \hat{X}^k, \Delta \hat{S}^k, \Delta \hat{\tau}_k)}{\|\phi(\hat{X}^k, \hat{S}^k, \hat{\tau}_k)\|_F} \\ &\stackrel{(3.5)}{=} -\|\phi(\hat{X}^k, \hat{S}^k, \hat{\tau}_k)\|_F. \end{aligned} \quad (3.14)$$

### 3.2. Globales und lokales Konvergenzverhalten

Angenommen, die Schrittweite kann nicht in einer endlichen Anzahl von Schritten berechnet werden. Dann gilt

$$\|\phi(\hat{X}^k + \alpha_2^t \Delta \hat{X}^k, \hat{S}^k + \alpha_2^t \Delta \hat{S}^k, \hat{\tau}_k + \alpha_2^t \Delta \hat{\tau}_k)\|_F > (1 - \hat{\sigma} \alpha_2^t) \beta \hat{\tau}_k$$

für alle  $t \in \mathbb{N}$ . Wegen  $\beta \hat{\tau}_k > \|\phi(\hat{X}^k, \hat{S}^k, \hat{\tau}_k)\|_F$  impliziert dies

$$\|\phi(\hat{X}^k + \alpha_2^t \Delta \hat{X}^k, \hat{S}^k + \alpha_2^t \Delta \hat{S}^k, \hat{\tau}_k + \alpha_2^t \Delta \hat{\tau}_k)\|_F > (1 - \hat{\sigma} \alpha_2^t) \|\phi(\hat{X}^k, \hat{S}^k, \hat{\tau}_k)\|_F$$

oder äquivalent

$$\frac{\|\phi(\hat{X}^k + \alpha_2^t \Delta \hat{X}^k, \hat{S}^k + \alpha_2^t \Delta \hat{S}^k, \hat{\tau}_k + \alpha_2^t \Delta \hat{\tau}_k)\|_F - \|\phi(\hat{X}^k, \hat{S}^k, \hat{\tau}_k)\|_F}{\alpha_2^t} > -\hat{\sigma} \|\phi(\hat{X}^k, \hat{S}^k, \hat{\tau}_k)\|_F$$

für alle  $t \in \mathbb{N}$ . Grenzübergang  $t \rightarrow \infty$  liefert unter Benutzung von (3.14)

$$-\|\phi(\hat{X}^k, \hat{S}^k, \hat{\tau}_k)\|_F = \psi'(\hat{X}^k, \hat{S}^k, \hat{\tau}_k)(\Delta \hat{X}^k, \Delta \hat{S}^k, \Delta \hat{\tau}_k) \geq -\hat{\sigma} \|\phi(\hat{X}^k, \hat{S}^k, \hat{\tau}_k)\|_F.$$

Wegen  $\hat{\sigma} \in (0, 1)$  impliziert dies  $\|\phi(\hat{X}^k, \hat{S}^k, \hat{\tau}_k)\|_F = 0$ . Daher gilt

$$\|\phi(\hat{X}^k + \alpha_2^t \Delta \hat{X}^k, \hat{S}^k + \alpha_2^t \Delta \hat{S}^k, \hat{\tau}_k + \alpha_2^t \Delta \hat{\tau}_k)\|_F \rightarrow 0 \quad \text{für } t \rightarrow \infty,$$

wogegen

$$(1 - \hat{\sigma} \alpha_2^t) \beta \hat{\tau}_k \rightarrow \beta \hat{\tau}_k \quad \text{für } t \rightarrow \infty$$

gilt. Dies ist ein Widerspruch zur Annahme, dass die Berechnung von  $\hat{\eta}_k$  nicht nach endlich vielen Schritten abbricht. Daher ist die Schrittweitenstrategie in Schritt (S.2) wohldefiniert.

Die Aussagen bezüglich der Umgebung  $\mathcal{N}(\beta)$  folgen sofort aus den Aufdatierungsregeln in Algorithmus 3.1.  $\square$

### 3.2. Globales und lokales Konvergenzverhalten

In diesem Abschnitt wollen wir sowohl das globale wie auch das lokale Konvergenzverhalten von Algorithmus 3.1 untersuchen. Dazu geben wir als erstes das globale Konvergenzverhalten an.

**Satz 3.7** *Wenn die durch Algorithmus 3.1 erzeugte Folge  $\{W^k\} = \{(X^k, \lambda^k, S^k)\}$  einen Häufungspunkt besitzt, dann konvergiert die Folge  $\{\tau_k\}$  gegen Null.*

### 3. Ein Glättungsverfahren

**Beweis:** Da die Folge  $\{\tau_k\}$  monoton fällt und durch Null nach unten beschränkt ist, konvergiert die Folge gegen ein  $\tau_* \geq 0$ . Ist  $\tau_* = 0$ , so sind wir fertig.

Sei daher angenommen, dass  $\tau_* > 0$  gilt. Dann folgt aus den Aufdatierungsregeln in Schritt (S.1) von Algorithmus 3.1 sofort

$$\hat{W}^k = W^k, \quad \hat{\tau}_k = \tau_k \quad \text{und} \quad \eta_k = 1 \quad (3.15)$$

für alle hinreichend großen  $k \in \mathbb{N}$ . Durch Übergang auf eine Teilfolge können wir daher ohne Beschränkung der Allgemeinheit annehmen, dass (3.15) für alle  $k \in \mathbb{N}$  gilt. Dann folgt aus den Aufdatierungsregeln in Schritt (S.2)

$$\tau_k = \tau_0 \prod_{j=0}^{k-1} (1 - \hat{\sigma}\hat{\eta}_j).$$

Wegen  $\tau_k \rightarrow \tau_* > 0$  nach Voraussetzung folgt daraus  $\lim_{k \rightarrow \infty} \hat{\eta}_k = 0$ . Daher erfüllt die Schrittweite  $\hat{\rho}_k := \hat{\eta}_k/\alpha_2$  für alle  $k \in \mathbb{N}$  hinreichend groß nicht die Ungleichung (3.6), d.h. es gilt

$$\|\phi(\hat{X}^k + \hat{\rho}_k \Delta \hat{X}^k, \hat{S}^k + \hat{\rho}_k \Delta \hat{S}^k, \hat{\tau}_k + \hat{\rho}_k \Delta \hat{\tau}_k)\|_F > (1 - \hat{\sigma}\hat{\rho}_k)\beta\hat{\tau}_k. \quad (3.16)$$

für alle diese  $k \in \mathbb{N}$ .

Sei nun  $W^* = (X^*, \lambda^*, S^*)$  ein Häufungspunkt der Folge  $\{W^k\}$  und  $\{W^k\}_K$  eine gegen  $W^*$  konvergente Teilfolge. Da  $\tau_* > 0$  gilt, folgt aus (3.15) und Lemma 3.5, dass die zugehörige Folge  $\{(\Delta \hat{W}^k, \Delta \hat{\tau}_k)\}_K$  gegen ein  $(\Delta W^*, \Delta \tau_*) = (\Delta X^*, \Delta \lambda^*, \Delta S^*, \Delta \tau_*)$  konvergiert. Dabei ist  $(\Delta W^*, \Delta \tau_*)$  eine Lösung des linearen Systems

$$\nabla \Theta(W^*, \tau_*)(\Delta \hat{W}, \Delta \hat{\tau}) = -\Theta(W^*, \tau_*) + \begin{pmatrix} 0 \\ (1 - \hat{\sigma})\tau_* \end{pmatrix}, \quad (3.17)$$

vergleiche (3.5). Insbesondere ist die Folge  $\{(\Delta \hat{W}^k, \Delta \hat{\tau}_k)\}_K$  beschränkt. Unter Benutzung von  $\{\hat{\rho}_k\}_K \rightarrow 0$ , (3.15), (3.16) sowie der Stetigkeit der Funktion  $\phi$  ergibt Grenzübergang  $k \rightarrow \infty$  auf der Teilfolge  $K$

$$\|\phi(X^*, S^*, \tau_*)\|_F \geq \beta\tau_*. \quad (3.18)$$

Auf der anderen Seite erhalten wir aus (3.15), (3.16), der Definition der Umgebung  $\mathcal{N}(\beta)$  und Satz 3.6

$$\begin{aligned} \|\phi(X^k + \hat{\rho}_k \Delta X^k, S^k + \hat{\rho}_k \Delta S^k, \tau_k + \hat{\rho}_k \Delta \tau_k)\|_F &> (1 - \hat{\sigma}\hat{\rho}_k)\beta\hat{\tau}_k \\ &= (1 - \hat{\sigma}\hat{\rho}_k)\beta\tau_k \end{aligned}$$



### 3.2. Globales und lokales Konvergenzverhalten

$$\geq (1 - \hat{\sigma}\hat{\rho}_k) \|\phi(X^k, S^k, \tau_k)\|_F$$

für alle  $k \in \mathbb{N}$  hinreichend groß. Mit Hinblick auf (3.15) impliziert dies

$$\frac{\|\phi(X^k + \hat{\rho}_k \Delta X^k, S^k + \hat{\rho}_k \Delta S^k, \tau_k + \hat{\rho}_k \Delta \tau_k)\|_F - \|\phi(X^k, S^k, \tau_k)\|_F}{\hat{\rho}_k} > -\hat{\sigma} \|\phi(X^k, S^k, \tau_k)\|_F.$$

Setzen wir  $\psi(X, S, \tau) := \|\phi(X, S, \tau)\|_F$  und benutzen wir (3.14), so gilt für  $k \rightarrow \infty$  auf der Teilfolge  $K$

$$-\|\phi(X^*, S^*, \tau_*)\|_F \geq -\hat{\sigma} \|\phi(X^*, S^*, \tau_*)\|_F,$$

da die Funktion  $\psi$  wegen  $\tau_* > 0$  stetig differenzierbar im Punkt  $(X^*, S^*, \tau_*)$  ist. Aus  $\hat{\sigma} \in (0, 1)$  folgt dann  $\|\phi(X^*, S^*, \tau_*)\|_F = 0$ , im Widerspruch zu (3.18).  $\square$

**Satz 3.8** *Jeder Häufungspunkt einer durch Algorithmus 3.1 erzeugten Folge  $\{W^k\} = \{(X^k, \lambda^k, S^k)\}$  ist eine Lösung der Optimalitätsbedingungen (2.6).*

**Beweis:** Sei  $W^* = (X^*, \lambda^*, S^*)$  ein Häufungspunkt der durch Algorithmus 3.1 erzeugten Folge  $\{W^k\} = \{(X^k, \lambda^k, S^k)\}$  und  $\{W^k\}_{k \in K}$  eine gegen  $W^*$  konvergente Teilfolge. Nach Satz 3.7 gilt  $\lim_{k \rightarrow \infty} \tau_k = 0$ . Da alle Iterierten nach Satz 3.6 in der Umgebung  $\mathcal{N}(\beta)$  liegen, folgt

$$\|\phi(X^*, S^*, 0)\|_F = \lim_{k \in K} \|\phi(X^k, S^k, \tau_k)\|_F \leq \lim_{k \in K} \beta \tau_k = 0.$$

Da alle Iterierten wegen (3.7) zulässig sind, impliziert dies und Lemma 2.11 bzw. Lemma 2.12, dass  $W^* = (X^*, \lambda^*, S^*)$  eine Lösung der Optimalitätsbedingungen (2.6) ist.  $\square$

Als Nächstes wollen wir die lokalen Konvergenzeigenschaften von Algorithmus 3.1 untersuchen. Unser Ziel ist es, zu zeigen, dass die Folge  $\{\tau_k\}$  unter geeigneten Voraussetzungen superlinear gegen Null konvergiert. Da dies von einigen Eigenschaften im Prädiktor-Schritt abhängt, wollen wir zunächst die folgende Annahme machen.

**Voraussetzung 3.9** *Die durch Algorithmus 3.1 erzeugte Folge  $\{\tau_k\}$  konvergiert gegen Null und es gilt*

$$\|(\Delta W^k, \Delta \tau_k)\| = O(\tau_k), \quad (3.19)$$

wobei  $(\Delta W^k, \Delta \tau_k)$  die in (3.4) berechnete Suchrichtung bezeichnet.

### 3. Ein Glättungsverfahren

Zur Rechtfertigung von Voraussetzung 3.9 bemerken wir zunächst, dass Satz 3.7 ein hinreichendes Kriterium für  $\{\tau_k\} \rightarrow 0$  liefert. Um die zweite Bedingung einzusehen, nehmen wir an, dass die Folge der Operatoren  $\nabla\Theta(W^k, \tau_k)^{-1}$  für  $k \rightarrow \infty$  beschränkt bleibt. Dann folgt aus dem linearen System (3.4), dass (3.19) gilt, falls die rechte Seite in (3.4) von der Ordnung  $O(\tau_k)$  ist. Dies wiederum ist ziemlich offensichtlich, denn die Zulässigkeit der Iterierten (vergleiche (3.7)) zusammen mit der Tatsache, dass alle Iterierten in der Umgebung  $\mathcal{N}(\beta)$  liegen (siehe auch Satz 3.6) zeigt, dass

$$\begin{aligned} \|\Theta(W^k, \tau_k)\| &= \sqrt{\|\phi(X^k, S^k, \tau_k)\|_F^2 + \tau_k^2} \\ &\leq \|\phi(X^k, S^k, \tau_k)\|_F + \tau_k \leq \beta\tau_k + \tau_k = O(\tau_k) \end{aligned}$$

gilt. Eine solche Bedingung gilt auch, falls wir die rechte Seite in (3.4) durch  $-\Theta(W^k, 0)$  ersetzen. Dann gilt nämlich wegen Korollar 2.24 und Satz 3.6

$$\begin{aligned} \|\Theta(W^k, 0)\| &= \|\phi(X^k, S^k, 0)\|_F \\ &\leq \|\phi(X^k, S^k, \tau_k) - \phi(X^k, S^k, 0)\|_F + \|\phi(X^k, S^k, \tau_k)\|_F \\ &\leq \kappa\sqrt{n}\tau_k + \beta\tau_k \\ &= O(\tau_k), \end{aligned}$$

wobei  $\kappa > 0$  die Konstante aus Lemma 2.23 ist. Insbesondere bleiben alle globalen und lokalen Konvergenzaussagen für Algorithmus 3.1 bestehen, wenn wir diese Modifikation der rechten Seite in (3.4) durchführen.

Um eine hinreichende Bedingung für Voraussetzung 3.9 anzugeben, führen wir die folgende Voraussetzung ein.

**Voraussetzung 3.10** Sei  $(X^*, \lambda^*, S^*)$  eine Lösung der Optimalitätsbedingungen (2.6) mit

- (a) (Strikte Komplementarität)  $X^* + S^* \succ 0$ ;
- (b) (Nichtdegeneriertheit) Für alle  $(\Delta X, \Delta\lambda, \Delta S) \in \mathcal{S}^{n \times n} \times \mathbb{R}^m \times \mathcal{S}^{n \times n}$  gilt die folgende Implikation:

$$\left. \begin{aligned} \sum_{i=1}^m \Delta\lambda_i A_i + \Delta S &= 0, \\ A_i \bullet \Delta X &= 0 \quad (i = 1, \dots, m), \\ X^* \Delta S + \Delta X S^* &= 0 \end{aligned} \right\} \implies (\Delta X, \Delta S) = (0, 0).$$

Beide Voraussetzungen sind mehr oder weniger Standard. Voraussetzung 3.10 (b) wurde von Kojima, Shida und Shindoh [50] eingeführt. Wie in [50]

### 3.2. Globales und lokales Konvergenzverhalten

bemerkt, hat Haerberly gezeigt, dass diese Voraussetzung (unter strikter Komplementarität) äquivalent ist zur primalen und dualen Nichtdegeneriertheit von Alizadeh, Haerberly und Overton [3].

Als Nächstes bringen wir eine Umformulierung von Voraussetzung 3.10 (b), welche im Wesentlichen in Lemma 6.2 von Kojima, Shida und Shindoh [51] zu finden ist. Jedoch wird in dieser Referenz strikte Komplementarität vorausgesetzt.

**Lemma 3.11** *Seien  $(X^*, \lambda^*, S^*)$  eine Lösung der Optimalitätsbedingungen (2.6) und  $(\Delta X, \Delta \lambda, \Delta S) \in \mathcal{S}^{n \times n} \times \mathbb{R}^m \times \mathcal{S}^{n \times n}$ . Dann gilt*

$$X^* \Delta S + \Delta X S^* = 0 \iff X^* \Delta S + \Delta S X^* + S^* \Delta X + \Delta X S^* = 0. \quad (3.20)$$

**Beweis:** Sei zunächst  $X^* \Delta S + \Delta X S^* = 0$ . Transponieren dieser Gleichung und anschließende Addition ergibt unter Berücksichtigung der Symmetrie der auftretenden Matrizen

$$X^* \Delta S + \Delta S X^* + S^* \Delta X + \Delta X S^* = 0, \quad (3.21)$$

also die rechte Seite von (3.20).

Sei daher jetzt (3.21) erfüllt. Wegen  $X^* S^* = 0$  und der Symmetrie von  $X^*$ ,  $S^*$  existiert nach Satz 2.5 eine gemeinsame Spektralzerlegung von  $X^*$  und  $S^*$ , d.h. es gilt  $X^* = Q D_{X^*} Q^T$  und  $S^* = Q D_{S^*} Q^T$  mit einer orthogonalen Matrix  $Q \in \mathbb{R}^{n \times n}$  und Diagonalmatrizen  $D_{X^*}$ ,  $D_{S^*}$ . Nach eventueller Umsortierung der Indizes können wir

$$D_{X^*} = \begin{pmatrix} D_\alpha & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{und} \quad D_{S^*} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & D_\gamma \end{pmatrix}$$

mit geeigneten Indexmengen  $\alpha$ ,  $\gamma$  und  $\beta := \{1, \dots, n\} \setminus (\alpha \cup \gamma)$  und positiv definiten (und daher insbesondere regulären) Diagonalmatrizen  $D_\alpha \in \mathbb{R}^{|\alpha| \times |\alpha|}$  und  $D_\gamma \in \mathbb{R}^{|\gamma| \times |\gamma|}$  schreiben. Wir setzen jetzt

$$\widetilde{\Delta S} := Q^T \Delta S Q =: \begin{pmatrix} \widetilde{\Delta S}_{\alpha\alpha} & \widetilde{\Delta S}_{\alpha\beta} & \widetilde{\Delta S}_{\alpha\gamma} \\ \widetilde{\Delta S}_{\alpha\beta}^T & \widetilde{\Delta S}_{\beta\beta} & \widetilde{\Delta S}_{\beta\gamma} \\ \widetilde{\Delta S}_{\alpha\gamma}^T & \widetilde{\Delta S}_{\beta\gamma}^T & \widetilde{\Delta S}_{\gamma\gamma} \end{pmatrix}$$

sowie

$$\widetilde{\Delta X} := Q^T \Delta X Q =: \begin{pmatrix} \widetilde{\Delta X}_{\alpha\alpha} & \widetilde{\Delta X}_{\alpha\beta} & \widetilde{\Delta X}_{\alpha\gamma} \\ \widetilde{\Delta X}_{\alpha\beta}^T & \widetilde{\Delta X}_{\beta\beta} & \widetilde{\Delta X}_{\beta\gamma} \\ \widetilde{\Delta X}_{\alpha\gamma}^T & \widetilde{\Delta X}_{\beta\gamma}^T & \widetilde{\Delta X}_{\gamma\gamma} \end{pmatrix}.$$

### 3. Ein Glättungsverfahren

Dann liefert Multiplikation der Gleichung (3.21) von links mit  $Q^T$  und von rechts mit  $Q$

$$\begin{aligned}
0 &= D_{X^*} \widetilde{\Delta S} + \widetilde{\Delta S} D_{X^*} + D_{S^*} \widetilde{\Delta X} + \widetilde{\Delta X} D_{S^*} \\
&= \begin{pmatrix} D_\alpha \widetilde{\Delta S}_{\alpha\alpha} & D_\alpha \widetilde{\Delta S}_{\alpha\beta} & D_\alpha \widetilde{\Delta S}_{\alpha\gamma} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} \widetilde{\Delta S}_{\alpha\alpha} D_\alpha & 0 & 0 \\ \widetilde{\Delta S}_{\alpha\beta}^T D_\alpha & 0 & 0 \\ \widetilde{\Delta S}_{\alpha\gamma}^T D_\alpha & 0 & 0 \end{pmatrix} + \\
&\quad \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ D_\gamma \widetilde{\Delta X}_{\alpha\gamma}^T & D_\gamma \widetilde{\Delta X}_{\beta\gamma}^T & D_\gamma \widetilde{\Delta X}_{\gamma\gamma} \end{pmatrix} + \begin{pmatrix} 0 & 0 & \widetilde{\Delta X}_{\alpha\gamma} D_\gamma \\ 0 & 0 & \widetilde{\Delta X}_{\beta\gamma} D_\gamma \\ 0 & 0 & \widetilde{\Delta X}_{\gamma\gamma} D_\gamma \end{pmatrix}.
\end{aligned}$$

Dies ist äquivalent zu

$$\begin{aligned}
0 &= D_\alpha \widetilde{\Delta S}_{\alpha\alpha} + \widetilde{\Delta S}_{\alpha\alpha} D_\alpha = L_{D_\alpha}[\widetilde{\Delta S}_{\alpha\alpha}], \\
0 &= D_\alpha \widetilde{\Delta S}_{\alpha\beta}, \\
0 &= D_\alpha \widetilde{\Delta S}_{\alpha\gamma} + \widetilde{\Delta X}_{\alpha\gamma} D_\gamma, \\
0 &= \widetilde{\Delta X}_{\beta\gamma} D_\gamma, \\
0 &= D_\gamma \widetilde{\Delta X}_{\gamma\gamma} + \widetilde{\Delta X}_{\gamma\gamma} D_\gamma = L_{D_\gamma}[\widetilde{\Delta X}_{\gamma\gamma}].
\end{aligned}$$

Unter Verwendung der positiven Definitheit von  $D_\alpha$  und  $D_\gamma$  folgt hieraus sofort

$$\widetilde{\Delta S} = \begin{pmatrix} 0 & 0 & \widetilde{\Delta S}_{\alpha\gamma} \\ 0 & \widetilde{\Delta S}_{\beta\beta} & \widetilde{\Delta S}_{\beta\gamma} \\ \widetilde{\Delta S}_{\alpha\gamma}^T & \widetilde{\Delta S}_{\beta\gamma}^T & \widetilde{\Delta S}_{\gamma\gamma} \end{pmatrix}, \quad \widetilde{\Delta X} = \begin{pmatrix} \widetilde{\Delta X}_{\alpha\alpha} & \widetilde{\Delta X}_{\alpha\beta} & \widetilde{\Delta X}_{\alpha\gamma} \\ \widetilde{\Delta X}_{\alpha\alpha}^T & \widetilde{\Delta X}_{\beta\beta} & 0 \\ \widetilde{\Delta X}_{\alpha\gamma}^T & 0 & 0 \end{pmatrix}$$

sowie  $0 = D_\alpha \widetilde{\Delta S}_{\alpha\gamma} + \widetilde{\Delta X}_{\alpha\gamma} D_\gamma$ . Damit ergibt sich schließlich

$$\begin{aligned}
X^* \Delta S + \Delta X S^* &= Q(D_{X^*} \widetilde{\Delta S} + \widetilde{\Delta X} D_{S^*}) Q^T \\
&= Q \begin{pmatrix} 0 & 0 & D_\alpha \widetilde{\Delta S}_{\alpha\gamma} + \widetilde{\Delta X}_{\alpha\gamma} D_\gamma \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} Q^T = 0.
\end{aligned}$$

Somit gilt die Äquivalenz (3.20).  $\square$

Der nächste Satz impliziert, dass Voraussetzung 3.9 erfüllt ist, falls die durch Algorithmus 3.1 erzeugten Iterierten  $(X^k, \lambda^k, S^k)$  gegen eine Lösung  $(X^*, \lambda^*, S^*)$  konvergieren, welche die Voraussetzungen 3.4 und 3.10 erfüllt.

### 3.2. Globales und lokales Konvergenzverhalten

Die Konvergenz der gesamten Folge gegen einen Lösungspunkt ist nicht sehr restriktiv, da die beiden Voraussetzungen 3.4 und 3.10 implizieren, dass der Punkt  $(X^*, \lambda^*, S^*)$  die einzige Lösung der Optimalitätsbedingungen (2.6) ist, vergleiche Korollar 3.13.

**Satz 3.12** *Angenommen, die Voraussetzungen 3.4 und 3.10 sind in der Lösung  $(X^*, \lambda^*, S^*)$  von (2.6) erfüllt. Dann ist die lineare Abbildung*

$$\nabla\Theta(X^*, \lambda^*, S^*, 0)$$

*bijektiv.*

**Beweis:** Wir betrachten nur den Fall, wo  $\phi$  die in (2.29) definierte geglättete Fischer-Burmeister-Funktion ist. Der Beweis für die geglättete Minimum-Funktion aus (2.30) kann analog erfolgen.

Wir definieren  $E := ((X^*)^2 + (S^*)^2)^{1/2}$ . Da die Lösung die Bedingung der strikten Komplementarität erfüllt, folgt  $0 \prec (X^* + S^*)^2 = (X^*)^2 + X^*S^* + S^*X^* + (S^*)^2 = (X^*)^2 + (S^*)^2$  und somit, dass  $E$  positiv definit ist. Daher impliziert Satz 2.27, dass die Funktion  $\Theta$  stetig differenzierbar im Punkt  $(X^*, \lambda^*, S^*, 0)$  ist.

Um einzusehen, dass  $\nabla\Theta(X^*, \lambda^*, S^*, 0)$  eine bijektive lineare Abbildung darstellt, müssen wir nur zeigen, dass sie injektiv ist. Daher betrachten wir die Gleichung

$$\nabla\Theta(X^*, \lambda^*, S^*, 0) (\Delta X, \Delta\lambda, \Delta S, \Delta\tau) = (0, 0, 0, 0)$$

und zeigen, dass  $(\Delta X, \Delta\lambda, \Delta S, \Delta\tau) = (0, 0, 0, 0)$  die einzige Lösung dieser Gleichung ist. Aus der letzten Zeile folgt sofort

$$\Delta\tau = 0.$$

Dieses und Satz 2.27 implizieren, dass die ersten drei Blockzeilen in der folgenden Weise geschrieben werden können:

$$0 = \sum_{i=1}^m \Delta\lambda_i A_i + \Delta S, \quad (3.22)$$

$$0 = A_i \bullet \Delta X \quad (i = 1, \dots, m), \quad (3.23)$$

$$0 = \Delta X + \Delta S - L_E^{-1}[X^* \Delta X + \Delta X X^* + S^* \Delta S + \Delta S S^*]. \quad (3.24)$$

Gleichung (3.24) impliziert

$$L_{E-X^*}[\Delta X] + L_{E-S^*}[\Delta S] = 0, \quad (3.25)$$

### 3. Ein Glättungsverfahren

vergleiche den Beweis zu Lemma 3.5. Benutzen wir jetzt die Tatsache, dass  $(X^*, \lambda^*, S^*)$  eine Lösung von (2.6) ist, so dass insbesondere  $X^*S^* = 0$  gilt und daher  $X^*$  und  $S^*$  kommutieren, so folgt, dass  $X^*$  und  $S^*$  nach Satz 2.5 eine gemeinsame Spektralzerlegung haben. Dies bedeutet, dass eine orthogonale Matrix  $Q \in \mathbb{R}^{n \times n}$  sowie positiv semidefinite Diagonalmatrizen  $D_X \in \mathbb{R}^{n \times n}$  und  $D_S \in \mathbb{R}^{n \times n}$  existieren mit

$$X^* = QD_XQ^T \quad \text{und} \quad S^* = QD_SQ^T.$$

Hiermit zeigt eine einfache Rechnung, dass  $E - X^* = S^*$  und  $E - S^* = X^*$  gilt. Somit kann (3.25) als

$$S^*\Delta X + \Delta X S^* + X^*\Delta S + \Delta S X^* = 0$$

geschrieben werden. Dies ist nach Lemma 3.11 äquivalent zu  $X^*\Delta S + \Delta X S^* = 0$ . Mit (3.22) und (3.23) folgt aus Voraussetzung 3.10 daher  $(\Delta X, \Delta S) = (0, 0)$ . Da die Matrizen  $A_i$  nach Voraussetzung 3.4 linear unabhängig sind, folgt aus (3.22) dann auch  $\Delta \lambda = 0$ . Damit ist alles gezeigt.  $\square$

Daraus ergibt sich sofort die Eindeutigkeit der Lösung.

