
„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

1	Präludium	1
2	Mathematische Grundlagen	4
3	Filtertechniken	8
3.1	Rayleigh-Ritz Verfahren	9
3.2	Konturintegration	14
3.3	Illustration	19
4	Rationales Filtern	21
4.1	Rayleigh-Ritz Iteration	21
4.2	Gauß'sche Quadratur	24
4.3	Approximation des Spektralprojektors	26
4.4	FEAST-Algorithmus	27
5	Numerische Experimente	30
5.1	Rayleigh-Ritz	30
5.2	Konturintegration	31
5.3	FEAST Algorithmus	33
5.4	Tools für MATLAB	33
A	Notationen	34
B	Sätze und Beweise	35
C	Nebenrechnungen	36

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 (λ, 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 Fußballs¹ ebenso wie beim Untersuchen des PageRanks² einer Website von Bedeutung. Entsprechend strotzt der Kanon von angebotenen numerischen Lösungsmethoden von Vielfalt und Virtuosität.³

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.

¹Hier 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.

² Siehe Abschnitt zwei in [Pag+99].

³Dies 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 Singulärwertzerlegung⁴ 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).

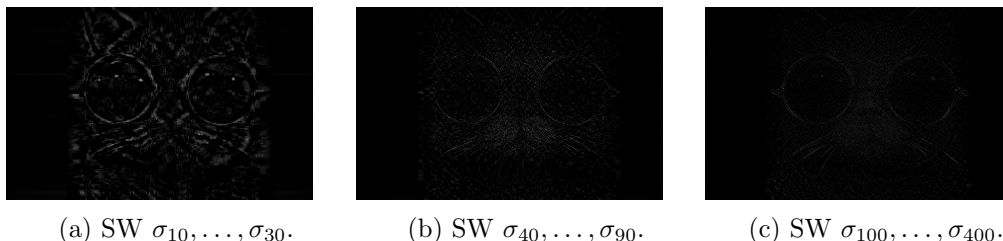


Abbildung 1.2: Rekonstruktionsversuch von Abbildung 1.1.

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.

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

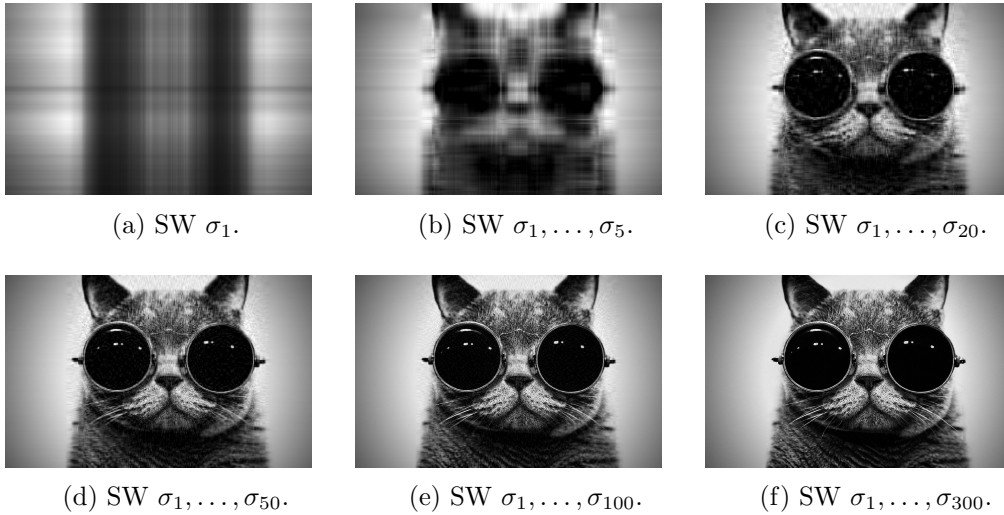


Abbildung 1.3: Approximation von Abbildung 1.1 mittels Singulärwertzerlegung.

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.¹ 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 außerdem auf die Angabe der Dimension verzichten. Im Spezialfall $B = I_n$ heißt (A, B) *gewöhnliches Eigenwertproblem der Dimension n* .

Ist ein passendes Paar (λ, 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 λ *Eigenwert von (A, B)* und x *Eigenvektor zum Eigenwert λ 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 *hermitescher 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.

¹Wir werden Matrizen in ihrer komponentenweisen Darstellung konsequent durch eckige Klammern „[“ und „]“ bezeichnen.

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 (λ, 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 Matrizen² und der Satz von der Existenz der Cholesky-Zerlegung³ 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

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

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. □

²Eine Formulierung mit Beweis ist im Anhang zu finden. Siehe Satz B.2.

³Siehe 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 (ONB) 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 $S_my = \tilde{x}$. Die Forderung (3.1) ist dann zu der Gleichung

$$S_m^H(AS_my - \tilde{\lambda}S_my) = 0$$

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 $(\tilde{\lambda}, S_my)$ 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, Algorithmus 4.5, 98])

- 1: Berechne eine ONB $\{s_i\}_{i=1:m}$ von \mathcal{S}_m und setze $S_m \leftarrow [s_i]_{i=1:m}$.
 - 2: Setze $\tilde{A} \leftarrow S_m^H A S_m$.
 - 3: Finde Eigenpaare von \tilde{A} und wähle die gewünschten k Paare $(\tilde{\lambda}_i, y_i)_{i=1:k}$, $k \leq m$ aus.
 - 4: Gib Ritz-Paare $(\tilde{\lambda}_i, S_m y_i)_{i=1:k}$ aus.
-

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, wie wir im vierten Kapitel sehen werden.

