

---

**„Weniger ist mehr.“**

**Eigenpaarfilterung am Beispiel hermitesch positiv  
definiter Eigenwertprobleme**

---

Dem Institut für Mathematik  
der Technischen Universität Berlin  
vorgelegte

**Bachelorarbeit**

von

**Thorsten Matthias Lucke**

Betreuer und Erstgutachter: Prof. Dr. Jörg Liesen  
Zweitgutachter: Prof. Dr. Christian Mehl



*Meiner Familie,  
der Thesisselbsthilfegruppe,  
der Hängemathe und dem 1 Café*

# Erklärung

Ich versichere, dass ich die vorliegende Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe. Ich reiche sie erstmals als Prüfungsleistung ein. Mir ist bekannt, dass ein Betrugsversuch mit der Note „nicht ausreichend“ geahndet wird und im Wiederholungsfall zum Ausschluss von der Erbringung weiterer Prüfungsleistungen führen kann.

Name:

Vorname:

Matrikelnummer:

Berlin, den

Signatur

# Danksagung

Obschon die vorliegende Arbeit der eigenen Feder entstammt, hätte dieses Werk kaum vollendet werden können, wären nicht einige Hürden beiseite geschafft worden. So sehe ich es als meine Pflicht, den Helfern des „Räumungsdienstes“ ein paar Worte der Ehrung zukommen zu lassen.

Der erste Dank gilt meinen Betreuern Prof. Dr. Jörg Liesen und Prof. Dr. Christian Mehl deren Sprechstunden stets eine Quelle der Inspiration waren. Immer hilfsbereit standen sie mir fachlich zur Seite und ließen mich von ihrer Erfahrung profitieren.

Ein weiterer Dank geht an die Mitarbeiterinnen und Mitarbeiter des studentischen Mathekaffees „1 Café“. Welche Qualen blieben mir durch die etlichen Teestündchen, Schokoriegel und zum Schlafen einladenden Sofas erspart! Im selben Atemzug danke ich den Mitgliedern der „Hängemathe“, deren kritisches Hinterfragen und Diskussionsfreudigkeit stets anregend war.

Schließlich richte ich meinen Dank an alle Lektoren, die akribisch jeden noch so hinterhältig versteckten Fehler entdeckt haben und damit das Lesen dieser Arbeit zu einem größeren Vergnügen machen.

# Inhaltsverzeichnis

|          |   |           |
|----------|---|-----------|
| <b>1</b> | <b>Präludium</b>                                | <b>1</b>  |
| <b>2</b> | <b>Mathematische Grundlagen</b>                 | <b>4</b>  |
| <b>3</b> | <b>Filtertechniken</b>                          | <b>8</b>  |
| 3.1      | Rayleigh-Ritz Verfahren . . . . .               | 9         |
| 3.2      | Konturintegration . . . . .                     | 14        |
| 3.3      | Illustration . . . . .                          | 16        |
| <b>4</b> | <b>Rationales Filtern</b>                       | <b>18</b> |
| 4.1      | Beschleunigte Rayleigh-Ritz Iteration . . . . . | 18        |
| 4.2      | Gauß'sche Quadratur . . . . .                   | 21        |
| 4.3      | Approximation des Spektralprojektors . . . . .  | 22        |
| <b>5</b> | <b>Numerische Experimente</b>                   | <b>25</b> |
| 5.1      | Rayleigh-Ritz . . . . .                         | 25        |
| 5.2      | Konturintegration . . . . .                     | 26        |
| 5.3      | FEAST Algorithmus . . . . .                     | 28        |
| 5.4      | Tools für MATLAB . . . . .                      | 28        |
| <b>A</b> | <b>Notationen</b>                               | <b>29</b> |
| <b>B</b> | <b>Sätze und Beweise</b>                        | <b>30</b> |

# Kapitel 1

## Präludium

Das Lösen von Eigenwertproblemen ist eine Standarddisziplin in der numerischen linearen Algebra. Die konkrete Problemstellung dabei lautet: Zu gegebenen Matrizen  $A, B \in \mathbb{C}^{n,n}$  sollen Paare  $(\lambda, x)$  mit  $\lambda \in \mathbb{C}$  und  $x \in \mathbb{C}^n \setminus \{0\}$  gefunden werden, welche der Gleichung

$$Ax = \lambda Bx \tag{1.1}$$

genügen. Solchen *Eigenwertgleichungen* begegnet man in ganz unterschiedlichen Kontexten. So sind sie beispielsweise bei der Bestimmung von Eigenfrequenzen oder dem Ermitteln von Fixpunkten beim Rotieren eines Fußballs<sup>1</sup> ebenso wie beim Untersuchen des PageRanks<sup>2</sup> einer Website von Bedeutung. Entsprechend strotzt der Kanon von angebotenen numerischen Lösungsmethoden von Vielfalt und Virtuosität.<sup>3</sup>

Nun mag der Fall eintreten, da es notwendig wird, lediglich eine Teilmenge aus der Menge aller Eigenpaare zu untersuchen. Betrachten wir zur Illustration folgende Abbildung.



Abbildung 1.1: „Mind Games of Ozzy“ von Andy Prokh.

Man stelle sich vor, man wolle dieses Bild komprimieren, um den benötigten Speicherplatz zu reduzieren. Dies lässt sich beispielsweise mit einer sogenannten *Singulärwertzerlegung* – von Numerikern liebevoll als das *Schweizer Taschenmesser der linearen Algebra* bezeichnet – bewerkstelligen.

---

<sup>1</sup>Hier wird auf den bekannten *Satz vom Fußball* angespielt. Dieser besagt, dass auf einem Fußball zwei Punkte existieren, die zu Spielbeginn und zur Halbzeit an der gleichen Stelle liegen – informell formuliert.

<sup>2</sup> Siehe Abschnitt zwei in [Pag+99].

<sup>3</sup>Dies bestätigt sich beispielsweise bei einem Blick in das Inhaltsverzeichnis von [Ste13].

Da es sich bei Abbildung 1.1 um ein Schwarz-Weiß-Bild handelt, lässt sie sich als eine Matrix  $M$  auffassen, deren Einträge die Grauwerte der einzelnen Pixel repräsentieren. Unter Anwendung des Satzes von der Singulärwertzerlegung<sup>4</sup> lässt sich nun das Bild der Katze durch eine Folge von Matrizen niedrigeren Ranges approximieren. Man berechnet hierfür die Wurzeln der von Null verschiedenen Eigenwerte von  $M^H M$  und verwendet dann diese sogenannten *Singulärwerte* um die Matrix  $M$  zu rekonstruieren. Auf weitere Details über den hinter diesem Verfahren stehenden Algorithmus wollen wir an dieser Stelle verzichten.

Die Matrix  $M$  hat in unserem Beispiel mehr als 700 Singulärwerte. Die Frage ist nun, ob wir uns von einigen Singulärwerten trennen können und dennoch eine akzeptable Bildqualität aufrecht erhalten können. Insbesondere ist wichtig zu wissen, welche dieser Werte entscheidend für die Rekonstruktion sind.

Dazu nehmen wir uns die  $n \in \mathbb{N}$  Singulärwerte  $\sigma_1, \dots, \sigma_n$  her. Ohne Einschränkung seien diese so nummeriert, dass  $\sigma_1 \geq \dots \geq \sigma_n$  gilt. Sollte dies nicht der Fall sein, nummerieren wir einfach um. Nun filtern wir nach eigenen Kriterien Teilmengen aus den Singulärwerten, beziehungsweise den Eigenwerten von  $M^H M$  heraus und rekonstruieren mit deren Hilfe die Abbildung 1.1.

Es sei an dieser Stelle darauf hingewiesen, dass die folgenden Beurteilungen der Bildqualität nicht mathematisch begründet sind. Die Eindrücke entstammen dem persönlichen Empfinden des Autors, sowie den Aussagen einer kleinen Testgruppe von Studierenden des Fachs Mathematik.

Beginnen wir mit einer mehr oder weniger willkürlichen Auswahl von Singulärwerten (SW).

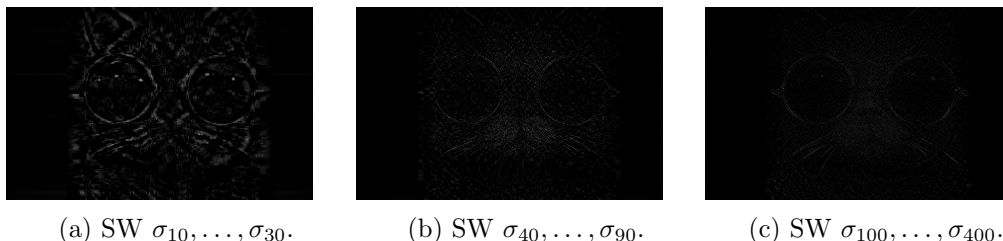


Abbildung 1.2: Rekonstruktionsversuch von Abbildung 1.1.

Die in Abbildung 1.2 gewählte Filtrierung erscheint etwas unglücklich, da das ursprüngliche Bild kaum wieder zu erkennen ist. Versuchen wir unser Glück daher mit einer anderen Wahl von Singulärwerten.

<sup>4</sup>Eine Formulierung ohne Beweis ist mit Satz ?? – zu finden im Anhang – gegeben.



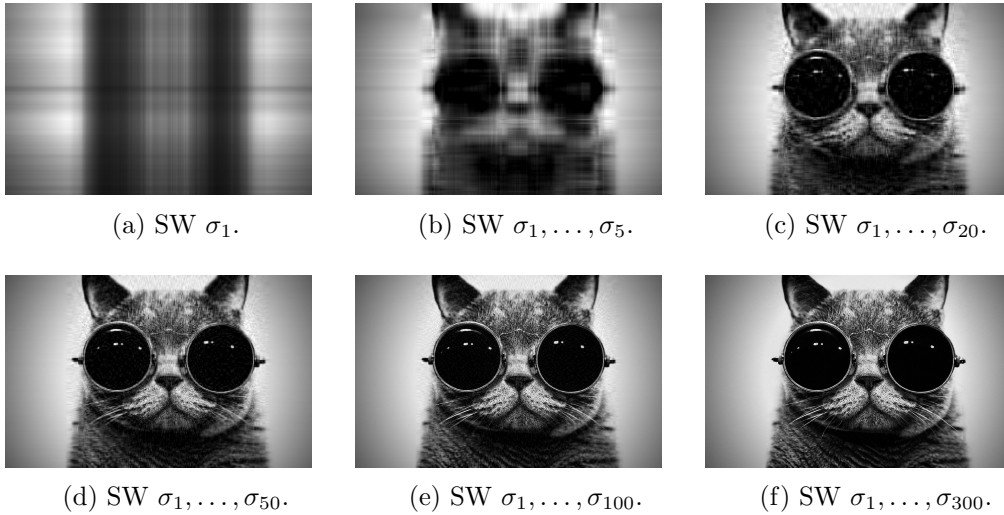


Abbildung 1.3: Approximation von Abbildung 1.1 mittels Singulärwertzerlegung.