**Korollar 3.13** *Sei  $(X^*, \lambda^*, S^*)$  eine Lösung der Optimalitätsbedingungen (2.6), welche die Voraussetzungen 2.2, 3.4 und 3.10 erfüllt. Dann ist  $(X^*, \lambda^*, S^*)$  die eindeutig bestimmte Lösung der Optimalitätsbedingungen.*

**Beweis:** Wegen Satz 3.12 ist die Abbildung  $\nabla\Theta(X^*, \lambda^*, S^*, 0)$  bijektiv. Aus Lemma 4.1.16 in Dennis und Schnabel [17] folgt dann, dass die Lösung lokal eindeutig ist. Da die Lösungsmenge der Optimalitätsbedingungen unter der Slater-Bedingung jedoch konvex ist, folgt daraus auch die globale Eindeutigkeit von  $(X^*, \lambda^*, S^*)$ .  $\square$

Es sei an dieser Stelle bemerkt, dass Satz 3.12 lediglich eine hinreichende Bedingung für Voraussetzung 3.9 ist. Da die Voraussetzungen in Satz 3.12 implizieren, dass die Lösungsmenge der Optimalitätsbedingungen (2.6) lediglich aus einem Punkt besteht, ist Satz 3.12 sehr restriktiv. Nun zeigen jedoch einige Resultate für lineare Programme und Komplementaritätsprobleme, dass Voraussetzung 3.9 auch unter schwächeren Voraussetzungen, welche nicht die Eindeutigkeit der Lösung implizieren, gilt. Vergleiche dazu beispielsweise Tseng [68]. Weiter lassen die Ausführungen in Kapitel 5 vermuten, dass sich sehr wahrscheinlich lokal schnelle Konvergenz für Algorithmus 3.1 auch ohne strikte Komplementarität beweisen lässt. Dann muss jedoch die Nichtdegeneriertheitsvoraussetzung geeignet verschärft werden (vergleiche Voraussetzung 5.10), was wiederum die Eindeutigkeit der Lösung impliziert.

### 3.2. Globales und lokales Konvergenzverhalten

Wir beginnen nun die lokale Konvergenzanalyse von Algorithmus 3.1 und starten mit dem folgendem technischen Lemma.

**Lemma 3.14** *Sei Voraussetzung 3.9 erfüllt. Dann gilt*

$$\|\phi(X^k + \Delta X^k, S^k + \Delta S^k, \tau_k + \Delta\tau_k)\|_F = o(\tau_k).$$

**Beweis:** Da  $\nabla\phi(X^k, S^k, \tau_k)(\Delta X^k, \Delta S^k, \Delta\tau_k) = -\phi(X^k, S^k, \tau_k)$  nach (3.4), folgt mit  $V^k = (X^k, S^k, \tau_k)$  und  $\Delta V^k = (\Delta X^k, \Delta S^k, \Delta\tau_k)$  aus dem Mittelwertsatz in Integralform

$$\begin{aligned} & \|\phi(X^k + \Delta X^k, S^k + \Delta S^k, \tau_k + \Delta\tau_k)\|_F \\ &= \|\phi(V^k + \Delta V^k)\|_F \\ &= \left\| \int_0^1 \nabla\phi(V^k + \eta\Delta V^k)(\Delta V^k) d\eta + \phi(V^k) \right\|_F \\ &= \left\| \int_0^1 \nabla\phi(V^k + \eta\Delta V^k)(\Delta V^k) d\eta - \nabla\phi(V^k)(\Delta V^k) \right\|_F \\ &\leq \int_0^1 \left\| \left[ \nabla\phi(V^k + \eta\Delta V^k) - \nabla\phi(V^k) \right](\Delta V^k) \right\|_F d\eta \\ &= o(\|\Delta V^k\|) \\ &= o(\|(\Delta X^k, \Delta S^k, \Delta\tau_k)\|), \end{aligned}$$

wobei die vorletzte Gleichung aus der stetigen Differenzierbarkeit von  $\phi$  folgt. Aus Voraussetzung 3.9 ergibt sich dann

$$\|\phi(X^k + \Delta X^k, S^k + \Delta S^k, \tau_k + \Delta\tau_k)\|_F = o(\tau_k),$$

also die Behauptung.  $\square$

Die Hauptarbeit, um zu zeigen, dass die Folge  $\{\tau_k\}$  lokal superlinear gegen Null konvergiert, geschieht in dem folgendem Lemma.

**Lemma 3.15** *Sei Voraussetzung 3.9 erfüllt und die Konstante  $\beta$  erfülle die Ungleichung  $\beta > \kappa\sqrt{n}$ , wobei  $\kappa$  die Konstante aus Lemma 2.23 bezeichnet. Dann konvergiert die Folge  $\{\eta_k\}$  gegen Null.*

**Beweis:** Sei  $\varepsilon > 0$  beliebig. Aus (3.4) folgt  $\Delta\tau_k = -\tau_k$  und zusammen mit Lemma 3.14 ergibt dies

$$\begin{aligned} \|\phi(X^k + \Delta X^k, S^k + \Delta S^k, 0)\|_F &= \|\phi(X^k + \Delta X^k, S^k + \Delta S^k, \tau_k + \Delta\tau_k)\|_F \\ &= o(\tau_k). \end{aligned}$$

### 3. Ein Glättungsverfahren

Daher existiert ein Index  $K_\varepsilon \in \mathbb{N}$ , so dass

$$\|\phi(X^k + \Delta X^k, S^k + \Delta S^k, 0)\|_F \leq \varepsilon \tau_k$$

für alle  $k \geq K_\varepsilon$  gilt. Dann folgt für alle  $\eta > 0$  und alle  $k \geq K_\varepsilon$

$$\begin{aligned} & \|\phi(X^k + \Delta X^k, S^k + \Delta S^k, \eta \tau_k)\|_F \\ & \leq \|\phi(X^k + \Delta X^k, S^k + \Delta S^k, 0)\|_F + \\ & \quad \|\phi(X^k + \Delta X^k, S^k + \Delta S^k, \eta \tau_k) - \phi(X^k + \Delta X^k, S^k + \Delta S^k, 0)\|_F \\ & \leq \varepsilon \tau_k + \kappa \sqrt{n} \eta \tau_k, \end{aligned}$$

wobei die letzte Ungleichung aus Korollar 2.24 folgt. Da  $\varepsilon \tau_k + \kappa \sqrt{n} \eta \tau_k \leq \beta \eta \tau_k$  für alle  $\eta \geq \frac{\varepsilon}{\beta - \kappa \sqrt{n}}$  gilt, folgt aus der Definition von  $\eta_k$ , dass  $\eta_k \alpha_1$  diese Ungleichung nicht erfüllt, d.h. es gilt  $\eta_k < \varepsilon / ((\beta - \kappa \sqrt{n}) \alpha_1)$ . Wegen  $\beta - \kappa \sqrt{n} > 0$  nach Voraussetzung und da  $\varepsilon > 0$  beliebig gewählt war, folgt  $\eta_k \rightarrow 0$ .  $\square$

Wir sind jetzt in der Lage, dass lokale Konvergenzverhalten von Algorithmus 3.1 anzugeben.

**Satz 3.16** *Unter der Voraussetzung 3.9 und  $\beta > \kappa \sqrt{n}$  gilt  $\tau_{k+1} = o(\tau_k)$ , d.h. der Glättungsparameter konvergiert lokal superlinear gegen Null.*

**Beweis:** Unter Benutzung von Lemma 3.15 und der Definitionen von  $\tau_{k+1}$  sowie  $\hat{\tau}_k$  in Algorithmus 3.1 ergibt sich  $\tau_{k+1} \leq \hat{\tau}_k = \eta_k \tau_k = o(\tau_k)$ , d.h.  $\tau_k \rightarrow 0$  superlinear.  $\square$

Wir beenden diesen Abschnitt mit der Bemerkung, dass Satz 3.16 auch gilt, wenn wir die rechte Seite in (3.4) durch  $-\Theta(W^k, 0)$  ersetzen. Dies folgt aus den Resultaten von Chen und Tseng aus [15] (so dass wir die Details hier weglassen) und folgt nicht direkt aus den vorhergehenden Resultaten, da die Lemmata 3.14 und 3.15 davon abhängen, dass die rechte Seite in (3.4) durch  $-\Theta(W^k, \tau_k)$  gegeben ist.

### 3.3. Eigenschaften der Newton-Systeme

In diesem Abschnitt wollen wir die praktische Berechnung der Lösung der Newton-Systeme (3.4) und (3.5) untersuchen. Etwas allgemeiner betrachten wir lineare Gleichungen der Form

$$\nabla \Theta(W, \tau)(\Delta W, \Delta \tau) = -\Theta(W, \nu) \tag{3.26}$$



mit  $W = (X, \lambda, S)$ ,  $\Delta W = (\Delta X, \Delta \lambda, \Delta S)$ ,  $\tau > 0$  und einem beliebigem Parameter  $\nu \in \mathbb{R}$ . In Algorithmus 3.1 tritt speziell  $\nu = \tau$  im Prädiktor-Schritt bzw.  $\nu = \hat{\sigma}\tau$  im Korrektor-Schritt auf. Wir haben jedoch auch erwähnt, dass im Prädiktor-Schritt die Wahl  $\nu = 0$  sinnvoll sein kann.

Wir werden in diesem gesamten Abschnitt die lineare Unabhängigkeit der Matrizen  $A_i$ ,  $i = 1, \dots, m$  voraussetzen. Dann ist Gleichung (3.26) eindeutig lösbar, da wegen  $\tau > 0$  der Operator  $\nabla\Theta(W, \tau)$  nach Satz 3.5 bijektiv ist.

### Minimum-Funktion

Wir werden jetzt zunächst den Fall der geglätteten Minimum-Funktion untersuchen, d.h.  $\Theta$  ist gegeben durch (3.2) und  $\phi$  ist die geglättete Minimum-Funktion aus (2.30). Den Fall der geglätteten Fischer-Burmeister-Funktion werden wir dann im Anschluss betrachten.

Definieren wir die Residuen

$$\begin{aligned} R_C &:= C - \sum_{j=1}^m \lambda_j A_j - S, \\ r_{b,i} &:= b_i - A_i \bullet X, \quad i = 1, \dots, m, \\ r_b &:= (r_{b,1}, \dots, r_{b,m})^T, \end{aligned}$$

so kann das Newton-System (3.26) geschrieben werden als

$$\sum_{j=1}^m \Delta \lambda_j A_j + \Delta S = R_C, \quad (3.27)$$

$$A_i \bullet \Delta X = r_{b,i}, \quad i = 1, \dots, m, \quad (3.28)$$

$$\nabla\phi(X, S, \tau)(\Delta X, \Delta S, \Delta\tau) = -\phi(X, S, \nu), \quad (3.29)$$

$$\Delta\tau = -\nu. \quad (3.30)$$

Benutzen wir Satz 2.27 (b) und beachten (3.30), so ergibt sich aus (3.29)

$$\begin{aligned} \Delta X + \Delta S - L_E^{-1}[(X - S)(\Delta X - \Delta S) + (\Delta X - \Delta S)(X - S) - 8\tau\nu I] \\ = -\phi(X, S, \nu), \end{aligned} \quad (3.31)$$

wobei natürlich  $E := ((X - S)^2 + 4\tau^2 I)^{1/2}$  die symmetrische und positiv definite Matrix aus Satz 2.27 ist. Anwendung des Lyapunov-Operators  $L_E$  auf beide Seiten von (3.31) ergibt nach geeigneter Umordnung der Terme

$$L_{E-(X-S)}[\Delta X] + L_{E+(X-S)}[\Delta S] + 8\tau\nu I = -L_E[\phi(X, S, \nu)].$$

Mit den Notationen

$$A_E := E - (X - S) \quad \text{und} \quad B_E := E + (X - S) \quad (3.32)$$

### 3. Ein Glättungsverfahren

kann dies umgeschrieben werden zu

$$L_{A_E}[\Delta X] + L_{B_E}[\Delta S] + 8\tau\nu I = -L_E[\phi(X, S, \nu)]. \quad (3.33)$$

Da die Matrizen  $A_E, B_E$  aus (3.32) beide symmetrisch und positiv definit sind (siehe auch den Beweis zu Satz 3.5), erhalten wir

$$\Delta X = -L_{A_E}^{-1} \left[ L_{B_E}[\Delta S] + L_E[\phi(X, S, \nu)] + 8\tau\nu I \right]. \quad (3.34)$$

Setzen wir hier  $\Delta S$  aus (3.27) ein und sortieren die Terme entsprechend um, so ergibt sich

$$\Delta X = \sum_{j=1}^m \Delta\lambda_j L_{A_E}^{-1} [L_{B_E}[A_j]] - L_{A_E}^{-1} \left[ L_{B_E}[R_C] + L_E[\phi(X, S, \nu)] + 8\tau\nu I \right].$$

Bilden wir jetzt das Skalarprodukt mit  $A_i$  für  $i = 1, \dots, m$  und benutzen wir (3.28), so ergibt sich

$$\sum_{j=1}^m \Delta\lambda_j L_{A_E}^{-1} [L_{B_E}[A_j]] \bullet A_i = r_{b,i} + L_{A_E}^{-1} \left[ L_{B_E}[R_C] + L_E[\phi(X, S, \nu)] + 8\tau\nu I \right] \bullet A_i \quad (3.35)$$

für  $i = 1, \dots, m$ . Diese Gleichung kann noch weiter umformuliert werden, wenn wir ausnutzen, dass der inverse Lyapunov-Operator nach Lemma 2.19 selbst-adjungiert ist. Dann ergibt sich

$$\sum_{j=1}^m \Delta\lambda_j L_{B_E}[A_j] \bullet L_{A_E}^{-1}[A_i] = r_{b,i} + \left( L_{B_E}[R_C] + L_E[\phi(X, S, \nu)] + 8\tau\nu I \right) \bullet L_{A_E}^{-1}[A_i] \quad (3.36)$$

für alle  $i = 1, \dots, m$ . Dies ist ein lineares Gleichungssystem in den Variablen  $\Delta\lambda \in \mathbb{R}^m$  mit der Koeffizientenmatrix  $M \in \mathbb{R}^{m \times m}$ , welche elementweise definiert ist durch

$$m_{ij} := L_{B_E}[A_j] \bullet L_{A_E}^{-1}[A_i], \quad i, j = 1, \dots, m. \quad (3.37)$$

Lösen wir dieses lineare Gleichungssystem, so können wir sofort  $\Delta S$  aus (3.27) berechnen. Man beachte, dass  $\Delta S$  offensichtlich symmetrisch ist, da sowohl  $R_C$  als auch alle  $A_i$  symmetrisch sind. Mit Blick auf (3.34) kann  $\Delta X$  als die Lösung einer Lyapunov-Gleichung mit symmetrischer rechter Seite berechnet werden. Daher ist  $\Delta X$  selbst symmetrisch, vergleiche auch Satz 2.17. Dies zeigt, dass unser Ansatz automatisch symmetrische Suchrichtungen erzeugt. Anders als bei den Inneren-Punkte-Methoden ist also keine zusätzliche Symmetrisierung des Newton-Systems bzw. der Lösung der Newton-Gleichung notwendig.

### 3.3. Eigenschaften der Newton-Systeme

Wir wollen jetzt als Nächstes die durch (3.37) definierte Matrix  $M = (m_{ij})$  genauer betrachten. Unser Ziel ist, zu zeigen, dass  $M$  symmetrisch und positiv definit ist. Die Symmetrie spielt eine große Rolle bei dem Aufwand zur Berechnung der Matrix  $M$ , da dann nur etwas mehr als die Hälfte der Matrixelemente berechnet werden muss. Die Symmetrie und positive Definitheit zusammen erlauben uns dann die Anwendung einer Cholesky-Zerlegung zur Lösung des linearen Gleichungssystems (3.36).

Überraschender Weise ist die positive Definitheit fast einfacher zu beweisen als die Symmetrie der Matrix  $M$ . Wir beginnen daher mit dem folgenden Resultat.

**Satz 3.17** *Sei Voraussetzung 3.4 erfüllt. Dann ist die Matrix  $M \in \mathbb{R}^{m \times m}$  mit den durch (3.37) gegebenen Einträgen  $m_{ij}$  positiv definit.*

**Beweis:** Sei  $d \in \mathbb{R}^m$  ein beliebiger Vektor. Dann gilt nach Lemma 2.19 (b) und (d) und Lemma 3.2

$$\begin{aligned}
 d^T M d &= \sum_{i,j=1}^m d_i d_j m_{ij} \\
 &= \sum_{i,j=1}^m d_i d_j (L_{B_E}[A_j] \bullet L_{A_E}^{-1}[A_i]) \\
 &= \sum_{i,j=1}^m L_{B_E}[d_j A_j] \bullet L_{A_E}^{-1}[d_i A_i] \\
 &= \sum_{i,j=1}^m (L_{A_E}^{-1} \circ L_{B_E})[d_j A_j] \bullet [d_i A_i] \\
 &= \sum_{j=1}^m (L_{A_E}^{-1} \circ L_{B_E})[d_j A_j] \bullet \left[ \sum_{i=1}^m d_i A_i \right] \\
 &= (L_{A_E}^{-1} \circ L_{B_E}) \left[ \sum_{j=1}^m d_j A_j \right] \bullet \left[ \sum_{i=1}^m d_i A_i \right] \\
 &\geq 0.
 \end{aligned}$$

Weiter impliziert Lemma 2.19 (d), dass Gleichheit nur im Fall  $\sum_{i=1}^m d_i A_i = 0$  auftreten kann. Dies impliziert aber sofort  $d_i = 0$  für alle  $i = 1, \dots, m$ , da die Matrizen  $A_1, \dots, A_m$  als linear unabhängig vorausgesetzt waren.  $\square$

Wir wollen jetzt noch zeigen, dass die Matrix  $M$  auch symmetrisch ist. Wir werden dabei Lemma 2.21 mit den Matrizen  $A = A_E$  und  $B = B_E$  anwenden und müssen daher zeigen, dass  $A_E$  und  $B_E$  kommutieren.

### 3. Ein Glättungsverfahren

**Lemma 3.18** Die zwei Matrizen  $A_E = E - (X - S)$  und  $B_E = E + (X - S)$  aus (3.32) mit  $E := ((X - S)^2 + 4\tau^2 I)^{1/2}$  kommutieren.

**Beweis:** Wir müssen  $A_E B_E = B_E A_E$  zeigen. Dazu wenden wir Satz 2.5 an und zeigen, dass die beiden Matrizen  $A_E$  und  $B_E$  eine gemeinsame Spektralzerlegung haben. Sei daher  $X - S = QDQ^T$  eine Spektralzerlegung der symmetrischen Matrix  $X - S$ . Dann ist  $(X - S)^2 = QD^2Q^T$  eine Spektralzerlegung von  $(X - S)^2$  und wir erhalten

$$(X - S)^2 + 4\tau^2 I = Q(D^2 + 4\tau^2 I)Q^T$$

sowie

$$E = ((X - S)^2 + 4\tau^2 I)^{1/2} = Q(D^2 + 4\tau^2 I)^{1/2}Q^T.$$

Daraus ergibt sich

$$A_E = E - (X - S) = Q((D^2 + 4\tau^2 I)^{1/2} - D)Q^T$$

und

$$B_E = E + (X - S) = Q((D^2 + 4\tau^2 I)^{1/2} + D)Q^T.$$

Dies zeigt, dass  $A_E$  und  $B_E$  eine gemeinsame Spektralzerlegung haben und somit nach Satz 2.5 kommutieren.  $\square$

Wir können jetzt die Symmetrie der Matrix  $M$  zeigen.

**Theorem 3.19** Die Matrix  $M = (m_{ij})$  mit den in (3.37) gegebenen Einträgen  $m_{ij}$  ist symmetrisch.

**Beweis:** Nach Lemma 3.18 kommutieren die beiden symmetrischen und positiv definiten Matrizen  $A_E = E - (X - S)$  und  $B_E = E + (X - S)$ . Nach Lemma 2.21 impliziert dies  $L_{A_E} \circ L_{B_E} = L_{B_E} \circ L_{A_E}$  oder, äquivalent,

$$L_{A_E}^{-1} \circ L_{B_E} = L_{B_E} \circ L_{A_E}^{-1}. \quad (3.38)$$

Daher kann (nach Lemma 2.19 (a), (b)) die Symmetrie von  $M$  umgeschrieben werden zu

$$\begin{aligned} m_{ij} = m_{ji} &\iff L_{B_E}[A_j] \bullet L_{A_E}^{-1}[A_i] = L_{B_E}[A_i] \bullet L_{A_E}^{-1}[A_j] \\ &\iff (L_{A_E}^{-1} \circ L_{B_E})[A_j] \bullet A_i = A_i \bullet (L_{B_E} \circ L_{A_E}^{-1})[A_j] \\ &\iff (L_{A_E}^{-1} \circ L_{B_E})[A_j] \bullet A_i = (L_{B_E} \circ L_{A_E}^{-1})[A_j] \bullet A_i. \end{aligned}$$

Die Symmetrie von  $M$  folgt daher aus (3.38).  $\square$

Nachdem wir jetzt gezeigt haben, dass die Newton-Gleichung (3.26) im Falle der Minimum-Funktion auf ein lineares Gleichungssystem mit symmetrischer und positiv definiten Koeffizientenmatrix zurückgeführt werden kann, wollen wir als Nächstes den Fall der Fischer-Burmeister-Funktion betrachten.

### Fischer-Burmeister-Funktion

Das Glättungsverfahren 3.1 kann auch auf eine andere Funktion  $\Theta$  angewendet werden, wenn wir die Minimum-Funktion aus (2.30) durch die geglättete Fischer-Burmeister-Funktion

$$\phi(X, S, \tau) := X + S - (X^2 + S^2 + 2\tau^2 I)^{1/2}$$

aus (2.29) ersetzen. Es stellt sich nun die Frage, welche der vorhergehenden Resultate in diesem Fall gültig bleiben.

Dazu beginnen wir wieder mit der Umformulierung des Newton-Systems (3.26) und erhalten aus Satz 2.27 (a) für die dritte Blockzeile (3.29)

$$\Delta X + \Delta S - L_E^{-1}[X\Delta X + \Delta X X + S\Delta S + \Delta S S - 4\tau\nu I] = -\phi(X, S, \nu),$$

wobei

$$E := (X^2 + S^2 + 2\tau^2 I)^{1/2} \quad (3.39)$$

natürlich die Matrix aus Satz 2.27 (a) bezeichnet. Nach einigen algebraischen Umformungen (ähnlich zu denen im Fall der Minimum-Funktion), erhalten wir

$$L_{A_E}[\Delta X] + L_{B_E}[\Delta S] + 4\tau\nu I = -L_E[\phi(X, S, \nu)] \quad (3.40)$$

mit

$$A_E := E - X \quad \text{und} \quad B_E := E - S.$$

Mit diesen Definitionen von  $A_E$  und  $B_E$  sieht man leicht ein, dass wir  $\Delta\lambda \in \mathbb{R}^m$  durch Lösung des linearen Gleichungssystems (3.36) mit der Matrix  $M = (m_{ij}) \in \mathbb{R}^{m \times m}$  und  $m_{ij}$  definiert durch (3.37) berechnen können. Da  $A_E$  und  $B_E$  aus (3.40) offensichtlich symmetrisch und positiv definit sind, folgt aus Satz 3.17, dass die Matrix  $M$  auch für die geglättete Fischer-Burmeister-Funktion positiv definit ist.

Jedoch sind die Matrizen  $A_E$  und  $B_E$  aus (3.40) im Allgemeinen nicht vertauschbar. Dies bedeutet, dass unserer Ansatz zum Beweis der Symmetrie von  $M$  nicht auf die geglättete Fischer-Burmeister-Funktion übertragbar ist. Tatsächlich stellt sich heraus, dass die Matrix  $M$  nicht mehr zwingend symmetrisch ist. Dies zeigt das folgende Gegenbeispiel.

### 3. Ein Glättungsverfahren

**Beispiel 3.20** Wähle  $n = 3$ ,  $m = 2$ ,  $\tau = 1$ ,  $\nu = 0$ ,  $\phi$  als geglättete Fischer-Burmeister-Funktion und

$$A_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Wir betrachten nun die Iterierten

$$X = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad S = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

Eine einfache Rechnung zeigt

$$E = (X^2 + S^2 + 2\tau^2 I)^{1/2} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Daher gilt

$$A_E = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad B_E = \begin{pmatrix} 2 & 0 & -1 \\ 0 & 1 & 0 \\ -1 & 0 & 2 \end{pmatrix}.$$

Dies ergibt

$$L_{B_E}[A_1] = \begin{pmatrix} 4 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad L_{B_E}[A_2] = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Weiter zeigt man leicht, dass

$$L_{A_E}^{-1}[A_1] = \begin{pmatrix} \frac{7}{24} & \frac{1}{12} & 0 \\ \frac{1}{12} & \frac{1}{24} & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad L_{A_E}^{-1}[A_2] = \begin{pmatrix} \frac{1}{24} & \frac{1}{12} & 0 \\ \frac{1}{12} & \frac{1}{24} & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

gilt, da

$$\begin{aligned} & L_{A_E}^{-1}[A_1] \cdot A_E + A_E \cdot L_{A_E}^{-1}[A_1] \\ &= \begin{pmatrix} \frac{7}{24} & \frac{1}{12} & 0 \\ \frac{1}{12} & \frac{1}{24} & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{7}{24} & \frac{1}{12} & 0 \\ \frac{1}{12} & \frac{1}{24} & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = A_1 \end{aligned}$$

ist. Analog zeigt man auch die Gültigkeit für  $L_{A_E}^{-1}[A_2]$ . Daraus folgt

$$M = \begin{pmatrix} \frac{7}{6} & \frac{1}{12} \\ \frac{1}{6} & \frac{7}{12} \end{pmatrix}.$$

Also ist  $M$  eine unsymmetrische Matrix.

Daraus folgt, dass der Aufwand zur Berechnung der Matrix  $M$  wesentlich geringer ist, wenn die geglättete Minimum-Funktion anstelle der geglätteten Fischer-Burmeister-Funktion benutzt wird.

Wir haben in diesem Abschnitt die Lyapunov-Formulierung der Newton-Gleichung (3.26) betrachtet und daraus ein lineares Gleichungssystem zur Berechnung von  $\Delta\lambda \in \mathbb{R}^m$  hergeleitet. Man kann die Newton-Gleichung jedoch auch zuerst in die übliche Matrix-Vektor-Form bringen. Wir werden dies (für den Fall  $\tau = 0$ , es ist aber offensichtlich, wie sich die Vorgehensweise verallgemeinern lässt) in Abschnitt 5.2 durchführen. Nutzt man die Blockstruktur des daraus resultierenden Gleichungssystems entsprechend aus, so ergibt sich wieder ein Gleichungssystem zur Berechnung von  $\Delta\lambda \in \mathbb{R}^m$ . In der Preprint-Version [44] der Arbeit [45] wurde gezeigt, dass die dabei auftretende Koeffizientenmatrix gerade die Matrix  $M$  ist. Die Aufstellung der Matrix  $M$  mittels der Matrix-Vektor-Formulierung benötigt ebenfalls  $O(n^3)$  Rechenoperationen, da die Inverse einer im Allgemeinen vollbesetzten Matrix berechnet werden muss. Dieser Ansatz liefert daher für die numerische Lösung der Newton-Gleichung keine neuen Ergebnisse und wird somit nur kurz am Ende von Abschnitt 5.2 ausgeführt. In Kapitel 5 wird die Matrix-Vektor-Formulierung jedoch eine wichtige Rolle spielen.

### 3.4. Numerische Resultate

Um die numerischen Eigenschaften des Glättungsverfahrens 3.1 zu testen, haben wir das Verfahren in `MATLAB` implementiert. Um den Programmieraufwand zu vereinfachen, haben wir die Datenstruktur, die Eingaberoutinen für die Testprobleme und einige Routinen aus dem Bereich der linearen Algebra aus dem `SDPT3`-Softwarepaket, Version 2.1, von Todd, Toh und Tütüncü benutzt.