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 . Auch das wird im vierten Kapitel anhand ausgewählter Beispiele vorgeführt.

Nun, da wir uns mit der Idee der Rayleigh-Ritz Methode vertraut gemacht haben, erweitern wir die obige Theorie auf Probleme der Art (A, B) für beliebiges B und betrachten entsprechend die Eigenwertgleichung

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

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

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

genügen. Durch die Wahl eines geeigneten Vektors $y \in \mathbb{C}^m$ kann die Näherungslösung \tilde{x} wie zuvor durch das Produkt $S_m y$ ersetzt werden, wobei die Spalten von S_m erneut eine ONB des Suchraums \mathcal{S}_m bilden.

Damit lässt sich die zu (3.4) äquivalente Forderung

$$S_m^H(AS_my - \tilde{\lambda}BS_my) = 0.$$

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

$$(S_m^H AS_m)y = \tilde{\lambda}(S_m^H BS_m)y. \quad (3.5)$$

mit $(\tilde{\lambda}, S_my)$ ein Ritz-Paar für das Eigenwertproblem (A, B) gewinnen.

Da mit der Matrix B ein weiterer Darsteller berücksichtigt werden muss, zieht dies als Konsequenz eine Anpassung des Algorithmus' 1 nach sich. Wir können diesen in der folgenden Manier abwandeln.

Algorithmus 2 Berechnung von Ritz-Paaren

- 1: Berechne ONB $\{s_i\}_{i=1:m}$ von \mathcal{S}_m und setze $S_m \leftarrow [s_i]_{i=1:m}$.
 - 2: Setze $\tilde{A} \leftarrow S_m^H AS_m$, $\tilde{B} \leftarrow S_m^H BS_m$.
 - 3: Finde Eigenpaare von (\tilde{A}, \tilde{B}) und wähle die gewünschten k Paare $(\tilde{\lambda}_i, y_i)_{i=1:k}$ für $k \leq m$ aus.
 - 4: Gib Ritz-Paare $(\tilde{\lambda}_i, S_my_i)_{i=1:k}$ aus.
-

Bei den beiden vorgestellten Algorithmen ist anzumerken, dass zunächst nicht von einer Eigenpaarfilterung im Sinne der Definition 3.1 gesprochen werden kann. Schließlich erhalten wir unter Umständen kein exaktes Eigenpaar sondern eben nur einen Ersatz. Um eine zufriedenstellende Qualität dieses Ersatzes gewährleisten zu können, 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. Damit lassen sich zu berücksichtigende Toleranzen, beispielsweise eine maximale euklidische Abweichung der Ritz-Werte zu den Eigenwerten oder Vorgaben bezüglich des Winkels zwischen Eigenraum und dem Unterraum der Ritz-Vektoren, einhalten. 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{S}_m unter A dazu führt, dass jedes Ritz-Paar bezüglich \mathcal{S}_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 wir Hermitizität und positive Definitheit von B fordern – sei ein m -dimensionaler Unterraum $\mathcal{S}_m \subseteq \mathbb{C}^n$ gegeben. Ist dieser invariant unter $B^{-1}A$, so ist jedes Ritzpaar von $B^{-1}A$ bezüglich \mathcal{S}_m auch ein Eigenpaar von (A, B) .

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

$$B^{-1}AS_m = S_mV_m \quad (3.6)$$

erfüllt. Insbesondere folgt unter Ausnutzung der Orthogonalität der Spalten von S_m die Identität $S_m^H B^{-1}AS_m = V_m$. Sind nun $\lambda \in \mathbb{C}$ und $y \in \mathbb{C}^m \setminus \{0\}$ so gewählt, dass (λ, S_my) ein Ritz-Paar von $B^{-1}A$ ist, so folgt aus

$$S_m^H B^{-1}AS_my = V_my$$

mit (3.6) die Gleichung

$$B^{-1}AS_my = S_mV_my = \lambda S_my$$

und schließlich auch die Behauptung durch Umstellen. \square

Aus dem eben bewiesenen Resultat können wir nun für das gewöhnliche Eigenwertproblem unmittelbar das folgende Korollar ableiten.

Korollar 3.3. Ist $A \in \mathbb{C}^{n,n}$ und $\mathcal{S}_m \subseteq \mathbb{C}^n$ 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 Eigenwertproblem (A, B) mit der zugehörigen Eigenwertgleichung

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

Die Aussage folgt dann aus dem vorigen Satz. \square

Wenden wir uns einer weiteren Kuriosität zu. Zwischen dem Rayleigh-Ritz 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 $\mathcal{S}_m \subseteq \mathbb{C}^n$ ein m -dimensionaler Unterraum. Sei weiter $\{s_i\}_{i=1:m} \subseteq \mathcal{S}_m$ eine Basis B -orthonormaler Vektoren, das heißt, die Matrix $S_m := [s_i]_{i=1:m} \in \mathbb{C}^{n,m}$ erfülle die Gleichung $S_m^H B S_m = I_m$. Dann ist die von der Matrix $P := S_m S_m^H B \in \mathbb{C}^{n,n}$ induzierte lineare Abbildung

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

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

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

Beweis. Nach Konstruktion gilt $p(\mathbb{C}^n) \subset \mathcal{S}_m$. Aus der B -Orthogonalität von \mathcal{S}_m folgt für alle $x \in \mathbb{C}^n$

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

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

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

Es gilt demnach $x - p(x) \perp_B \mathcal{U}_m$. Damit ist gezeigt, dass p eine B -orthogonale Projektion auf \mathcal{S}_m ist. 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.¹ \square

Mit Hilfe der im Satz eingeführten Projektionsmatrix P , lässt sich die Gleichung (3.5) 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

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

Die Behauptung, es würde ein Zusammenhang zu Projektionen bestehen, wurde also in der Tat zu Recht aufgestellt.