Dieser Versuch wirkt überzeugender. Bereits bei Figur 1.3e ist ein genauer Blick erforderlich, um die Makel der Komprimierung zu erkennen. Bei der Verwendung der ersten 300 SW ist es nahezu unmöglich einen Unterschied zum Original festzustellen. Es ist also zu überlegen, ob man sich mit der Qualität der Figur 1.3f zufrieden geben möchte und statt Abbildung 1.1 nicht einfach dieses abspeichert.

Nun ist das Approximieren von Bildern ein sehr spezieller Fall des Filterns von Eigenwerten und die Berechnung der Singulärwertzerlegung nicht immer zweckmäßig oder möglich. Daher setzt sich diese Arbeit mit Alternativen auseinander, die dem Filtern dienlich sind. Dabei werden neben den mathematischen Ideen dieser Alternativen auch Möglichkeiten der Implementation vorgestellt.

Im dritten Kapitel werden zunächst zwei Methoden präsentiert, die als Werkzeuge zum Filtern von Eigenpaaren dienlich sind. Daran wird sich nach einer kurzen analytischen Illustration dieser Verfahren eine Diskussion anschließen, bei der untersucht wird, wie es um die praktische Umsetzbarkeit der Verfahren bestellt ist. Wir werden feststellen, dass sich die Algorithmen kombinieren und modifizieren lassen und zu einem Verfahren führen, welches in der Literatur als FEAST-Algorithmus gehandelt wird. Zum Abschluss begeben wir uns in das „Numerik-Labor“ und werden die Grenzen der Algorithmen austesten.

Bevor es konkreter wird, erinnert der folgende Abschnitt an einige mathematische Grundlagen, die als helfende Handreichung das Lesen dieser Schrift mehr zur Freude, denn zur Schikane machen soll.

## Kapitel 2

# Mathematische Grundlagen

Wie versprochen, widmet sich das hiesige Kapitel mit der Wiederholung und Einführung wichtigen Vokabulars, welches im Rahmen dieser Arbeit häufig zum Einsatz kommen wird. Von zentraler Bedeutung werden dabei Inhalte der numerischen linearen Algebra sein.

Um es nicht bei einer bloßen Auflistung von Definitionen zu belassen, werden zudem Resultate präsentiert, auf die wir später zurückgreifen werden. Obschon sich der Autor bemüht hat, in der Literatur gängige Notation zu benutzen, bittet er den verständnisvollen Leser bei Unklarheiten im Anhang A nachzuschlagen.

Mit den Buchstaben  $m$  und  $n$  werden wir – sofern nicht anders vermerkt – zwei natürliche Zahlen bezeichnen. Dabei wollen wir die Null aus den natürlichen Zahlen  $\mathbb{N}$  ausgeschlossen wissen. Falls die Null zugelassen ist, schreiben wir explizit  $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$ . Des Weiteren bezeichnen wir wie üblich mit  $\mathbb{R}$  die Menge der reellen Zahlen, mit  $\mathbb{C}$  die Menge der komplexen Zahlen und entsprechend mit  $\mathbb{R}^{m,n}$  und  $\mathbb{C}^{m,n}$  die Mengen der reell- beziehungsweise komplexwertigen Matrizen. Wir unterschlagen  $n$  im Fall  $n = 1$ .

In Anlehnung an die in *MATLAB* verwendete Syntax, werden wir Gebrauch von der Notation  $i = m : n$  an Stelle von  $i = m, m + 1, \dots, n - 1, n$  machen. Diese Kurzschreibweise wird etwa bei der Einführung von Indizierungen zum Einsatz kommen. So lässt sich beispielsweise eine Menge  $\{x_1, x_2, \dots, x_n\}$  mit hier nicht näher bestimmten Elementen kurz durch  $\{x_k\}_{k=1:n}$  ausdrücken.

Es sei nun  $A$  eine quadratische, komplexwertige Matrix, also  $A \in \mathbb{C}^{n,n}$ . Diese wird als *hermitesch* bezeichnet, falls sie die Identität  $A = A^H$  erfüllt und ist *positiv definit*, sofern für alle Vektoren  $x \in \mathbb{C}^n \setminus \{0\}$  die Abschätzung

$$x^H A x > 0$$

gilt. Folglich werden wir eine Matrix *hermitesch positiv definit* (HPD) nennen, wenn sie sowohl hermitesch als auch positiv definit ist.

Ist  $A$  solch eine HPD-Matrix, dann nennen wir eine Menge von Vektoren  $\{x_k\}_{k=1:m} \subseteq \mathbb{C}^n$  *orthonormal bezüglich  $A$*  oder schlicht:  *$A$ -orthonormal*, falls

$$x_i^H A x_j = \delta_{i,j} := \begin{cases} 1 & \text{wenn } i = j \\ 0 & \text{wenn } i \neq j \end{cases}$$

für alle  $i, j = 1 : m$  gilt. Allgemeiner heißt eine Matrix  $X \in \mathbb{C}^{n,n}$  *orthogonal bezüglich  $A$*  oder  *$A$ -orthogonal*, falls sie

$$X^H A X = I_n := [\delta_{i,j}]_{i,j=1:n}$$

erfüllt.<sup>1</sup> Aus der Hermitizität von  $A$  und  $I_n$  folgt natürlich ebenfalls  $X A X^H = I_n$ . Für den Fall  $A = I_n$  ignorieren wir in der Formulierung den Bezug zu  $A$  und sprechen lediglich von Orthogonalität beziehungsweise Orthonormalität.

Sei im Folgenden  $A$  wieder eine beliebige komplexwertige Matrix. Neben dieser betrachten wir nun noch eine weitere Matrix  $B \in \mathbb{C}^{n,n}$ . Unter einem  *$n$ -dimensionalen Eigenwertproblem* wollen wir die Aufgabe verstehen, Paare  $(\lambda, x) \in \mathbb{C} \times \mathbb{C}^n \setminus \{0\}$  zu finden, die der Gleichung

$$A x = \lambda B x \tag{2.1}$$

genügen. Wenn klar ist, dass von solch einem Problem die Rede ist, werden wir anstelle des Wortes „Eigenwertproblem“ die Notation  $(A, B)$  als Bezeichnung verwenden und ausßerdem auf die Angabe der Dimension verzichten. Im Spezialfall  $B = I_n$  heißt  $(A, B)$  *gewöhnliches Eigenwertproblem der Dimension  $n$* .

Ist ein passendes Paar  $(\lambda, x)$  gefunden, welches die *Eigenwertgleichung* (2.1) löst, so nennen wir dieses *Eigenpaar von  $(A, B)$*  oder kurz: *Eigenpaar*, falls bekannt ist, von welchem Eigenwertproblem die Rede ist. Dabei heißt  $\lambda$  *Eigenwert von  $(A, B)$*  und  $x$  *Eigenvektor zum Eigenwert  $\lambda$  von  $(A, B)$* . Auch hier werden wir auf die Angabe des Eigenwertproblems verzichten, wenn der Kontext dies gestattet. Beim gewöhnlichen Eigenwertproblem sehen wir von der Paar-Schreibweise ab und sprechen direkt von Eigenpaaren, Eigenwerten und Eigenvektoren von  $A$ . Die Menge aller Eigenwerte, das sogenannte *Spektrum*, bezeichnen wir mit  $\Lambda(A, B)$  im allgemeinen und mit  $\Lambda(A)$  im gewöhnlichen Fall.

Wie auf der Titelseite angekündigt, wird sich diese Arbeit überwiegend mit der Behandlung *hermitesch positiv definiter Eigenwertprobleme* (HPD-Eigenwertprobleme) befassen. Damit seien fürderhin Eigenwertprobleme  $(A, B)$  bezeichnet, bei denen wir uns mit einer hermiteschen Matrix  $A$  sowie einer HPD-Matrix  $B$  konfrontiert sehen. Eigenwertprobleme dieser Art besitzen eine Reihe nützlicher Eigenschaften, an denen wir uns zu einem späteren Zeitpunkt in der Arbeit bedienen werden. Von besonderer Bedeutung wird dabei das folgende Resultat sein.

---

<sup>1</sup>Wir werden Matrizen in ihrer komponentenweisen Darstellung konsequent durch eckige Klammern „[“ und „]“ beenden.

**Satz 2.1.** Ist  $(A, B)$  ein HPD-Eigenwertproblem der Dimension  $n$ , so sind alle zugehörigen Eigenwerte reell. Außerdem existiert eine  $B$ -orthogonale Matrix  $X \in \mathbb{C}^{n,n}$  sowie eine Diagonalmatrix  $\Lambda \in \mathbb{R}^{n,n}$  mit

$$A = BX\Lambda X^{-1}. \quad (2.2)$$

*Beweis.* Sei  $(\lambda, x)$  ein Eigenpaar von  $(A, B)$ . Der positiven Definitheit von  $B$  wegen, gilt nach Definition  $x^H B x > 0$ . Aus

$$\lambda(x^H B x) = x^H A x = x^H A^H x = (Ax)^H x = \overline{\lambda}(Bx)^H x = \overline{\lambda}(x^H B x)$$

folgt dann die Gleichheit  $\lambda = \overline{\lambda}$  und somit  $\lambda \in \mathbb{R}$ .

Um die Existenz der im Satz angegebenen Faktorisierung (2.2) zu zeigen, ziehen wir unterstützend den Beweis aus [Par98, Theorem 15.3.2, 344 f.] zurate und ergänzen diesen zur Vervollständigung durch weitere Argumente.

Zunächst nutzen wir die Hermitizität und die positive Definitheit von  $B$  aus. Mit diesen beiden Eigenschaften garantieren uns der Spektralsatz für hermitesche Matrizen<sup>2</sup> und der Satz von der Existenz der Cholesky-Zerlegung<sup>3</sup> eine Faktorisierung der Art

$$B = X_B \Lambda_B^2 X_B^H. \quad (2.3)$$

Dabei ist  $\Lambda_B \in \mathbb{R}^{n,n}$  eine diagonale Matrix und  $X_B \in \mathbb{C}^{n,n}$  orthogonal. Wir definieren uns nun eine weitere Matrix

$$M := \Lambda_B^{-1} X_B^H A X_B \Lambda_B^{-1},$$

welche wegen

$$M^H = (\Lambda_B^{-1} X_B^H A X_B \Lambda_B^{-1})^H = (\Lambda_B^{-1})^H X_B^H A^H (X_B^H)^H (\Lambda_B^{-1})^H = \Lambda_B^{-1} X_B^H A X_B \Lambda_B^{-1} = M$$

hermitesch ist. Erneut finden wir mit Hilfe des Spektralsatzes eine Zerlegung der Form

$$M = X_M \Lambda_M X_M$$

von  $M$ . Setzen wir nun  $X := X_M \Lambda_M X_M^H$ , erhalten wir die Identitäten

$$\begin{aligned} X^H A X &= X_M^H \Lambda_B^{-1} X_B^H A X_B \Lambda_B^{-1} X_M = X_M^H M X_M = \Lambda_M \\ X^H B X &= X_M^H \Lambda_B^{-1} X_B^H B X_B \Lambda_B^{-1} X_M \stackrel{(2.3)}{=} X_M^H X_M = I_n. \end{aligned}$$

Also ist  $X$  schon  $B$ -orthogonal. Schließlich folgt Gleichung (2.2) aus

$$X^{-1} B^{-1} A X = X^{-1} B^{-1} (X^H)^{-1} X^H A X = \Lambda_M.$$

durch Umstellen. □

<sup>2</sup>Eine Formulierung mit Beweis ist im Anhang zu finden. Siehe Satz B.2.