Dabei haben wir die Schrittweitenparameter als  $\alpha_1 = \alpha_2 = 0.5$  gewählt. Der Zentrierungsparameter  $\hat{\sigma}$  wird dynamisch während der Iteration so wie von Engelke und Kanzow [20] im Falle linearer Programme vorgeschlagen aufdatiert.

Die Lösung der Newton-Systeme erfolgt wie in Abschnitt 3.3, so dass wir uns jetzt noch über die Berechnung eines geeigneten Startvektors  $(X^0, \lambda^0, S^0)$

### 3. Ein Glättungsverfahren

Gedanken machen müssen: Dazu nennen wir einen Vektor  $(X, \lambda, S)$  zulässig für die Optimalitätsbedingungen (2.6), wenn er primal und dual zulässig ist, d.h. wenn er die linearen Gleichungen  $A_i \bullet X = b_i$ ,  $i = 1, \dots, m$  und  $\sum_{i=1}^m \lambda_i A_i + S = C$  erfüllt. Man beachte, dass wir ausdrücklich nicht verlangen, dass  $X \succeq 0$  oder  $S \succeq 0$  für einen zulässigen Punkt gilt. In Algorithmus 3.1 haben wir den Startvektor  $(X^0, \lambda^0, S^0)$  als zulässig in diesem Sinne vorausgesetzt.

Zur Berechnung eines zulässigen Startvektors definieren wir die symmetrische Matrix  $\mathcal{A} \in \mathbb{R}^{m \times m}$  durch

$$\mathcal{A}_{ij} = A_i \bullet A_j \quad \text{für alle } i, j = 1, \dots, m$$

und berechnen  $y^0$  als Lösung des linearen Gleichungssystems  $\mathcal{A}y = b$ . Dann definieren wir

$$X^0 := \sum_{i=1}^m y_i^0 A_i$$

und berechnen  $\lambda^0$  aus  $\mathcal{A}\lambda = (A_1 \bullet C, \dots, A_m \bullet C)^T$ . Schließlich setzen wir

$$S^0 := C - \sum_{i=1}^m \lambda_i^0 A_i$$

und erhalten so einen offensichtlich zulässigen Startpunkt  $(X^0, \lambda^0, S^0)$ .

Nach der Berechnung des Startpunktes setzen wir die verbleibenden Parameter in Algorithmus 3.1 wie folgt:

$$\tau_0 = \|\phi(X^0, S^0, 0)\| / 5 \quad \text{und} \quad \beta = \max\{2.1 \cdot \sqrt{n}, 1.5 \cdot \|\phi(X, S, \tau_0)\| / \tau_0\}.$$

Wir beenden die Iteration, wenn  $\tau_k/n < 10^{-6}$  gilt (man beachte, dass wir die Zentralen-Pfad-Bedingungen mit  $\tau^2$  parametrisiert haben) und das Maß für die Zulässigkeit

$$\max \left\{ \frac{\| [b_i - A_i \bullet X^k]_{i=1}^m \|_2}{\max\{1, \|b\|_2\}}, \frac{\|C - S^k - \sum_{i=1}^m \lambda_i^k A_i\|_F}{\max\{1, \|C\|_2\}} \right\}$$

kleiner als  $10^{-10}$  ist. Der Grund dafür, dass wir  $\tau_k$  durch  $n$  dividieren, basiert auf der Tatsache, dass  $\|\phi(X^k, S^k, 0)\|_F = O(\tau_k)$  gilt. Da wir erreichen wollen, dass  $\|\phi(X^k, S^k, 0)\|_F$  klein wird, ist es plausibel, abzubrechen, falls  $\tau_k$  klein ist. Da wir aber die Frobenius-Norm nehmen, wird es mit größerer Dimension  $n$  immer schwieriger, den Term  $\|\phi(X^k, S^k, 0)\|_F$  zu verkleinern. Daher haben wir, um das Abbruchkriterium mehr oder weniger unabhängig von der Dimension  $n$  der auftretenden Matrizen  $X^k$  und  $S^k$  zu machen, das oben erwähnte Abbruchkriterium benutzt.



Tabelle 3.1.: Durchschnittliche Iterationsanzahl für semidefinite Programme kleiner Dimension.

Problem	$n$	$m$	AHO	HKM	NT	Min	FB
random	10	10	8.2	13.5	12.6	6.5	7.0
Norm min	20	6	8.0	9.3	10.1	6.7	8.0
Cheby	20	11	7.9	9.8	9.9	5.9	7.7
Maxcut	10	10	7.4	8.2	8.4	5.5	6.4
ETP	20	10	11.8	14.7	12.2	10.8	13.6
Lovasz	10	$\approx 25$	7.6	8.7	8.7	9.1	9.3
LogCheby	60	6	10.4	10.9	11.1	10.8	11.1
ChebyC	40	11	7.6	8.5	9.0	5.2	5.5

Theoretisch ist das Maß für die Zulässigkeit immer Null während des gesamten Verfahrens. Numerisch ist die Situation jedoch ein wenig anders. Während die duale Zulässigkeit kein Problem darstellt (da  $S^k$  so definiert wird, dass die duale Zulässigkeit Null ist), so haben wir manchmal Probleme mit der primalen Zulässigkeit festgestellt. Um die primale Zulässigkeit zu verbessern, haben wir daher eine Projektionstechnik benutzt, wie sie auch SDPT<sup>3</sup> verwendet: Nachdem wir die Newton-Richtung  $(\Delta X, \Delta \lambda, \Delta S, \Delta \tau)$  berechnet haben, prüfen wir, ob die Ungleichung

$$\|[A_i \bullet (X + \Delta X)]_{i=1}^m - b\| > \|[A_i \bullet X]_{i=1}^m - b\|$$

gilt. Ist diese Ungleichung erfüllt, so ersetzen wir  $\Delta X$  durch seine orthogonale Projektion auf den Raum  $\{U \in \mathcal{S}^{n \times n} \mid A_i \bullet U = 0, i = 1, \dots, m\}$ . Als Konsequenz daraus erhalten wir, dass die primale Zulässigkeit für alle Testprobleme nahe der Maschinengenauigkeit bleibt.

Mit dem SDPT<sup>3</sup>-Programmpaket werden acht Testprobleme mitgeliefert. Die Ergebnisse, sowohl für die Minimum-Funktion (Min) als auch die Fischer-Burmeister-Funktion (FB), für unterschiedliche Dimensionen dieser Testprobleme werden in den Tabellen 3.1–3.3 dargestellt. Um diese mit denen von Inneren-Punkte-Methoden zu vergleichen, ist die Iterationszahl für eine unzulässige Pfad-Verfolgungs-Methode aus SDPT<sup>3</sup> für die drei populärsten Suchrichtungen AHO, HKM und NT mit angegeben worden. Wir verweisen auf Toh, Todd und Tütüncü [66] für weitere Details. Da alle Testprobleme von Zufallszahlen abhängen, geben wir nur die durchschnittlichen Iterationszahlen über jeweils zehn Testläufe an.

Für die meisten Testprobleme benötigt Algorithmus 3.1, sowohl mit der Minimum-Funktion als auch mit der Fischer-Burmeister-Funktion, weniger Iterationen als alle Inneren-Punkte-Methoden. Jedoch ist unser Abbruchkriterium von dem der Inneren-Punkte-Methoden verschieden und nicht direkt

### 3. Ein Glättungsverfahren

Tabelle 3.2.: Durchschnittliche Iterationsanzahl für semidefinite Programme mittlerer Dimension.

Problem	$n$	$m$	AHO	HKM	NT	Min	FB
random	20	20	10.2	14.4	13.1	8.9	10.4
Norm min	40	11	8.5	10.1	10.7	7.7	8.7
Cheby	40	21	7.7	9.7	10.0	6.1	7.6
Maxcut	21	21	8.2	9.6	9.6	6.3	7.1
ETP	40	20	12.6	16.7	13.3	14.0	16.5
Lovasz	21	$\approx 105$	9.7	10.1	10.4	12.7	14.6
LogCheby	120	11	12.3	13.2	13.1	13.5	16.4
ChebyC	80	21	8.3	9.2	9.4	6.1	6.9

Tabelle 3.3.: Durchschnittliche Iterationsanzahl für semidefinite Programme großer Dimension.

Problem	$n$	$m$	AHO	HKM	NT	Min	FB
random	50	50	10.4	15.6	13.7	10.9	13.0
Norm min	100	26	9.4	10.7	11.2	8.8	9.2
Cheby	100	27	9.3	10.4	11.4	7.1	8.3
Maxcut	50	50	9.0	10.0	10.5	6.7	7.7
ETP	100	50	13.7	18.4	15.1	19.1	18.4
Lovasz	30	$\approx 220$	10.3	10.6	10.7	15.3	17.0
LogCheby	300	51	13.6	14.0	13.7	13.6	13.3
ChebyC	200	41	9.0	9.8	10.0	6.8	8.2

vergleichbar. Weiter muss erwähnt werden, dass eine Iteration in Algorithmus 3.1 im Allgemeinen teurer ist als eine Iteration der Inneren-Punkte-Methoden, da wir die Quadratwurzel einer Matrix berechnen müssen. Auf jeden Fall sind die angegebenen Resultate von Algorithmus 3.1 deutlich besser als die von Chen und Tseng [15] für eine ähnliche Methode.

Vergleichen wir die Ergebnisse für die Minimum-Funktion mit denen der Fischer-Burmeister-Funktion, so stellt sich heraus, dass die Minimum-Funktion in fast allen Fällen weniger Iterationen benötigt. Die Minimum-Funktion ist daher, nicht nur aus den in Abschnitt 3.3 genannten Gründen, der Fischer-Burmeister-Funktion vorzuziehen.

Schließlich enthalten die Tabellen 3.4 und 3.5 einige Resultate, wenn Algorithmus 3.1 auf einige Probleme der Testproblemsammlung SDPLIB (siehe Borchers [8]) angewendet wird. In diesen Tabellen geben wir für jedes Testproblem die Anzahl der Iterationen, den letzten Wert des Glättungsparameters  $\tau$ , die relative Dualitätslücke sowie das Maß der Zulässigkeit bei Abbruch des Verfahrens an. Man beachte, dass für viele Testprobleme die Dualitätslücke

negativ ist. Dies liegt daran, dass die Iterierten nicht zwingend positiv semidefinit sein müssen. Dies ist auch der Grund dafür, dass wir ein anderes Abbruchkriterium als die Inneren-Punkte-Methoden benutzen müssen.

Es stellt sich wieder heraus, dass die Iterationszahlen für die Minimum-Funktion niedriger als für die Fischer-Burmeister-Funktion sind.

### 3. Ein Glättungsverfahren

Tabelle 3.4.: Numerische Resultate von Algorithmus 3.1 mit der Minimum-Funktion für ausgewählte Probleme aus SDPLIB [8].

Problem	$n$	$m$	Iter.	$\tau$	Dualitätslücke	Zulässigkeit
arch0	335	174	44	5.9e-05	-1.037695e-05	3.499271e-13
arch2	335	174	43	9.4e-05	5.195117e-05	6.384054e-13
arch4	335	174	47	1.3e-04	7.104246e-05	4.250081e-13
arch8	335	174	78	1.1e-04	-3.542500e-06	9.052898e-13
gpp100	100	101	18	9.9e-05	-3.042351e-06	2.123789e-15
gpp124-1	124	125	19	1.0e-04	-1.912270e-05	1.195466e-14
gpp124-2	124	125	19	7.0e-05	-2.615799e-06	2.176315e-15
gpp124-3	124	125	16	1.1e-04	-2.025964e-06	1.887616e-15
gpp124-4	124	125	20	6.5e-05	-4.630696e-06	1.331054e-14
gpp250-1	250	250	19	2.1e-04	-1.294246e-05	2.476407e-14
gpp250-2	250	250	17	1.9e-04	-5.718274e-06	2.130110e-14
gpp250-3	250	250	16	1.7e-04	-3.424270e-06	1.258957e-14
gpp250-4	250	250	17	2.2e-04	-2.061362e-06	3.875924e-14
mcp100	100	100	10	1.8e-06	-2.068888e-09	6.683366e-16
mcp124-1	124	124	15	2.6e-05	-5.421892e-09	5.389812e-16
mcp124-2	124	124	10	8.6e-05	-2.899952e-07	7.948236e-16
mcp124-3	124	124	9	2.9e-05	-8.401942e-08	6.089117e-16
mcp124-4	124	124	9	5.8e-07	-1.672955e-09	7.768182e-16
mcp250-1	250	250	14	5.7e-05	-9.849429e-08	9.333300e-16
mcp250-2	250	250	11	1.1e-04	-4.597849e-07	1.003759e-15
mcp250-3	250	250	11	8.7e-05	-1.589781e-07	1.081043e-15
mcp250-4	250	250	11	4.9e-05	-1.219282e-07	1.007461e-15
mcp500-1	500	500	26	3.7e-04	-5.734057e-07	1.087701e-15
mcp500-2	500	500	14	1.4e-04	-3.508263e-07	1.429353e-15
mcp500-3	500	500	11	3.7e-04	-1.295995e-06	1.526598e-15
mcp500-4	500	500	10	5.6e-05	-7.792705e-08	1.574678e-15
theta1	50	104	13	3.9e-05	-1.307451e-07	6.261965e-17
theta2	100	498	15	1.7e-05	-1.766186e-07	1.049632e-14
theta3	150	1106	15	7.8e-05	-1.009075e-06	1.998401e-15
theta4	200	1949	15	9.5e-06	-1.006602e-07	3.996803e-15
truss1	13	6	8	3.3e-09	-3.003989e-09	3.621438e-15
truss2	133	58	13	1.3e-05	-7.869353e-06	2.209316e-14
truss3	31	27	14	5.0e-06	-5.614971e-10	2.660288e-15
truss4	19	12	7	1.3e-05	-3.397473e-05	1.324462e-15
truss5	331	208	16	2.1e-04	-6.840872e-07	1.803378e-14
truss6	451	172	21	2.5e-04	-2.430952e-04	4.601362e-13
truss7	301	86	25	5.9e-05	-2.316906e-07	3.795188e-13
truss8	628	496	20	1.7e-04	-4.805725e-06	3.038575e-14

Tabelle 3.5.: Numerische Resultate von Algorithmus 3.1 mit der Fischer-Burmeister-Funktion für ausgewählte Probleme aus SDPLIB [8].

Problem	$n$	$m$	Iter.	$\tau$	Dualitätslücke	Zulässigkeit
arch0	335	174	56	1.6e-04	-5.828533e-04	9.734233e-13
arch2	335	174	53	5.7e-05	3.098419e-03	1.372540e-12
arch4	335	174	66	3.0e-05	2.853574e-03	3.654509e-12
arch8	335	174	181	3.2e-04	-4.010495e-02	1.642867e-11
gpp100	100	101	19	7.2e-05	-6.845674e-03	5.970729e-15
gpp124-1	124	125	20	7.0e-05	-4.643835e-03	8.425464e-15
gpp124-2	124	125	19	1.2e-04	-1.285198e-02	1.321856e-15
gpp124-3	124	125	20	8.1e-05	-1.753912e-02	5.085961e-15
gpp124-4	124	125	21	1.1e-04	-1.150693e-01	1.001059e-14
gpp250-1	250	251	25	6.9e-05	-4.223426e-03	7.218274e-15
gpp250-2	250	251	17	1.9e-04	-5.002356e-02	1.515276e-14
gpp250-3	250	251	19	1.6e-04	-1.101880e-01	2.148630e-15
gpp250-4	250	251	18	1.2e-04	-2.464797e-01	3.582583e-14
mcp100	100	100	11	7.4e-05	-2.915273e-03	6.606322e-16
mcp124-1	124	124	19	4.9e-05	-1.516797e-04	7.565126e-16
mcp124-2	124	124	11	5.2e-05	-1.705338e-03	7.641071e-16
mcp124-3	124	124	13	3.2e-05	-9.698575e-04	7.326564e-16
mcp124-4	124	124	12	1.3e-05	-7.742095e-04	8.927513e-16
mcp250-1	250	250	19	5.4e-05	-2.991996e-03	8.904056e-16
mcp250-2	250	250	13	1.3e-04	-2.656172e-02	9.989181e-16
mcp250-3	250	250	11	1.6e-04	-4.376059e-02	1.055638e-15
mcp250-4	250	250	11	1.5e-04	-3.618552e-02	1.093721e-15
mcp500-1	500	500	27	1.6e-04	-2.328282e-03	1.073100e-15
mcp500-2	500	500	15	1.8e-04	-3.659477e-02	1.408674e-15
mcp500-3	500	500	16	2.1e-04	-2.198002e-01	1.484790e-15
mcp500-4	500	500	13	1.8e-04	-1.207274e-01	1.608995e-15
theta1	50	104	15	1.4e-05	-3.141141e-04	6.661338e-16
theta2	100	498	16	8.3e-05	-2.155491e-03	2.775558e-15
theta3	150	1106	16	1.4e-04	-1.327784e-02	2.800719e-15
theta4	200	1949	17	1.5e-04	-3.249694e-02	3.881648e-15
truss1	13	6	8	2.0e-07	-2.297970e-05	1.076292e-15
truss2	133	58	13	8.5e-06	-1.214543e-02	1.998902e-14
truss3	31	27	17	1.1e-06	-1.648927e-09	3.660882e-15
truss4	19	12	8	2.6e-06	-4.137745e-04	1.911007e-15
truss5	331	208	14	1.2e-04	-9.369707e-04	2.518953e-14
truss6	451	172	28	8.2e-05	-2.232924e-04	3.395744e-13
truss7	301	86	30	2.5e-05	-4.717313e-05	4.954012e-13
truss8	628	496			Zu wenig Hauptspeicher	

### *3. Ein Glättungsverfahren*

## 4. Trust-Region-Verfahren

Nachdem wir im letzten Kapitel durch eine geeignete Schrittweitenstrategie im Korrektor-Schritt globale Konvergenz erzielen konnten, wollen wir in diesem Kapitel eine Globalisierung mittels eines Trust-Region-Ansatzes untersuchen. Ziel ist es wiederum, eine Lösung der Optimalitätsbedingungen (2.6), d.h. von

$$\begin{aligned} \sum_{i=1}^m \lambda_i A_i + S &= C, \\ A_i \bullet X &= b_i \quad \forall i = 1, \dots, m, \\ X \succeq 0, S \succeq 0, XS &= 0. \end{aligned} \tag{4.1}$$

zu bestimmen. Wir haben bereits in Satz 2.13 gezeigt, dass diese äquivalent sind zu dem System

$$\Phi(X, \lambda, S) = 0$$

mit

$$\Phi(X, \lambda, S) := \begin{pmatrix} \sum_{i=1}^m \lambda_i A_i + S - C \\ A_i \bullet X - b_i \quad (i = 1, \dots, m) \\ \phi(X, S) \end{pmatrix}.$$

Dabei bezeichnet  $\phi$  die Fischer-Burmeister-Funktion aus (2.10). Aus Gründen, die wir später genauer erläutern werden, betrachten wir in diesem Kapitel nicht die (geglättete) Minimum-Funktion, sondern nur die (geglättete) Fischer-Burmeister-Funktion. Wir erweitern die Funktion  $\Phi$  jetzt wieder um die Glättungsvariable  $\tau$  und definieren

$$\Theta(W) := \Theta(X, \lambda, S, \tau) := \begin{pmatrix} \sum_{i=1}^m \lambda_i A_i + S - C \\ A_i \bullet X - b_i \quad (i = 1, \dots, m) \\ \phi(X, S, \tau) \\ e^{c\tau} - 1 \end{pmatrix}$$

für  $W = (X, \lambda, S, \tau) \in \mathcal{S}^{n \times n} \times \mathbb{R}^m \times \mathcal{S}^{n \times n} \times \mathbb{R}$ . Dabei bezeichnet  $\phi$  die geglättete Fischer-Burmeister-Funktion aus (2.29) und  $c > 0$  einen beliebigen, positiven Parameter. Wegen

$$\tau = 0 \iff e^{c\tau} - 1 = 0$$

erhalten wir sofort, dass der Vektor  $(X^*, \lambda^*, S^*)$  die Optimalitätsbedingungen (4.1) genau dann erfüllt, wenn  $W^* = (X^*, \lambda^*, S^*, \tau_*)$  eine Lösung des nichtlinearen Gleichungssystems  $\Theta(W^*) = 0$  ist. Es sei an dieser Stelle bemerkt, dass

#### 4. Trust-Region-Verfahren

wir die letzte Blockzeile von  $\Theta$  wie im letzten Kapitel auch durch  $\tau$  anstelle von  $e^{c\tau} - 1$  hätten definieren können. Da wir im Algorithmus jedoch  $\tau \neq 0$  garantieren wollen, wird sich unsere Definition als günstiger erweisen.

Das folgende Lemma zeigt, dass die Abbildung  $\nabla\Theta(X, \lambda, S, \tau)$  unter geeigneten Voraussetzungen invertierbar ist.

**Lemma 4.1** *Sei Voraussetzung 3.4 erfüllt und  $(X, \lambda, S, \tau) \in \mathcal{S}^{n \times n} \times \mathbb{R}^m \times \mathcal{S}^{n \times n} \times \mathbb{R}_{++}$  gegeben. Dann ist die lineare Abbildung  $\nabla\Theta(X, \lambda, S, \tau)$  bijektiv.*

**Beweis:** Es genügt wiederum zu zeigen, dass das System

$$\begin{aligned} \sum_{i=1}^m \Delta\lambda_i A_i + \Delta S &= 0, \\ A_i \bullet \Delta X &= 0, \quad i = 1, \dots, m, \\ \nabla\phi(X, S, \tau)(\Delta X, \Delta S, \Delta\tau) &= 0, \\ ce^{c\tau} \Delta\tau &= 0 \end{aligned}$$

nur die Lösung  $(\Delta X, \Delta\lambda, \Delta S, \Delta\tau) = (0, 0, 0, 0)$  hat. Aus der letzten Zeile ergibt sich wegen  $\tau > 0$  und  $c > 0$  sofort  $\Delta\tau = 0$ . Folgen wir daher dem Beweis zu Lemma 3.5, so ergibt sich auch  $(\Delta X, \Delta\lambda, \Delta S) = (0, 0, 0)$  und somit die Behauptung.  $\square$

Um einen Trust-Region-Ansatz anwenden zu können, müssen wir eine geeignete Zielfunktion definieren. Dazu benötigen wir zunächst das folgende Resultat von Tseng [67, Lemma 6.3].

**Lemma 4.2 (Tseng [67])** *Sei  $\phi$  die Fischer-Burmeister-Funktion aus (2.10) und  $\psi : \mathcal{S}^{n \times n} \times \mathcal{S}^{n \times n} \rightarrow \mathbb{R}$  definiert durch*

$$\psi(X, S) := \frac{1}{2} \|\phi(X, S)\|_F^2.$$

Dann gelten die folgenden Aussagen:

- (a) *Für alle  $X, S \in \mathcal{S}^{n \times n}$  gilt  $\psi(X, S) \geq 0$  und  $\psi(X, S) = 0$  genau dann, wenn  $X \succeq 0$ ,  $S \succeq 0$  und  $XS = 0$  gilt.*
- (b)  *$\psi$  ist stetig differenzierbar in jedem Punkt  $(X, S) \in \mathcal{S}^{n \times n} \times \mathcal{S}^{n \times n}$  mit*

$$\begin{aligned} \nabla_X \psi(X, S) &= L_E^{-1}[\phi(X, S)](E - X) + (E - X)L_E^{-1}[\phi(X, S)] \\ \nabla_S \psi(X, S) &= L_E^{-1}[\phi(X, S)](E - S) + (E - S)L_E^{-1}[\phi(X, S)] \end{aligned}$$

*und  $E := (X^2 + S^2)^{1/2}$ . Insbesondere gilt  $\phi(X, S) \in \mathcal{S}_E^{n \times n}$ .*



(c) Für alle  $X, S \in \mathcal{S}^{n \times n}$  gilt

$$\nabla_X \psi(X, S) \bullet \nabla_S \psi(X, S) \geq \|\phi(X, S) L_E^{-1}[\phi(X, S)]\|_F^2$$

$$\text{mit } E := (X^2 + S^2)^{1/2}.$$

Man beachte, dass eine entsprechende Aussage für die Minimum-Funktion aus (2.14) nicht gilt. Hat sich die Minimum-Funktion im letzten Kapitel noch als vorteilhaft gegenüber der Fischer-Burmeister-Funktion herausgestellt, da sich der Aufwand zur Aufstellung und Lösung der Newton-Systeme in diesem Fall verringert, so ist diese für die folgenden Überlegungen ungeeignet. Wir werden daher in diesem ganzen Kapitel nur mit der Fischer-Burmeister-Funktion und ihrer geglätteten Variante arbeiten.

Motiviert durch das vorhergehende Resultat definieren wir die Gütefunktion (engl. *merit function*)

$$\theta(X, \lambda, S, \tau) := \frac{1}{2} \|\Theta(X, \lambda, S, \tau)\|^2. \quad (4.2)$$

Dann können wir nach Lemma 4.2 (a) die Optimalitätsbedingungen schreiben als unrestringiertes Optimierungsproblem

$$\min \theta(W), \quad W \in \mathcal{S}^{n \times n} \times \mathbb{R}^m \times \mathcal{S}^{n \times n} \times \mathbb{R}. \quad (4.3)$$

Weiter ist nach Lemma 4.2 (b) die Funktion  $\theta$  stetig differenzierbar in allen Punkten  $(X, \lambda, S, \tau)$ . Ist sogar  $\Theta$  im Punkt  $W = (X, \lambda, S, \tau)$  differenzierbar (was insbesondere für  $\tau \neq 0$  der Fall ist), so gilt

$$\nabla \theta(W)(\Delta W) = \Theta(W) \bullet \nabla \Theta(W)(\Delta W). \quad (4.4)$$

Dabei bezeichnet  $\bullet$  das Skalarprodukt im  $\mathcal{S}^{n \times n} \times \mathbb{R}^m \times \mathcal{S}^{n \times n} \times \mathbb{R}$ , welches definiert ist durch

$$(X^1, \lambda^1, S^1, \tau_1) \bullet (X^2, \lambda^2, S^2, \tau_2) := X^1 \bullet X^2 + (\lambda^1)^T \lambda^2 + S^1 \bullet S^2 + \tau_1 \tau_2.$$

Das nächste Resultat zeigt, dass schon jeder stationäre Punkt von  $\theta$  eine Lösung der Optimalitätsbedingungen (4.1) darstellt. Dafür benötigen wir wiederum Voraussetzung 3.4 der linearen Unabhängigkeit der Matrizen  $A_i$ , welche theoretisch jedoch keine Einschränkung darstellt.

**Lemma 4.3** *Sei Voraussetzung 3.4 erfüllt und  $\theta$  die durch (4.2) gegebene Funktion, wobei  $\phi$  die in (2.29) definierte geglättete Fischer-Burmeister-Funktion bezeichnet. Ist  $(X^*, \lambda^*, S^*, \tau_*)$  ein stationärer Punkt von  $\theta$ , d.h. gilt*

$$\nabla \theta(X^*, \lambda^*, S^*, \tau_*) = 0,$$

*so ist  $\theta(X^*, \lambda^*, S^*, \tau_*) = 0$ . Der Punkt  $(X^*, \lambda^*, S^*, \tau_*)$  ist daher eine Lösung der Optimalitätsbedingungen (4.1).*

#### 4. Trust-Region-Verfahren

**Beweis:** Sei  $W^* = (X^*, \lambda^*, S^*, \tau_*)$  ein stationärer Punkt von  $\theta$ . Angenommen, es ist  $\tau_* \neq 0$ . Da  $W^*$  ein stationärer Punkt von  $\theta$  ist, folgt

$$0 = \nabla\theta(W^*)(\Delta W) = \Theta(W^*) \bullet \nabla\Theta(W^*)(\Delta W)$$

für alle  $\Delta W$ . Dies ist aber nur möglich, falls  $\Theta(W^*) = 0$ , da die Abbildung  $\nabla\Theta(W^*)$  nach Lemma 4.1 bijektiv ist. Somit gilt  $\tau_* = 0$ , ein Widerspruch. Daher ist doch  $\tau_* = 0$ .