Nach diesen theoretischen Ausführungen fragt sich der anwendungsorientierte Leser zu Recht, welcher Nutzen aus den obigen Einsichten gezogen werden kann. Dafür wenden wir uns erneut dem Rayleigh-Ritz Verfahren zu und betrachten nun den Spezialfall eines hermitesch positiv definiten Eigenwertproblems (A, B) .

¹Für eine genauere Ausführung sei der Leser auf Rechnung C.1 im Anhang verwiesen.

Angenommen, uns stünde mit dem Suchraum \mathcal{S}_m bereits ein von Eigenvektoren aufgespannter Unterraum zur Verfügung. Dann sichert uns Satz 2.1 die Existenz einer Matrix $S_m := [s_i]_{i=1:m} \in \mathbb{C}^{n,m}$ mit B -orthonormalen Spalten aus Eigenvektoren von (A, B) und $\text{Bild}(S_m) = \mathcal{S}_m$, sowie die Existenz einer reellen Diagonalmatrix $\Lambda_m := \text{diag}(\lambda_i)_{i=1:m} \in \mathbb{R}^{m,m}$ mit

$$AS_m = BS_m\Lambda_m.$$

Dabei entsprechen die Diagonaleinträge von Λ_m gerade den zu den Spalten von S_m korrespondierenden Eigenwerten von (A, B) . Folglich wäre das Bild des Spektralprojektors $P = S_m S_m^H B$ nach Konstruktion wegen

$$B^{-1}AP = B^{-1}(AS_m)S_m^H B = B^{-1}(BS_m\Lambda)S_m^H B = S_m\Lambda S_m^H B = \left(\sum_{i=1}^m \lambda_i s_i s_i^H \right) \cdot B$$

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{S}_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 $S_m S_m^H$ mit Methoden der Funktionentheorie berechnbar ist.

3.2 Konturintegration

Betrachten wir das n -dimensionale Eigenwertproblem (A, B) . Zu einer gegebenen Teilmenge Λ_1 vom Spektrum des besagten Problems wollen wir den Spektralprojektor konstruieren, welcher Vektoren auf den von den zu Λ_1 korrespondierenden Eigenvektoren aufgespannten Unterraum \mathcal{S}_1 projiziert. Dazu schränken wir uns vorerst auf das gewöhnliche Eigenwertproblem (A, I_n) ein und orientieren uns in den folgenden Ausführungen an der Dramaturgie von [Lie13, Abschnitt 4.9]. Dabei nehmen wir uns die Freiheit, die Notation an die Bedürfnisse dieser Arbeit anzupassen und gelegentlich von den vorgeschlagenen Schreibweisen abzuweichen.

Wir wollen annehmen, dass die Zahl m der Dimension von \mathcal{S}_1 entspricht. Seien Matrizen $A_1 \in \mathbb{C}^{m,m}$, $A_2 \in \mathbb{C}^{n-m, n-m}$ und $S \in \mathbb{C}^{n,n}$ derart vorgegeben, dass $\Lambda(A_1) = \Lambda_1$ einerseits und $\Lambda_2 := \Lambda(A_2) \cap \Lambda_1 = \emptyset$ andererseits gelten und eine Faktorisierung der Art

$$A = S \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix} S^{-1} \quad (3.7)$$

existiert.² Dann können wir zwei Matrizen $S_1 \in \mathbb{C}^{n,m}$ und $S_2 \in \mathbb{C}^{n, n-m}$ finden mit $\text{Bild}(S_1) = \mathcal{S}_1$, welche die Partitionierung $S = [S_1, S_2]$ erlaubt.

²Diese Zerlegung kann beispielsweise mit Hilfe der Jordan'schen Normalform erreicht werden. Siehe Satz B.4 im Anhang.

Durch Definition der Matrix

$$P_1 := S \begin{bmatrix} I_m & 0 \\ 0 & 0 \end{bmatrix} S^{-1} \quad (3.8)$$

erhalten wir dann eine Projektion auf \mathcal{S}_1 : Nach Konstruktion gilt $P_1 x \in \mathcal{S}_1$ für alle $x \in \mathbb{C}^n$ und $P_1^2 = P_1$. Andererseits liefert die Matrix

$$P_2 := I_n - P_1$$

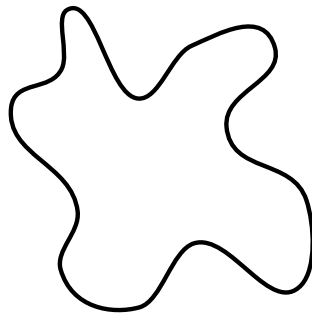
eine Projektion auf $\mathcal{S}_2 := \mathbb{C}^n \setminus \mathcal{S}_1$. Da sich jeder Vektor $x \in \mathbb{C}^n$ in

$$x = P_1 x + (I_n - P_1)x$$

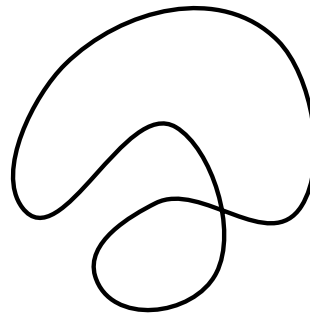
dekomponieren lässt und trivialerweise $P_2 x \perp \mathcal{S}_2^\perp$ gilt, können wir schließen, dass P_1 orthogonal bezüglich \mathcal{S}_2^\perp auf \mathcal{S}_1 abbildet.