<sup>3</sup>Siehe Satz B.3 im Anhang.

Die Existenz der Faktorisierung (2.3) werden wir später nutzen, um bei hermitesch positiv definiten Eigenwertproblem  $(A, B)$  anstelle von (2.1) das Lösen von

$$AX = BXA \tag{2.4}$$

verlangen zu können. Diese Notation wird häufig bei der Ausformulierung von Algorithmen in Erscheinung treten.

Fahren wir fort mit einer beliebig gewählten Matrix  $A \in \mathbb{C}^{m,n}$ . Diese induziert bekannterweise eine lineare Abbildung  $\mathcal{A}: \mathbb{C}^n \rightarrow \mathbb{C}^m$  vermöge  $\mathcal{A}(x) := Ax$ . Wir werden fortan nicht mehr zwischen der Abbildung  $\mathcal{A}$  und der zugehörigen Matrix unterscheiden und wollen  $A$  sowohl als Element des Vektorraumes  $\mathbb{C}^{m,n}$  als auch als lineare Abbildung zwischen  $\mathbb{C}^m$  und  $\mathbb{C}^n$  begreifen. Demnach sind Notationen wie  $\text{Bild}(A)$  und ähnliche abbildungsrelevante Schreibweisen wohldefiniert.

Nehmen wir uns zu guter Letzt einen  $m$ -dimensionalen Unterraum  $\mathcal{U} \subseteq \mathbb{C}^n$  her und bemühen erneut eine Matrix  $A \in \mathbb{C}^{n,n}$ . Der eben eingeführte Unterraum heißt *A-invariant*, falls die Aussage  $A\mathcal{U} \subseteq \mathcal{U}$  gilt. Ist dann  $U \in \mathbb{C}^{n,m}$  eine Matrix mit  $\text{Bild}(U) = \mathcal{U}$ , so existiert folglich eine Matrix  $V \in \mathbb{C}^{m,m}$  mit  $AU = UV$ .

Ausgestattet mit den eben besprochenen Definitionen und Resultaten, können wir nun mit dem Hauptteil dieser Arbeit beginnen.

## Kapitel 3

# Filtertechniken

In der Einleitung dieser Arbeit wurde bereits häufiger vom *Filtern* gesprochen. Bevor wir uns mit Techniken des Filtern beschäftigen können, wollen wir die Intuition hinter diesem Begriff der Form halber sauber definieren.

**Definition 3.1** (Filter). *Es sei  $M$  eine nicht leere Menge. Eine Abbildung der Gestalt*

$$\mathfrak{F}: M \rightarrow \{0, 1\}, \quad m \mapsto \mathfrak{F}(m)$$

*heißt Filterfunktion oder Filter auf der Menge  $M$ . Die Zuordnung  $m \mapsto \mathfrak{F}(m)$  heißt Filtern von  $m$  aus  $M$ .*

Damit können wir uns nun in diesem Kapitel der Frage widmen, ob und wie man aus der Menge aller Eigenpaare eines Eigenwertproblems mit der zugehörigen Eigenwertgleichung

$$Ax = \lambda Bx$$

eine gewünschte Teilmenge auswählen kann. Dabei wollen wir an die Matrizen  $A$  und  $B$  zunächst keine weiteren Anforderungen stellen.

Ist  $M$  entweder die Potenzmenge des Spektrums oder die Menge aller Eigenpaare von  $(A, B)$  und  $\mathfrak{F}$  eine Filterfunktion auf  $M$ , so wollen wir vereinbaren, dass wir stets auf der Suche nach der Urbildmenge  $\mathfrak{F}^{-1}(\{1\})$  sind. Wollen wir etwa die Menge  $G$  aller ganzzahligen Eigenwerte eines gewissen Eigenwertproblems finden, so gelte  $\mathfrak{F}^{-1}(\{1\}) = G$ .

Die Methoden zur Bestimmung von Teilmengen von Eigenpaaren oder Eigenwerten sind ebenso vielfältig, wie die Möglichkeiten das gesamte Spektrum und entsprechende Eigenvektoren zu bestimmen. Wir wollen aber im Rahmen dieser Arbeit aus dem Katalog der Verfahren lediglich zwei auswählen und besprechen: *Rayleigh-Ritz Verfahren* und *Konturintegration*.

### 3.1 Rayleigh-Ritz Verfahren

Die Idee dieser Klasse von Verfahren ist das Approximieren des von den gesuchten Eigenvektoren aufgespannten Unterraums. Bevor wir uns mit beliebigen Eigenwertproblemen  $(A, B)$  auseinandersetzen, betrachten wir den Spezialfall  $B = I_n$  und wenden uns der entsprechenden Eigenwertgleichung

$$Ax = \lambda x$$

zu. Sei im Folgenden  $\mathcal{S}_m \subseteq \mathbb{C}^n$  ein  $m$ -dimensionaler Unterraum. Dieser zunächst nicht näher bestimmte *Suchraum* wird als Grundlage für das Verfahren gewählt. Gesucht sind nun Paare  $(\tilde{\lambda}, \tilde{x}) \in \mathbb{C} \times \mathcal{S}_m \setminus \{0\}$ , welche wir als approximierte Lösungen des Eigenproblems verstehen wollen, die die Eigenschaft

$$s^H(A\tilde{x} - \tilde{\lambda}\tilde{x}) = 0 \quad (3.1)$$

für alle  $s \in \mathcal{S}_m$  erfüllen. Das *Residuum*  $(A\tilde{x} - \tilde{\lambda}\tilde{x})$  soll also orthogonal auf dem Suchraum stehen. Paare, die diesem Anliegen nachkommen, werden auch *Ritz-Paare* bezüglich des Suchraums  $\mathcal{S}_m$  genannt.

Wir wollen nun annehmen, dass mit der Menge  $\{s_i\}_{i=1:m} \subseteq \mathbb{C}^n$  eine Basis orthonormaler Vektoren des Unterraums  $\mathcal{S}_m$  gegeben ist. Setzen wir dann  $S_m := [s_i]_{i=1:m} \in \mathbb{C}^{n,m}$ , so muss wegen  $\tilde{x} \in \mathcal{S}_m$  ein Vektor  $y \in \mathbb{C}^m$  existieren, mit  $S_m y = \tilde{x}$ . Die Forderung (3.1) ist dann zu der Gleichung

$$S_m^H(A S_m y - \tilde{\lambda} S_m y) = 0$$

äquivalent. Als direkte Konsequenz dieser Umformulierung erhalten wir unter Ausnutzung der Orthogonalität der Spalten von  $S_m$  mit

$$(S_m^H A S_m) y = \tilde{\lambda} y \quad (3.2)$$

die Eigenwertgleichung zum Eigenwertproblem  $(S_m^H A S_m, I_m)$ . Jede Lösung  $(\tilde{\lambda}, y)$  von (3.2) liefert dann nach Konstruktion mit  $(\tilde{\lambda}, S_m y)$  ein Ritz-Paar des gewöhnlichen Eigenwertproblems bezüglich des Suchraums  $\mathcal{S}_m$ .

In Abhängigkeit von der Wahl des Suchraumes variiert natürlich Güte der Approximation. Ist etwa  $\mathcal{S}_m$  durch Eigenvektoren aufgespannt, so ist jedes Ritz-Paar schon ein Eigenpaar. Wir werden später aus den Untersuchungen des Eigenwertproblems für beliebiges  $B$  folgern, dass bereits im Falle der Invarianz von  $\mathcal{S}_m$  unter  $A$  jedes Ritz-Paar von  $A$  bezüglich  $\mathcal{S}_m$  bereits ein Eigenpaar von  $A$  ist. An dieser Stelle begnügen wir uns vorerst mit dieser erfreulichen Botschaft.

Die eben skizzierte Vorgehensweise zur Berechnung von Ritz-Paaren des gewöhnlichen Eigenwertproblems lässt sich algorithmisch wie folgt zusammenfassen.

---

**Algorithmus 1** Berechnung von Ritz-Paaren (Vgl. [Saa11, S. 98])

---

- 1: Berechne eine Orthonormalbasis  $\{s_i\}_{i=1:m}$  von  $\mathcal{S}_m$  und setze  $S_m \leftarrow [s_i]_{i=1:m}$ .
  - 2: Setze  $\tilde{A} \leftarrow S_m^H A S_m$  und löse  $\tilde{A}Y = Y\Lambda$  in  $Y$  und  $\Lambda$ .
  - 3: Setze  $\tilde{X} \leftarrow S_m Y$  und gib Ritz-Paare  $(\lambda_i, \tilde{x}_i)_{i=1:m}$  aus.
- 

An dieser Stelle sei an die in Kapitel 2 mit Gleichung (2.4) eingeführte Notation erinnert. Es entsprechen also die Spalten von  $Y \in \mathbb{C}^{n,m}$  gewissen Eigenvektoren von  $(\tilde{A}, I_m)$ , während auf der Diagonalen von  $\Lambda \in \mathbb{C}^{m,m}$  die zugehörigen Eigenwerte zu finden sind.

Die Idee dieser Methode ist also simpel: Transformiere das Eigenwertproblem mit einer gewissen Matrix in ein anderes Eigenwertproblem und benutze dessen Lösungen, um Ritz-Paare des ursprünglichen Problems zu erhalten. Doch wozu die Mühe, das ursprüngliche Eigenwertproblem in ein anderes Eigenwertproblem zu überführen? Zwar gelingt es, aus dem transformierten Problem (3.2) Ritz-Paare zu extrahieren, aber wäre nicht auch denkbar, sämtliche Eigenpaare von  $A$  zu approximieren und die zum Unterraum  $\mathcal{S}_m$  korrespondierende Teilmenge direkt auszuwählen?

Dies mag in Einzelfällen in der Tat sinnvoller sein. Sprechen wir allerdings von Matrixdimensionen hinreichender Größe, ist eine vollständige Berechnung aller Eigenpaare mitunter ein sehr zeitintensives Vergnügen. Bei genauerer Betrachtung der Gleichung (3.2) fällt auf, dass die Matrix  $(S_m^H A S_m) \in \mathbb{C}^{m,m}$  im Falle  $n \gg m$  ein mitunter deutlich kleineres Format hat, als die Matrix  $A$  im ursprünglichen Problem. Man darf hier also erwarten, dass die benötigte Laufzeit zur Bestimmung der Ritz-Paare mit dem Algorithmus 1 geringer ist, als beim Approximieren sämtlicher Eigenpaare von  $A$ . Dies wird im vierten Kapitel anhand ausgewählter Beispiele vorgeführt.