Wegen  $\nabla\theta(W^*) = 0$  verschwinden alle partiellen Ableitungen, d.h.

$$0 = \nabla_X \theta(X^*, \lambda^*, S^*, 0) = \sum_{i=1}^m (A_i \bullet X^* - b_i) A_i + \frac{1}{2} \nabla_X \|\phi(X^*, S^*, 0)\|_F^2, \quad (4.5)$$

$$0 = \nabla_{\lambda_j} \theta(X^*, \lambda^*, S^*, 0) = \left( \sum_{i=1}^m \lambda_i^* A_i + S^* - C \right) \bullet A_j, \quad j = 1, \dots, m, \quad (4.6)$$

$$0 = \nabla_S \theta(X^*, \lambda^*, S^*, 0) = \sum_{i=1}^m \lambda_i^* A_i + S^* - C + \frac{1}{2} \nabla_S \|\phi(X^*, S^*, 0)\|_F^2. \quad (4.7)$$

Nach Lemma 4.2 existiert  $\nabla_{(X,S)} \|\phi(X, S, 0)\|_F^2$  für alle  $X, S \in \mathcal{S}^{n \times n}$ . Setzen wir jetzt  $E := ((X^*)^2 + (S^*)^2)^{1/2}$ , so folgt aus Lemma 4.2 (c) zusammen mit (4.5)–(4.7)

$$\begin{aligned} & \|\phi(X^*, S^*, 0) L_E^{-1}[\phi(X^*, S^*, 0)]\|_F^2 \\ & \leq \left( \frac{1}{2} \nabla_X \|\phi(X^*, S^*, 0)\|_F^2 \right) \bullet \left( \frac{1}{2} \nabla_S \|\phi(X^*, S^*, 0)\|_F^2 \right) \\ & = \left( \sum_{j=1}^m (A_j \bullet X^* - b_j) A_j \right) \bullet \left( \sum_{i=1}^m \lambda_i^* A_i + S^* - C \right) \\ & = \sum_{j=1}^m (A_j \bullet X^* - b_j) \left( A_j \bullet \left( \sum_{i=1}^m \lambda_i^* A_i + S^* - C \right) \right) \\ & = 0, \end{aligned}$$

also

$$\phi(X^*, S^*, 0) L_E^{-1}[\phi(X^*, S^*, 0)] = 0.$$

Dies bedeutet aber  $\phi(X^*, S^*, 0) = 0$  nach Satz 2.20. Dann liefert aber Lemma 4.2 (b)

$$0 = \frac{1}{2} \nabla_{(X,S)} \|\phi(X^*, S^*, 0)\|_F^2,$$

so dass die primale und duale Zulässigkeit direkt aus (4.5) und (4.7) sowie der linearen Unabhängigkeit der Matrizen  $A_i$  folgt.  $\square$

Da jede Lösung des unrestringierten Optimierungsproblems (4.3) gleichzeitig ein stationärer Punkt von  $\theta$  ist, folgt:

**Korollar 4.4** *Ein Punkt  $(X^*, \lambda^*, S^*, \tau_*)$  ist ein stationärer Punkt von  $\theta$  genau dann, wenn  $\theta(X^*, \lambda^*, S^*, \tau_*) = 0$  gilt.*

## 4.1. Trust-Region-Algorithmus

Wir wollen nun die Ergebnisse des letzten Abschnitts anwenden, um ein Trust-Region-Verfahren zur Lösung der Optimalitätsbedingungen (4.1) zu entwickeln. Dafür kombinieren wir das Newton-Verfahren für das nichtlineare Gleichungssystem  $\Theta(W) = 0$  mit einem Trust-Region-Ansatz für nichtlineare Gleichungssysteme, wie er beispielsweise von Dennis und Schnabel [17] oder Kanzow und Zupke [49] beschrieben wird. Wir geben zunächst den Algorithmus formal an.

### Algorithmus 4.5

(S. 0) *(Initialisierung)*

Wähle  $W^0 = (X^0, \lambda^0, S^0, \tau_0) \in \mathcal{S}^{n \times n} \times \mathbb{R}^m \times \mathcal{S}^{n \times n} \times \mathbb{R}$  mit  $\tau_0 \neq 0$ ,  $\varepsilon \geq 0$ ,  $\alpha, \gamma \in (0, 1)$ ,  $0 < \rho_1 < \rho_2$ ,  $\Delta_0, \Delta_{\min} > 0$  sowie  $0 < \sigma_1 < 1 < \sigma_2$  und setze  $k := 0$ .

(S. 1) *(Abbruchkriterium)*

Ist  $\|\nabla\theta(W^k)\| \leq \varepsilon$ : STOPP.

(S. 2) *(Prädiktor-Schritt)*

Sei  $\Delta W^k = (\Delta X^k, \Delta \lambda^k, \Delta S^k, \Delta \tau_k) \in \mathcal{S}^{n \times n} \times \mathbb{R}^m \times \mathcal{S}^{n \times n} \times \mathbb{R}$  eine Lösung des Newton-Systems

$$\nabla\Theta(W^k)(\Delta W) = -\Theta(W^k). \quad (4.8)$$

Ist  $\theta(W^k + \Delta W^k) \leq \gamma\theta(W^k)$ , so nennen wir die Iteration erfolgreich, setzen

$$W^{k+1} := W^k + \Delta W^k, \quad \Delta_{k+1} := \max\{\Delta_{\min}, \Delta_k\}$$

und gehen zu (S. 4). Ansonsten gehe zu Schritt (S. 3).

(S. 3) *(Korrektor-Schritt, Trust-Region-Schritt)*

Berechne eine Lösung  $\Delta \hat{W}^k$  des Trust-Region-Teilproblems

$$\min Q_k(\Delta \hat{W}) \quad \text{u.d.N.} \quad \|\Delta \hat{W}\| \leq \Delta_k \quad (4.9)$$

#### 4. Trust-Region-Verfahren

mit

$$Q_k(\Delta\hat{W}) := \frac{1}{2} \|\Theta(W^k) + \nabla\Theta(W^k)(\Delta\hat{W})\|^2$$

und berechne

$$\alpha_k := \begin{cases} 1 & , \text{ falls } \tau_k + \hat{\Delta}\tau_k \neq 0 \\ \alpha & , \text{ sonst} \end{cases}, \quad r_k := \frac{\theta(W^k) - \theta(W^k + \alpha_k\Delta\hat{W}^k)}{\theta(W^k) - Q_k(\alpha_k\Delta\hat{W}^k)}.$$

Ist  $r_k \geq \rho_1$ , so nennen wir die Iteration erfolgreich und setzen

$$W^{k+1} := W^k + \alpha_k\Delta\hat{W}^k,$$

ansonsten war die Iteration nicht erfolgreich und wir setzen

$$W^{k+1} := W^k.$$

- (a) Ist  $r_k < \rho_1$ , so setze  $\Delta_{k+1} := \sigma_1\Delta_k$ .
- (b) Ist  $r_k \in [\rho_1, \rho_2)$ , so setze  $\Delta_{k+1} := \max\{\Delta_{\min}, \Delta_k\}$ .
- (c) Ist  $r_k \geq \rho_2$ , so setze  $\Delta_{k+1} := \max\{\Delta_{\min}, \sigma_2\Delta_k\}$ .

(S.4) Setze  $k \leftarrow k + 1$  und gehe zu (S.1).

Bevor wir das vorgestellte Trust-Region-Verfahren formal untersuchen, wollen wir zunächst ein paar allgemeine Bemerkungen machen. Das Abbruchkriterium in Schritt (S.1) ist motiviert durch Lemma 4.3. Im Prädiktor-Schritt (S.2) führen wir einen Schritt des ganz normalen Newton-Verfahrens für nichtlineare Gleichungen aus. Tritt dabei eine geeignete Reduktion des Funktionswertes der Gütefunktion  $\theta$  ein, so akzeptieren wir den vollen Newton-Schritt und datieren den Trust-Region-Radius auf. In diesem Fall überspringen wir den Korrektor-Schritt.

Beim vorhergehend betrachteten Prädiktor-Korrektor-Glättungsverfahren haben wir teilweise die rechte Seite der Gleichung (4.8) durch

$$-\Theta(X^k, \lambda^k, S^k, 0),$$

d.h.  $\tau_k$  in der rechten Seite durch Null ersetzt. Wir bemerkten, dass alle (lokalen) Konvergenzaussagen bei dieser Modifikation erhalten bleiben. In dem jetzt zu betrachtenden Trust-Region-Verfahren macht eine solche Modifikation keinen Sinn, da dann keine Reduktion des Glättungsparameters  $\tau_k$  stattfindet. Im Gegensatz zum Glättungsverfahren haben wir hier keine nachgeschaltete Schrittlängenbestimmung, welche eine geeignete Reduktion von  $\tau_k$  garantieren würde.

Im Korrektor-Schritt berechnen wir eine Lösung einer geeigneten quadratischen Approximation der Gütefunktion  $\theta$ . Die Berechnung der Schrittweite  $\alpha_k$  ist notwendig, um im gesamten Iterationsverlauf  $\tau_k \neq 0$  und somit die stetige Differenzierbarkeit der Funktion  $\Theta$  zu garantieren. Man beachte jedoch, dass  $\tau_k$  nicht zwingend positiv sein muss und auch nicht unbedingt, wie im Glättungsverfahren aus Kapitel 3, monoton fällt.

Die Aufdatierung des Trust-Region-Parameters hängt von dem Quotienten  $r_k$  ab, welcher ein Maß für die Qualität des quadratischen Modells auf der Kugel  $\{W^k + \Delta W \mid \|\Delta W\| \leq \Delta_k\}$  darstellt.

Wir haben bereits bemerkt, dass die Schrittweite  $\alpha_k$  im Korrektor-Schritt  $\tau_{k+1} \neq 0$  nach dem Korrektor-Schritt garantiert. Das folgende Resultat zeigt, dass dies auch der Fall ist, falls der Prädiktor-Schritt erfolgreich war.

**Lemma 4.6** *Sei  $W^k = (X^k, \lambda^k, S^k, \tau_k)$  mit  $\tau_k \neq 0$  gegeben und  $\Delta W^k = (\Delta X^k, \Delta \lambda^k, \Delta S^k, \Delta \tau_k)$  eine Lösung der Newton-Gleichung (4.8). Dann gilt  $\tau_k + \Delta \tau_k > 0$ .*

**Beweis:** Die Funktion  $f(\tau) := e^{c\tau} - 1$  mit  $c > 0$  ist strikt konvex. Daher gilt

$$f(\tau_k + \Delta \tau_k) > f(\tau_k) + f'(\tau_k)\Delta \tau_k$$

oder äquivalent

$$e^{c(\tau_k + \Delta \tau_k)} > e^{c\tau_k} + ce^{c\tau_k}\Delta \tau_k. \quad (4.10)$$

Nun folgt aber aus der letzten Blockzeile in (4.8), dass  $\Delta \tau_k$  so berechnet wird, dass die rechte Seite in (4.10) gleich Eins ist. Dann gilt  $e^{c(\tau_k + \Delta \tau_k)} > 1$ , was wegen  $c > 0$  äquivalent ist zu  $\tau_k + \Delta \tau_k > 0$ .  $\square$

Für die folgende Konvergenzanalyse nehmen wir an, dass der Abbruchparameter  $\varepsilon$  gleich Null ist und dass der Algorithmus 4.5 nicht in endlich vielen Schritten in einer Lösung abbricht. Dann ist die folgende, auf Powell [58] zurückgehende, Ungleichung im Zusammenhang mit Trust-Region-Verfahren wohlbekannt, vergleiche auch [13, 49].

**Lemma 4.7 (Powell [58])** *Sei  $\Delta \hat{W}^k$  eine Lösung des Trust-Region-Teilproblems (4.9). Dann gilt die Ungleichung*

$$\theta(W^k) - Q_k(\Delta \hat{W}^k) \geq \frac{1}{2} \|\nabla \theta(W^k)\| \min \left\{ \Delta_k, \frac{\|\nabla \theta(W^k)\|}{\|\nabla \Theta(W^k)\|^2} \right\}. \quad (4.11)$$

#### 4. Trust-Region-Verfahren

**Beweis:** Da  $\Delta\hat{W}^k$  das globale Minimum des Trust-Region-Teilproblems (4.9) darstellt, gilt für alle zulässigen  $\Delta\hat{W}$

$$\begin{aligned}
\theta(W^k) - Q_k(\Delta\hat{W}^k) &\geq \theta(W^k) - Q_k(\Delta\hat{W}) \\
&= \frac{1}{2} \|\Theta(W^k)\|^2 - \frac{1}{2} \|\Theta(W^k) + \nabla\Theta(W^k)(\Delta\hat{W})\|^2 \\
&= -\Theta(W^k) \bullet \nabla\Theta(W^k)(\Delta\hat{W}) - \frac{1}{2} \|\nabla\Theta(W^k)(\Delta\hat{W})\|^2 \\
&\stackrel{(4.4)}{=} -\nabla\theta(W^k)(\Delta\hat{W}) - \frac{1}{2} \|\nabla\Theta(W^k)(\Delta\hat{W})\|^2 \\
&\geq -\nabla\theta(W^k)(\Delta\hat{W}) - \frac{1}{2} \|\nabla\Theta(W^k)\|^2 \|\Delta\hat{W}\|^2.
\end{aligned} \tag{4.12}$$

Ist nun  $\Delta_k \|\nabla\Theta(W^k)\|^2 \leq \|\nabla\theta(W^k)\|$ , so betrachten wir speziell den zulässigen Vektor

$$\Delta\hat{W} := -\frac{\Delta_k}{\|\nabla\theta(W^k)\|} \nabla\theta(W^k)$$

und erhalten aus (4.12) die Ungleichung

$$\begin{aligned}
\theta(W^k) - Q_k(\Delta\hat{W}^k) &\geq \Delta_k \|\nabla\theta(W^k)\| - \frac{1}{2} \Delta_k^2 \|\nabla\Theta(W^k)\|^2 \\
&\geq \frac{1}{2} \Delta_k \|\nabla\theta(W^k)\|.
\end{aligned} \tag{4.13}$$

Ist andererseits  $\Delta_k \|\nabla\Theta(W^k)\|^2 > \|\nabla\theta(W^k)\|$ , dann ist der Vektor

$$\Delta\hat{W} := -\frac{1}{\|\nabla\Theta(W^k)\|^2} \nabla\theta(W^k)$$

wiederum zulässig für (4.9) und aus (4.12) folgt

$$\theta(W^k) - Q_k(\Delta\hat{W}^k) \geq \frac{\|\nabla\theta(W^k)\|^2}{\|\nabla\Theta(W^k)\|^2} - \frac{1}{2} \frac{\|\nabla\theta(W^k)\|^2}{\|\nabla\Theta(W^k)\|^2} = \frac{1}{2} \frac{\|\nabla\theta(W^k)\|^2}{\|\nabla\Theta(W^k)\|^2}. \tag{4.14}$$

Kombinieren wir (4.13) und (4.14), so ergibt sich mit

$$\theta(W^k) - Q_k(\Delta\hat{W}^k) \geq \frac{1}{2} \|\nabla\theta(W^k)\| \min \left\{ \Delta_k, \frac{\|\nabla\theta(W^k)\|}{\|\nabla\Theta(W^k)\|^2} \right\}$$

die Behauptung. □

Man beachte, dass immer  $\nabla\Theta(W^k) \neq 0$  gilt. Ansonsten wäre nämlich  $\nabla\theta(W^k) = 0$  nach (4.4) und somit Algorithmus 4.5 in Schritt (S.1) abgebrochen. Daher ist der Nenner auf der rechten Seite von (4.11) immer von Null verschieden.

Das nächste Resultat ist nur eine Variante von Lemma 4.7.

**Lemma 4.8** *Sei  $\Delta\hat{W}^k$  eine Lösung der Trust-Region-Teilproblems (4.9) und  $\alpha \in [0, 1]$ . Dann gilt die Ungleichung*

$$\theta(W^k) - Q_k(\alpha\Delta\hat{W}^k) \geq \frac{\alpha}{2} \|\nabla\theta(W^k)\| \min \left\{ \Delta_k, \frac{\|\nabla\theta(W^k)\|}{\|\nabla\Theta(W^k)\|^2} \right\}.$$

**Beweis:** Wegen  $\alpha^2 \leq \alpha$  ergibt sich

$$\begin{aligned} & \theta(W^k) - Q_k(\alpha\Delta\hat{W}^k) \\ &= \theta(W^k) - \theta(W^k) - \alpha\nabla\theta(W^k)(\Delta\hat{W}^k) - \frac{\alpha^2}{2} \|\nabla\Theta(W^k)\Delta\hat{W}^k\|^2 \\ &\geq -\alpha\nabla\theta(W^k)(\Delta\hat{W}^k) - \frac{\alpha}{2} \|\nabla\Theta(W^k)(\Delta\hat{W}^k)\|^2 \\ &= \alpha(\theta(W^k) - Q_k(\Delta\hat{W}^k)). \end{aligned}$$

Damit folgt die Behauptung aus Lemma 4.7. □

Als Konsequenz aus den Lemmata 4.6 und 4.8 ergibt sich, dass der Trust-Region-Algorithmus 4.5 auf ein beliebiges semidefinites Programm angewendet werden kann.

**Korollar 4.9** *Algorithmus 4.5 ist wohldefiniert.*

**Beweis:** Wir müssen nur noch zeigen, dass der Nenner

$$\theta(W^k) - Q_k(\alpha_k\Delta\hat{W}^k)$$

in der Definition von  $r_k$  in Schritt (S.3) für alle  $k \in \mathbb{N}$  von Null verschieden ist. Angenommen, dies ist nicht der Fall. Dann ist  $\nabla\theta(W^k) = 0$  wegen Lemma 4.8. Dann hätte der Algorithmus aber in Schritt (S.1) gestoppt, ein Widerspruch. □

Wir wollen diesen Abschnitt mit einem Resultat beenden, das zeigt, dass die Glättungsvariable  $\tau_k$  streng monoton fällt, falls der Prädiktor-Schritt irgendwann für alle  $k$  erfolgreich ist.

#### 4. Trust-Region-Verfahren

**Lemma 4.10** Sei  $W^k = (X^k, \lambda^k, S^k, \tau_k)$  mit  $\tau_k > 0$  gegeben und  $\Delta W^k = (\Delta X^k, \Delta \lambda^k, \Delta S^k, \Delta \tau_k)$  eine Lösung der Newton-Gleichung (4.8). Dann gilt  $\Delta \tau_k \in (-\tau_k, 0)$  und daher  $\tau_k + \Delta \tau_k \in (0, \tau_k)$ .

**Beweis:** Aus der letzten Blockzeile in (4.8) folgt (unabhängig von allen anderen Variablen)

$$\Delta \tau_k = \frac{1 - e^{c\tau_k}}{ce^{c\tau_k}}. \quad (4.15)$$

Dann folgt  $\Delta \tau_k < 0$  aus  $e^{c\tau_k} > 1$  für alle  $\tau_k > 0$ , da die Exponentialfunktion streng monoton ist. Man beachte dabei, dass  $c > 0$  vorausgesetzt wurde.

Weiter folgt aus (4.15)

$$\begin{aligned} \Delta \tau_k &= \frac{1}{ce^{c\tau_k}} \left( 1 - \sum_{i=0}^{\infty} \frac{(c\tau_k)^i}{i!} \right) \\ &= -\frac{1}{ce^{c\tau_k}} \sum_{i=1}^{\infty} \frac{(c\tau_k)^i}{i!} \\ &= -\frac{c\tau_k}{ce^{c\tau_k}} \sum_{i=0}^{\infty} \frac{(c\tau_k)^i}{(i+1)!} \\ &> -\frac{c\tau_k}{ce^{c\tau_k}} \sum_{i=0}^{\infty} \frac{(c\tau_k)^i}{i!} \\ &= -\tau_k, \end{aligned}$$

wobei wir ein weiteres Mal  $\tau_k > 0$  benutzt haben. □

## 4.2. Globale und lokale Konvergenzeigenschaften

In diesem Abschnitt wollen wir die globalen und lokalen Konvergenzresultate des Trust-Region-Verfahrens 4.5 präsentieren. Wir beginnen die globale Konvergenzanalyse mit dem folgenden Resultat, das für vektorwertige Funktionen beispielsweise in Geiger und Kanzow [24] zu finden ist.

**Lemma 4.11** Sei  $\{W^k\}$  eine durch Algorithmus 4.5 erzeugte Folge und  $\{W^k\}_K$  eine gegen  $W^*$  konvergente Teilfolge. Ist dann  $W^*$  kein stationärer Punkt von  $\theta$ , so gilt

$$\liminf_{k \rightarrow \infty, k \in K} \Delta_k > 0.$$



## 4.2. Globale und lokale Konvergenzeigenschaften

**Beweis:** Wir definieren die Indexmenge

$$\bar{K} := \{k - 1 \mid k \in K\}.$$

Dann konvergiert die Teilfolge  $\{W^{k+1}\}_{k \in \bar{K}}$  gegen  $W^*$ . Wir müssen nun zeigen, dass

$$\liminf_{k \rightarrow \infty, k \in \bar{K}} \Delta_{k+1} > 0 \quad (4.16)$$

gilt. Angenommen, die Ungleichung (4.16) ist nicht erfüllt. Durch eventuellen Übergang auf eine Teilfolge können wir daher annehmen, dass

$$\lim_{k \rightarrow \infty, k \in \bar{K}} \Delta_{k+1} = 0 \quad (4.17)$$

gilt. In Hinblick auf die Aufdatierungsregeln für den Trust-Region-Radius in Schritt (S.3) (man beachte, dass die untere Schranke  $\Delta_{\min}$  hier entscheidend eingeht) impliziert dies, dass alle Iterationen  $k \in \bar{K}$  für hinreichend großes  $k$  nicht erfolgreich sind. Insbesondere ist der Newton-Schritt (S.2) für hinreichend großes  $k$  nicht erfolgreich. Es existieren daher unendlich viele Trust-Region-Schritte mit  $W^k = W^{k+1}$  und

$$r_k < \rho_1 \quad (4.18)$$

für alle  $k \in \bar{K}$  mit  $k \geq k_0$  für ein genügend großes  $k_0 \in \mathbb{N}$ . Da  $\{W^{k+1}\}_{k \in \bar{K}}$  nach Voraussetzung gegen  $W^*$  konvergiert, impliziert dies außerdem, dass die Teilfolge  $\{W^k\}_{k \in \bar{K}}$  ebenfalls gegen  $W^*$  konvergiert. Vielmehr gilt  $W^* = W^{k_0}$  und somit  $\tau_* \neq 0$ . Wegen  $\Delta_{k+1} = \sigma_1 \Delta_k$  erhalten wir aus (4.17)

$$\lim_{k \rightarrow \infty, k \in \bar{K}} \Delta_k = 0. \quad (4.19)$$

Da der Grenzwert nach Voraussetzung kein stationärer Punkt der Funktion  $\theta$  ist, existiert eine Konstante  $\beta_1 > 0$  mit

$$\|\nabla\theta(W^k)\| \geq \beta_1 \quad (4.20)$$

für alle  $k \in \bar{K}$  groß genug. Da die Funktion  $\Theta$  stetig differenzierbar in einer Umgebung von  $W^*$  ist (beachte  $\tau_* > 0$ ), existiert weiter eine Konstante  $\beta_2 > 0$  mit

$$\|\nabla\Theta(W^k)\|^2 \leq \beta_2 \quad (4.21)$$

für alle  $k \in \bar{K}$  groß genug. Benutzen wir jetzt Lemma 4.8, so erhalten wir aus (4.19)–(4.21) für alle  $k \in \bar{K}$  hinreichend groß

$$\theta(W^k) - Q_k(\alpha_k \Delta \hat{W}^k) \geq \frac{\alpha_k}{2} \|\nabla\theta(W^k)\| \min \left\{ \Delta_k, \frac{\|\nabla\theta(W^k)\|}{\|\nabla\Theta(W^k)\|^2} \right\}$$

#### 4. Trust-Region-Verfahren

$$\begin{aligned}
&\geq \frac{\alpha_k}{2} \beta_1 \min \left\{ \Delta_k, \frac{\beta_1}{\beta_2} \right\} \\
&= \frac{\alpha_k}{2} \beta_1 \Delta_k \\
&\geq \frac{\alpha_k}{2} \beta_1 \|\Delta \hat{W}^k\|.
\end{aligned}$$

Da  $\theta$  wegen Lemma 4.2 stetig differenzierbar ist, existiert für jedes  $k \in \mathbb{N}$  ein Vektor

$$V^k = W^k + \mu_k \Delta \hat{W}^k, \quad \mu_k \in (0, \alpha_k),$$

so dass

$$\theta(W^k + \alpha_k \Delta \hat{W}^k) = \theta(W^k) + \nabla \theta(V^k)(\alpha_k \Delta \hat{W}^k). \quad (4.22)$$

Offensichtlich gilt  $\{V^k\}_{k \in \bar{K}} \rightarrow W^*$ , da  $\|\Delta \hat{W}^k\| \leq \Delta_k$  und die Folge  $\mu_k$  beschränkt ist. Es folgt dann aus (4.20)–(4.22)

$$\begin{aligned}
&|r_k - 1| \\
&= \left| \frac{\theta(W^k) - \theta(W^k + \alpha_k \Delta \hat{W}^k)}{\theta(W^k) - Q_k(\alpha_k \Delta \hat{W}^k)} - 1 \right| \\
&= \left| \frac{Q_k(\alpha_k \Delta \hat{W}^k) - \theta(W^k + \alpha_k \Delta \hat{W}^k)}{\theta(W^k) - Q_k(\alpha_k \Delta \hat{W}^k)} \right| \\
&= \frac{\left| \nabla \theta(W^k)(\alpha_k \Delta \hat{W}^k) + \frac{1}{2} \|\nabla \theta(W^k)(\alpha_k \Delta \hat{W}^k)\|^2 - \nabla \theta(V^k)(\alpha_k \Delta \hat{W}^k) \right|}{\theta(W^k) - Q_k(\alpha_k \Delta \hat{W}^k)} \\
&\leq \frac{2}{\alpha_k \beta_1 \|\Delta \hat{W}^k\|} \left| (\nabla \theta(W^k) - \nabla \theta(V^k))(\alpha_k \Delta \hat{W}^k) + \frac{1}{2} \|\nabla \theta(W^k)(\alpha_k \Delta \hat{W}^k)\|^2 \right| \\
&\leq \frac{2}{\beta_1 \|\Delta \hat{W}^k\|} \left( \|\nabla \theta(W^k) - \nabla \theta(V^k)\| \|\Delta \hat{W}^k\| + \frac{\alpha_k}{2} \|\nabla \theta(W^k)\|^2 \|\Delta \hat{W}^k\|^2 \right) \\
&\leq \frac{1}{\beta_1} (2 \|\nabla \theta(W^k) - \nabla \theta(V^k)\| + \alpha_k \beta_2 \|\Delta \hat{W}^k\|) \rightarrow 0
\end{aligned}$$

für  $k \rightarrow \infty$ ,  $k \in \bar{K}$ . Dies zeigt  $r_k \rightarrow 1$ , ein Widerspruch zu (4.18).  $\square$

Das folgende Lemma ist eine wichtige Konsequenz aus Lemma 4.11.

**Lemma 4.12** *Sei  $\{W^k\}$  eine durch Algorithmus 4.5 erzeugte Folge. Dann gibt es unendlich viele erfolgreiche Iterationen.*

**Beweis:** Angenommen, die Anzahl der erfolgreichen Iterationen ist endlich. Dann existiert ein Index  $k_0$ , so dass  $r_k < \rho_1$  und  $W^k = W^{k+1}$  für alle  $k \geq k_0$  gilt. Dies bedeutet aber  $W^k \rightarrow W^{k_0}$  und  $\Delta_k \rightarrow 0$ . Wegen  $\nabla \theta(W^{k_0}) \neq 0$

(ansonsten hätte der Algorithmus gestoppt) ist dies ein Widerspruch zu Lemma 4.11.  $\square$

Wir sind nun in der Lage, das globale Konvergenzverhalten von Algorithmus 4.5 anzugeben und zu beweisen.