Der Projektor P_1 aus (3.8) lässt sich mit Hilfe eines Integrals darstellen. Bevor wir dazu kommen, benötigen wir zusätzliches Vokabular.

Definition 3.5 (Einfache geschlossene Kurve). *Es sei $S := \{x \in \mathbb{R}^2 : \|x\|_2 = 1\}$ und $\gamma: S \rightarrow \mathbb{R}^2$ ein Homöomorphismus, das heißt, eine stetige Bijektion mit stetiger Umkehrabbildung. Dann heißt $\Gamma := \text{Bild}(\gamma)$ einfache geschlossene Kurve oder auch Jordan-Kurve.*



(a) Einfache geschl. Kurve.



(b) Keine Jordan-Kurve.

Abbildung 3.1: Skizze verschiedener Kurven.

Insbesondere sind Jordan-Kurven also geschlossene und überschneidungsfreie Kurven.

Da wir über solche Jordan-Kurven integrieren wollen, führen wir noch den Begriff der *orientierten Kurve* ein. Dabei wollen wir es jedoch der Einfachheit halber bei einer informellen Beschreibung belassen. Eine Jordan-Kurve heißt *positiv orientiert*, falls sie in ihrer Parametrisierung im mathematisch positiven Sinne „durchlaufen“ wird (Vgl. Abb. 3.2).

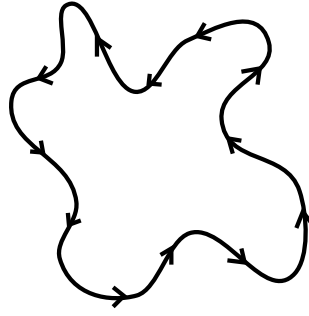


Abbildung 3.2: Skizze einer positiv orientierten Jordan-Kurve.

Wir können nun den Gedanken aus [Lie13, Abschnitt 4.9] weiter folgen und betrachten für $t \subseteq \mathbb{C}^n$ die durch

$$G: \Omega \setminus \Lambda(A) \rightarrow \mathbb{C}^{n,n}, t \mapsto (tI_n - A)^{-1}$$

definierte *Resolventenfunktion* – auch *Green-Funktion* genannt (Vgl. [Pol09]). Angenommen, die Menge Λ_1 liegt im Inneren eines von einer positiv orientierten Jordan-Kurve Γ umschlossenen Gebiets und Λ_2 im Inneren des Komplementes dieses Gebiets.

Dann gilt die Identität

$$P_1 = \frac{1}{2\pi\iota} \int_{\Gamma} G(t) dt \quad (3.9)$$

beziehungsweise unter Ausnutzung der Faktorisierung (3.7)

$$P_1 = \frac{1}{2\pi\iota} S \begin{bmatrix} \int_{\Gamma} (tI_m - A_1)^{-1} dt & 0 \\ 0 & \int_{\Gamma} (tI_{n-m} - A_2)^{-1} dt \end{bmatrix} S^{-1}. \quad (3.10)$$

Hierbei ist das Integral eintragsweise zu verstehen und ι als imaginäre Einheit zu lesen. Da die Menge Λ_2 nach Vorgabe außerhalb des umrandeten Gebietes liegt, folgt aus dem Cauchy'schen Integralsatz³

$$\int_{\Gamma} (tI_{n-m} - A_2)^{-1} dt = 0_{n-m}.$$

Zur weiteren Argumentation benötigen wir den folgenden Satz.

Satz 3.6. Ist $M \in \mathbb{C}^{n,n}$ mit $\|M\|_2 < 1$, so ist $I_n - M$ invertierbar und es gilt

$$(I_n - M)^{-1} = \sum_{j=0}^{\infty} M^j$$

Beweis. (Vgl. [loa00, Abschnitt 2.3.4]) Zunächst zur Regularität: Angenommen, die Matrix $I_n - M$ wäre nicht invertierbar. Dann muss ein von Null verschiedener Vektor $x \in \mathbb{C}^n$ existieren, mit $(I_n - M)x = 0$. Wir erhalten insbesondere $x = Mx$. Die induzierte 2-Norm für Matrizen ist bekanntermaßen submultiplikativ. Es gilt daher $\|x\|_2 \leq \|M\|_2 \|x\|_2$. Dann führt aber die Konsequenz $\|M\|_2 \geq 1$ zum Widerspruch.

Nun zur Potenzreihenentwicklung: Nutzen wir erneut die Submultiplikativität der 2-Norm,

³Siehe Satz B.5 im Anhang.

so erhalten wir für alle $k \in \mathbb{N}$ die Abschätzung $\|M^k\|_2 \leq \|M\|_2^k$ und damit $\lim_{k \rightarrow \infty} M^k = 0_n$ wegen $\|M\|_2 < 1$. Aus

$$\sum_{j=0}^{\infty} M^j (I_n - M) = \lim_{k \rightarrow \infty} \sum_{j=0}^k M^j (I_n - M) = \lim_{k \rightarrow \infty} I_n - M^{k+1} = I_n$$

folgt dann durch Umstellen die Behauptung. \square

Korollar 3.7. Mit den obigen Notationen folgt für alle $t \in \mathbb{C}$ mit $|t| > \|A_1\|_2$ die Identität