Nun, da wir der Existenz und Funktionsweise der Rayleigh-Ritz-Methode gewahr wurden, erweitern wir die obige Theorie auf das verallgemeinerte Eigenwertproblem

$$Ax = \lambda Bx. \quad (3.3)$$

Dabei gehen wir ganz analog zum gewöhnlichen Eigenwertproblem vor und betrachten wieder einen  $m$ -dimensionalen Suchraum  $\mathcal{U}_m \subseteq \mathbb{C}^n$ . Gefunden werden sollen dieses Mal Paare  $(\tilde{\lambda}, \tilde{x}) \in \mathbb{C} \times \mathcal{U}_m \setminus \{0\}$  die der Orthogonalitätsbedingung

$$A\tilde{x} - \tilde{\lambda}B\tilde{x} \perp \mathcal{U}_m \quad (3.4)$$

genügen. Durch die Wahl eines geeigneten Vektors  $y \in \mathbb{C}^m$  kann die Näherungslösung  $\tilde{x}$  wie zuvor durch das Produkt  $U_m y$  ersetzt werden, wobei die Spalten von  $U_m$  erneut eine



ONB des Suchraums  $\mathcal{U}_m$  bilden. Folglich lässt sich die zu (3.4) äquivalente Forderung

$$U_m^H(AU_my - \tilde{\lambda}BU_my) = 0.$$

aufstellen. Wie bereits beim gewöhnlichen Eigenwertproblem, lässt sich nun aus jeder Lösung  $(\tilde{\lambda}, y)$  von

$$(U_m^H AU_m)y = \tilde{\lambda}(U_m^H BU_m)y. \quad (3.5)$$

mit  $(\tilde{\lambda}, U_my)$  ein Ritz-Paar für das verallgemeinerte Eigenwertproblem gewinnen.

Da mit der Matrix  $B$  ein weiterer Darsteller auf der Eigenproblembühne berücksichtigt werden muss, zieht dies als Konsequenz eine Anpassung des Algorithmus' ?? nach sich. Wir können diesen in der folgenden Manier abwandeln.

---

**Algorithmus 2** Berechnung von Ritz-Paaren

---

- 1: Berechne ONB  $\{u_i\}_{i=1:m}$  von  $\mathcal{U}_m$  und setze  $U_m \leftarrow [u_i]_{i=1:m}$ .
  - 2: Setze  $\tilde{A} \leftarrow U_m^H AU_m$ ,  $\tilde{B} \leftarrow U_m^H BU_m$  und löse  $\tilde{A}Y = \tilde{B}Y\Lambda$  in  $Y$  und  $\Lambda$ .
  - 3: Setze  $\tilde{X} \leftarrow U_m Y$  und gib Ritz-Paare  $(\lambda_i, \tilde{x}_i)_{i=1:m}$  aus.
- 

Nun mag der Fall eintreten, dass bei der Berechnung der Ritz-Paare gewisse Toleranzen berücksichtigt werden müssen. Es könnten etwa maximale euklidische Abweichungen der Ritz-Werte zu den Eigenwerten vorgegeben sein. Auch Vorgaben bezüglich des Winkels zwischen Eigenraum und dem Unterraum der Ritz-Vektoren sind denkbar. Dann ist es erforderlich die Wahl des Suchraums nicht dem Geschick alleine zu überlassen, sondern eine gegebenenfalls weniger geschickte Wahl algorithmisch kompensieren zu können. Dies kann zum Beispiel durch das Einführen einer Iterationsvorschrift gelingen, die in jedem Iterationsschritt den Suchraum ändert. Wir werden diese Idee zu einem späteren Zeitpunkt in dieser Arbeit weiterverfolgen.

Wir wollen uns nun vorerst von der Algorithmik verabschieden und einige Beobachtungen festhalten. Eingangs wurde behauptet, dass die Invarianz des Suchraumes  $\mathcal{U}_m$  unter  $A$  dazu führt, dass jedes Ritz-Paar bezüglich  $\mathcal{U}_m$  ein Eigenpaar von  $A$  ist. Doch warum ist das so? Dies zu beantworten verpflichtet sich der folgende Satz.

**Satz 3.2.** Neben zwei Matrizen  $A, B \in \mathbb{C}^{n,n}$  – wobei  $B$  eine HPD-Matrix ist – sei für  $m \in \mathbb{N}$  mit  $m \leq n$  ein  $m$ -dimensionaler Unterraum  $\mathcal{U}_m \subseteq \mathbb{C}_n$  gegeben. Ist dieser invariant unter  $B^{-1}A$ , so ist jedes Ritzpaar von  $B^{-1}A$  bezüglich  $\mathcal{U}_m$  auch ein Eigenpaar von  $(A, B)$ .

*Beweis.* Beginnen wir mit einer ONB  $\{u_i\}_{i=1:m} \subseteq \mathcal{U}_m$  des Unterraums  $\mathcal{U}_m$  und setzen wie bisher  $U_m := [u_i]_{i=1:m} \in \mathbb{C}^{n,m}$ . Aufgrund der Invarianz von  $\mathcal{U}_m$  unter  $B^{-1}A$  muss eine Matrix  $V_m \in \mathbb{C}^{m,m}$  existieren, welche die Gleichung

$$B^{-1}AU_m = U_m V_m \quad (3.6)$$

erfüllt. Insbesondere folgt unter Ausnutzung der Orthogonalität der Spalten von  $U_m$  die Identität  $U_m^H B^{-1}AU_m = V_m$ . Sind nun  $\lambda \in \mathbb{C}$  und  $y \in \mathbb{C}^m \setminus \{0\}$  so gewählt, dass  $(\lambda, U_my)$

ein Ritz-Paar von  $B^{-1}A$  ist, so folgt aus

$$U_m^H B^{-1} A U_m y = V_m y$$

mit (3.6) die Gleichung

$$B^{-1} A U_m y = U_m V_m y = \lambda U_m y$$

und schließlich auch die Behauptung durch Umstellen.  $\square$

Da das gewöhnliche Eigenwertproblem ein Spezialfall des allgemeinen Eigenwertproblems ist, können wir aus dem eben bewiesenen Resultat unmittelbar das folgende Korollar ableiten.

**Korollar 3.3.** Ist  $A \in \mathbb{C}^{n,n}$  und  $\mathcal{U}_m \subseteq \mathbb{C}^n$  ein  $m$ -dimensionaler  $A$ -invarianter Unterraum, so ist jedes Ritz-Paar von  $A$  bezüglich  $\mathcal{U}_m$  ein Eigenpaar von  $A$ .

*Beweis.* Betrachte für  $B := I_n$  das verallgemeinerte Eigenwertproblem

$$Ax = \lambda Bx = \lambda x.$$

Die Aussage folgt dann aus dem vorigen Satz.  $\square$

Es wird sich zu einem späteren Zeitpunkt herausstellen, dass diese Erkenntnis aus algorithmischer Sicht höchst nützlich ist. Die Rayleigh-Ritz-Methode lässt sich nämlich in ein iteratives Verfahren umwandeln, welches in jedem Schritt den Suchraum ändert. Ist die Suchraumiterierte irgendwann einmal  $A$ -invariant, so kann der Algorithmus abgebrochen werden.

Wenden wir uns einer weiteren Kuriosität zu. Zwischen obigem Verfahren und dem Konzept der Projektion besteht ein enger Zusammenhang. Um zu dieser Einsicht zu gelangen, bedarf es ein wenig Vorbereitung.

**Satz 3.4.** Es sei  $B \in \mathbb{C}^{n,n}$  eine HPD-Matrix und für  $m \in \mathbb{N}$  mit  $m \leq n$  sei ein  $m$ -dimensionaler Unterraum  $\mathcal{U}_m \subseteq \mathbb{C}^n$  gegeben. Sei weiter  $\{u_i\}_{i=1:m} \subseteq \mathcal{U}_m$  eine Basis  $B$ -orthonormaler Vektoren, das heißt, die Matrix  $U_m := [u_i]_{i=1:m} \in \mathbb{C}^{n,m}$  erfülle die Gleichung  $U_m^H B U_m = I_m$ . Dann ist die von der Matrix  $P := U_m U_m^H B \in \mathbb{C}^{n,n}$  induzierte lineare Abbildung

$$p: \mathbb{C}^n \rightarrow \mathbb{C}^n, x \mapsto U_m U_m^H B x$$

eine  $B$ -orthogonale Projektion auf den Unterraum  $\mathcal{U}_m$ . Außerdem gilt für alle  $x \in \mathbb{C}^n$  die Identität

$$\|x - p(x)\|_B = \min_{y \in \mathcal{U}_m} \|x - y\|_B.$$

*Beweis.* Aus der  $B$ -Orthogonalität von  $U_m$  folgt für alle  $x \in \mathbb{C}^n$

$$p^2(x) = U_m U_m^H B U_m U_m^H B x = U_m U_m^H B x = p(x).$$

Ist nun  $y \in \mathcal{U}_m$  und  $x \in \mathbb{C}^n$ , dann folgt wegen  $p(y) = y$  und  $B = B^H$  auch

$$\begin{aligned}\langle y, x - p(x) \rangle_B &= y^H (Bx - BU_m U_m^H Bx) \\ &= y^H (Bx - B^H U_m U_m^H Bx) \\ &= y^H Bx - p(y)^H Bx = 0.\end{aligned}$$

Es gilt demnach  $x - p(x) \perp_B \mathcal{U}_m$ . Die Optimierungsaufgabe wird schließlich wegen

$$\begin{aligned}\|x - y\|_B^2 &= \|x - p(x) + p(x) - y\|_B^2 \\ &= \|x - p(x)\|_B^2 + \|p(x) - y\|_B^2 \\ &\geq \|x - p(x)\|_B^2\end{aligned}$$

gelöst. Dabei ist bei der zweiten Gleichheit zu berücksichtigen, dass  $x - p(x) \in \mathcal{U}_m^{\perp_B}$  und  $p(x) - y \in \mathcal{U}_m$  gilt.  $\square$

Mit Hilfe der im Satz eingeführten Projektionsmatrix  $P$ , lässt sich die Gleichung (3.5) – zu finden auf Seite 11 – umformulieren. Durch Linksmultiplikation mit  $B^H U_m$  und dem Einschub der Identität  $U_m^H B U_m$  gilt nämlich

$$B^H U_m (U_m^H A U_m) (U_m^H B U_m) y = \tilde{\lambda} B^H U_m (U_m^H B U_m) (U_m^H B U_m) y$$

und folglich

$$P^H A P U_m y = \tilde{\lambda} P^H B P U_m y.$$

Wenn wir uns an dieser Stelle erinnern, dass  $\tilde{x}$  durch  $U_m y$  ersetzt wurde, so erhalten wir

$$P^H A P \tilde{x} = \tilde{\lambda} P^H B P \tilde{x}.$$

Im Fall der gewöhnlichen Eigenwertgleichung erhalten wir speziell

$$P A P \tilde{x} = \tilde{\lambda} \tilde{x}.$$

Wenden wir uns nun nochmal dem Rayleigh-Ritz-Verfahren zu und betrachten das verallgemeinerte Eigenwertproblem für eine hermitesche Matrix  $A \in \mathbb{C}^{n,n}$  sowie eine HPD-Matrix  $B \in \mathbb{C}^{n,n}$ . Angenommen, uns stünde mit dem Suchraum  $\mathcal{U}_m$  bereits ein von Eigenvektoren aufgespannter Unterraum zur Verfügung. Dann existiert eine Matrix  $U_m := [u_i]_{i=1:m} \in \mathbb{C}^{n,m}$  mit  $B$ -orthonormalen Spalten aus Eigenvektoren von  $(A, B)$  und  $\text{Bild}(U_m) = \mathcal{U}_m$ , sowie eine reelle Diagonalmatrix  $\Lambda_m := \text{diag}(\lambda_i)_{i=1:m} \in \mathbb{C}^{m,m}$  mit

$$A U_m = B U_m \Lambda_m.$$

Dabei entsprechen die Diagonaleinträge von  $\Lambda_m$  gerade den zu den Spalten von  $U_m$  korrespondierenden Eigenwerten von  $(A, B)$ . Folglich wäre das Bild des Spektralprojektors