**Satz 4.13** *Sei Voraussetzung 3.4 erfüllt und  $\{W^k\}$  eine durch Algorithmus 4.5 erzeugte Folge. Dann ist jeder Häufungspunkt von  $\{W^k\}$  ein stationärer Punkt von  $\theta$  und daher wegen Lemma 4.3 eine Lösung der Optimalitätsbedingungen (4.1).*

**Beweis:** Sei  $W^*$  ein Häufungspunkt der Folge  $\{W^k\}$  und  $\{W^k\}_{k \in K}$  eine gegen  $W^*$  konvergente Teilfolge. Dann gilt  $W^{k+1} = W^k$  für alle nicht erfolgreichen Iterationen. Da es unendlich viele erfolgreiche Iterationen gibt, können wir ohne Beschränkung der Allgemeinheit annehmen, dass alle Iterationen  $W^k$  mit  $k \in K$  erfolgreich sind.

Wir unterscheiden jetzt zwei Fälle: Es gibt unendlich viele erfolgreiche Prädiktor-Schritte mit  $k \in K$  oder nicht. Falls es unendlich viele erfolgreiche Prädiktor-Schritte mit  $k \in K$  gibt, können wir annehmen, dass alle  $k \in K$  erfolgreiche Newton-Schritte sind. Da die gesamte Folge  $\{\theta(W^k)\}$  monoton fällt und nach unten beschränkt ist und für einen erfolgreichen Newton-Schritt  $\theta(W^{k+1}) \leq \gamma\theta(W^k)$  gilt, folgt dann  $\theta(W^*) = 0$ . Nach Korollar 4.4 ist  $W^*$  somit ein stationärer Punkt von  $\theta$ .

Nehmen wir daher jetzt an, dass es nur endlich viele erfolgreiche Prädiktor-Schritte mit  $k \in K$  gibt und dass  $\nabla\theta(W^*) \neq 0$ . Durch eventuellen Übergang auf eine Teilfolge können wir annehmen, dass alle Iterationen  $k \in K$  Korrektor-Schritte sind. Da  $\theta$  stetig differenzierbar ist und  $\nabla\theta(W^*) \neq 0$  gilt, gibt es eine Konstante  $\beta_1 > 0$  mit

$$\|\nabla\theta(W^k)\| \geq \beta_1$$

für alle  $k \in K$ . Weiter gilt (eventuell nach einem Übergang auf eine weitere Teilfolge)  $\nabla\Theta(W^k) \rightarrow H^k$  mit einem  $H^k \in \partial\Theta(W^*)$ . Hierbei bezeichnet  $\partial\Theta(W^*)$  die verallgemeinerte Jacobi-Matrix von  $\Theta$ , vergleiche Clarke [12, Abschnitt 2.6]. Die verallgemeinerte Jacobi-Matrix ist nach [12, Proposition 2.6.2] von oben halbstetig (engl. *upper semicontinuous*). Daher existiert ein  $\beta_2 > 0$  mit

$$\|\nabla\Theta(W^k)\|^2 \leq \beta_2$$

#### 4. Trust-Region-Verfahren

für alle  $k \in K$ . Da alle Iterationen  $k \in K$  erfolgreich sind, gilt  $r_k \geq \rho_1$  für alle  $k \in K$  und daher wegen Lemma 4.8

$$\begin{aligned} \theta(W^k) - \theta(W^{k+1}) &\geq \rho_1(\theta(W^k) - Q_K(\alpha_k \Delta \hat{W}^k)) \\ &\geq \frac{\alpha_k}{2} \rho_1 \|\nabla \theta(W^k)\| \min \left\{ \Delta_k, \frac{\|\nabla \theta(W^k)\|}{\|\nabla \Theta(W^k)\|^2} \right\} \\ &\geq \frac{\alpha}{2} \rho_1 \beta_1 \min \left\{ \Delta_k, \frac{\beta_1}{\beta_2} \right\} \end{aligned} \quad (4.23)$$

für alle  $k \in K$ . Da die gesamte Folge  $\{\theta(W^k)\}$  offensichtlich monoton fallend und nach unten beschränkt ist, ist sie konvergent. Dies und (4.23) impliziert

$$\begin{aligned} \frac{1}{2} \rho_1 \beta_1 \alpha \sum_{k \in K} \min \left\{ \Delta_k, \frac{\beta_1}{\beta_2} \right\} &\leq \sum_{k \in K} (\theta(W^k) - \theta(W^{k+1})) \\ &\leq \sum_{k=0}^{\infty} (\theta(W^k) - \theta(W^{k+1})) \\ &< \infty. \end{aligned}$$

Daraus folgt  $\Delta_k \rightarrow 0$ , ein Widerspruch zu Lemma 4.11.  $\square$

Als Nächstes wollen wir zeigen, dass das Trust-Region-Verfahren 4.5 unter gewissen Voraussetzungen lokal superlinear konvergent ist. Wir benötigen dazu das folgende Lemma, welches auf Moré und Sorensen [54, Lemma 4.10] zurückgeht. Wir benutzen hier jedoch eine für uns praktischere Formulierung, wie sie beispielsweise auch in Geiger und Kanzow [24, Lemma 9.6] angegeben ist.

**Lemma 4.14 (Moré, Sorensen [54])** *Angenommen,  $W^*$  ist ein isolierter Häufungspunkt der (nicht notwendig durch Algorithmus 4.5 erzeugten) Folge  $\{W^k\}$  mit*

$$\{\|W^{k+1} - W^k\|\}_K \rightarrow 0$$

*für jede gegen  $W^*$  konvergente Teilfolge  $\{W^k\}_K$ . Dann konvergiert bereits die gesamte Folge  $\{W^k\}$  gegen  $W^*$ .*

In dem folgenden Lemma setzen wir wieder voraus, dass die Lösung  $W^* = (X^*, \lambda^*, S^*, \tau_*)$  der Optimalitätsbedingungen strikt komplementär und nicht-degeneriert ist.

## 4.2. Globale und lokale Konvergenzeigenschaften

**Lemma 4.15** *Sei Voraussetzung 3.4 erfüllt und  $W^*$  eine Lösung der Optimalitätsbedingungen, welche Voraussetzung 3.10 erfüllt. Dann existieren ein  $\delta > 0$  und eine Konstante  $\mu > 0$ , so dass*

$$\mu \|\Delta W\|^2 \leq \frac{1}{2} \|\nabla\Theta(W)(\Delta W)\|^2$$

für alle  $\Delta W$  und alle  $W = (X, \lambda, S, \tau)$  mit  $\|W - W^*\| \leq \delta$  und  $(X^2 + S^2 + 2\tau^2 I)^{1/2} \succ 0$  gilt.

**Beweis:** Angenommen, die Behauptung gilt nicht. Dann existiert für jedes  $k \in \mathbb{N}$  ein  $W^k = (X^k, \lambda^k, S^k, \tau_k)$  mit  $((X^k)^2 + (S^k)^2 + 2\tau^2 I)^{1/2} \succ 0$  und ein  $\Delta W^k$  mit

$$\{W^k\} \rightarrow W^* \quad \text{und} \quad \|\nabla\Theta(W^k)(\Delta W^k)\|^2 \leq \frac{1}{k} \|\Delta W^k\|^2.$$

Ohne Beschränkung der Allgemeinheit können wir dabei annehmen, dass  $\|\Delta W^k\| = 1$  gilt. Dann besitzt die Folge  $\{\Delta W^k\}$  eine gegen einen Häufungspunkt  $\Delta W^* \neq 0$  konvergente Teilfolge  $\{\Delta W^k\}_{k \in K}$ . Aus  $\nabla\Theta(W^k) \rightarrow \nabla\Theta(W^*)$  folgt

$$\|\nabla\Theta(W^*)(\Delta W^*)\| \leq 0.$$

In Anbetracht der Tatsache, dass die Abbildung  $\nabla\Theta(W^*)$  injektiv ist und die Voraussetzung 3.10 erfüllt, gilt  $\Delta W^* = 0$ , ein Widerspruch.  $\square$

Um die lokale Konvergenzaussage zu beweisen, benötigen wir das folgende Lemma, das auf Facchinei und Soares [21] zurückgeht und auch in Kanzow und Qi [48] gefunden werden kann.

**Lemma 4.16 (Facchinei, Soares [21])** *Sei Voraussetzung 3.4 erfüllt und  $W^*$  eine Lösung von  $\Theta(W) = 0$ , welche die Voraussetzung 3.10 erfüllt. Weiter seien  $\{W^k\}$  und  $\{\Delta W^k\}$  zwei (nicht notwendig durch Algorithmus 4.5 erzeugte) Folgen mit*

$$\{W^k\} \rightarrow W^* \quad \text{und} \quad \|W^k + \Delta W^k - W^*\| = o(\|W^k - W^*\|).$$

Dann gilt

$$\|\Theta(W^k + \Delta W^k)\| = o(\|\Theta(W^k)\|).$$

Nach diesen Vorbereitungen kommen wir jetzt zu unserem lokalen Konvergenzresultat.

**Satz 4.17** *Seien die Voraussetzungen 2.2 und 3.4 erfüllt,  $\{W^k\}$  eine durch Algorithmus 4.5 erzeugte Folge und  $W^*$  ein Häufungspunkt von  $\{W^k\}$ , welcher Voraussetzung 3.10 genügt. Dann gelten die folgenden Aussagen:*

#### 4. Trust-Region-Verfahren

(a) Die gesamte Folge  $\{W^k\}$  konvergiert gegen  $W^*$ .

(b) Die Folge  $\{W^k\}$  konvergiert superlinear gegen  $W^*$ .

**Beweis:** Sei  $W^*$  ein Häufungspunkt der Folge  $\{W^k\}$ , welcher Voraussetzung 3.10 genügt.

(a): Wir zeigen (a), indem wir die Voraussetzungen von Lemma 4.14 verifizieren. Nach Korollar 3.13 ist unter den Voraussetzungen 2.2, 3.4 und 3.10  $W^*$  die einzige Lösung der Optimalitätsbedingungen (4.1) und daher ein isolierter Häufungspunkt der Folge  $\{W^k\}$ .

Sei nun  $\{W^k\}_K$  irgendeine gegen  $W^*$  konvergente Teilfolge. Da  $W^*$  Voraussetzung 3.10 genügt, existiert wegen Lemma 4.1 aus Stetigkeitsgründen eine Konstante  $C > 0$ , so dass

$$\|\nabla\Theta(W^k)^{-1}\| \leq C \quad (4.24)$$

für alle  $k \in K$  groß genug gilt.

Ist der Prädiktor-Schritt für ein  $k \in K$  erfolgreich, so gilt

$$\|\Delta W^k\| = \|\nabla\Theta(W^k)^{-1}\Theta(W^k)\| \leq C \|\Theta(W^k)\| \quad (4.25)$$

für alle hinreichend großen  $k \in K$ .

Ist der Prädiktor-Schritt nicht erfolgreich, so gilt für die Lösung  $\Delta\hat{W}^k$  des Trust-Region-Teilproblems (4.9)

$$\begin{aligned} \theta(W^k) + \nabla\theta(W^k)(\Delta\hat{W}^k) + \frac{1}{2} \|\nabla\Theta(W^k)(\Delta\hat{W}^k)\|^2 \\ = Q_k(\Delta\hat{W}^k) \leq Q_k(0) = \theta(W^k) \end{aligned}$$

und daher

$$\frac{1}{2} \|\nabla\Theta(W^k)(\Delta\hat{W}^k)\|^2 \leq -\nabla\theta(W^k)(\Delta\hat{W}^k) \leq \|\nabla\theta(W^k)\| \|\Delta\hat{W}^k\|. \quad (4.26)$$

Aus Lemma 4.15 folgt die Existenz einer Konstanten  $\mu > 0$  mit

$$\mu \|\Delta\hat{W}^k\|^2 \leq \frac{1}{2} \|\nabla\Theta(W^k)(\Delta\hat{W}^k)\|^2$$

für alle  $k \in K$  hinreichend groß. Hieraus und aus (4.26) folgt

$$\|\Delta\hat{W}^k\| \leq \frac{1}{\mu} \|\nabla\theta(W^k)\| \quad (4.27)$$

## 4.2. Globale und lokale Konvergenzeigenschaften

für alle  $k \in K$  hinreichend groß. Setzen wir jetzt  $\Delta\tilde{W}^k := \Delta W^k$ , falls der Prädiktor-Schritt erfolgreich ist, und  $\Delta\tilde{W}^k := \Delta\hat{W}^k$  sonst, so ergibt eine Kombination der Ungleichungen (4.25) und (4.27)

$$\|\Delta\tilde{W}^k\| \leq \max \left\{ C \|\Theta(W^k)\|, \frac{1}{\mu} \|\nabla\theta(W^k)\| \right\} \quad (4.28)$$

für alle  $k \in K$  hinreichend groß. Da  $\{W^k\}_{k \in K}$  eine gegen  $W^*$  konvergente Teilfolge ist und  $W^*$  nach dem globalen Konvergenzsatz 4.13 eine Lösung der Optimalitätsbedingungen ist, gilt  $\nabla\theta(W^k) \rightarrow \nabla\theta(W^*) = 0$  nach Korollar 4.4 sowie  $\Theta(W^k) \rightarrow \Theta(W^*) = 0$ . Damit folgt aus (4.28) sofort  $\|\Delta\tilde{W}^k\| \rightarrow 0$  und somit

$$\|W^{k+1} - W^k\| \rightarrow 0$$

wegen  $\|W^{k+1} - W^k\| \leq \|\Delta\tilde{W}^k\|$ . Aussage (a) folgt dann aus Lemma 4.14.

(b): Da  $W^*$  Voraussetzung 3.10 erfüllt und (a) gilt, existiert eine Konstante  $C > 0$ , so dass (4.24) für alle  $k \in \mathbb{N}$  groß genug erfüllt ist.

Da  $W^*$  eine Lösung des nichtlinearen Gleichungssystems  $\Theta(W) = 0$  ist, erhalten wir aus (4.24),  $\{W^k\} \rightarrow W^*$  und der stetigen Differenzierbarkeit von  $\Theta$  für die Newton-Richtung  $\Delta W_N^k = -\nabla\Theta(W^k)^{-1}\Theta(W^k)$

$$\begin{aligned} & \|W^k + \Delta W_N^k - W^*\| \\ &= \|-\nabla\Theta(W^k)^{-1}(\Theta(W^k) - \Theta(W^*) - \nabla\Theta(W^k)(W^k - W^*))\| \\ &\leq C \|\Theta(W^k) - \Theta(W^*) - \nabla\Theta(W^k)(W^k - W^*)\| \\ &= o(\|W^k - W^*\|). \end{aligned}$$

Lemma 4.16 impliziert daher

$$\|\Theta(W^k + \Delta W_N^k)\| = o(\|\Theta(W^k)\|)$$

und somit

$$\theta(W^k + \Delta W_N^k) = o(\theta(W^k)).$$

Insbesondere gilt

$$\theta(W^k + \Delta W_N^k) \leq \gamma\theta(W^k)$$

für alle  $k \in \mathbb{N}$  hinreichend groß. Daher sind für alle  $k \in \mathbb{N}$  hinreichend groß der Prädiktor-Schritt erfolgreich und das Trust-Region-Verfahren 4.5 reduziert sich auf das lokale Newton-Verfahren, welches unter den gegebenen Voraussetzungen bekanntermaßen superlinear konvergent ist, vergleiche beispielsweise Geiger und Kanzow [24, Satz 9.2] und Nocedal und Wright [57, Theorem 11.2].  $\square$

### 4.3. Numerische Resultate

Auch hier haben wir, um das numerische Verhalten zu testen, das Trust-Region-Verfahren 4.5 in MATLAB implementiert. Um den Aufwand zur Erstellung der Programme zu verringern, haben wir wiederum die Datenstruktur, die Eingaberoutinen für die Testprobleme und einige lineare Algebra-Routinen aus dem SDPT<sup>3</sup>-Softwarepaket [66] benutzt.

Die Berechnung des Startpunktes erfolgt genau so wie beim Glättungsverfahren 3.1 (vergleiche Abschnitt 3.4). Die Parameter des Trust-Region-Verfahrens wurden gewählt als

$$\begin{array}{llll} c = 0.1, & \sigma_1 = 0.5, & \Delta_0 = 5, & \rho_1 = 10^{-4}, \\ \gamma = 0.8, & \sigma_2 = 2, & \Delta_{\min} = 10^{-2}, & \rho_2 = 0.75. \end{array}$$

Die numerische Lösung der Newton-Gleichung (4.8) haben wir bereits in Abschnitt 3.3 diskutiert, zur (inexakten) Lösung des Trust-Region-Teilproblems (4.9) haben wir einen so genannten *double dogleg step*, wie er beispielsweise in Dennis und Schnabel [17, Abschnitt 6.4.2] beschrieben wird, implementiert.

Nach Korollar 4.4 gilt  $\nabla\theta(W) = 0$  genau dann, wenn  $\theta(W) = 0$  ist. Wir haben daher das Abbruchkriterium in Schritt (S.1) dahingehend modifiziert, dass wir abbrechen, wenn

$$\theta(W^k) \leq n\varepsilon \quad \text{mit} \quad \varepsilon := 10^{-10}$$

gilt. Die Multiplikation von  $\varepsilon$  mit  $n$  dient wiederum dazu, dass Abbruchkriterium mehr oder weniger unabhängig von der Dimension des Problems zu machen. Man vergleiche hierzu auch die Ausführungen in Abschnitt 3.4. Ferner brechen wir ab, falls die maximale Iterationsanzahl von  $k_{\max} = 300$  erreicht wird. In diesen Fällen betrachten wir die Probleme als nicht gelöst.

In den Tabellen 4.1–4.3 sind die durchschnittlichen Iterationszahlen des Trust-Region-Verfahrens für die acht Testprobleme aus SDPT<sup>3</sup> angegeben. Diese Zahlen sind wieder jeweils aus zehn Testläufen für jede Probleminstanz berechnet worden. Da bei einer nicht erfolgreichen Iteration die Iterierte  $W^k$  sich nicht ändert und nur der Trust-Region-Radius verkleinert wird, liegt der wesentliche Rechenaufwand in den erfolgreichen Schritten. Wir geben daher auch die durchschnittliche Anzahl der erfolgreichen Iterationen an. Weiter stellt sich heraus, dass das Trust-Region-Verfahren einige Testprobleminstanzen nicht löst, so dass auch für jedes Testproblem die Anzahl der korrekt gelösten Probleme mit ausgegeben wird und die Durchschnitte nur aus den jeweils gelösten Problemen berechnet werden.

Vergleichen wir die Zahl der erfolgreichen Iterationen mit den Iterationszahlen von Algorithmus 3.1, so zeigt sich, dass das Trust-Region-Verfahren durchgehend schlechtere Ergebnisse liefert. Da auch noch einige Testproblem nicht gelöst werden, ist das Verfahren deutlich schlechter als das Prädiktor-Korrektor-Verfahren 3.1.



Tabelle 4.1.: Durchschnittliche Iterationsanzahl für semidefinite Programme kleiner Dimension.

Problem	$n$	$m$	Iterationen	erfolgreiche Iterationen	gelöste Probleme
random	10	10	8.7	8.4	10
Norm min	20	6	10.1	9.1	10
Cheby	20	11	11.9	9.0	10
Maxcut	10	10	7.9	7.2	10
ETP	20	10	42.8	32.2	10
Lovasz	10	$\approx 25$	17.6	11.6	10
LogCheby	60	6	21.0	15.3	10
ChebyC	40	11	9.8	7.5	10

Tabelle 4.2.: Durchschnittliche Iterationsanzahl für semidefinite Programme mittlerer Dimension.

Problem	$n$	$m$	Iterationen	erfolgreiche Iterationen	gelöste Probleme
random	20	20	15.0	13.7	10
Norm min	40	11	11.4	10.9	10
Cheby	40	21	12.0	9.2	10
Maxcut	21	21	11.3	9.6	10
ETP	40	20	78.3	51.8	9
Lovasz	21	$\approx 105$	51.9	43.9	10
LogCheby	120	11	40.1	24.9	10
ChebyC	80	21	10.0	7.7	10

Tabelle 4.3.: Durchschnittliche Iterationsanzahl für semidefinite Programme großer Dimension.

Problem	$n$	$m$	Iterationen	erfolgreiche Iterationen	gelöste Probleme
random	50	50	21.7	20.5	10
Norm min	100	26	12.9	12.6	10
Cheby	100	27	13.2	10.3	10
Maxcut	50	50	12.9	12.2	10
ETP	100	50	78.2	50.1	8
Lovasz	30	$\approx 220$	50.8	43.7	9
LogCheby	300	51	118.7	66.0	10
ChebyC	200	41	17.0	11.1	10

#### 4. *Trust-Region-Verfahren*

## 5. Ein nichtglattes Newton-Verfahren

In diesem Abschnitt wollen wir ein nichtglattes Newton-Verfahren zur Lösung der Optimalitätsbedingungen (2.6) auf seine lokalen Konvergenzeigenschaften untersuchen. Ziel ist es, die schon in den vorherigen Abschnitten benutzte Voraussetzung der Nichtdegeneriertheit so zu modifizieren, dass wir lokal schnelle (sprich: lokal quadratische) Konvergenz ohne strikte Komplementarität verifizieren können.

Nach dem Wissensstand des Autors ist bisher weder für die Inneren-Punkte-Methoden, wie sie zum Beispiel in [30, 3, 70, 65, 50, 56] behandelt werden, noch für die in dieser Arbeit behandelten Glättungsverfahren (vergleiche [15, 62, 42, 64]) bekannt, dass für diese lokal schnelle Konvergenz ohne strikte Komplementarität gilt. Einzige Ausnahme ist die Arbeit von Sun, Sun und Qi [63], wo lokal quadratische Konvergenz eines nichtglatten Newton-Verfahrens für semidefinite Komplementaritätsprobleme ohne die Voraussetzung der strikten Komplementarität gezeigt wird. Jedoch ist die dort benötigte Voraussetzung der positiven Definitheit einer gewissen Matrix nie erfüllt, wenn die Resultate auf den Spezialfall linearer semidefiniter Programme angewendet werden. In der revidierten Fassung [64] der Arbeit [63] benutzen die Autoren einen anderen Ansatz (welcher Ähnlichkeiten zu dem von uns gleich betrachteten aufweist) und spezialisieren die Resultate für semidefinite Programme. Sie setzen dann jedoch wieder die strikte Komplementarität im Lösungspunkt voraus, um lokal schnelle Konvergenz des betrachteten nichtglatten Newton-Verfahrens zu beweisen.

### 5.1. Algorithmus

Wir wollen daher jetzt die lokale Konvergenz eines nichtglatten Newton-Verfahrens zur Lösung der Optimalitätsbedingungen (2.6) betrachten. Wie in Kapitel 2 beschrieben, werden wir die Optimalitätsbedingungen in ein nichtglattes Gleichungssystem überführen. Dazu betrachten wir in diesem Abschnitt ausschließlich die in (2.14) definierte Minimum-Funktion, welche für  $X, S \in \mathcal{S}^{n \times n}$  definiert ist durch

$$\phi(X, S) = X + S - ((X - S)^2)^{1/2}.$$

Wir werden jetzt, da wir ein nichtglattes Newton-Verfahren betrachten wollen, nur diese Funktion selber und nicht ihre geglättete Approximation betrachten.

## 5. Ein nichtglattes Newton-Verfahren

Um im Folgenden die Formeln zu vereinfachen, definieren wir die matrixwertige Betragsfunktion  $|\cdot| : \mathcal{S}^{n \times n} \rightarrow \mathcal{S}^{n \times n}$  durch

$$|A| := (A^2)^{1/2}.$$

Diese Definition stimmt auf den reellen Zahlen  $\mathbb{R}$  mit der üblichen Definition des Absolutbetrages überein. Mit dieser Funktion lässt sich die Minimum-Funktion schreiben als

$$\phi(X, S) := X + S - |X - S|. \quad (5.1)$$

Die Minimum-Funktion hat eine Reihe interessanter Eigenschaften, welche wir größtenteils bereits in Kapitel 2 bewiesen haben. Diese sind in dem folgenden Resultat zusammengefasst.

**Lemma 5.1** *Sei  $\phi$  wie in (5.1) definiert. Dann gelten die folgenden Aussagen für beliebige symmetrische Matrizen  $X, S \in \mathcal{S}^{n \times n}$ :*

(a)  $\phi$  erfüllt die Äquivalenz

$$\phi(X, S) = 0 \iff X \succeq 0, S \succeq 0, XS = 0. \quad (5.2)$$

(b) (Sun, Sun [62]) Die Matrix  $E = |X - S| = ((X - S)^2)^{1/2}$  ist regulär (oder, äquivalent, positiv definit) genau dann, wenn die Abbildung  $(A, B) \mapsto |A - B|$  stetig differenzierbar im Punkt  $(A, B) = (X, S)$  ist.

(c) Die Matrix  $E := |X - S|$  ist genau dann regulär, wenn  $\phi$  im Punkt  $(X, S)$  stetig differenzierbar (im Sinne von Fréchet) ist. In diesem Fall gilt

$$\nabla \phi(X, S)(U, V) = U + V - L_E^{-1}[(X - S)(U - V) + (U - V)(X - S)].$$

**Beweis:** Die Aussage (a) wurde bereits in Lemma 2.12 gezeigt. Teil (b) folgt aus Sun und Sun [62]. Teil (c) folgt aus (b), wobei wir die Formel für die Ableitung bereits in Satz 2.27 gezeigt haben.  $\square$

Wir benutzen die Minimum-Funktion jetzt wieder, um die Optimalitätsbedingungen (2.6) in ein nichtglattes Gleichungssystem

$$\Theta(X, \lambda, S) = 0 \quad (5.3)$$

zu überführen, wobei  $\Theta : \mathcal{S}^{n \times n} \times \mathbb{R}^m \times \mathcal{S}^{n \times n} \rightarrow \mathcal{S}^{n \times n} \times \mathbb{R}^m \times \mathcal{S}^{n \times n}$  definiert ist durch

$$\Theta(X, \lambda, S) := \begin{pmatrix} \sum_{i=1}^m \lambda_i A_i + S - C \\ A_i \bullet X - b_i \quad (i = 1, \dots, m) \\ \phi(X, S) \end{pmatrix}.$$

Wenden wir ein nichtglattes Newton-Verfahren auf das System (5.3) an, so erhalten wir eine Iterationsvorschrift der Form

$$W^{k+1} := W^k + \Delta W^k, \quad k = 0, 1, 2, \dots$$

wobei wir wiederum die Abkürzung

$$W^k := (X^k, \lambda^k, S^k)$$

benutzt haben und  $\Delta W^k$  eine Lösung der linearen Gleichung  $H^k \Delta W^k = -\Theta(W^k)$  ist. Dabei ist  $H^k$  ein Element des B-Subdifferentials von  $\Theta$  in  $W^k$ , d.h.  $H^k \in \partial_B \Theta(W^k)$ , wobei

$$\partial_B \Theta(W) := \{H \mid \exists \{W^k\} \subset D_\Theta : W^k \rightarrow W, \nabla \Theta(W^k) \rightarrow H\}$$

und  $D_\Theta$  die Menge von Punkten  $W$  bezeichnet, in denen  $\Theta$  differenzierbar ist, vergleiche Clarke [12]. Man beachte, dass die Menge  $\partial_B \Theta(W)$  der Subgradienten immer nichtleer ist, da  $\Theta$  lokal Lipschitz-stetig ist. Weiter gilt  $\partial_B \Theta(W^k) = \{\nabla \Theta(W^k)\}$  in allen stetig differenzierbaren Punkten  $W^k$  von  $\Theta$ .

Wir geben jetzt das nichtglatte Newton-Verfahren an, welches wir im Folgenden genauer untersuchen wollen.

### Algorithmus 5.2

(S.0) Wähle  $W^0 := (X^0, \lambda^0, S^0) \in \mathcal{S}^{n \times n} \times \mathbb{R}^m \times \mathcal{S}^{n \times n}$ ,  $\varepsilon \geq 0$  und setze  $k := 0$ .

(S.1) Ist  $\|\Theta(W^k)\| \leq \varepsilon$ : STOPP.

(S.2) Wähle  $H^k \in \partial_B \Theta(W^k)$  und berechne  $\Delta W^k = (\Delta X^k, \Delta \lambda^k, \Delta S^k)$  als Lösung des linearen Systems  $H^k \Delta W = -\Theta(W^k)$ .

(S.3) Setze  $W^{k+1} := W^k + \Delta W^k$ ,  $k \leftarrow k + 1$  und gehe zu Schritt (S.1).