$$(tI_m - A_1)^{-1} = \frac{1}{t} \sum_{j=0}^{\infty} (t^{-1} A_1)^j.$$

Beweis. Man mache sich die Identität

$$(tI_m - A_1)^{-1} = \frac{1}{t} (I_m - t^{-1} A_1)^{-1}$$

klar und benutze obigen Satz. \square

Wenden wir uns nun wieder der Gleichung (3.10) zu. Es bleibt zu zeigen, dass

$$\frac{1}{2\pi i} \int_{\Gamma} (tI_m - A_1)^{-1} dt = I_m \quad (3.11)$$

gilt. Dazu halten wir uns nachwievor an die Ausführungen von [Lie13, Abschnitt 4.9]. Wie von den Autoren vorgeschlagen, setzen wir

$$F(A_1, t) := \frac{1}{t} \sum_{j=1}^{\infty} (t^{-1} A_1)^j$$

und erhalten mit Hilfe des Korollars 3.7

$$(tI_m - A_1)^{-1} = \frac{1}{t} I_m + F(A_1, t). \quad (3.12)$$

Nach Konstruktion sind alle Eigenwerte von A_1 beziehungsweise die Singularitäten von $(tI_m - A_1)^{-1}$ im Inneren von Γ . Der Wert des Integrals auf der linken Seite von Gleichung (3.11) verändert sich nicht, wenn wir anstelle von Γ über eine andere Jordan-Kurve, die das von Γ umschlossene Gebiet enthält, integrieren. Daher können wir als Integrationskontur $\hat{\Gamma}$ ebenso gut den Kreis um den Nullpunkt mit hinreichend großem Radius $r > \|A_1\|_2$ wählen.

Für $t \in [0, 2\pi]$ ist die Abbildung $t \mapsto re^{it}$ eine Parametrisierung von $\widehat{\Gamma}$. Diese Kurve ist dann positiv orientiert und wir erhalten wegen

$$\int_{\widehat{\Gamma}} \frac{1}{t} dt = \int_0^{2\pi} \frac{1}{re^{it}} ire^{it} dt = 2\pi i$$

und unter Ausnutzung von Gleichung (3.12)

$$\int_{\Gamma} (tI_m - A_1)^{-1} dt = \int_{\widehat{\Gamma}} \frac{1}{t} I_m dt + \int_{\widehat{\Gamma}} F(A_1, t) dt = 2\pi i \cdot I_m + \int_{\widehat{\Gamma}} F(A_1, t) dt.$$

Betrachten wir nun für $t \in \widehat{\Gamma}$ den Ausdruck $F(A_1, t)$ etwas genauer. Es gilt

$$\|F(A_1, t)\|_2 = \left\| \frac{1}{t} \sum_{j=1}^{\infty} \left(\frac{1}{t} A_1 \right)^j \right\|_2 \leq \frac{1}{|t|} \sum_{j=1}^{\infty} \left(\frac{\|A_1\|_2}{|t|} \right)^j = \frac{\|A_1\|_2}{|t|^2} \sum_{j=0}^{\infty} \left(\frac{\|A_1\|_2}{|t|} \right)^j.$$

Nach Konstruktion gilt $\|A_1\|_2 < r \leq |t|$. Mit der Kenntnis über den Grenzwert der geometrischen Reihe erhalten wir daher die Abschätzung

$$\|F(A_1, t)\|_2 \leq \frac{\|A_1\|_2}{r} \frac{1}{1 - \|A_1\|_2/r}.$$

Da diese Abschätzung für alle $r > \|A_1\|_2$ gilt und die rechte Seite für $r \rightarrow \infty$ den Wert Null annimmt, muss daher

$$\frac{1}{2\pi i} \int_{\Gamma} (tI_m - A_1)^{-1} dt = \frac{1}{2\pi i} \left(2\pi i \cdot I_m + \int_{\widehat{\Gamma}} F(A_1, t) dt \right) = I_m$$

und somit auch Gleichung (3.9) gelten.

Welche Schlussfolgerungen lassen sich hieraus für ein Eigenwertproblem (A, B) ziehen, bei der eine beliebige Matrix B vorgegeben ist? Diese Frage soll zumindest für HPD-Eigenwertprobleme beantwortet werden.

Sei also (A, B) ein HPD-Eigenwertproblem. Dann ist B invertierbar und jede Singularität der Greenfunktion $G_1(t) := (tI_n - B^{-1}A)^{-1}$ ist auch eine Singularität der Funktion $G_2(t) := (tB - A)^{-1}$. Ist wie eben Λ_1 eine Teilmenge des Spektrums, Γ eine umschließende Jordan-Kurve und die Zerlegung

$$B^{-1}A = S \begin{bmatrix} B_1^{-1}A_1 & 0 \\ 0 & B_2^{-1}A_2 \end{bmatrix} S^{-1}$$

analog zu Gleichung 3.7 gegeben, so schließt man auf ähnliche Weise

$$\frac{1}{2\pi i} \int_{\Gamma} (tI_n - B^{-1}A)^{-1} dt = \frac{1}{2\pi i} \int_{\Gamma} (tB - A)^{-1} dt = P_1.$$

3.3 Illustration

Bevor wir uns vom laufenden Kapitel verabschieden, soll die Theorie der vorangegangenen Abschnitte an einem konkreten Beispiel vorgeführt werden. Dabei werden wir mit Hilfe der Konturintegration zunächst den Spektralprojektor berechnen. Ist dies erledigt, reduzieren wir wie im Rayleigh-Ritz Verfahren die Dimension des Eigenwertproblems und konstruieren ausgehend vom transformierten Problem Eigenpaare des ursprünglichen Problems.