$P = U_m U_m^H B$  nach Konstruktion wegen

$$B^{-1}AP = B^{-1}(AU_m)U_m^H B = B^{-1}(BU_m\Lambda)U_m^H B = U_m\Lambda U_m^H B = \left( \sum_{i=1}^m \lambda_i u_i u_i^H \right) \cdot B$$

ein  $(B^{-1}A)$ -invarianter Unterraum. Daher erscheint es wegen Satz 3.2 sinnvoll, den Spektralprojektor  $P$  in das Rayleigh-Ritz-Verfahren zu integrieren. Da jedoch der Unterraum  $\mathcal{U}_m$  im Allgemeinen nicht bekannt ist, stellt sich die Frage, ob es überhaupt sinnvoll ist nach dem Projektor zu suchen. Wir werden jedoch alsbald feststellen, dass die Matrix  $U_m U_m^H$  analytisch berechnbar ist.

## 3.2 Konturintegration

Die folgenden Zeilen präsentieren eine Vorgehensweise, die es gestattet, den im vorigen Abschnitt eingeführten Spektralprojektor analytisch zu konstruieren. Wir werden zu einem späteren Zeitpunkt untersuchen, in wie weit die Numerik in der Lage ist, dieses Verfahren algorithmisch umzusetzen.

Zunächst ist es erforderlich einige Begriffe und Notationen einzuführen. Für zwei Zahlen  $z_1, z_2 \in \mathbb{C}$  sei mit der Menge

$$[z_1, z_2] := \{(1-t)z_1 + tz_2 \in \mathbb{C} \mid t \in [0, 1]\}$$

die direkte Verbindungsstrecke zwischen  $z_1$  und  $z_2$  bezeichnet. Im Spezialfall  $z_1, z_2 \in \mathbb{R}$  stellen wir uns daher unter  $[z_1, z_2]$  ein reellwertiges Intervall eingebettet in die komplexe Zahlenebene vor. Des Weiteren benötigen wir sogenannte *Jordan-Kurven*. Diese definieren wir im Kontext dieser Arbeit wie folgt.

**Definition 3.5.** *Es sei  $S := \{z \in \mathbb{C} : |z| = 1\}$ . Eine Menge  $\Gamma \subset \mathbb{C}$  heißt Jordan-Kurve, falls ein Homöomorphismus  $\gamma: S \rightarrow \Gamma$  existiert mit  $\Gamma = \text{Bild}(\gamma)$ . Dabei ist ein Homöomorphismus eine stetige Bijektion mit stetiger Umkehrabbildung.*

Insbesondere sind Jordan-Kurven also geschlossene und überschneidungsfreie Kurven. Ausgestattet mit diesen neuen Vokabeln kann die nun Präsentation des angekündigten Verfahrens beginnen. Im weiteren Verlauf seien  $A, B \in \mathbb{C}^{n,n}$  hermitesche Matrizen und  $B$  zusätzlich positiv definit.

Wir wollen annehmen, dass zwei Zahlen  $\lambda_1, \lambda_2 \in \mathbb{R}$  mit  $\lambda_1 < \lambda_2$  vorgegeben sind, die das abgeschlossene Intervall  $I := [\lambda_1, \lambda_2]$  definieren. Wie bisher sollen Paare der Gestalt  $(\lambda, x) \in I \times \mathbb{C}^n \setminus \{0\}$  ermittelt werden, welche der verallgemeinerten Eigenwertgleichung

$$Ax = \lambda Bx \tag{3.7}$$

genügen. Der Ausgangspunkt zur Bestimmung der gesuchten Eigenpaare ist die durch

$$\begin{aligned} G: \Omega &\rightarrow \mathbb{C}^{n,n} \\ \omega &\mapsto (\omega B - A)^{-1} \end{aligned}$$

definierte *Green-Funktion*, wobei  $\Omega \subseteq \mathbb{C}$  eine vom Spektrum von  $B^{-1}A$  disjunkte Teilmenge der komplexen Zahlen ist.<sup>1</sup> Diese Funktion  $G$  wird nun über eine Jordan-Kurve  $\Gamma$ , die um das vorgegebene Intervall  $I$  „herumläuft“, in der Gestalt

$$\frac{1}{2\pi\iota} \int_{\Gamma} G(\omega) \, d\omega \tag{3.8}$$

integriert.<sup>2</sup> Das Integral ist hierbei eintragsweise zu verstehen.

Wir wollen nun annehmen, dass es genau  $m \in \mathbb{N}$  Eigenwerte gibt, die im Inneren von  $I$  liegen. Dann kann man nachweisen, dass sich (3.8) durch eine Matrix  $U_m = [u_i]_{i=1:m} \in \mathbb{C}^{n,m}$  mit  $B$ -orthonormalen Spalten in

$$\frac{1}{2\pi\iota} \int_{\gamma} G(\omega) \, d\omega = U_m U_m^H$$

faktorisieren lässt, wobei  $u_i$  für jedes  $i = 1 : m$  ein zu einem in  $I$  liegenden Eigenwert korrespondierender Eigenvektor ist.<sup>3</sup> Als direkte Konsequenz erhalten wir folglich mit dem Produkt

$$\left( \frac{1}{2\pi\iota} \int_{\gamma} G(\omega) \, d\omega \right) \cdot B$$

den Spektralprojektor  $U_m U_m^H B$ .

---

<sup>1</sup>Wäre  $\Omega$  nicht disjunkt vom Spektrum von  $B^{-1}A$ , so gäbe es ein  $\omega \in \Omega$ , welches zur Singularität von  $(\omega B - A)$  führte.

<sup>2</sup>Mit  $\iota$  ist im Folgenden stets die imaginäre Einheit bezeichnet, also  $\iota = \sqrt{-1}$ .

<sup>3</sup>Für einen Beweis des Spezialfalls  $B = I$  sei auf [IG96, 64f] und [Lie13, S. 222-226] verwiesen.

### 3.3 Illustration

Bevor wir uns vom zweiten Kapitel verabschieden, sollen die Erkenntnisse der vorangegangenen Abschnitte an einem Beispiel vorgeführt werden. Dabei werden wir ausgehend von einem verallgemeinerten Eigenwertproblem gewisse Eigenwerte filtern und die Konturintegration mit der Ritz-Methode kombinieren.

Dafür wenden wir uns den beiden Matrizen  $A, B \in \mathbb{C}^{3,3}$  zu, welche durch

$$A := \text{diag}(3, 1, 4) \text{ und } B := \text{diag}(1, 5, 9)$$

gegeben sind. Dann ist  $A$  hermitesch und  $B$  eine HPD Matrix. Durch einfaches Nachrechnen überprüft man, dass das verallgemeinerte Eigenwertproblem (3.7) durch Vektoren  $x_1 \in \text{Span}_{\mathbb{C}}\{e_1\}$  mit zugehörigem Eigenwert  $\lambda_1 = 3$ , Vektoren  $x_2 \in \text{Span}_{\mathbb{C}}\{e_2\}$  mit zugehörigem Eigenwert  $\lambda_2 = 1/5$ , sowie Vektoren  $x_3 \in \text{Span}_{\mathbb{C}}\{e_3\}$  mit zugehörigem Eigenwert  $\lambda_3 = 4/9$  gelöst wird.<sup>4</sup>

Vergessen wir für den Moment, dass uns die Eigenpaare bekannt sind und versuchen mit den Methoden, die in den vorangegangenen Abschnitten vorgestellt wurden, die Eigenpaare auf  $I = [-1, 1]$  zu bestimmen. Gemäß Abschnitt 3.2 wählen wir daher als Integrationskontur  $\Gamma$  das Bild der Funktion

$$\gamma: [0, 2\pi] \rightarrow \mathbb{C}^n, \varphi \mapsto e^{i\varphi} \quad (3.9)$$

und integrieren die Green-Funktion

$$G(\omega) = \text{diag}(\omega - 3, 5\omega - 1, 9\omega - 4)^{-1} = \text{diag}\left(\frac{1}{\omega - 3}, \frac{1}{5\omega - 1}, \frac{1}{9\omega - 4}\right)$$

darüber. Nach Definition 2.4 ist  $\Gamma$  eine Jordan-Kurve. Da nach Konstruktion keiner der Eigenwerte im Bild von  $\gamma$  liegt, ist  $G$  auf der gesamten Kontur wohldefiniert. Folgen wir also weiter dem Abschnitt 3.2, so erhalten wir

$$\frac{1}{2\pi i} \int_{\Gamma} G \, ds = \frac{1}{2\pi i} \int_0^{2\pi} G(\gamma(\omega)) \cdot \gamma'(\omega) \, d\omega = \text{diag}\left(0, \frac{1}{5}, \frac{1}{9}\right)$$

Diese Matrix lässt sich mit den  $B$ -orthonormalen Vektoren  $u_1 = 1/\sqrt{5} \cdot e_2$  und  $u_2 = 1/3 \cdot e_3$  wie gewünscht in

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1/6 & 0 \\ 0 & 0 & 1/9 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 1/\sqrt{6} & 0 \\ 0 & 1/3 \end{bmatrix} \begin{bmatrix} 0 & 1/\sqrt{6} & 0 \\ 0 & 0 & 1/3 \end{bmatrix} =: U_2 U_2^H$$

faktorisieren. Nachdem dies geschafft ist, lassen wir uns von der in Abschnitt 3.1 diskutierten Vorgehensweise inspirieren und reduzieren das Problem zunächst mit Hilfe einer

<sup>4</sup>Hier bezeichnen  $e_1, e_2, e_3 \in \mathbb{C}^n$  die kanonischen Einheitsvektoren.

vollrangigen Matrix  $V \in \mathbb{C}^{3,2}$  auf ein Problem kleinerer Dimension. Zu diesem Zwecke wählen wir

$$V := \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$$

und benutzen die Matrix  $Q := U_2 U_2^H V$  um das Ausgangsproblem vermöge  $\tilde{A} := Q^H A Q$  und  $\tilde{B} := Q^H B Q$  auf das Problem

$$\tilde{A}y = \begin{bmatrix} 1/36 & 0 \\ 0 & 1/9 \end{bmatrix} y = \mu \begin{bmatrix} 5/36 & 0 \\ 0 & 1/4 \end{bmatrix} y = \mu \tilde{B}y$$

der Dimension  $(2 \times 2)$  zu transformieren. Die Eigenwerte  $1/5$  und  $4/9$  lassen sich leicht ablesen und stimmen wie erwartet mit den sich auf  $[-1, 1]$  befindlichen Eigenwerten des eingangs formulierten Problems überein. Passende Eigenvektoren sind Elemente aus  $\text{Span}_{\mathbb{C}}\{[1 \ 0]^T\}$  und  $\text{Span}_{\mathbb{C}}\{[0 \ 1]^T\}$ .