Man beachte, dass die Lösung  $\Delta W^k = (\Delta X^k, \Delta \lambda^k, \Delta S^k)$  des linearen Systems in Schritt (S.2) automatisch symmetrische Matrizen  $\Delta X^k$  und  $\Delta S^k$  liefert (vergleiche hierzu auch die Ausführungen in Abschnitt 3.3).

Die lokalen Konvergenzeigenschaften dieses Verfahrens sind in dem folgenden Satz wiedergegeben. Im Beweis wird die Eigenschaft, dass die auftretenden Funktionen stark halbglatt (engl.: *strongly semismooth*) sind, verwendet. Da wir diese Eigenschaft nur an dieser Stelle benötigen, verzichten wir auf eine formale Definition und eine Darstellung der Eigenschaften und verweisen stattdessen auf die im Beweis angegebene Literatur.

## 5. Ein nichtglattes Newton-Verfahren

**Satz 5.3** Sei  $W^* := (X^*, \lambda^*, S^*)$  eine Lösung der Optimalitätsbedingungen (2.6) und seien alle Elemente  $H^* \in \partial_B \Theta(W^*)$  des B-Subdifferentials von  $\Theta$  im Punkt  $W^*$  regulär. Dann ist der Algorithmus 5.2 lokal quadratisch konvergent.

**Beweis:** Nach Sun und Sun [62, Theorem 4.13] ist die Minimum-Funktion  $\phi$  stark halbglatt (vergleiche auch [60, 59] für die Definition und einige Eigenschaften von stark halbglatten Funktionen). Daher ist  $\Theta$  stark halbglatt. Die lokal quadratische Konvergenz folgt dann aus Theorem 3.1 in Qi [59].  $\square$

Mit Blick auf Satz 5.3 ist das Ziel dieses Kapitels, geeignete Bedingungen zu formulieren, unter denen alle Elemente des B-Subdifferentials  $\partial_B \Theta(W^*)$  regulär sind. Dies geschieht in Unterabschnitt 5.3 nach einigen vorhergehenden Bemerkungen im folgenden Unterabschnitt.

### 5.2. Matrix-Vektor-Formulierung der Newton-Systeme

In diesem gesamten Unterabschnitt werden wir annehmen, dass die Funktion  $\Theta$  im aktuellen Punkt  $(X, \lambda, S) \in \mathcal{S}^{n \times n} \times \mathbb{R}^m \times \mathcal{S}^{n \times n}$  stetig differenzierbar ist. Nach Lemma 5.1 (b), (c) garantiert diese Voraussetzung, dass die Matrix  $X - S$  regulär ist. Im Unterabschnitt 5.3 werden wir diese Voraussetzung wieder fallen lassen.

Da  $\Theta$  im Punkt  $(X, \lambda, S)$  stetig differenzierbar ist, ist im nichtglatten Newton-Verfahren zur Lösung von  $\Theta(X, \lambda, S) = 0$  im aktuellen Punkt die lineare Gleichung

$$\nabla \Theta(X, \lambda, S)(\Delta X, \Delta \lambda, \Delta S) = -\Theta(X, \lambda, S) \quad (5.4)$$

zu lösen. Für unsere Betrachtungen in Unterabschnitt 5.3 ist es nützlich, dieses lineare System in der üblichen Matrix-Vektor-Formulierung zu schreiben. Wir wollen daher jetzt eine entsprechende Formulierung herleiten.

Mit  $\Theta$  ist auch die Minimum-Funktion  $\phi$  im Punkt  $(X, \lambda, S)$  stetig differenzierbar. Daher folgt aus Lemma 5.1 (c), dass die Matrix

$$E := |X - S| = ((X - S)^2)^{1/2} \quad (5.5)$$

positiv definit, also insbesondere regulär ist. Somit ist der zugehörige Lyapunov-Operator  $L_E$  invertierbar. Definieren wir jetzt die Residuen

$$\begin{aligned} R_C &:= C - \sum_{j=1}^m \lambda_j A_j - S, \\ r_{b,i} &:= b_i - A_i \bullet X, \quad i = 1, \dots, m, \\ r_b &:= (r_{b,1}, \dots, r_{b,m})^T, \end{aligned}$$

## 5.2. Matrix-Vektor-Formulierung der Newton-Systeme

so kann das Newton-System (5.4) geschrieben werden als

$$\sum_{j=1}^m \Delta \lambda_j A_j + \Delta S = R_C, \quad (5.6)$$

$$A_i \bullet \Delta X = r_{b,i}, \quad i = 1, \dots, m, \quad (5.7)$$

$$\nabla \phi(X, S)(\Delta X, \Delta S) = -\phi(X, S). \quad (5.8)$$

Um dieses System im üblichen Matrix-Vektor-Format zu schreiben, müssen wir Matrizen in Vektoren überführen. Für eine beliebige (nicht zwingend symmetrische) Matrix  $A \in \mathbb{R}^{n \times n}$  kann dies mit Hilfe des Operators  $\text{vec} : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{n^2}$  durchgeführt werden, welcher definiert ist durch

$$\text{vec}(A) := (a_{11}, a_{21}, \dots, a_{n1}, a_{12}, a_{22}, \dots, a_{n2}, \dots, a_{nn})^T \in \mathbb{R}^{n^2}.$$

Der Operator  $\text{vec}$  schreibt also einfach die Spalten der Matrix  $A$  untereinander in einen langen Vektor der Länge  $n^2$ . Im Falle einer symmetrischen Matrix benötigen wir jedoch nicht alle Elemente, da die nötige Information bereits im unteren linken (bzw. oberen rechten) Dreieck der Matrix enthalten ist. Es genügt daher, nur die untere, linke Hälfte von  $A$  zu betrachten. Der entsprechende Operator  $\text{svec} : \mathcal{S}^{n \times n} \rightarrow \mathbb{R}^{\frac{n(n+1)}{2}}$  ist definiert durch

$$\text{svec}(A) := (a_{11}, \sqrt{2}a_{21}, \dots, \sqrt{2}a_{n1}, a_{22}, \sqrt{2}a_{32}, \dots, \sqrt{2}a_{n2}, \dots, a_{nn})^T.$$

Der Faktor  $\sqrt{2}$  vor den Nichtdiagonalelementen hat den Grund, dass  $\text{svec}$  dann konsistent mit den inneren Produkten ist, d.h. dass

$$A \bullet B = \text{svec}(A)^T \text{svec}(B) \quad \forall A, B \in \mathcal{S}^{n \times n} \quad (5.9)$$

gilt.

Die hier gemachte Definition von  $\text{svec}$  stammt, ebenso wie die folgenden Definitionen des Kronecker- bzw. symmetrischen Kronecker-Produkts, aus den Arbeiten [3, 65] und hat sich in der Literatur als Standard durchgesetzt. Es sei jedoch erwähnt, dass es auf die Reihenfolge, in der die Elemente in den Vektor geschrieben werden, eigentlich nicht ankommt. So könnte man  $\text{svec}(A)$  auch definieren als

$$(a_{11}, \sqrt{2}a_{21}, a_{22}, \sqrt{2}a_{31}, \sqrt{2}a_{32}, a_{33}, \dots, \sqrt{2}a_{n(n-1)}, a_{nn})^T$$

Diese Definition hat aus praktischer Sicht den Vorteil, dass sie sich einfacher auf Blockdiagonalmatrizen verallgemeinern lässt. Aus diesem Grund wird diese Anordnung der Elemente auch in den neueren Versionen der SDPT<sup>3</sup>-Software von Toh, Todd und Tütüncü [66] verwendet. Für theoretische Untersuchungen ist die Anordnung aber unerheblich und wir bleiben daher bei der in der Literatur üblichen Definition.

## 5. Ein nichtglattes Newton-Verfahren

Nach der Einführung der Operatoren  $\text{vec}$  und  $\text{svec}$  stellt sich als Nächstes die Frage, wie ein gewöhnliches Matrixprodukt mit Hilfe von  $\text{vec}$  und  $\text{svec}$  ausgedrückt werden kann. Hierzu definieren wir für zwei (nicht notwendig symmetrische) Matrizen  $G, K \in \mathbb{R}^{n \times n}$  das so genannte *Kronecker-Produkt* durch

$$G \otimes K := [g_{ij}K] \in \mathbb{R}^{n^2 \times n^2}.$$

Dann kann die Identität

$$(G \otimes K) \text{vec}(H) = \text{vec}(KHG^T), \quad H \in \mathbb{R}^{n \times n}$$

leicht verifiziert werden. Analog definieren wir das *symmetrische Kronecker-Produkt* durch

$$(G \otimes_s K) \text{svec}(H) := \frac{1}{2} \text{svec}(KHG^T + GHK^T), \quad H \in \mathcal{S}^{n \times n}. \quad (5.10)$$

Einige Eigenschaften des symmetrischen Kronecker-Produkts sind in dem folgenden Lemma zusammengefasst. Die Beweise dafür sind elementar. Die Aussagen (a) bis (d) können in [3, 65] gefunden werden, während Teil (e) im Wesentlichen nur eine Umformulierung von Aussage (d) ist und sich mit den Rechenregeln aus [3, 65] leicht beweisen lässt.

**Lemma 5.4** *Seien  $G, K \in \mathbb{R}^{n \times n}$ . Dann hat das durch (5.10) definierte symmetrische Kronecker-Produkt  $\otimes_s$  die folgenden Eigenschaften:*

- (a)  $G \otimes_s K = K \otimes_s G$ .
- (b)  $(G \otimes_s K)^T = K^T \otimes_s G^T$ .
- (c) Sind  $G$  und  $K$  symmetrisch positiv definit, so auch  $G \otimes_s K$ .
- (d) Sind  $G, K$  zwei vertauschbare symmetrische Matrizen mit Eigenwerten  $\sigma_1, \dots, \sigma_n$  und  $\mu_1, \dots, \mu_n$ , und sei  $q_1, \dots, q_n$  eine gemeinsame Menge von orthonormalen Eigenvektoren, dann sind die  $n(n+1)/2$  Eigenwerte von  $G \otimes_s K$  gegeben durch

$$\frac{1}{2}(\sigma_i \mu_j + \mu_i \sigma_j), \quad 1 \leq j \leq i \leq n$$

mit dazugehörigen orthonormalen Eigenvektoren  $v_{ij}$ ,  $1 \leq j \leq i \leq n$ , welche durch

$$v_{ij} := \begin{cases} \text{svec}(q_i q_i^T), & \text{falls } i = j, \\ \frac{1}{\sqrt{2}} \text{svec}(q_i q_j^T + q_j q_i^T), & \text{falls } j < i \end{cases}$$

definiert sind.



## 5.2. Matrix-Vektor-Formulierung der Newton-Systeme

(e) Sind  $G, K$  zwei vertauschbare symmetrische Matrizen mit simultanen Spektralzerlegungen  $G = QD_GQ^T$  und  $K = QD_KQ^T$  mit einer orthogonalen Matrix  $Q \in \mathbb{R}^{n \times n}$  und Diagonalmatrizen  $D_G, D_K \in \mathbb{R}^{n \times n}$ , dann gilt  $G \otimes_s K = (Q \otimes_s Q)(D_G \otimes_s D_K)(Q \otimes_s Q)^T$ . Insbesondere ist  $Q \otimes_s Q$  eine orthogonale Matrix.

Wir betrachten nun das Newton-System (5.4), d.h. wir betrachten das durch die Gleichungen (5.6)–(5.8) gegebene System. Die Umformulierung der ersten beiden Gleichungen (5.6) und (5.7) ist elementar und geschieht (unter Beachtung von (5.9)) genauso wie in [65]. Dies resultiert in den Gleichungen

$$\mathcal{A}^T \Delta \lambda + \text{svec}(\Delta S) = \text{svec}(R_C) \quad (5.11)$$

und

$$\mathcal{A} \text{svec}(\Delta X) = r_b, \quad (5.12)$$

wobei

$$\mathcal{A} := (\text{svec}(A_1), \dots, \text{svec}(A_m))^T \in \mathbb{R}^{m \times \frac{n(n+1)}{2}}. \quad (5.13)$$

Es bleibt also die Umformulierung der dritten Blockzeile (5.8). Benutzen wir Lemma 5.1 (b) und die Definition von  $E$  in (5.5), so kann (5.8) geschrieben werden als

$$\Delta X + \Delta S - L_E^{-1}[(X - S)(\Delta X - \Delta S) + (\Delta X - \Delta S)(X - S)] = -\phi(X, S). \quad (5.14)$$

Wie in Abschnitt 3.3 beschrieben, ist dies äquivalent zu

$$L_{A_E}[\Delta X] + L_{B_E}[\Delta S] = -L_E[\phi(X, S)]. \quad (5.15)$$

mit

$$A_E := E - (X - S) \quad \text{und} \quad B_E := E + (X - S). \quad (5.16)$$

Anwendung von  $\frac{1}{2} \text{svec}$  auf beiden Seiten von (5.15) liefert dann

$$\frac{1}{2} \text{svec}(L_{A_E}[\Delta X]) + \frac{1}{2} \text{svec}(L_{B_E}[\Delta S]) = -\frac{1}{2} \text{svec}(L_E[\phi(X, S)]).$$

Aus der Definition (5.10) ergibt sich weiter für alle symmetrischen Matrizen  $A, H \in \mathcal{S}^{n \times n}$  die Identität

$$\frac{1}{2} \text{svec}(L_A[H]) = \frac{1}{2} \text{svec}(AHI + IHA) = (I \otimes_s A) \text{svec}(H).$$

Setzen wir daher

$$\mathcal{E} := I \otimes_s A_E \quad \text{und} \quad \mathcal{F} := I \otimes_s B_E, \quad (5.17)$$

## 5. Ein nichtglattes Newton-Verfahren

so erhalten wir

$$\mathcal{E} \operatorname{svec}(\Delta X) + \mathcal{F} \operatorname{svec}(\Delta S) = -(I \otimes_s E) \operatorname{svec}(\phi(X, S)). \quad (5.18)$$

Da  $E$  und  $I$  beide positiv definit sind, folgt aus Lemma 5.4 (c), dass die Matrix  $(I \otimes_s E)$  positiv definit und somit regulär ist. Somit kann (5.18) umgeschrieben werden zu

$$(I \otimes_s E)^{-1} \mathcal{E} \operatorname{svec}(\Delta X) + (I \otimes_s E)^{-1} \mathcal{F} \operatorname{svec}(\Delta S) = -\operatorname{svec}(\phi(X, S)). \quad (5.19)$$

Zusammenfassend ergibt sich aus (5.11), (5.12) sowie (5.19) das folgende Resultat.

**Satz 5.5** *Sei  $(X, \lambda, S)$  gegeben, so dass  $X - S$  regulär ist. Dann ist der Vektor  $(\Delta X, \Delta \lambda, \Delta S) \in \mathcal{S}^{n \times n} \times \mathbb{R}^m \times \mathcal{S}^{n \times n}$  genau dann eine Lösung des Newton-Systems (5.4), wenn der Vektor  $(\operatorname{svec}(\Delta X), \Delta \lambda, \operatorname{svec}(\Delta S))$  eine Lösung des linearen Gleichungssystems*

$$\begin{pmatrix} 0 & \mathcal{A}^T & \mathcal{I} \\ \mathcal{A} & 0 & 0 \\ (I \otimes_s E)^{-1} \mathcal{E} & 0 & (I \otimes_s E)^{-1} \mathcal{F} \end{pmatrix} \begin{pmatrix} \operatorname{svec}(\Delta X) \\ \Delta \lambda \\ \operatorname{svec}(\Delta S) \end{pmatrix} = \begin{pmatrix} \operatorname{svec}(R_C) \\ r_b \\ -\operatorname{svec}(\phi(X, S)) \end{pmatrix} \quad (5.20)$$

ist.

Das lineare Gleichungssystem (5.20) hat eine starke Ähnlichkeit zu dem in Zusammenhang mit Inneren-Punkte-Methoden auftretenden System in [65]. Man beachte jedoch, dass die Matrizen  $\mathcal{E}$  und  $\mathcal{F}$  hier eine andere Bedeutung haben.

Wie bereits in Abschnitt 3.3 angekündigt, wollen wir einen kurzen Blick auf die praktische Lösbarkeit des linearen Gleichungssystems (5.20) werfen. Unter Ausnutzung der Blockstruktur ergibt eine einfache Rechnung das lineare Gleichungssystem

$$\mathcal{A} \mathcal{E}^{-1} \mathcal{F} \mathcal{A} \Delta \lambda = r_b + \mathcal{A} \mathcal{E}^{-1} ((I \otimes_s E) \operatorname{svec}(\phi(X, S)) + \mathcal{F} \operatorname{svec}(R_C))$$

zur Berechnung von  $\Delta \lambda \in \mathbb{R}^m$ . Wie in [44] gezeigt wurde, ist die dabei auftretende Koeffizientenmatrix  $\mathcal{A} \mathcal{E}^{-1} \mathcal{F} \mathcal{A}$  identisch mit der Matrix  $M = (m_{ij})$  aus (3.37). Die Suchrichtungen  $\Delta X$  und  $\Delta S$  lassen sich dann aus den Formeln

$$\operatorname{svec}(\Delta S) = \operatorname{svec}(R_C) - \mathcal{A}^T \Delta \lambda$$

sowie

$$\operatorname{svec}(\Delta X) = -\mathcal{E}^{-1} ((I \otimes_s E) \operatorname{svec}(\phi(X, S)) + \mathcal{F} \operatorname{svec}(\Delta S))$$

berechnen. Der wesentliche Aufwand bei dieser Vorgehensweise ist die Berechnung der Inversen  $\mathcal{E}^{-1}$  der im Allgemeinen vollbesetzten Matrix  $\mathcal{E} = I \otimes_s A_E$ .

### 5.3. Lokale Konvergenz

Sei nun  $(X^*, \lambda^*, S^*)$  eine Lösung der Optimalitätsbedingungen (2.6). Um die folgende Herleitung zu motivieren, nehmen wir zunächst an, dass die folgenden Bedingungen, welche schon aus der Konvergenztheorie der vorhergehenden Abschnitte bekannt sind, erfüllt sind. Insbesondere nehmen wir an, dass die Lösung strikt komplementär ist.

**Voraussetzung 5.6** Sei  $(X^*, \lambda^*, S^*)$  eine Lösung der Optimalitätsbedingungen (2.6).

(A.1) *(Lineare Unabhängigkeit)*

Die Matrizen  $A_1, \dots, A_m$  sind linear unabhängig.

(A.2) *(Nichtdegeneriertheit)*

Die folgende Implikation gilt für jedes  $(\Delta X, \Delta \lambda, \Delta S) \in \mathcal{S}^{n \times n} \times \mathbb{R}^m \times \mathcal{S}^{n \times n}$ :

$$\left. \begin{array}{l} \sum_{i=1}^m \Delta \lambda_i A_i + \Delta S = 0, \\ A_i \bullet \Delta X = 0 \quad (i = 1, \dots, m), \\ X^* \Delta S + \Delta X S^* = 0 \end{array} \right\} \implies (\Delta X, \Delta S) = (0, 0).$$

(A.3) *(Strikte Komplementarität)*

$X^* + S^* \succ 0$ .

Diese Bedingungen sind Standard-Voraussetzungen, um lokal schnelle Konvergenz für viele Algorithmen zur Lösung von semidefiniten Programmen zu zeigen. Man vergleiche auch Abschnitt 3.2.

Als Nächstes geben wir die entsprechende Matrix-Vektor-Formulierung dieser Bedingungen an.

**Lemma 5.7** Sei  $(X^*, \lambda^*, S^*)$  eine Lösung der Optimalitätsbedingungen (2.6). Dann gelten die folgenden Aussagen.

(a) Voraussetzung (A.1) ist äquivalent dazu, dass die Matrix  $\mathcal{A}$  aus (5.13) vollen Rang hat.

(b) Voraussetzung (A.2) ist für  $(\Delta X, \Delta \lambda, \Delta S) \in \mathcal{S}^{n \times n} \times \mathbb{R}^m \times \mathcal{S}^{n \times n}$  äquivalent zu der Implikation

$$\left. \begin{array}{l} \mathcal{A}^T \Delta \lambda + \text{svec}(\Delta S) = 0, \\ \mathcal{A} \text{svec}(\Delta X) = 0, \\ (I \otimes_s X^*) \text{svec}(\Delta S) + \dots \\ (I \otimes_s S^*) \text{svec}(\Delta X) = 0 \end{array} \right\} \implies \begin{cases} \text{svec}(\Delta X) = 0, \\ \text{svec}(\Delta S) = 0. \end{cases}$$

## 5. Ein nichtglattes Newton-Verfahren

(c) Voraussetzung (A.2) ist für  $(\Delta X, \Delta \lambda, \Delta S) \in \mathcal{S}^{n \times n} \times \mathbb{R}^m \times \mathcal{S}^{n \times n}$  äquivalent zu der Implikation

$$\left. \begin{aligned} \mathcal{A}^T \Delta \lambda + \text{svec}(\Delta S) &= 0, \\ \mathcal{A} \text{svec}(\Delta X) &= 0, \\ V^*(I \otimes_s D_{S^*})(V^*)^T \text{svec}(\Delta X) + \dots \\ V^*(I \otimes_s D_{X^*})(V^*)^T \text{svec}(\Delta S) &= 0 \end{aligned} \right\} \implies \begin{cases} \text{svec}(\Delta X) = 0, \\ \text{svec}(\Delta S) = 0. \end{cases}$$

Hier bezeichnen  $X^* = QD_{X^*}Q^T$  und  $S^* = QD_{S^*}Q^T$  die simultane Spektralzerlegung der zwei kommutierenden Matrizen  $X^*$ ,  $S^*$  und  $V^* := Q \otimes_s Q$ .

**Beweis:** (a) Dies folgt direkt aus der Definition der Matrix  $\mathcal{A}$ .

(b) Anwendung des svec-Operators liefert sofort

$$\sum_{i=1}^m \Delta \lambda_i A_i + \Delta S = 0 \iff \mathcal{A}^T \Delta \lambda + \text{svec}(\Delta S) = 0$$

und

$$A_i \bullet \Delta X = 0, \quad i = 1, \dots, m \iff \mathcal{A} \text{svec}(\Delta X) = 0.$$

Es bleibt also noch die Äquivalenz

$$X^* \Delta S + \Delta X S^* = 0 \iff (I \otimes_s X^*) \text{svec}(\Delta S) + (I \otimes_s S^*) \text{svec}(\Delta X) = 0 \quad (5.21)$$

zu zeigen. Dazu nehmen wir zunächst an, dass  $X^* \Delta S + \Delta X S^* = 0$  gilt. Unter Benutzung der Tatsache, dass  $\Delta X$  und  $\Delta S$  symmetrisch sind, ergibt Transponieren dieser Gleichung  $\Delta S X^* + S^* \Delta X = 0$ . Addition dieser beiden Gleichungen und Anwendung von  $\frac{1}{2} \text{svec}$  auf die resultierende Gleichung sowie Benutzung von (5.10) ergibt die Formulierung auf der rechten Seite von (5.21). Umgekehrt sei

$$(I \otimes_s X^*) \text{svec}(\Delta S) + (I \otimes_s S^*) \text{svec}(\Delta X) = 0.$$

Mit Blick auf (5.10) ist dies äquivalent zu

$$X^* \Delta S + \Delta S X^* + S^* \Delta X + \Delta X S^* = 0. \quad (5.22)$$

Nach Lemma 3.11 ist dies jedoch wieder äquivalent zu  $X^* \Delta S + \Delta X S^* = 0$ . Somit gilt (5.21).

(c) Die Äquivalenz der ersten beiden Gleichungen wurde schon in Teil (b) gezeigt. Wir müssen daher nur noch zeigen, dass

$$X^* \Delta S + \Delta X S^* = 0$$

genau dann gilt, wenn

$$V^*(I \otimes_s D_{S^*})(V^*)^T \text{svec}(\Delta X) + V^*(I \otimes_s D_{X^*})(V^*)^T \text{svec}(\Delta S) = 0 \quad (5.23)$$

ist. Sei dazu zunächst  $X^*\Delta S + \Delta X S^* = 0$ . Unter Benutzung der Symmetrie von  $\Delta X$  und  $\Delta S$  ergibt Transponieren dieser Gleichung und anschließende Addition

$$X^*\Delta S + \Delta X S^* + S^*\Delta X + \Delta X S^* = 0. \quad (5.24)$$

Multiplikation von links mit  $Q^T$  und von rechts mit  $Q$  sowie Anwendung von  $\frac{1}{2} \text{svec}$  auf beiden Seiten liefert

$$\begin{aligned} 0 &= \frac{1}{2} \text{svec} [D_{X^*} Q^T \Delta S Q + Q^T \Delta S Q D_{X^*} + D_{S^*} Q^T \Delta X Q + Q^T \Delta X Q D_{S^*}] \\ &= (I \otimes_s D_{X^*}) \text{svec}(Q^T \Delta S Q) + (I \otimes_s D_{S^*}) \text{svec}(Q^T \Delta X Q) \\ &\stackrel{(5.10)}{=} (I \otimes_s D_{X^*})(Q^T \otimes_s Q^T) \text{svec}(\Delta S) + (I \otimes_s D_{S^*})(Q^T \otimes_s Q^T) \text{svec}(\Delta X) \\ &= (I \otimes_s D_{X^*})(Q \otimes_s Q)^T \text{svec}(\Delta S) + (I \otimes_s D_{S^*})(Q \otimes_s Q)^T \text{svec}(\Delta X) \\ &= (I \otimes_s D_{X^*})(V^*)^T \text{svec}(\Delta S) + (I \otimes_s D_{S^*})(V^*)^T \text{svec}(\Delta X). \end{aligned}$$

Multiplikation mit der orthogonalen und somit regulären Matrix  $V^*$  liefert somit (5.23).

Sei nun umgekehrt (5.23) erfüllt. Dann ergibt sich aus den vorhergehenden Überlegungen die Gleichung (5.24), da wir von da an nur Äquivalenzumformungen vorgenommen haben. Mit Lemma 3.11 folgt daher wiederum  $X^*\Delta S + \Delta X S^* = 0$  und somit die behauptete Äquivalenz.  $\square$

Teil (b) von Lemma 5.7 werden wir in dem unten stehenden Regularitätsergebnis benutzen, während Teil (c) hier angegeben wurde, um einen besseren Vergleich zwischen der Bedingung (A.2) und einer später einzuführenden Voraussetzung zu haben.

Wir zeigen als Nächstes, wie das vorhergehende Lemma genutzt werden kann, um die Regularität der Matrix aus Satz 5.5 in einer strikt komplementären Lösung zu zeigen. Der folgende Satz ist nichts anderes als die vektorwertige Formulierung von Satz 3.12. Diese Formulierung ist jedoch für unsere folgenden Betrachtungen praktischer.

**Satz 5.8** *Sei  $(X^*, \lambda^*, S^*)$  eine Lösung der Optimalitätsbedingungen (2.6), welche Voraussetzungen (A.1)–(A.3) erfüllt. Weiter definiere die Matrizen (vergleiche auch (5.16), (5.17))*

$$E^* := |X^* - S^*|,$$

5. Ein nichtglattes Newton-Verfahren

$$\begin{aligned}\mathcal{E}^* &:= I \otimes_s A_{E^*} \quad \text{mit} \quad A_{E^*} := E^* - (X^* - S^*), \\ \mathcal{F}^* &:= I \otimes_s B_{E^*} \quad \text{mit} \quad B_{E^*} := E^* + (X^* - S^*).\end{aligned}$$

Dann existiert  $(I \otimes_s E^*)^{-1}$  und die Matrix

$$\begin{pmatrix} 0 & \mathcal{A}^T & \mathcal{I} \\ \mathcal{A} & 0 & 0 \\ (I \otimes_s E^*)^{-1} \mathcal{E}^* & 0 & (I \otimes_s E^*)^{-1} \mathcal{F}^* \end{pmatrix}$$

(vergleiche (5.20)) ist regulär.