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 sich das HPD-Eigenwertproblem (A, B) 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 $\lambda_3 = 4/9$ lösen lässt.⁴

Vergessen wir für den Moment, dass uns die Eigenpaare bekannt sind und versuchen mit den Methoden, welche wir in den vorangegangenen Abschnitten besprochen haben, die Eigenpaare auf dem reellen Intervall $I = [-1, 1]$ zu bestimmen. Gemäß Abschnitt 3.2 wählen wir daher als Integrationskontur Γ das Bild der Funktion

$$\gamma: [0, 2\pi] \rightarrow \mathbb{C}^n, \varphi \mapsto e^{i\varphi} \quad (3.13)$$

und integrieren darüber die Green-Funktion

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

Nach Definition 3.5 ist Γ eine positiv orientierte Jordan-Kurve. Da nach Konstruktion keiner der Eigenwerte im Bild von γ liegt, ist G auf der gesamten Kurve wohldefiniert. Folgen wir also weiter dem Abschnitt 3.2, so erhalten wir

$$\frac{1}{2\pi i} \int_{\Gamma} G \, ds = \frac{1}{2\pi i} \int_0^{2\pi} G(\gamma(t)) \cdot \gamma'(t) \, dt = \text{diag}\left(0, \frac{1}{5}, \frac{1}{9}\right)$$

Diese Matrix lässt sich mit den B -orthonormalen Vektoren $x_1 = 1/\sqrt{6} \cdot e_2$ und $x_2 = 1/3 \cdot e_3$ wie gewünscht in der Art

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1/6 & 0 \\ 0 & 0 & 1/9 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 1/\sqrt{6} & 0 \\ 0 & 1/3 \end{bmatrix} \begin{bmatrix} 0 & 1/\sqrt{6} & 0 \\ 0 & 0 & 1/3 \end{bmatrix} = x_1 x_1^H + x_2 x_2^H =: X X^H$$

darstellen.

⁴Hier bezeichnen $e_1, e_2, e_3 \in \mathbb{C}^3$ die Einheitsvektoren, also $e_1 = [1, 0, 0]$, $e_2 = [0, 1, 0]$ und $e_3 = [0, 0, 1]$.

Nachdem dies geschafft ist, lassen wir uns von der in Abschnitt 3.1 diskutierten Vorgehensweise inspirieren und reduzieren das Problem zunächst mit Hilfe einer vollrangigen Matrix $S \in \mathbb{C}^{3,2}$ auf ein Problem kleinerer Dimension. Zu diesem Zwecke wählen wir

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

und benutzen die Matrix $P := XX^H S$ um das Ausgangsproblem vermöge $\tilde{A} := P^H A P$ und $\tilde{B} := P^H B P$ 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 (2×2) 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

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

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

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

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

Kapitel 4

Rationales Filtern

Nachdem das vorangegangene Kapitel 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 Rayleigh-Ritz Iteration

Bei der Behandlung des Rayleigh-Ritz Verfahrens wurde angedeutet, dass das Einbinden einer geeigneten Iterationsvorschrift dabei helfen kann, die Güte von errechneten Ritz-Paaren zu verbessern. Hierbei ist „Güte“ natürlich in Abhängigkeit vom Kontext zu bewerten. Im Folgenden wollen wir uns genauer mit dieser ominösen Iterationsvorschrift auseinander setzen und beginnen die Herleitung bei einem sehr einfach umsetzbaren Verfahren zur Bestimmung von Eigenpaaren.

Ausgangspunkt für unsere Betrachtungen ist ein gewöhnliches Eigenwertproblem mit einer von Null verschiedenen hermiteschen Matrix $A \in \mathbb{C}^{n,n}$. Bei der sogenannten *Potenzmethode* wird ausgehend von einem Startvektor $y_{(0)} \in \mathbb{C}^n \setminus \{0\}$ 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. Man kann zeigen, dass die Folge der Iterierten gegen einen zum betragsmäßig größten Eigenwert gehörenden Eigenvektor konvergiert.¹

¹Ein Beweis hierzu ist im Anhang zu finden. Siehe Satz B.6.

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. Dazu wählen diesmal eine Startmatrix $Y_0 \in \mathbb{C}^{n,m}$ vollen Ranges und berechnen die k -te Iterierte mit

$$Y_{(k)} = A^k Y_{(0)}.$$

Bei der Normalisierung ist allerdings Vorsicht geboten: In seinen Abhandlungen über Unterraumiterationen merkt Y. Saad in [Saa11, Abschnitt 5.1] an, dass es beim ungeschickten Normalisieren vorkommen kann, dass die Spalten von $Y_{(k)}$ zunehmend ihre lineare Unabhängigkeit verlieren.²

Anstatt jede Spalte von $Y_{(k)}$ separat zu normalisieren, wird eine QR-Zerlegung³ bemüht. Diese führt in der Tat zu einer Normalisierung. Ist nämlich $Y_{(k)} = QR$ mit $Q = [q_i]_{i=1:m}$ die QR-Zerlegung, so gilt $\text{Bild}(Y_{(k)}) = \text{Bild}(Q)$ und $\|q_i\|_2 = 1$ für $i = 1 : m$. Folglich stellt folgender Algorithmus eine Verallgemeinerung der Potenzmethode dar.