Es bleibt die Rücktransformation der Eigenvektoren. Dazu wählen wir aus kosmetischen Gründen die Eigenvektoren

$$y_1 := \begin{bmatrix} \sqrt{6} \\ 0 \end{bmatrix} \text{ und } y_2 := \begin{bmatrix} 0 \\ 3 \end{bmatrix}$$

und setzen  $Y = [y_1 \ y_2]$ . Wir erhalten dann

$$QY = \begin{bmatrix} 0 & 0 \\ 1/6 & 0 \\ 0 & 1/9 \end{bmatrix} \begin{bmatrix} \sqrt{6} & 0 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 1/\sqrt{6} & 0 \\ 0 & 1/3 \end{bmatrix} = U_2.$$

Eine Bemerkung zum Abschluss: Erinnern wir uns an Satz 3.2, so ist es nicht verwunderlich, dass diese Rechnung exakte Eigenpaare liefert. Schließlich hat  $Q$  vollen Rang und das Bild der von  $Q$  induzierten linearen Abbildung ist invariant unter  $B^{-1}A$ .

## Kapitel 4

# Rationales Filtern

Nachdem das vorangegangene Kapitel verschiedene Ideen zum Filtern von Eigenpaaren theoretisch beleuchtet hat, werden wir uns nun mit der Frage der praktischen Umsetzbarkeit beschäftigen. Im Mittelpunkt wird dabei die Konstruktion geeigneter Suchräume stehen, welche im Rayleigh-Ritz-Verfahren zum Einsatz kommen sollen. Es wird sich zeigen, dass die Konturintegration hierbei ein nützliches Hilfsmittel darstellt.

### 4.1 Beschleunigte Rayleigh-Ritz Iteration

Zu Beginn des zweiten Kapitels wurde mit dem Rayleigh-Ritz-Verfahren eine Methode vorgestellt, mit dessen Hilfe Ritz-Paare bezüglich eines gewissen Suchraums bestimmt werden können. Es wurde angedeutet, dass das Einbinden einer geeigneten Iterationsvorschrift zur Verringerung der Abweichungen von Ritz-Paaren zu tatsächlichen Eigenpaaren führen kann. Dies soll im Folgenden genauer ausgeführt werden.

1

Ausgangspunkt für unsere Betrachtungen ist wie bisher das hermitesche Eigenwertproblem

$$Ax = \lambda Bx. \tag{4.1}$$

Betrachten wir zunächst für den Spezialfall  $B = I$  eine sehr leicht umsetzbare Methode um Eigenpaare zu bestimmen – die sogenannte *Potenzmethode*. Ausgehend von einem Startvektor  $y_{(0)} \in \mathbb{C}^n \setminus \{0\}$  wird in jeder Iteration der Vektor

$$y_{(k+1)} = \frac{1}{\|A^{k+1}y_{(0)}\|} A^{k+1}y_{(0)}$$

oder dazu äquivalent

$$y_{(k+1)} = \frac{1}{\|Ay_{(k)}\|} Ay_{(k)}$$

berechnet. Dieser Vorgang wird wiederholt, bis gewisse Abbruchkriterien erfüllt sind. Um einzusehen, dass die Folge der Iterierten gegen den zum betragsmäßig größten Eigenwert

---

<sup>1</sup>Um besser zu verstehen orientieren wir uns bei der Einführung des Algs an der Dramaturgie von Saad blabla 115ff



gehörenden Eigenvektor konvergiert, begnügen wir uns an dieser Stelle mit folgender Plausibilitätsbetrachtung, welche in ähnlicher Form in [Ste13, S. 56] zu finden ist.

Seien  $(\lambda_i, x_i)_{i=1:n}$  die Eigenpaare von  $A$ . Aufgrund der Hermitizität bilden dann die Eigenvektoren eine Basis des  $\mathbb{C}^n$ . Folglich existieren komplexwertige Skalare  $\alpha_1, \dots, \alpha_n$  mit

$$y_{(0)} = \sum_{i=1}^n \alpha_i x_i.$$

Wenden wir nun die  $k$ -te Potenz von  $A$  auf  $y_{(0)}$  an, ergibt sich daher wegen  $A^k x_i = \lambda_i^k x_i$

$$A^k y_{(0)} = \sum_{i=1}^n \alpha_i \lambda_i^k x_i. \quad (4.2)$$

Wir wollen ohne Einschränkung annehmen, dass  $|\lambda_1| \geq |\lambda_i|$  für alle  $i$  mit  $1 < i \leq n$  gilt. Gegebenenfalls nummerieren wir die Eigenpaare und Skalare um. Dann wird für größer werdende  $k$  die rechte Seite von (4.2) durch den Term  $\alpha_1 \lambda_1^k x_1$  dominiert. In Kombination mit der Normierung führt dies letztlich zur behaupteten Approximierung.<sup>2</sup>

Da uns daran gelegen ist, nicht nur mit einzelnen Elementen des  $\mathbb{C}^n$  zu arbeiten sondern mit Matrizen zu hantieren, müssen wir ein allgemeinere Form der Potenzmethode betrachten. In seinen Abhandlungen über Unterraumiterationen schlägt Y. Saad in [Saa11, 118 f.] folgende Methode vor.<sup>3</sup>

---

**Algorithmus 3** Rayleigh-Ritz-Iteration
 

---

- 1: Wähle  $m$  Zufallsvektoren  $V_{(0)} \leftarrow [v_i]_{i=1:m} \in \mathbb{C}^{n,m}$ . Setze  $k \leftarrow 1$  und wähle initialen Exponenten  $p_{(1)} \in \mathbb{N}$ .
  - 2: **repeat**
  - 3:   Setze  $\hat{Q}_{(k)} \leftarrow A^{p_{(k)}} V_{(k-1)}$ .
  - 4:   Orthonormalisiere die Spalten von  $\hat{Q}_{(k)}$  zu  $Q_{(k)}$ .
  - 5:   Setze  $\tilde{A}_{(k)} \leftarrow Q_{(k)}^H A Q_{(k)}$ .
  - 6:   Berechne die Schurvektoren  $\tilde{X}_{(k)} \leftarrow [x_i]_{i=1:m}$  von  $\tilde{A}$  mit Hilfe des QR-Algorithmus’.
  - 7:   Setze  $V_{(k)} \leftarrow Q_{(k)} \tilde{X}_{(k)}$ .
  - 8:    $k \leftarrow k + 1$ .
  - 9: **until** Verfahren konvergiert.
- 

Hierbei ist  $(p_{(k)})_{k \in \mathbb{N}}$  eine Folge von natürlichzahligen Exponenten auf die wir später eingehen werden. Zuvor begutachten wir die Zeilen vier bis sieben. Hier ist eine deutliche Nähe zu dem Rayleigh-Ritz-Verfahren zu erkennen. Als Suchraum wird die lineare Hülle der Spaltenvektoren von  $\hat{Q}_{(k)}$  benutzt und die Berechnung von Ritz-Vektoren erfolgt in Schritt sechs – schließlich stimmen unserer Voraussetzungen wegen die Schurvektoren gerade mit den Eigenvektoren von  $\tilde{A}_{(k)}$  überein. Wir verzichten an dieser Stelle auf Ausführungen über die Funktionsweise des *QR-Algorithmus*’. Der interessierte Leser wende sich zum Beispiel

---

<sup>2</sup>Ein formaler Beweis ist im Anhang zu finden.

<sup>3</sup>Der Algorithmus wurde vom Autor frei in die deutsche Sprache übersetzt und die Notation im Sinne der Konsistenz dieser Arbeit angepasst.

an [Ste13, 55 ff.].

Nun zur Folge der Exponenten. Im Spezialfall, dass  $p_{(k)} = k$  für alle  $k \in \mathbb{N}$  gilt, ist obiger Algorithmus stark mit der *Potenzmethode* zur Bestimmung von Eigenpaaren verwandt, ja sogar eine Verallgemeinerung davon.<sup>4</sup> Würden die Exponenten so gewählt werden, müsste entsprechend oft der dritte Schritt des Algorithmus<sup>5</sup> durchgeführt werden. Y. Saad [Saa11, S. 116] merkt an, dass es aufgrund der hohen Kosten dieser Berechnung sinnvoll sein kann, die Wahl der Exponenten flexibler zu gestalten.

man will nicht dauern orthogonalisieren, daher blabla s. 116 in saad

---

**Algorithmus 4** Beschleunigtes iteratives Rayleigh-Ritz-Verfahren

---

- 1: Wähle  $m$  Zufallsvektoren  $V_{(0)} \leftarrow [v_i]_{i=1:m} \in \mathbb{C}^{n,m}$ . Setze  $k \leftarrow 1$ .
  - 2: **repeat**
  - 3:   Approximiere den Unterraumprojektor:  $Q_{(k)} \leftarrow \rho(B^{-1}A)V_{(k-1)}$
  - 4:   Reduziere die Dimension:  $\tilde{A}_{(k)} \leftarrow Q_{(k)}^H A Q_{(k)}$ ,  $\tilde{B}_{(k)} \leftarrow Q_{(k)}^H B Q_{(k)}$ .
  - 5:   Löse das transformierte Problem  $\tilde{A}_{(k)} \tilde{X}_{(k)} = \tilde{B}_{(k)} \tilde{X}_{(k)} \tilde{\Lambda}_{(k)}$  in  $\tilde{X}_{(k)}$  und  $\tilde{\Lambda}_{(k)}$ .
  - 6:   Setze  $V_{(k)} \leftarrow Q_{(k)} \tilde{X}_{(k)}$ .
  - 7:    $k \leftarrow k + 1$ .
  - 8: **until** Abbruchkriterium ist erfüllt.
- 

Die obigen Fragen sind also leicht beantwortet: In dieser Variante des Rayleigh-Ritz-Verfahrens wird solange über den Suchraum iteriert, bis die ermittelten Ritz-Paare gewisse Anforderungen erfüllen. Es drängt sich jedoch eine weitere Frage auf. Welche Funktion erfüllt das Polynom  $\rho$ ?

Im Kontext dieses Algorithmus<sup>6</sup> wird  $\rho$  auch als *Filter* oder *Beschleuniger* bezeichnet. Von dessen Wahl hängt nämlich ab, ob und wie gut Eigenpaare approximiert werden: Sei  $[\lambda_1, \lambda_2]$  ein reelles Intervall, auf dem  $l \in \mathbb{N}$  Eigenwerte und die korrespondierenden Eigenvektoren gefunden werden können. Ist nun  $\rho(B^{-1}A)$  der Spektralprojektor,  $m = l$  und hat die Matrix  $Q_{(1)} = \rho(B^{-1}A)V_{(0)}$  vollen Rang, so konvergiert der Algorithmus ?? in einer Iteration (Vgl. [TP16, S. 356]). Dies folgt unter Ausnutzung der Invarianz des Bildes von  $Q_{(1)}$  unter  $B^{-1}A$  aus dem Satz 3.2.