**Beweis:** Da Voraussetzung (A.3) gilt, ist die Matrix  $E^*$  positiv definit. Daher ist  $I \otimes_s E^*$  nach Lemma 5.4 (c) ebenfalls positiv definit. Somit existiert die Inverse  $(I \otimes_s E^*)^{-1}$  und wir sind fertig, wenn wir zeigen, dass die Matrix

$$\begin{pmatrix} 0 & \mathcal{A}^T & \mathcal{I} \\ \mathcal{A} & 0 & 0 \\ \mathcal{E}^* & 0 & \mathcal{F}^* \end{pmatrix}$$

regulär ist. Sei dazu  $(\text{svec}(\Delta X), \Delta\lambda, \text{svec}(\Delta S))$  mit

$$\begin{pmatrix} 0 & \mathcal{A}^T & \mathcal{I} \\ \mathcal{A} & 0 & 0 \\ \mathcal{E}^* & 0 & \mathcal{F}^* \end{pmatrix} \begin{pmatrix} \text{svec}(\Delta X) \\ \Delta\lambda \\ \text{svec}(\Delta S) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

gegeben. Blockweise lässt sich dies schreiben als

$$\begin{aligned}\mathcal{A}^T \Delta\lambda + \text{svec}(\Delta S) &= 0, \\ \mathcal{A} \text{svec}(\Delta X) &= 0, \\ \mathcal{E}^* \text{svec}(\Delta X) + \mathcal{F}^* \text{svec}(\Delta S) &= 0.\end{aligned}\tag{5.25}$$

Wegen  $X^* S^* = 0$  kommutieren die beiden Matrizen  $X^*$ ,  $S^*$ . Folglich existieren eine orthogonale Matrix  $Q^* \in \mathbb{R}^{n \times n}$  und zwei Diagonalmatrizen  $D_{X^*}, D_{S^*} \succeq 0$  mit

$$X^* = Q^* D_{X^*} (Q^*)^T \quad \text{und} \quad S^* = Q^* D_{S^*} (Q^*)^T.\tag{5.26}$$

Wegen

$$X^* S^* = 0 \iff D_{X^*} D_{S^*} = 0$$

gilt

$$\begin{aligned}E^* &= |X^* - S^*| \\ &= ((X^* - S^*)^2)^{1/2}\end{aligned}$$

$$\begin{aligned}
 &= Q^* ((D_{X^*} - D_{S^*})^2)^{1/2} (Q^*)^T \\
 &= Q^* (D_{X^*} + D_{S^*}) (Q^*)^T \\
 &= X^* + S^*.
 \end{aligned}$$

Daraus folgt

$$A_{E^*} = E^* - (X^* - S^*) = 2S^* \quad \text{und} \quad B_{E^*} = E^* + (X^* - S^*) = 2X^*$$

und somit

$$\mathcal{E}^* = I \otimes_s A_{E^*} = 2(I \otimes_s S^*) \quad \text{und} \quad \mathcal{F}^* = I \otimes_s B_{E^*} = 2(I \otimes_s X^*).$$

Daher kann die letzte Blockzeile von (5.25) umgeschrieben werden zu

$$(I \otimes_s X^*) \text{svec}(\Delta S) + (I \otimes_s S^*) \text{svec}(\Delta X) = 0.$$

Zusammen mit den ersten beiden Zeilen von (5.25) folgt aus Voraussetzung (A.2) und Lemma 5.7 (b), dass  $\text{svec}(\Delta X) = 0$  und  $\text{svec}(\Delta S) = 0$  gilt. Dies impliziert  $\mathcal{A}^T \Delta \lambda = 0$  nach (5.25). Da die Spalten von  $\mathcal{A}^T$  nach Voraussetzung 5.6 (a) linear unabhängig sind, folgt hieraus  $\Delta \lambda = 0$ .  $\square$

Falls Voraussetzung (A.3) im Punkt  $W^* = (X^*, \lambda^*, S^*)$  nicht gilt, so ist die Matrix  $E^*$  nicht mehr positiv definit. Es gibt daher keinen Grund, dass die Matrix  $I \otimes_s E^*$  nichtsingulär ist. Es ist trotzdem möglich, die vorhergehende Diskussion auf diesen Fall zu erweitern. Hierzu sei daran erinnert, dass jedes Element  $H^* \in \partial_B \Theta(W^*)$  als Grenzwert  $H^* = \lim_{k \rightarrow \infty} \nabla \Theta(W^k)$  dargestellt werden kann, wobei  $W^k := (X^k, \lambda^k, S^k)$  eine gegen  $W^*$  konvergente Teilfolge ist, so dass  $\Theta$  in jedem Punkt  $W^k$  differenzierbar ist. Wegen Lemma 5.1 (b), (c) ist  $\Theta$  genau dann stetig differenzierbar im Punkt  $W^k$ , wenn  $X^k - S^k$  regulär ist. Aber genau diese Voraussetzung haben wir für die Aussagen in Abschnitt 5.2 gemacht. Mit der Notation (siehe auch (5.5), (5.16), (5.17))

$$\begin{aligned}
 E^k &:= |X^k - S^k|, \\
 \mathcal{E}^k &:= I \otimes_s A_{E^k} \quad \text{mit} \quad A_{E^k} := E^k - (X^k - S^k), \\
 \mathcal{F}^k &:= I \otimes_s B_{E^k} \quad \text{mit} \quad B_{E^k} := E^k + (X^k - S^k)
 \end{aligned}$$

ergibt sich daher aus Satz 5.5, dass die übliche Matrix-Vektor-Formulierung der Jacobi-Matrix  $\nabla \Theta(W^k)$  durch

$$\begin{pmatrix} 0 & \mathcal{A}^T & \mathcal{I} \\ \mathcal{A} & 0 & 0 \\ (I \otimes_s E^k)^{-1} \mathcal{E}^k & 0 & (I \otimes_s E^k)^{-1} \mathcal{F}^k \end{pmatrix} \quad (5.27)$$

## 5. Ein nichtglattes Newton-Verfahren

gegeben ist. Wir müssen jetzt die möglichen Grenzelemente von Matrizen dieser Bauart finden. Dafür müssen wir die letzte Blockzeile genauer betrachten und somit das Konvergenzverhalten der zwei Matrizen

$$(I \otimes_s E^k)^{-1} \mathcal{E}^k = (I \otimes_s E^k)^{-1} (I \otimes_s A_{E^k})$$

und

$$(I \otimes_s E^k)^{-1} \mathcal{F}^k = (I \otimes_s E^k)^{-1} (I \otimes_s B_{E^k})$$

untersuchen. Sei dazu

$$X^k - S^k = Q^k \Pi^k (Q^k)^T \quad \text{mit} \quad \Pi^k = \text{diag}(\pi_1^k, \dots, \pi_n^k)$$

eine Spektralzerlegung der symmetrischen Matrix  $X^k - S^k$  und

$$|\Pi^k| = \text{diag}(|\pi_1^k|, \dots, |\pi_n^k|).$$

Dann folgt sofort

$$\begin{aligned} E^k &= Q^k |\Pi^k| (Q^k)^T, \\ A_{E^k} &= Q^k (|\Pi^k| - \Pi^k) (Q^k)^T, \\ B_{E^k} &= Q^k (|\Pi^k| + \Pi^k) (Q^k)^T. \end{aligned}$$

Sei  $q_i^k$  die  $i$ -te Spalte von  $Q^k$  und definiere die orthogonale Matrix  $V^k$  durch

$$V^k := (\dots, v_{ij}^k, \dots)_{1 \leq j \leq i \leq n},$$

wobei die Spalten  $v_{ij}^k \in \mathbb{R}^{n(n+1)/2}$  durch

$$v_{ij}^k := \begin{cases} \text{svec}(q_i^k (q_i^k)^T), & \text{falls } i = j, \\ \frac{1}{\sqrt{2}} \text{svec}(q_i^k (q_j^k)^T + q_j^k (q_i^k)^T), & \text{falls } j < i \end{cases}$$

gegeben sind, vergleiche Lemma 5.4. Dann liefert Lemma 5.4 (d)

$$\begin{aligned} I \otimes_s E^k &= V^k \text{diag} \left( \dots, \frac{1}{2} (|\pi_i^k| + |\pi_j^k|), \dots \right) (V^k)^T, \\ I \otimes_s A_{E^k} &= V^k \text{diag} \left( \dots, \frac{1}{2} (|\pi_i^k| + |\pi_j^k| - \pi_i^k - \pi_j^k), \dots \right) (V^k)^T, \\ I \otimes_s B_{E^k} &= V^k \text{diag} \left( \dots, \frac{1}{2} (|\pi_i^k| + |\pi_j^k| + \pi_i^k + \pi_j^k), \dots \right) (V^k)^T. \end{aligned}$$



Man beachte, dass alle Diagonalelemente  $\pi_i^k$  der Matrix  $\Pi^k$  von Null verschieden sind, da die Regularität von  $E^k$  vorausgesetzt wurde. Daher gilt  $|\pi_i^k| + |\pi_j^k| \neq 0$  für alle  $1 \leq j \leq i \leq n$  und wir erhalten

$$\begin{aligned} (I \otimes_s E^k)^{-1}(I \otimes_s A_{E^k}) &= V^k \operatorname{diag} \left( \dots, \frac{|\pi_i^k| + |\pi_j^k| - \pi_i^k - \pi_j^k}{|\pi_i^k| + |\pi_j^k|}, \dots \right) (V^k)^T \\ &=: V^k \Sigma_-^k (V^k)^T, \\ (I \otimes_s E^k)^{-1}(I \otimes_s B_{E^k}) &= V^k \operatorname{diag} \left( \dots, \frac{|\pi_i^k| + |\pi_j^k| + \pi_i^k + \pi_j^k}{|\pi_i^k| + |\pi_j^k|}, \dots \right) (V^k)^T \\ &=: V^k \Sigma_+^k (V^k)^T. \end{aligned}$$

Mit diesen Umformulierungen können wir die Matrix (5.27) schreiben als

$$\begin{pmatrix} 0 & \mathcal{A}^T & \mathcal{I} \\ \mathcal{A} & 0 & 0 \\ V^k \Sigma_-^k (V^k)^T & 0 & V^k \Sigma_+^k (V^k)^T \end{pmatrix}.$$

Durch Übergang auf eine Teilfolge können wir ohne Beschränkung der Allgemeinheit annehmen, dass die orthogonale (und daher beschränkte) Matrixfolge  $\{Q^k\}$  (und somit auch  $\{\Pi^k\}$ ) konvergiert mit

$$Q^* := \lim_{k \rightarrow \infty} Q^k \quad \text{und} \quad \Pi^* := \lim_{k \rightarrow \infty} \Pi^k.$$

Dann ist  $Q^*$  wieder eine orthogonale Matrix, welche die Identität  $X^* - S^* = Q^* \Pi^* (Q^*)^T$  erfüllt, d.h.  $Q^*$ ,  $\Pi^*$  korrespondieren zu einer Spektralzerlegung der symmetrischen Matrix  $X^* - S^*$ . Die entsprechende Teilfolge von  $V^k$  konvergiert gegen

$$V^* := (\dots, v_{ij}^*, \dots)_{1 \leq j \leq i \leq n}, \quad (5.28)$$

wobei

$$v_{ij}^* := \begin{cases} \operatorname{svec}(q_i^* (q_i^*)^T), & \text{falls } i = j, \\ \frac{1}{\sqrt{2}} \operatorname{svec}(q_i^* (q_j^*)^T + q_j^* (q_i^*)^T), & \text{falls } j < i \end{cases} \quad (5.29)$$

und  $q_i^*$  die  $i$ -te Spalte von  $Q^*$  bezeichnet, vergleiche Lemma 5.4 (d). Weiter ist leicht einzusehen, dass alle Diagonalelemente der Matrizen  $\Sigma_-^k$  und  $\Sigma_+^k$  in dem Intervall  $[0, 2]$  liegen und dass  $\Sigma_-^k + \Sigma_+^k = 2I$  gilt. Wir können also wiederum ohne Beschränkung der Allgemeinheit annehmen, dass

$$\Sigma_-^k \rightarrow \Sigma_-^* \quad \text{und} \quad \Sigma_+^k \rightarrow \Sigma_+^*$$

für zwei Diagonalmatrizen  $\Sigma_-^*$ ,  $\Sigma_+^*$  gilt. Daraus folgt

$$\begin{aligned} (I \otimes_s E^k)^{-1}(I \otimes_s A_{E^k}) &\rightarrow V^* \Sigma_-^* (V^*)^T, \\ (I \otimes_s E^k)^{-1}(I \otimes_s B_{E^k}) &\rightarrow V^* \Sigma_+^* (V^*)^T. \end{aligned} \quad (5.30)$$

## 5. Ein nichtglattes Newton-Verfahren

Um die genaue Struktur der Diagonalmatrizen  $\Sigma_-^*$  und  $\Sigma_+^*$  zu untersuchen, schreiben wir  $\Pi^* = \text{diag}(\pi_1^*, \dots, \pi_n^*)$  und definieren die Indexmengen

$$\begin{aligned}\alpha &:= \{i \mid \pi_i^* > 0\} = \{i \mid \lambda_i(X^*) > 0, \lambda_i(S^*) = 0\}, \\ \beta &:= \{i \mid \pi_i^* = 0\} = \{i \mid \lambda_i(X^*) = 0, \lambda_i(S^*) = 0\}, \\ \gamma &:= \{i \mid \pi_i^* < 0\} = \{i \mid \lambda_i(X^*) = 0, \lambda_i(S^*) > 0\},\end{aligned}\tag{5.31}$$

wobei  $\lambda_i(X^*)$  bzw.  $\lambda_i(S^*)$  die Eigenwerte der Matrix  $X^*$  bzw.  $S^*$  in geeigneter Reihenfolge bezeichnet. Dann zeigt eine einfache Rechnung

$$\frac{|\pi_i^k| + |\pi_j^k| - \pi_i^k - \pi_j^k}{|\pi_i^k| + |\pi_j^k|} \rightarrow \sigma_{ij}^- \quad \text{mit} \quad \sigma_{ij}^- \begin{cases} > 0, & \text{falls } i \in \gamma \text{ oder } j \in \gamma, \\ \in [0, 2], & \text{falls } i \in \beta \text{ und } j \in \beta, \\ 0, & \text{sonst} \end{cases}$$

und analog

$$\frac{|\pi_i^k| + |\pi_j^k| + \pi_i^k + \pi_j^k}{|\pi_i^k| + |\pi_j^k|} \rightarrow \sigma_{ij}^+ \quad \text{mit} \quad \sigma_{ij}^+ \begin{cases} > 0, & \text{falls } i \in \alpha \text{ oder } j \in \alpha, \\ \in [0, 2], & \text{falls } i \in \beta \text{ und } j \in \beta, \\ 0, & \text{sonst.} \end{cases}$$

Insbesondere gilt

$$\begin{aligned}\sigma_{ij}^- &= 0 \text{ falls } (i, j) \in (\alpha \times \alpha) \cup (\alpha \times \beta) \cup (\beta \times \alpha), \\ \sigma_{ij}^+ &= 0 \text{ falls } (i, j) \in (\beta \times \gamma) \cup (\gamma \times \beta) \cup (\gamma \times \gamma).\end{aligned}$$

Weiter sei bemerkt, dass die Grenzwerte zwar nicht eindeutig bestimmt sind für alle Indexpaare  $(i, j) \in \beta \times \beta$ , dass aber wenigstens eines der beiden Elemente  $\sigma_{ij}^-$ ,  $\sigma_{ij}^+$  immer positiv ist.

Zusammenfassend ergibt sich damit das folgende Resultat.

**Satz 5.9** *Sei  $(X^*, \lambda^*, S^*)$  eine Lösung der Optimalitätsbedingungen (2.6) und  $X^* - S^* = Q^* \Pi^* (Q^*)^T$  eine Spektralzerlegung von  $X^* - S^*$ . Weiter seien die Indexmengen  $\alpha$ ,  $\beta$ ,  $\gamma$  wie in (5.31) definiert. Dann kann jedes Element des B-Subdifferentials  $\partial_B \Theta(X^*, \lambda^*, S^*)$  in der üblichen Matrix-Vektor-Schreibweise geschrieben werden als*

$$\begin{pmatrix} 0 & \mathcal{A}^T & \mathcal{I} \\ \mathcal{A} & 0 & 0 \\ V^* \Sigma_-^* (V^*)^T & 0 & V^* \Sigma_+^* (V^*)^T \end{pmatrix}.$$

Dabei ist  $V^*$  die Matrix aus (5.28), (5.29) und die Diagonalmatrizen

$$\Sigma_-^* = \text{diag}(\dots, \sigma_{ij}^-, \dots)_{1 \leq j \leq i \leq n},$$

$$\Sigma_+^* = \text{diag}(\dots, \sigma_{ij}^+, \dots)_{1 \leq j \leq i \leq n}$$

haben die folgenden Eigenschaften:

$$\begin{aligned} \Sigma_-^* \succeq 0, \quad \Sigma_+^* \succeq 0, \quad \Sigma_-^* + \Sigma_+^* &= 2I, \\ \sigma_{ij}^- &= 0 \text{ falls } (i, j) \in (\alpha \times \alpha) \cup (\alpha \times \beta) \cup (\beta \times \alpha), \\ \sigma_{ij}^+ &= 0 \text{ falls } (i, j) \in (\beta \times \gamma) \cup (\gamma \times \beta) \cup (\gamma \times \gamma). \end{aligned}$$

Durch diese Beobachtungen motiviert, führen wir als Nächstes eine leicht modifizierte Nichtdegeneriertheitsbedingung ein.

**Voraussetzung 5.10** Seien  $(X^*, \lambda^*, S^*)$  eine Lösung der Optimalitätsbedingungen (2.6),  $X^* - S^* = Q^* \Pi^* (Q^*)^T$  eine Spektralzerlegung von  $X^* - S^*$ ,  $\alpha$ ,  $\beta$ ,  $\gamma$  die Indexmengen aus (5.31) und die zugehörige Matrix  $V^*$  wie in (5.28) und (5.29) definiert.

(A.4) Für beliebige Diagonalmatrizen

$$\begin{aligned} \Sigma_- &= \text{diag}(\dots, \sigma_{ij}^-, \dots)_{1 \leq j \leq i \leq n}, \\ \Sigma_+ &= \text{diag}(\dots, \sigma_{ij}^+, \dots)_{1 \leq j \leq i \leq n} \end{aligned}$$

mit

$$\begin{aligned} \Sigma_- \succcurlyeq 0, \quad \Sigma_+ \succcurlyeq 0, \quad \Sigma_- + \Sigma_+ \succ 0, \\ \sigma_{ij}^- &= 0 \text{ falls } (i, j) \in (\alpha \times \alpha) \cup (\alpha \times \beta) \cup (\beta \times \alpha), \\ \sigma_{ij}^+ &= 0 \text{ falls } (i, j) \in (\beta \times \gamma) \cup (\gamma \times \beta) \cup (\gamma \times \gamma) \end{aligned}$$

gilt die folgende Implikation für jeden Vektor  $(\text{svec}(\Delta X), \Delta \lambda, \text{svec}(\Delta S))$ :

$$\left. \begin{aligned} \mathcal{A}^T \Delta \lambda + \text{svec}(\Delta S) &= 0 \\ \mathcal{A} \text{svec}(\Delta X) &= 0, \\ V^* \Sigma_- (V^*)^T \text{svec}(\Delta X) + \dots & \\ V^* \Sigma_+ (V^*)^T \text{svec}(\Delta S) &= 0 \end{aligned} \right\} \implies \begin{cases} \text{svec}(\Delta X) = 0, \\ \text{svec}(\Delta S) = 0. \end{cases}$$

Um den Unterschied zwischen Voraussetzung (A.4) und der Nichtdegeneriertheit (A.2) einzusehen, bemerken wir zunächst einmal, dass die genauen Werte der Diagonalmatrizen  $\Sigma_-$  und  $\Sigma_+$  aus Voraussetzung (A.4) uninteressant sind. Es ist einzig wichtig, ob ein gegebenes Diagonalelement Null ist oder nicht. Dies und Lemma 5.7 (c) zeigt, dass Voraussetzung (A.2) mit  $\Sigma_- = I \otimes_s D_{S^*}$  und  $\Sigma_+ = I \otimes_s D_{X^*}$  in Voraussetzung (A.4) enthalten ist. Dabei haben wir die Notation aus Lemma 5.7 benutzt. Ist jedoch die Indexmenge  $\beta$  nichtleer, so kann man leicht einsehen, dass die Voraussetzung  $\Sigma_- + \Sigma_+ \succ 0$

## 5. Ein nichtglattes Newton-Verfahren

in diesem Fall nicht für alle Diagonalelemente mit  $(i, j) \in \beta \times \beta$  gilt. Dies ist der Hauptunterschied zwischen der modifizierten Nichtdegeneriertheit (A.4) und der Nichtdegeneriertheitsvoraussetzung (A.2). Voraussetzung (A.4) ist also eine stärkere Voraussetzung als (A.2). Jedoch können wir in diesem Fall lokal quadratische Konvergenz ohne strikte Komplementarität beweisen. Dies ist die wichtigste Konsequenz aus dem folgenden Satz.

**Satz 5.11** *Sei  $W^* = (X^*, \lambda^*, S^*)$  eine Lösung der Optimalitätsbedingungen (2.6), welche die Voraussetzungen (A.1) und (A.4) erfüllt. Dann sind alle Elemente  $H^* \in \partial_B \Theta(W^*)$  invertierbar.*

**Beweis:** Nach Konstruktion haben alle Elemente  $H^* \in \partial_B \Theta(W^*)$  in der üblichen Matrix-Vektor-Schreibweise eine Matrixform wie in Satz 5.9. Die beiden Diagonalmatrizen  $\Sigma_-^*$  und  $\Sigma_+^*$  aus Satz 5.9 erfüllen die Voraussetzung (A.4), so dass wir analog zu dem Beweis zu Satz 5.8 erhalten, dass die Matrix unter den Voraussetzungen (A.1) und (A.4) nicht singulär ist.  $\square$

Als direkte Konsequenz aus den Sätzen 5.3 und 5.11 erhalten wir die folgende Aussage.

**Satz 5.12** *Sei  $(X^*, \lambda^*, S^*)$  eine Lösung der Optimalitätsbedingungen (2.6), welche den Voraussetzungen (A.1) und (A.4) genügen. Dann ist Algorithmus 5.2 lokal quadratisch konvergent.*

Das nächste Resultat folgt ebenfalls aus Satz 5.11 zusammen mit Proposition 2.5 aus [59], siehe auch Korollar 3.13.

**Korollar 5.13** *Sei Voraussetzung 2.2 erfüllt und  $(X^*, \lambda^*, S^*)$  eine Lösung der Optimalitätsbedingungen (2.6), welche den Voraussetzungen (A.1) und (A.4) genügt. Dann ist  $(X^*, \lambda^*, S^*)$  die eindeutig bestimmte Lösung der Optimalitätsbedingungen (2.6).*

## 5.4. Numerische Beispiele

Es stellt sich nun die Frage, ob es tatsächlich Beispiele gibt, die die modifizierte Nichtdegeneriertheitsvoraussetzung (A.4) erfüllen und ob man tatsächlich lokal schnelle Konvergenz erhält. Wir werden daher die in diesem Kapitel entwickelte Theorie an zwei kleinen Beispielen illustrieren. Das erste Beispiel stammt von Alizadeh, Haerberly und Overton [2] und stellt ein semidefinites Programm mit einer eindeutigen Lösung dar, welche nicht strikt komplementär ist. Es zeigt sich jedoch, dass die modifizierte Nichtdegeneriertheitsbedingung (A.4) im Lösungspunkt erfüllt ist.

**Beispiel 5.14** Sei  $n = 3$ ,  $m = 3$  mit  $b = (1, 0, 0)^T$  und

$$C = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad A_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Dann ist Voraussetzung (A.1) offensichtlich erfüllt. Weiter hat das zugehörige semidefinite Programm die eindeutige Lösung

$$X^* = \text{diag}(1, 0, 0), \quad \lambda^* = (0, 0, 0)^T, \quad S^* = \text{diag}(0, 0, 1),$$

welche nicht die Bedingung (A.3) der strikten Komplementarität erfüllt. Jedoch ist die so genannte *primale-duale Nichtdegeneriertheit* von Alizadeh, Haeberly und Overton [2] erfüllt, woraus die Existenz und Eindeutigkeit der Lösung folgt, vergleiche [2]. Wir wollen jetzt nachrechnen, dass Voraussetzung (A.4) erfüllt ist. Dazu sei

$$X^* - S^* = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} = Q^* \Pi^* (Q^*)^T \quad \text{mit} \quad Q^* := I \in \mathbb{R}^{3 \times 3}$$

eine Spektralzerlegung von  $X^* - S^*$ . Dann gilt

$$V^* = (v_{11}^*, v_{21}^*, v_{31}^*, v_{22}^*, v_{32}^*, v_{33}^*) = I \in \mathbb{R}^{6 \times 6}.$$

Seien  $\Sigma_-$ ,  $\Sigma_+$  zwei wie in Voraussetzung (A.4) gegebene Diagonalmatrizen. Die Indextmengen aus (5.31) sind gegeben durch  $\alpha = \{1\}$ ,  $\beta = \{2\}$ ,  $\gamma = \{3\}$  und es folgt

$$\begin{aligned} \Sigma_- &= \text{diag}(0, 0, \sigma_{31}^-, \sigma_{22}^-, \sigma_{32}^-, \sigma_{33}^-), \\ \Sigma_+ &= \text{diag}(\sigma_{11}^+, \sigma_{21}^+, \sigma_{31}^+, \sigma_{22}^+, 0, 0) \end{aligned} \tag{5.32}$$

mit gewissen reellen Zahlen, welche die folgenden Bedingungen erfüllen:

$$\begin{aligned} \sigma_{31}^-, \sigma_{32}^-, \sigma_{33}^- &> 0, \\ \sigma_{11}^+, \sigma_{21}^+, \sigma_{31}^+ &> 0, \\ \sigma_{22}^-, \sigma_{22}^+ &\geq 0, \quad \sigma_{22}^- + \sigma_{22}^+ > 0. \end{aligned} \tag{5.33}$$

Nun seien  $\Delta X, \Delta S \in \mathcal{S}^{n \times n}$  und  $\Delta \lambda \in \mathbb{R}^m$  mit

$$\begin{aligned} \mathcal{A}^T \Delta \lambda + \text{svec}(\Delta S) &= 0, \quad \mathcal{A} \text{svec}(\Delta X) = 0, \\ V^* \Sigma_- (V^*)^T \text{svec}(\Delta X) + V^* \Sigma_+ (V^*)^T \text{svec}(\Delta S) &= 0, \end{aligned} \tag{5.34}$$

wobei

$$\mathcal{A} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \sqrt{2} & 1 & 0 & 0 \\ 0 & \sqrt{2} & 0 & 0 & 0 & 1 \end{pmatrix},$$

## 5. Ein nichtglattes Newton-Verfahren

vergleiche (5.13). Dann ist das System (5.34) äquivalent zu

$$\begin{aligned}
\Delta\lambda_1 + \Delta S_{11} &= 0, & \Delta\lambda_3 + \Delta S_{21} &= 0, \\
\Delta\lambda_2 + \Delta S_{31} &= 0, & \Delta\lambda_2 + \Delta S_{22} &= 0, \\
\Delta S_{32} &= 0, & \Delta\lambda_3 + \Delta S_{33} &= 0, \\
\Delta X_{11} &= 0, & 2\Delta X_{31} + \Delta X_{22} &= 0, \\
2\Delta X_{21} + \Delta X_{33} &= 0, & \sigma_{11}^+ \Delta S_{11} &= 0, \\
\sigma_{21}^+ \Delta S_{21} &= 0, & \sigma_{31}^- \Delta X_{31} + \sigma_{31}^+ \Delta S_{31} &= 0, \\
\sigma_{22}^- \Delta X_{22} + \sigma_{22}^+ \Delta S_{22} &= 0, & \sigma_{32}^- \Delta X_{32} &= 0, \\
\sigma_{33}^- \Delta X_{33} &= 0.
\end{aligned}$$

Es folgt nun sofort

$$\begin{aligned}
\Delta X_{11} &= \Delta X_{21} = \Delta X_{32} = \Delta X_{33} = 0, \\
\Delta S_{11} &= \Delta S_{21} = \Delta S_{32} = \Delta S_{33} = 0, \\
\Delta\lambda_1 &= \Delta\lambda_3 = 0,
\end{aligned}$$

so dass wir nur noch  $\Delta X_{31} = \Delta X_{22} = 0$  und  $\Delta S_{31} = \Delta S_{22} = 0$  als Konsequenz aus

$$\Delta\lambda_2 + \Delta S_{31} = 0, \quad \Delta\lambda_2 + \Delta S_{22} = 0, \quad (5.35)$$

$$2\Delta X_{31} + \Delta X_{22} = 0, \quad \sigma_{31}^- \Delta X_{31} + \sigma_{31}^+ \Delta S_{31} = 0, \quad (5.36)$$

$$\sigma_{22}^- \Delta X_{22} + \sigma_{22}^+ \Delta S_{22} = 0 \quad (5.37)$$

zeigen müssen. Aus (5.35) erhalten wir  $\Delta S_{31} = \Delta S_{22}$ . Dies, (5.36), (5.37) sowie  $\sigma_{31}^+ > 0$  impliziert

$$\begin{aligned}
0 &= \sigma_{22}^- \Delta X_{22} + \sigma_{22}^+ \Delta S_{22} \\
&= -2\sigma_{22}^- \Delta X_{31} + \sigma_{22}^+ \Delta S_{31} \\
&= -2\sigma_{22}^- \Delta X_{31} - \frac{\sigma_{31}^-}{\sigma_{31}^+} \sigma_{22}^+ \Delta X_{31} \\
&= - \left( 2\sigma_{22}^- + \frac{\sigma_{31}^-}{\sigma_{31}^+} \sigma_{22}^+ \right) \Delta X_{31}.
\end{aligned}$$

Multiplikation mit  $-\sigma_{31}^+ \neq 0$  liefert daher

$$\underbrace{(2\sigma_{31}^+ \sigma_{22}^- + \sigma_{31}^- \sigma_{22}^+)}_{>0 \text{ nach (5.33)}} \Delta X_{31} = 0.$$

Daraus folgt  $\Delta X_{31} = 0$  und somit  $\Delta X = \Delta S = 0$ . Also ist Voraussetzung (A.4) erfüllt.