Algorithmus 3 Verallgemeinerte Potenzmethode (Vgl. [Saa11, Algorithmus 5.1, 115])

- 1: Wähle linear unabhängige Vektoren $\{y_i\}_{i=1:m}$ und setze $Y_{(0)} \leftarrow [y_i]_{i=1:m}$ und $k \leftarrow 1$
 - 2: **repeat**
 - 3: Setze $Y_{(k)} \leftarrow AY_{(k-1)}$ und berechne QR-Zerlegung $Y_{(k)} = QR$.
 - 4: Setze $Y_{(k)} \leftarrow Q$ und $k \leftarrow k + 1$.
 - 5: **until** Verfahren konvergiert.
-

Saad weist darauf hin, dass die Kosten der Berechnung der QR-Zerlegung sehr hoch werden können. Da der von den Spalten von $Y_{(k)}$ aufgespannte Unterraum gleich dem von den Spalten von $A^k Y_{(0)}$ aufgespannten Unterraum ist, schlägt Saad daher folgende Abwandlung des eben vorgestellten Algorithmus' vor.

Algorithmus 4 Gebrauch variabler Exponenten (Vgl. [Saa11, Algorithmus 5.2, 116])

- 1: Wähle linear unabhängige Vektoren $\{y_i\}_{i=1:m}$, setze $Y \leftarrow [y_i]_{i=1:m}$ und wähle initialen Exponenten k .
 - 2: **repeat**
 - 3: Setze $S \leftarrow A^k Y$ und orthonormalisiere S zu \hat{S} .
 - 4: Setze $Y \leftarrow \hat{S}$.
 - 5: Wähle neuen Exponenten k .
 - 6: **until** Verfahren konvergiert.
-

Auch hier ist zu beachten, dass im Falle der Wahl eines sehr großen Exponenten die Unabhängigkeit der Spalten von S nicht mehr gewährleistet werden kann. Wir wollen an dieser Stelle auf Konvergenz- und Laufzeitanalysen der eben vorgestellten Algorithmen verzichten und kommen schließlich zur vielfach angekündigten Iterationsvorschrift, welche sich aus der Potenzmethode ableitet.

²Dies ist unmittelbar einzusehen, wenn man bedenkt, dass jede Spalte gegen einen dominanten Eigenvektor konvergiert.

³Eine Formulierung des Satzes über die Existenz der QR-Zerlegung ist im Anhang zu finden. Siehe Satz B.7. Für weitere Ausführungen siehe auch [Ste13, 55 ff.].

Algorithmus 5 Iteratives Rayleigh-Ritz Verfahren (Vgl. [Saa11, Algorithmus 5.3, 118])

- 1: Wähle linear unabhängige Vektoren $\{y_i\}_{i=1:m}$, setze $Y \leftarrow [y_i]_{i=1:m}$ und wähle initialen Exponenten k .
 - 2: **repeat**
 - 3: Setze $S \leftarrow A^k Y$.
 - 4: Orthonormalisiere die Spalten von S und setze $\tilde{A} \leftarrow S^H A S$.
 - 5: Berechne Eigenvektoren $\tilde{X} \leftarrow [\tilde{x}_i]_{i=1:m}$ von \tilde{A} .
 - 6: Setze $Y \leftarrow S \tilde{X}$.
 - 7: Wähle neuen Exponenten k .
 - 8: **until** Verfahren konvergiert.
-

Zunächst ein Wort zur fünften Zeile. Hier wurde das Berechnen von Schurvektoren – so wie es in der oben zitierten Quelle vorgeschlagen wird – durch das Berechnen von Eigenvektoren ersetzt. Dies ist möglich, weil A nach Vereinbarung ein hermitesche Matrix ist und somit unitär diagonalisiert werden kann. Es ist daher nicht nötig zwischen Eigenvektoren und Schurvektoren zu unterscheiden. Wie genau diese Eigenvektoren berechnet werden, wollen wir im Rahmen dieser Arbeit nicht genauer erläutern.

Die Wurzeln des eben erarbeiteten Algorithmus' sind deutlich zu erkennen. In den Zeilen vier bis sechs wird das Rayleigh-Ritz Verfahren benutzt. Anstelle von Ritz-Paaren werden allerdings lediglich Ritz-Vektoren berechnet. In jeder Iteration wird wie beim Potenzverfahren ein neuer Exponent festgelegt und somit ein neuer Suchraum $\mathcal{S} = \text{Bild}(A^k Y)$ vorgegeben. Erläuterungen zum Konvergenzverhalten sind in [Saa11, Abschnitt 5] zu finden. Zur Berechnung von Eigenpaaren eines HPD-Eigenwertproblems (A, B) bei dem B nicht mehr der Identität entspricht, müssen die Zeilen vier bis sechs gemäß Algorithmus 2 angepasst werden.

Es ist möglich, den zuletzt eingeführten Algorithmus weiter abzuwandeln. Dazu betrachten wir wieder ein HPD-Eigenwertproblem (A, B) . Zu diesem Duo gesellt sich nun mit $\rho(B^{-1}A)$ ein Polynom in $B^{-1}A$, welches wir benutzen, um das iterative Rayleigh-Ritz Verfahren wie folgt zu ändern.

Algorithmus 6 Beschleunigte Rayleigh-Ritz Iteration (Vgl. [TP16, Algorithmus A])

- 1: Wähle m linear unabhängige Vektoren $Y_{(0)} \leftarrow [y_i]_{i=1:m} \in \mathbb{C}^{n,m}$. Setze $k \leftarrow 1$.
 - 2: **repeat**
 - 3: Approximiere den Unterraumprojektor: $P_{(k)} \leftarrow \rho(B^{-1}A)Y_{(k-1)}$
 - 4: Reduziere die Dimension: $\tilde{A}_{(k)} \leftarrow P_{(k)}^H A P_{(k)}$, $\tilde{B}_{(k)} \leftarrow P_{(k)}^H B P_{(k)}$.
 - 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 $Y_{(k)} \leftarrow P_{(k)} \tilde{X}_{(k)}$.
 - 7: $k \leftarrow k + 1$.
 - 8: **until** Abbruchkriterium ist erfüllt.
-

Diese Methode ähnelt stark dem Algorithmus 5. Die Zeilen vier bis sechs entsprechen erneut dem Rayleigh-Ritz Verfahren, aber die Berechnung des Suchraums $\mathcal{S} := \text{Bild}(P_{(k)})$ geht anders vonstatten.