Da der Spektralprojektor in den meisten Fällen unbekannt sein dürfte, liegt die Idee nahe, ihn wenigstens zu approximieren. Tang und Polizzi [TP16, S. 356] merken an, dass dies gut funktioniert, falls  $\rho$  eine durch *Gauß-Legendre-Quadratur* konstruierte rationale Funktion ist. Um den Gedankengang der Autoren nachvollziehen zu können, wird sich das folgende Intermezzo mit der Auffrischung des Konzeptes von Quadraturformeln und rationalen Funktionen beschäftigen. Im Anschluss fahren wir mit der Konstruktion des Projektors fort.

---

<sup>4</sup>Vgl. [Saa11, 115 ff.] und [JK15, 1 f.].

## 4.2 Gauß'sche Quadratur

Um ein Integral numerisch zu approximieren, bedient man sich sogenannter *Quadraturformeln*. Dazu betrachten wir eine stetige Funktion  $f: \mathbb{R} \rightarrow \mathbb{R}$ , welche wir auf einem gegebenen Intervall  $I := [a, b] \subset \mathbb{R}$  integrieren wollen.<sup>5</sup> Zu gegebenen Stützpunkten  $(x_i, f(x_i))_{i=0:n}$  auf  $I \times \mathbb{R}$  sei  $p_n$  das zugehörige *Interpolationspolynom* vom Grad  $n$ , also ein Polynom, welches  $p(x_i) = f(x_i)$  für alle  $i$  mit  $0 \leq i \leq n$  erfüllt.

Dann bezeichnen wir die Näherung

$$Q_n(f) := \int_a^b p_n(x) \, dx = (b-a) \sum_{i=0}^n \omega_i f(x_i) \quad (4.3)$$

als *interpolatorische Quadraturformel*. Dabei gilt

$$\omega_k = \int_0^1 \prod_{j=0, j \neq k}^n \frac{t - t_j}{t_k - t_j} \, dt, \quad t_j = \frac{x_j - a}{b - a}.$$

Die Qualität der Approximation, also die Abweichung vom exakten Integral, hängt maßgeblich von der Wahl und Anzahl der Stützpunkte ab. Wollten wir beispielsweise das Integral einer konstanten Funktion berechnen, so erschiene es wenig plausibel, anstelle der direkten Berechnung ein Polynom vom Grad 69 auf 70 Stützstellen für die Approximation zu bemühen.

Bei der Anwendung von Gauß-Legendre-Quadraturen ergibt sich die Wahl der Stützpunkte durch die Berechnung von Nullstellen von Polynomen, die in einer Orthogonalitätsbeziehung zueinander stehen. Wir werden gleich formal formulieren, wie dies zu verstehen ist.

Ausgangspunkt für die Integration ist nun eine stetige Funktion  $f$ , die eine Faktorisierung in zwei stetige Funktionen  $g$  und  $\omega$  der Art

$$f = g \cdot \omega$$

besitzt, wobei  $\omega$  auf dem Integrationsintervall  $[a, b]$  positiv sein soll.<sup>6</sup> Ziel ist daher die Berechnung von

$$\int_a^b g(x) \omega(x) \, dx. \quad (4.4)$$

Wir fordern nun, dass (4.3) mit (4.4) für alle Polynome bis zum Grad  $(2n-1)$  übereinstimmt. Dazu betrachten wir die Standardbasis  $\{x^i\}_{i=1:(2n-1)}$  auf dem Raum der Polynome

<sup>5</sup>Im Allgemeinen ist die Stetigkeit von  $f$  nicht zwingend erforderlich. Wir werden uns hier der Einfachheit halber auf stetige Funktionen einschränken.

<sup>6</sup>Im unstetigen Fall darf die Funktion  $\omega$  in höchstens endlich vielen Punkten negativ sein.

vom Grad  $2n - 1$ . Dann landen wir unweigerlich bei dem Gleichungssystem

$$\sum_{j=1}^n \omega_j x_j^k = \int_a^b x^k \omega(x) \, dx \text{ mit } k = 0, \dots, 2n - 1.$$

Man kann zeigen, dass die Lösung dieses Systems durch Nullstellen eines Polynoms gegeben ist, welches durch ein Gram-Schmidt-Orthogonalisierungsverfahren bezüglich des Skalarproduktes

$$\langle p, q \rangle_\omega := \int_a^b p(t)q(t)\omega(t) \, dt$$

konstruiert wurde. Das heißt konkret: Ausgehend vom Polynom  $p_0 \equiv 1$  ist

$$p_n(x) := x^n - \sum_{j=0}^{n-1} \frac{\langle x^n, p_j \rangle_\omega}{\langle p_j, p_j \rangle_\omega} p_j(x)$$

gerade dasjenige Polynom, durch dessen Nullstellen das obige Gleichungssystem gelöst wird. Sind nun  $x_1, \dots, x_n$  die Nullstellen dieses  $n$ -ten Orthogonalitätspolynoms, so heißt die numerische Integrationsformel

$$Q_n(f) = \sum_{j=1}^n \omega_j f(x_j) \text{ mit } \omega_j = \langle L_j, 1 \rangle_\omega = \int_a^b L_j(x) \rho(x) \, dx$$

*Gauß'sche Quadraturformel der  $n$ -ten Ordnung.* Dabei ist

$$L_j(x) = \prod_{k \neq j=1}^n \frac{x - x_k}{x_k - x_j}.$$

## 4.3 Approximation des Spektralprojektors

Nachdem wir diese Kurzzusammenfassung des Gauß-Quadratur-Verfahrens diskutiert haben, widmen wir uns wieder der Approximation des Spektralprojektors  $P = X_k X_k^H B$ . Dabei übernehmen wir die am Anfang dieses Kapitels eingeführte Notation, sowie die an die Matrizen  $A$  und  $B$  gestellten Voraussetzungen.

Wenden wir uns daher wieder dem reellen Intervall  $I := [\lambda_1, \lambda_2]$  zu. Das Ziel ist die Konstruktion einer rationalen Funktion  $\varrho: \mathbb{C} \rightarrow \mathbb{C}$  mit  $\varrho(\mathbb{R}) \subseteq \mathbb{R}$ , die auf  $I$  näherungsweise der Indikatorfunktion von  $I$  entspricht. Dazu bemühen wir die Cauchy'sche Integraldarstellung der Indikatorfunktion und wandeln diese mit Hilfe numerischer Quadraturformeln in die gewünschte rationale Funktion  $\varrho$  um.

Zuächst zur Indikatorfunktion: Ist  $c \in \mathbb{R}$  der Mittelpunkt des Intervalls  $I$  und  $r$  der Abstand des Mittelpunktes zum Rand des Intervalls, dann entspricht die Menge

$$\mathcal{C} := \{z \in \mathbb{C} : |z - c| = r\}$$

gerade einer Sphäre mit Radius  $r$  um  $c$ . Mit dem Cauchy'schen Integralsatz lässt sich zeigen, dass im Falle  $z \notin \mathcal{C}$

$$\frac{1}{2\pi\iota} \int_{\mathcal{C}} \frac{1}{\omega - z} d\omega = \begin{cases} 1 & \text{falls } |z - c| < r \\ 0 & \text{falls } |z - c| > r \end{cases}$$

gilt. Dieses Integral gilt es nun zu approximieren. **wie genau?** (Vgl. [Jäge, S. 20]).

Wir werden das eben diskutierte Integral mit einer Gauß-Legendre Quadraturformel annähern. Dazu ist es nötig die Kontur  $\mathcal{C}$  so zu parametrisieren, dass eine Transformation der Integrationsgrenzen auf das Intervall  $[-1, 1]$  möglichst einfach möglich ist. Zu diesem Zweck sei

$$\begin{aligned} \gamma: [-1, 3] &\rightarrow \mathbb{C} \\ t &\mapsto c + re^{t\frac{\pi}{2}(1+t)} \end{aligned}$$

als Parametrisierung der Sphäre gewählt. Die Ableitung von  $\gamma$  ist dann für jedes  $t \in [-1, 3]$  durch

$$\gamma'(t) = \iota \frac{\pi}{2} r e^{t\frac{\pi}{2}(1+t)}$$

gegeben. Mit den bekannten Regeln der Integration erhalten wir somit für alle  $z \notin \mathcal{C}$

$$\begin{aligned} \frac{1}{2\pi\iota} \int_{\mathcal{C}} \frac{1}{\omega - z} d\omega &= \frac{1}{2\pi\iota} \int_{-1}^3 \frac{\gamma'(t)}{\gamma(t) - z} dt \\ &= \frac{1}{2\pi\iota} \left( \int_{-1}^1 \frac{\gamma'(t)}{\gamma(t) - z} dt + \int_1^3 \frac{\gamma'(t)}{\gamma(t) - z} dt \right) \\ &= \frac{1}{2\pi\iota} \left( \int_{-1}^1 \frac{\gamma'(t)}{\gamma(t) - z} dt + \int_{-1}^1 \frac{\gamma'(2-t)}{\gamma(2-t) - z} dt \right) \\ &= \frac{1}{2\pi\iota} \int_{-1}^1 \left( \frac{\gamma'(t)}{\gamma(t) - z} + \frac{\overline{\gamma'(t)}}{\overline{\gamma(t)} - z} \right) dt \end{aligned}$$

wobei  $\overline{\gamma(t)}$  und  $\overline{\gamma'(t)}$  die komplexen Konjugationen von  $\gamma(t)$  beziehungsweise  $\gamma'(t)$  bezeichnen.

Für  $q \in \mathbb{N}$  mit **wie groß ist  $q$  genau?** seien  $(w_j, t_j)_{j=1:q}$  die für die Gauß-Legendre-Quadratur benötigten Gewichte und Diskretisierungspunkte. Dann setzen wir

$$\rho(z) := \frac{1}{2\pi\iota} \sum_{j=1}^q \left( \frac{w_j \cdot \gamma'(t_j)}{\gamma(t_j) - z} - \frac{w_j \cdot \overline{\gamma'(t_j)}}{\overline{\gamma(t_j)} - z} \right)$$

und erhalten nach der Substitution  $\gamma(t_j) := \gamma_j$  und  $\sigma_j := w_j \gamma'(t_j)/(2\pi\iota)$  die gewünschte rationale Funktion

$$\varrho: \mathbb{C} \rightarrow \mathbb{C}, z \mapsto \sum_{j=1}^q \left( \frac{\sigma_j}{\gamma_j - z} + \frac{\overline{\sigma_j}}{\overline{\gamma_j} - z} \right)$$

zur Approximation der Indikatorfunktion. Hierbei ist bemerkenswert, dass die rationale Funktion bereits in Partialbruchzerlegung vorliegt. Setzen wir schließlich  $B^{-1}A$  in die rationale Funktion ein, so erhalten wir

$$\begin{aligned}\varrho(B^{-1}A) &= \sum_{k=1}^q \sigma_k(\gamma_k I - B^{-1}A)^{-1} + \sum_{k=1}^q \overline{\sigma_k}(\overline{\gamma_k} I - B^{-1}A)^{-1} \\ &= \sum_{k=1}^q \sigma_k(\gamma_k B - A)^{-1} B + \sum_{k=1}^q \overline{\sigma_k}(\overline{\gamma_k} B - A)^{-1} B\end{aligned}$$

und folglich

$$\varrho(B^{-1}A)V = \sum_{k=1}^q \sigma_k(\gamma_k B - A)^{-1} B V + \sum_{k=1}^q \overline{\sigma_k}(\overline{\gamma_k} B - A)^{-1} B V$$

für eine Matrix  $V \in \mathbb{C}^{n,q}$ .