Tabelle 5.1.: Numerische Ergebnisse für Beispiel 5.14

$k$	rel. Dualitätslücke	$\ \Theta(W^k)\ $	$\ W^k - W^*\ $
0	0.000000e+00	8.833707e-01	9.106836e-01
1	5.377397e-02	1.779061e-01	1.965437e-01
2	3.438035e-05	5.372091e-03	5.556943e-03
3	3.270881e-08	7.046664e-05	1.814441e-04
4	2.105543e-09	4.864869e-06	1.260400e-05
5	1.043839e-11	2.860393e-07	7.432387e-07

Um zu sehen, dass wir wirklich lokal schnelle Konvergenz erhalten, wenden wir das Prädiktor-Korrektor-Newton-Verfahren aus Kapitel 3 auf dieses Beispiel an. Dieses Verfahren ist, grob gesprochen, eine globalisierte Variante von Algorithmus 5.2. Die zugehörigen numerischen Resultate sind in Tabelle 5.1 angegeben. Die Spalten enthalten den Iterationszähler, den Absolutbetrag der relativen Dualitätslücke zwischen der primalen und dualen Zielfunktion, die Norm der Funktion  $\Theta$  in der aktuellen Iterierten  $W^k = (X^k, \lambda^k, S^k)$  sowie den Abstand von  $W^k$  zur Lösung  $W^* = (X^*, \lambda^*, S^*)$ . Man beachte, dass die Dualitätslücke im Startpunkt zwar Null ist, wir aber nicht in einem Optimum sind, da die Iterierte nicht zulässig ist.

Das zweite Beispiel stammt von Kojima, Shida und Shindoh [50]. Auch dieses Beispiel hat eine eindeutige Lösung, welche nicht strikt komplementär ist. Hier aber ist auch Voraussetzung (A.4) nicht erfüllt.

**Beispiel 5.15** Sei  $n = 3$ ,  $m = 2$  mit  $b = (-1, 0)^T$  und

$$C = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad A_1 = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & 0 & 0.5 \\ 0 & -0.5 & 0 \\ 0.5 & 0 & -1 \end{pmatrix}.$$

Voraussetzung (A.1) ist auch hier offensichtlich wieder erfüllt. Das zugehörige semidefinite Programm hat die eindeutige Lösung

$$X^* = \text{diag}(1, 0, 0), \quad \lambda^* = (0, 0)^T, \quad S^* = \text{diag}(0, 0, 1),$$

so dass die Bedingung der strikten Komplementarität verletzt ist. Ähnlich wie beim letzten Beispiel erhalten wir  $V^* = I$ ,  $\alpha = \{1\}$ ,  $\beta = \{2\}$ , sowie  $\gamma = \{3\}$ . Seien daher  $\Sigma_-$ ,  $\Sigma_+$  zwei Matrizen wie in Voraussetzung (A.4), welche (5.32) und (5.33) erfüllen. Nun sei  $(\Delta X, \Delta \lambda, \Delta S)$  ein Vektor, so dass (5.34) gilt mit

$$\mathcal{A} = \begin{pmatrix} -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1/\sqrt{2} & -0.5 & 0 & -1 \end{pmatrix}.$$

## 5. Ein nichtglattes Newton-Verfahren

Tabelle 5.2.: Numerische Ergebnisse für Beispiel 5.15

$k$	rel. Dualitätslücke	$\ \Theta(W^k)\ $	$\ W^k - W^*\ $
0	0.000000e+00	8.002975e-01	1.327358e+00
1	1.370620e-01	3.130563e-01	6.265862e-01
2	1.772473e-03	7.147265e-03	2.687633e-02
3	2.393343e-05	4.128274e-04	1.526783e-03
4	7.823405e-08	3.397737e-05	1.207830e-04
5	5.080087e-10	2.105440e-06	7.471116e-06

Komponentenweise kann dies geschrieben werden als

$$\begin{aligned}
 -\Delta\lambda_1 + \Delta S_{11} &= 0, & \Delta S_{21} &= 0, \\
 \frac{1}{\sqrt{2}}\Delta\lambda_2 + \sqrt{2}\Delta S_{31} &= 0, & -\frac{1}{2}\Delta\lambda_2 + \Delta S_{22} &= 0, \\
 \Delta S_{32} &= 0, & -\Delta\lambda_2 + \Delta S_{33} &= 0, \\
 -\Delta X_{11} &= 0, & \Delta X_{31} - \frac{1}{2}\Delta X_{22} - \Delta X_{33} &= 0, \\
 \sigma_{11}^+ \Delta S_{11} &= 0, & \sigma_{21}^+ \Delta S_{21} &= 0, \\
 \sigma_{31}^- \Delta X_{31} + \sigma_{31}^+ \Delta S_{31} &= 0, & \sigma_{22}^- \Delta X_{22} + \sigma_{22}^+ \Delta S_{22} &= 0, \\
 \sigma_{32}^- \Delta X_{32} &= 0, & \sigma_{33}^- \Delta X_{33} &= 0
 \end{aligned}$$

mit gewissen Zahlen  $\sigma_{ij}^-$ ,  $\sigma_{ij}^+$ , welche (5.33) erfüllen. Dies ist ein homogenes lineares Gleichungssystem mit 14 Gleichungen und 14 Unbekannten. Da die beiden Gleichungen  $\Delta S_{21} = 0$  und  $\sigma_{21}^+ \Delta S_{21} = 0$  jedoch linear abhängig sind, hat dieses System eine von Null verschiedene Lösung. Also ist Voraussetzung (A.4) nicht erfüllt.

Obwohl Beispiel 5.15 nicht Voraussetzung (A.4) erfüllt, stellt sich heraus, dass das Prädiktor-Korrektor-Newton-Verfahren aus Kapitel 3, angewendet auf dieses Beispiel, immer noch lokal schnell konvergiert (vergleiche Tabelle 5.2). Daraus kann man schließen, dass Voraussetzung (A.4) zwar hinreichend, jedoch nicht zwingend notwendig für lokal quadratische Konvergenz ist.

Die beiden beschriebenen Beispiele benutzten beide  $n = 3$  für ein semi-definites Programm mit einer eindeutigen Lösung, welche nicht die strikte Komplementarität erfüllt. Es stellt sich daher jetzt die Frage, ob wir unsere Theorie nicht auch an einem Beispiel mit kleinerer Dimension  $n = 2$  illustrieren können. Wegen dem nächsten Resultat ist dies nicht möglich, wenn die Slater-Bedingung 2.2 erfüllt ist.



**Lemma 5.16** *Ist  $n = 2$  und die Regularitätsbedingung 2.2 von Slater für (2.6) erfüllt, so besitzen die Optimalitätsbedingungen (2.6) stets eine strikt komplementäre Lösung.*

**Beweis:** Zunächst einmal besitzt (2.6) unter der Slater-Bedingung stets eine Lösung. Sei daher  $(X^*, \lambda^*, S^*)$  irgendeine Lösung der Optimalitätsbedingungen. Angenommen, diese Lösung erfüllt nicht die strikte Komplementarität. Dann kann man einfach einsehen, dass  $X^* = 0$  oder  $S^* = 0$  gilt. Sei nun  $(\hat{X}, \hat{\lambda}, \hat{S})$  der die Slater-Bedingung erfüllende Punkt. Ist  $X^* = 0$ , so folgt, dass  $(X^*, \hat{\lambda}, \hat{S})$  eine strikt komplementäre Lösung von (2.6) ist. Ist andererseits  $S^* = 0$ , so ist  $(\hat{X}, \lambda^*, S^*)$  eine Lösung von (2.6), welche strikte Komplementarität erfüllt. In beiden Fällen können wir also eine strikt komplementäre Lösung finden.  $\square$

## 5. Ein nichtglattes Newton-Verfahren

## 6. Zusammenfassung

Die in dieser Arbeit behandelten theoretischen Resultate und numerischen Experimente zeigen, dass die Glättungsverfahren für semidefinite Programme zum Teil sehr gute Ergebnisse liefern.

Die geglätteten NCP-Funktionen wurden zwar schon von Chen und Tseng [15] zur Lösung von semidefiniten Programmen und zur Entwicklung eines Glättungsverfahrens eingesetzt, die Äquivalenz mit den Zentralen-Pfad-Bedingungen ist jedoch erstmals in der gemeinsamen Arbeit [42] des Autors mit Prof. Kanzow erwähnt. Den Glättungsparameter  $\tau$  als Variable (und nicht als Parameter) zu betrachten, hat sich numerisch als großer Vorteil herausgestellt. Weiter werden bei den Glättungsverfahren automatisch symmetrische Suchrichtungen erzeugt, was bei Inneren-Punkte-Methoden nicht der Fall ist.

Für das in Kapitel 3 vorgestellte Prädiktor-Korrektor-Verfahren kann man unter Standardvoraussetzungen globale und lokal schnelle Konvergenz beweisen. Die theoretischen Untersuchungen zur Lösung der Newton-Gleichung zeigen, dass die Minimum-Funktion aus Gründen der Effizienz der Fischer-Burmeister-Funktion vorzuziehen ist. Die numerischen Ergebnisse, welche zumindest vergleichbar mit den Inneren-Punkte-Methoden sind, zeigen auch ein numerisch besseres Verhalten der Minimum-Funktion. Aus diesen Gründen ist meines Erachtens zur Lösung von semidefiniten Programmen die Minimum-Funktion der Fischer-Burmeister-Funktion vorzuziehen.

Bei dem vorgestellten Trust-Region-Verfahren musste aus den in Kapitel 4 genannten Gründen jedoch wieder auf die Fischer-Burmeister-Funktion zurückgegriffen werden. Der vorgestellte Algorithmus hat zwar theoretisch gute Eigenschaften, die numerischen Ergebnisse können jedoch nicht überzeugen. Das Verfahren ist daher aus der Sicht des Autors keine gute Alternative.

Die in Kapitel 5 gemachten Untersuchungen zeigen das erste Mal lokal quadratische Konvergenz eines auf einem Glättungsansatz beruhenden Verfahrens zur Lösung von semidefiniten Programmen, ohne dass die strikte Komplementarität im Lösungspunkt vorausgesetzt werden muss. Dafür wurde die Nichtdegeneriertheitsbedingung geeignet verschärft. Es ist jedoch noch nicht ganz klar, wie diese (in der Matrix-Vektor-Formulierung angegebene) Bedingung in der üblichen Lyapunov-Schreibweise formuliert werden kann. Eine entsprechende Konvergenzrate ohne strikte Komplementarität ist dem Autor bisher für keine Innere-Punkte-Methode bekannt.

## 6. Zusammenfassung

## Symbolverzeichnis

$ \cdot $	Absolutbetrag einer symmetrischen Matrix: $ A  = (A^2)^{1/2}$ .
$\bullet$	Skalarprodukt im $\mathbb{R}^{n \times n}$ oder $\mathbb{R}^{n \times n} \times \mathbb{R}^m \times \mathbb{R}^{n \times n} \times \mathbb{R}$ .
diag	Diagonalmatrix.
$\star$	Hadamard-Produkt zweier Matrizen: $A \star B = (a_{ij}b_{ij})_{i,j=1,\dots,n}$ .
$\otimes$	Kronecker-Produkt zweier Matrizen.
$\otimes_s$	Symmetrisches Kronecker-Produkt zweier Matrizen.
$\lambda_{\max}$	Größter Eigenwert einer symmetrischen Matrix.
$\lambda_{\min}$	Kleinster Eigenwert einer symmetrischen Matrix.
$\mathbb{N}$	Menge der natürlichen Zahlen.
$\ \cdot\ _2$	Euklidische Vektornorm: $\ x\ _2 = \sqrt{x^T x}$ .
$\ \cdot\ _2$	Spektralnrm einer Matrix: $\ A\ _2 = \sqrt{\lambda_{\max}(A^T A)}$ .
$\ \cdot\ $	Norm auf dem $\mathbb{R}^{n \times n} \times \mathbb{R}^m \times \mathbb{R}^{n \times n}$ , $\mathbb{R}^{n \times n} \times \mathbb{R}^m \times \mathbb{R}^{n \times n} \times \mathbb{R}$ bzw. $\mathbb{R}^{n \times n} \times \mathbb{R}^{n \times n} \times \mathbb{R}$ .
$\ \cdot\ _F$	Frobenius-Norm einer Matrix: $\ A\ _F = \sqrt{A \bullet A}$ .
$\ \cdot\ $	Zu $\ \cdot\ $ gehörige Operatornorm.
$A \succ 0$	Die Matrix $A$ ist symmetrisch und positiv definit.
$A \succeq 0$	Die Matrix $A$ ist symmetrisch und positiv semidefinit.
$\mathbb{R}$	Menge der reellen Zahlen.
$\mathbb{R}^{n \times n}$	Raum der reellen $n \times n$ -Matrizen.
$\mathbb{R}_+$	Menge der nichtnegativen reellen Zahlen.
$\mathbb{R}_{++}$	Menge der positiven reellen Zahlen.
$\mathcal{S}^{n \times n}$	Raum der symmetrischen $n \times n$ -Matrizen.
$\mathcal{S}_+^{n \times n}$	Raum der symmetrischen und positiv semidefiniten $n \times n$ -Matrizen.
$\mathcal{S}_{++}^{n \times n}$	Raum der symmetrischen und positiv definiten $n \times n$ -Matrizen.
tr	Spur einer Matrix: $\text{tr}(A) = \sum_{i=1}^m a_{ii}$ .

## *Symbolverzeichnis*

## Literaturverzeichnis

- [1] F. ALIZADEH: *Interior point methods in semidefinite programming with applications to combinatorial optimization*. SIAM Journal on Optimization 5 (1995), S. 13–51.
- [2] F. ALIZADEH, J.-P. HAEBERLY UND M.L. OVERTON: *Complementarity and nondegeneracy in semidefinite programming*. Mathematical Programming 77 (1997), S. 111–128.
- [3] F. ALIZADEH, J.-P. HAEBERLY UND M.L. OVERTON: *Primal-dual interior-point methods for semidefinite programming: Convergence rates, stability and numerical results*. SIAM Journal on Optimization 8 (1998), S. 746–768.
- [4] A. BEN-TAL, M. KOČVARA, A. NEMIROVSKI UND J. ZOWE: *Free material design via semidefinite programming: The multiload case with contact conditions*. SIAM Review 42 (2000), S. 695–715.
- [5] A. BEN-TAL UND A. NEMIROVSKI: *Robust truss topology design via semidefinite programming*. SIAM Journal on Optimization 7 (1997), S. 991–1016.
- [6] A. BEN-TAL UND A. NEMIROVSKI: *Lectures on Modern Convex Optimization*. MPS-SIAM Series on Optimization, SIAM, Philadelphia, PA, 2001.
- [7] Å. BJÖRK UND S. HAMMARLING: *A Schur method for the square root of a matrix*. Linear Algebra and its Applications 52/53 (1983), S. 127–140.
- [8] B. BORCHERS: *SDPLIB 1.2, A library of semidefinite programming test problems*. Optimization Methods and Software 11 (1999), S. 597–611.
- [9] W. BUNSE UND A. BUNSE-GERSTNER: *Numerische lineare Algebra*. Teubner, Stuttgart 1985.
- [10] J.V. BURKE UND S. XU: *A non-interior predictor-corrector path-following method for LCP*. In Reformulation: Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods, M. Fukushima und L. Qi

- (Hrsg.), Kluwer Academic Publishers, Dordrecht, Netherlands, 1999, S. 45–63.
- [11] J.V. BURKE UND S. XU: *A non-interior predictor-corrector path following algorithm for the monotone linear complementarity problem*. Mathematical Programming 87 (2000), S. 113–130.
- [12] F.H. CLARKE: *Optimization and Nonsmooth Analysis*. John Wiley & Sons, New York, NY, 1983 (Nachdruck von SIAM, Philadelphia, PA, 1990).
- [13] A.R. CONN, N.I.M. GOULD UND P.L. TOINT: *Trust-Region Methods*. MPS-SIAM Series on Optimization, SIAM, Philadelphia, PA, 2000.
- [14] R. CORREA UND H. RAMIREZ: *A global algorithm for nonlinear semidefinite programming*. Research Report 4672, INRIA, Le Chesnay Cedex, France, 2002.
- [15] X. CHEN UND P. TSENG: *Non-interior continuation methods for solving semidefinite complementarity problems*. Mathematical Programming 95 (2003), S. 431–474.
- [16] J.W. DEMMEL: *Applied Numerical Linear Algebra*. SIAM, Philadelphia, PA, 1997.
- [17] J.E. DENNIS UND R.B. SCHNABEL: *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*. Prentice-Hall, Englewood Cliffs, NJ, 1983 (Nachdruck bei SIAM, Philadelphia).
- [18] S. ENGELKE UND C. KANZOW: *Improved smoothing-type methods for the solution of linear programs*. Numerische Mathematik 90 (2002), S. 487–507.
- [19] S. ENGELKE UND C. KANZOW: *On the solution of linear programs by Jacobian smoothing methods*. Annals of Operations Research 103 (2001), S. 49–70.
- [20] S. ENGELKE UND C. KANZOW: *Predictor-corrector smoothing methods for linear programs with a more flexible update of the smoothing parameter*. Computational Optimization and Applications 23 (2002), S. 299–320.
- [21] F. FACCHINEI UND J. SOARES: *A new merit function for nonlinear complementarity problems and a related algorithm*. SIAM Journal on Optimization 7 (1997), S. 225–247.



- [22] A. FISCHER: *A special Newton-type optimization method*. Optimization 24 (1992), S. 269–284.
- [23] A. FORSGREN: *Optimality conditions for nonconvex semidefinite programming*. Mathematical Programming 88 (2000), S. 105–128.
- [24] C. GEIGER UND C. KANZOW: *Numerische Verfahren zur Lösung unrestringierter Optimierungsaufgaben*. Springer, Berlin, Heidelberg 1999.
- [25] C. GEIGER UND C. KANZOW: *Theorie und Numerik restringierter Optimierungsaufgaben*. Springer, Berlin, Heidelberg 2001.
- [26] M.X. GOEMANS: *Semidefinite programming in combinatorial optimization*. Mathematical Programming 79 (1997), S. 143–161.
- [27] M.X. GOEMANS UND D.P. WILLIAMSON: *Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming*. Journal of the ACM 42 (1995), S. 1115–1145.
- [28] M. GRÖTSCHEL, L. LOVÁSZ UND A. SCHRIJVER: *Geometric Algorithms and Combinatorial Optimization*. Springer, Berlin, Heidelberg, New-York 1988.
- [29] L. HAN, J.C. TRINKLE UND Z.X. LI: *Grasp analysis as linear matrix inequality problems*. IEEE Transactions on Robotics and Automation 16 (2000), S. 663–674.
- [30] C. HELMBERG, F. RENDL, R. VANDERBEI UND H. WOLKOWICZ: *An interior-point method for semidefinite programming*. SIAM Journal on Optimization 6 (1996), S. 342–361.
- [31] N.J. HIGHAM: *Computing real square roots of a real matrix*. Linear Algebra and its Applications 88/89 (1987), S. 405–430.
- [32] N.J. HIGHAM: *Stable iterations for the matrix square root*. Numerical Algorithms 15 (1997), S. 227–242.
- [33] N.J. HIGHAM. Private Kommunikation, 2001.
- [34] C.W.J. HOL, C.W. SCHERER, E.G. VAN DER MECHE UND O.H. BOSGRA: *A nonlinear SDP approach to fixed-order controller synthesis and comparison with two other methods applied to an active suspension system*. European Journal of Control 9 (2003), S. 11–26.

- [35] R. A. HORN UND C. R. JOHNSON: *Matrix Analysis*. Cambridge University Press, 1985.
- [36] R. A. HORN UND C. R. JOHNSON: *Topics in Matrix Analysis*. Cambridge University Press, 1991.
- [37] F. JARRE: *An interior method for nonconvex semidefinite programs*. Optimization and Engineering 1 (2000), S. 347–372.
- [38] F. JARRE: *Some aspects of nonlinear semidefinite programs*. System Modeling and Optimization XX, F.W. Sachs and R. Tichatschke (eds.), Kluwer Academic Publishers, 2003.
- [39] F. JARRE UND J. STOER: *Optimierung*. Springer, Berlin, Heidelberg, New York, 2004.
- [40] H. JIANG: *Smoothed Fischer-Burmeister equation methods for the complementarity problem*. Technical Report, Department of Mathematics, University of Melbourne, Melbourne, Australia, June 1997.
- [41] C. KANZOW: *Some noninterior continuation methods for linear complementarity problems*. SIAM Journal on Matrix Analysis and Applications 17 (1996), S. 851–868.
- [42] C. KANZOW UND C. NAGEL: *Semidefinite programs: New search directions, smoothing-type methods, and numerical results*. SIAM Journal on Optimization 13 (2002), S. 1–23.
- [43] C. KANZOW UND C. NAGEL: *Corrigendum: Semidefinite programs: New search directions, smoothing-type methods, and numerical results*. Erscheint in SIAM Journal on Optimization.
- [44] C. KANZOW UND C. NAGEL: *Some practical aspects of a Newton-type method for semidefinite programs*. Preprint 243, Institut für Angewandte Mathematik und Statistik, Universität Würzburg, Würzburg, Germany, November 2001.
- [45] C. KANZOW UND C. NAGEL: *Some structural properties of a Newton-type method for semidefinite programs*. Erscheint in Journal of Optimization Theory and Applications.
- [46] C. KANZOW UND C. NAGEL: *Quadratic convergence of a nonsmooth Newton-type method for semidefinite programs without strict complementarity*. Preprint 250, Institut für Angewandte Mathematik und Statistik, Universität Würzburg, Würzburg, Germany, Juli 2003.

- [47] C. KANZOW, C. NAGEL UND M. FUKUSHIMA: *Successive linearization methods for nonlinear semidefinite programs*. Preprint 252, Institut für Angewandte Mathematik und Statistik, Universität Würzburg, Würzburg, Germany, August 2003.
- [48] C. KANZOW UND H.-D. QI, *A QP-free constrained Newton-type method for variational inequality problems*. *Mathematical Programming* 85 (1999), S. 81–106.
- [49] C. KANZOW UND M. ZUPKE: *Inexact trust-region methods for nonlinear complementarity problems*. In: M. FUKUSHIMA UND L. QI (Hrsg.): *Reformulation - Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods*. Kluwer Academic Press, Dordrecht, The Netherlands, 1999, S. 211–233.
- [50] M. KOJIMA, M. SHIDA, UND S. SHINDOH: *Local convergence of predictor-corrector infeasible-interior-point algorithms for SDPs and SDLCPs*. *Mathematical Programming* 80 (1998), S. 129–160.
- [51] M. KOJIMA, M. SHIDA, UND S. SHINDOH: *A predictor-corrector interior-point algorithm for the semidefinite linear complementarity problem using the Alizadeh-Haeberly-Overton search direction*. *SIAM Journal on Optimization* 9 (1999), S. 444–465.
- [52] M. KOČVARA UND M. STINGL: *PENNON: A code for convex nonlinear and semidefinite programming*. *Optimization Methods and Software* 18 (2003), S. 317–333.
- [53] Y.Y. LU: *A Padé approximation method for square roots of symmetric positive definite matrices*. *SIAM Journal on Matrix Analysis and Applications* 19 (1998), S. 833–845.
- [54] J.J. MORÉ UND D.C. SORENSEN: *Computing a trust-region step*. *SIAM Journal on Scientific and Statistical Computing* 4 (1983), S. 553–572.
- [55] Y. NESTEROV UND A. NEMIROVSKII: *Interior-Point Polynomial Algorithms in Convex Programming*. SIAM, Philadelphia, PA, 1994.
- [56] Y. NESTEROV UND M.J. TODD: *Primal-dual interior-point methods for self-scaled cones*. *SIAM Journal on Optimization* 8 (1998), S. 324–364.
- [57] J. NOCEDAL UND S.J. WRIGHT: *Numerical Optimization*. Springer Series in Operations Research, Springer, New York, Berlin, Heidelberg, 1999.

- [58] M.J.D. POWELL: *Convergence properties of a class of minimization algorithms*. In: O.L. MANGASARIAN, R.R. MEYER UND S.M. ROBINSON (Hrsg.): *Nonlinear Programming 2*. Academic Press, New York, NY, 1975, S. 1–27.
- [59] L. QI: *Convergence analysis of some algorithms for solving nonsmooth equations*. *Mathematics of Operations Research* 18 (1993), S. 227–244.
- [60] L. QI UND J. SUN: *A nonsmooth version of Newton's method*. *Mathematical Programming* 58 (1993), S. 353–367.
- [61] A. SHAPIRO: *First and second order analysis of nonlinear semidefinite programs*. *Mathematical Programming* 77 (1997), S. 301–320.
- [62] D. SUN UND J. SUN: *Semismooth matrix valued functions*. *Mathematics of Operations Research* 27 (2002), S. 150–169.
- [63] J. SUN, D. SUN UND L. QI: *From strong semismoothness of the squared smoothing matrix function to semidefinite complementarity problems*. Applied Mathematics Report AMR 00/20, School of Mathematics, University of New South Wales, Sydney, Australia, October 2000.
- [64] J. SUN, D. SUN UND L. QI: *Quadratic convergence of a smoothing Newton method for nonsmooth matrix equations and its applications in semidefinite optimization problems*. Preprint, Department of Decision Sciences, National University of Singapore, Singapore, April 2002 (Revidierte Fassung von [63]).
- [65] M.J. TODD, K.C. TOH UND R.H. TÜTÜNCÜ: *On the Nesterov-Todd direction in semidefinite programming*. *SIAM Journal on Optimization* 8 (1998), S. 769–796.
- [66] K.C. TOH, M.J. TODD UND R.H. TÜTÜNCÜ: *SDPT3 — a Matlab software package for semidefinite programming, version 2.1*. *Optimization Methods and Software* 11 (1999), S. 545–581.
- [67] P. TSENG: *Merit functions for semi-definite complementarity problems*. *Mathematical Programming* 83 (1998), S. 159–185.
- [68] P. TSENG: *Error bounds and superlinear convergence analysis of some Newton-type methods in optimization*. In: *Nonlinear Optimization and Related Topics*, G. D. Pillo und F. Giannessi (Hrsg.), Kluwer Academic Publishers, 2000, S. 445–462.

- [69] L. VANDENBERGHE UND S. BOYD: *Semidefinite programming*. SIAM Review 38 (1996), S. 49–95.
- [70] Y. ZHANG: *On extending some primal-dual interior-point algorithm from linear programming to semidefinite programming*. SIAM Journal on Optimization 8 (1998), S. 365–386.