## Kapitel 5

# Numerische Experimente

Im abschließenden Kapitel werden wir die Methoden, die in den Kapiteln zwei und drei vorgestellt wurden, naiv implementieren. Dabei werden wir uns auf reelle, symmetrische Eigenwertprobleme beschränken. Im Folgenden untersuchen wir also das Problem

$$Ax = \lambda Bx$$

für zwei symmetrische Matrizen  $A, B \in \mathbb{R}^{n,n}$  und fordern außerdem die positive Definitheit von  $B$ . Desweiteren setzen wir voraus, dass der Input des Users sinnvoll ist und den für die Funktionsfähigkeit der Algorithmen notwendigen Voraussetzungen genügt.

Die Implementation erfolgt durchweg mit *Octave*, ist aber ohne Weiteres in *MATLAB* übersetzbar.

### 5.1 Rayleigh-Ritz

Beginnen wir mit dem Rayleigh-Ritz-Verfahren (Algorithmus ??). Ähnlich dem Pseudocode, bezeichnen wir mit  $U$  den Suchraum und verzichten dabei im Code auf den Index  $m$ . Zur Orthogonalisierung des Suchraumes benutzen wir die bereits in Octave implementierte Funktion

$$\text{orth}(U)$$

und die `eig`-Funktion zum Lösen des reduzierten Eigenwertproblems. Eine Implementation könnte also wie folgt aussehen.

```
1 %% Eingabe: Matrizen A, B des Eigenwertproblems,
2 %%           Suchraum U
3 %% Ausgabe: Matrix X mit Ritz-Vektoren,
4 %%           Zeilenvektor L mit Ritz-Werten
5
6 %% Funktionsaufruf: [X, L] = rayleighRitz(A, B, U)
7
8 function[X,lambda] = rayleighRitz(A,B,U);
9
```

```

10 % orthonormiere U
11 U = orth(U);
12
13 % transformiere Problem
14 A = U'*A*U; B = U'*B*U;
15
16 % loese reduziertes Problem
17 [L,Y] = eig(A,B);
18
19 % berechne Ritz-Paare
20 X = U*Y;
21 L = diag(L);

```

Algorithmus 5.1: Implementation des Rayleigh-Ritz Verfahrens zur Bestimmung eines Ritz-Paares.

## 5.2 Konturintegration

Zunächst wählen wir also für  $k \in \mathbb{N}$  eine vollrangige Zufallsmatrix  $Y_k$ , um die Transformationsmatrix  $Q$  zu ermitteln. Dazu führen wir die GREEN-Funktion

$$\text{funGreen} = @(w) \text{inv}(w.*B-A)$$

ein und wählen eine passende Integrationskontur  $C$ . Um es einfach zu halten, wird diese in der Implementation die Menge

$$C = \left\{ z \in \mathbb{C} : \left\| z - \frac{\lambda_{\max} - \lambda_{\min}}{2} \right\|_1 = 1 \right\}$$

sein, wobei  $\lambda_{\min}$  und  $\lambda_{\max}$  die vorgegebenen Intervallgrenzen bezeichnen. Die entsprechenden Quadraturpunkte werden im Quellcode durch die Variablen

$$n0, n1, n2, n3$$

repräsentiert.

Ist  $Q$  bestimmt, benutzen wir die MATLAB interne Eigenwertfunktion

$$\text{eig}(Aq, Bq)$$

um die RITZ-Paare des transformierten Problems zu berechnen. Schließlich berechnen wir die Eigenwerte und Eigenvektoren des ursprünglichen Problems. Die Matrix  $\Phi_k$  wird im Quelltext mit  $v_q$  bezeichnet.

Wir wollen annehmen, dass der Input des Users die Voraussetzungen des Problems (3.7) erfüllt und die Anzahl  $k \in \mathbb{N}$  der Eigenwerte im Intervall  $[\lambda_{\min}, \lambda_{\max}]$  bekannt ist. Eine naive Implementation könnte daher wie folgt aussehen:



```

1  %% Eingabe: Matrizen A, B des Eigenwertproblems,
2  %%           Intervallgrenzen lmin, lmax,
3  %%           Anzahl k der erwarteten Eigenwerte
4  %% Ausgabe: Matrix V mit Eigenvektoren,
5  %%           Matrix D mit Eigenwerten auf der Diagonalen
6
7  %% Funktionsaufruf: [V, D] = feast(A, B, lmin, lmax, k)
8
9  function[V, D] = feast(A, B, lmin, lmax, k);
10
11 n = length(A);
12
13 % erzeuge Zufallsmatrix
14 Y = randn(n, k);
15
16 % generiere Green-Funktion
17 funGreen = @(w) inv( w.*B - A );
18
19 % berechne die Integrationsknotenpunkte
20 n0 = lmax;
21 n1 = (lmin + lmax)/2 + (lmax - (lmin + lmax)/2)*i;
22 n2 = lmin;
23 n3 = (lmin + lmax)/2 - (lmax - (lmin + lmax)/2)*i;
24
25 % Integrationskontur
26 C = [n1 n2 n3];
27
28 % berechne Transformationsmatrix Q
29 Q = (-1/(2*pi*i))*integral(funGreen, n0, n0,...
30     'Waypoints', C, 'ArrayValued', true) * Y;
31
32 % berechne Ritz-Paare des transformierten Problems
33 Aq = Q'*A*Q; Bq = Q'*B*Q;
34 [Vq, D] = eig(Aq, Bq);
35
36 % berechne die gesuchten Eigenvektoren
37 V = Q*Vq;

```

Algorithmus 5.2: Naive FEAST Implementation.

Überprüfen wir die Funktionalität mit den Matrizen  $A$  und  $B$  aus dem Abschnitt 3.3, erhalten wir folgende Ausgabe:

```

>> [V, D] = feast(A, B, -3, 3, 2)

V =

-1.4142 - 0.0000i    0.0000 + 0.0000i
 0.0000 + 0.0000i    1.4142 - 0.0000i
-0.0000 + 0.0000i   -0.0000 + 0.0000i

```

```
D =  
  
    0.0000         0  
         0    2.0000  
  
>>
```

Die vorgestellte Umsetzung des FEAST-Algorithmus' lässt sich ohne Weiteres auf dynamische Matrixdimensionen übertragen. Da sich mit zunehmender Dimension unter Umständen auch die Kardinalität des Spektrums und die Anzahl der Eigenräume ändert, sollten gegebenenfalls Optimierungsmaßnahmen getroffen werden, wie etwa die Wahl der Knotenpunkte für die Konturintegration.

### 5.3 FEAST Algorithmus

### 5.4 Tools für MATLAB

## Anhang A

# Notationen

$I_n \mathbb{N}, \mathbb{N}_0 A^H$

# Anhang B

## Sätze und Beweise

**Satz B.1** (Singularwertzerlegung). Ist  $A \in \mathbb{C}^{m,n}$  eine  $r$ -rangige Matrix, so existieren unitäre Matrizen  $U \in \mathbb{C}^{m,m}$ ,  $V \in \mathbb{C}^{n,n}$  und  $\Sigma_+^2 = \text{diag}(\sigma_1^2, \dots, \sigma_r^2) \in \mathbb{C}^{r,r}$  mit der Eigenschaft  $\sigma_1 \geq \dots \geq \sigma_r > 0$ , die eine Faktorisierung der Art

$$A = U \Sigma V^H$$

ermöglichen. Dabei ist

$$\Sigma = \begin{bmatrix} \Sigma_+ & 0_{r,n-r} \\ 0_{m-r,r} & 0_{m-r,n-r} \end{bmatrix}.$$

nochmal bearbeiten mit den quadraten

**Satz B.2** (Spektralsatz für hermitesche Matrizen). bla

**Satz B.3** (Existenz der Cholesky-Zerlegung). bla

**Satz B.4.** Es sei  $A \in \mathbb{C}^{n,n}$  eine hermitesche Matrix mit Eigenpaaren  $(\lambda_i, x_i)_{i=1:n} \in \mathbb{R} \times \mathbb{C}^n \setminus \{0\}$  und  $n \geq 2$ . Es gelte weiter  $|\lambda_1| > |\lambda_j|$  für alle  $j > 1$ . Ist  $y_0 \in \mathbb{C}^n \setminus \{0\}$ , so konvergiert die Folge  $(y_k)_{k \in \mathbb{N}}$  mit

$$y_{(k+1)} = \frac{1}{\|A^{k+1}y_{(0)}\|} A^{k+1}y_{(0)}$$

gegen einen normierten Eigenvektor zum Eigenwert  $\lambda_1$

*Beweis.*

□

# Literatur

- [Bar15] Marc van Barel. “Designing rational filter functions for solving eigenvalue problems by contour integration.” In: (2015).
- [Ber16] Mario Berljafa. “Rational Krylov Toolbox for MATLAB.” In: (2016).
- [IG96] P. Lancaster und L. Rodman I. Gohberg. *Invariant subspaces of matrices with applications*. 1996.
- [Jäge] Klaus Jänich. *Funktionentheorie. Eine Einführung*. Irgendein Jahr.
- [JK15] E. Polizzi und P. Tang J. Kestyn. “FEAST eigensolver for non-hermitian problems”. In: (2015).
- [Lie13] Jörg Liesen. *Krylov subspace methods. Principles and analysis*. 2013.
- [Pag+99] Lawrence Page u. a. *The PageRank Citation Ranking: Bringing Order to the Web*. Technical Report 1999-66. Previous number = SIDL-WP-1999-0120. Stanford InfoLab, 1999. URL: <http://ilpubs.stanford.edu:8090/422/>.
- [Par98] Beresford N. Parlett. *The Symmetric Eigenvalue Problem*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1998. ISBN: 0-89871-402-8.
- [Pol09] Eric Polizzi. “Density-matrix-based algorithm for solving eigenvalue problems.” In: (2009).
- [Saa11] Y. Saad. *Numerical methods for large eigenvalue problems*. 2011?
- [Ste13] Gilbert W. Stewart. *Matrix algorithms*. Bd. 2: Eigensystems. 2013.
- [TP16] Ping Tak Peter Tang und Eric Polizzi. “Feast as a subspace iteration eigensolver accelerated by approximate spectral projection.” In: (2016).