Im Kontext dieses Algorithmus' wird ρ 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 $Y_{(1)} = \rho(B^{-1}A)Y_{(0)}$ vollen Rang, so konvergiert der Algorithmus 6 in einer Iteration (Vgl. [TP16, S. 356]). Dies folgt unter Ausnutzung der Invarianz des Bildes von $P_{(1)}$ unter $B^{-1}A$ aus dem Satz 3.2.

Da der Spektralprojektor in den meisten Fällen unbekannt sein dürfte, liegt die Idee nahe, ihn zu approximieren. Tang und Polizzi [TP16, S. 356] merken an, dass dies gut funktioniert, falls ρ 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. Dabei sehen wir von Beweisen und ausufernden Erläuterungen ab, da diese Thematik in den meisten Einführungsbüchern zur numerischen Mathematik ausführlich besprochen wird.⁴ Im Anschluss fahren wir mit der Konstruktion des Projektors fort.

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.⁵ 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

$$Q_n(f) := \int_a^b p_n(x) \, dx = (b-a) \sum_{i=0}^n \omega_i f(x_i) \quad (4.1)$$

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.

⁴Siehe zum Beispiel [Pla58, Abschnitt 6]

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

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 in Kürze 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 ω der Art

$$f = \omega \cdot g$$

besitzt, wobei ω auf dem Integrationsintervall $[a, b]$ positiv sein soll. Die Funktion ω wird auch als *Gewichtsfunktion* bezeichnet. Ziel ist also nun die Berechnung von

$$\int_a^b \omega(x)g(x) \, dx, \quad (4.2)$$

wobei wir zusätzlich fordern, dass (4.1) und (4.2) für alle Polynome bis zum Grad $(2n-1)$ übereinstimmen.

Dazu betrachten wir die Standardbasis $\{x^k\}_{k=0:(2n-1)}$ auf dem Raum der Polynome vom Grad $(2n-1)$. Dann landen wir unweigerlich bei dem Gleichungssystem

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

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

$$\langle p, q \rangle_\omega := \int_a^b p(x)q(x)\omega(x) \, dx$$

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 $\{x_j\}_{j=1:n}$ 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.

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

4.3 Approximation des Spektralprojektors

Sei (A, B) ein HPD-Eigenwertproblem und $I := [\lambda_1, \lambda_2]$ ein reelles Intervall. Sei weiter $\mathcal{X} \subseteq \mathbb{C}^n$ derjenige Unterraum, welcher von den Eigenvektoren aufgespannt wird, die zu den im Inneren von I befindlichen Eigenwerten korrespondieren. Wir wollen in diesem Abschnitt den Spektralprojektor approximieren, welcher den Vektorraum \mathbb{C}^n auf \mathcal{X} projiziert.

Die erste Ingredienz, die zum Gelingen dieser Approximation beiträgt, ist eine rationale Funktion $\varrho: \mathbb{C} \rightarrow \mathbb{C}$ mit $\varrho(\mathbb{R}) \subseteq \mathbb{R}$, welche auf I näherungsweise der Indikatorfunktion von I entspricht. Dafür bemühen wir die Cauchy'sche Integraldarstellung von der Indikatorfunktion und wandeln diese mit Hilfe numerischer Quadraturformeln in die gewünschte rationale Funktion ϱ um.

Zuächst zur Indikatorfunktion: Ist c der Mittelpunkt des Intervalls I und r der Abstand des Mittelpunktes zum Rand des Intervalls, dann entspricht die Menge

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

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

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

gilt. Um dieses Integral mit einer Gauß'schen Quadraturformel annähern zu können, benötigen wir eine Parametrisierung von \mathcal{C} . Hierfür ziehen wir [TP16] zurate und wählen demzufolge die Funktion

$$\gamma: [-1, 3] \rightarrow \mathbb{C}, x \mapsto c + re^{i\frac{\pi}{2}(1+x)}$$

als Parametrisierung der Sphäre \mathcal{C} . Wir erhalten dann für alle $z \notin \mathcal{C}$ die Gleichung

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

wobei $\overline{\gamma(x)}$ und $\overline{\gamma'(x)}$ die komplexen Konjugationen von $\gamma(x)$ beziehungsweise $\gamma'(x)$ bezeichnen.

Es seien nun $(w_j, x_j)_{j=1:m}$ die für die Gauß'sche Quadraturformel benötigten Gewichte und Diskretisierungspunkte. Dann setzen wir

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

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

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

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

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

und folglich

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

für eine Matrix $Y \in \mathbb{C}^{n,m}$. Sind A, B und Y reellwertig, so lässt sich dies zu

$$\varrho(B^{-1}A)Y = 2 \sum_{k=1}^q \Re (\sigma_k (\gamma_k B - A)^{-1} B Y) \quad (4.3)$$

vereinfachen. Dabei gebe die Funktion \Re den Realteil ihres Inputs an.

4.4 FEAST-Algorithmus

Im Jahr 2009 stellte Eric Polizzi in [Pol09] ein Verfahren vor, welches sich die Theorien der vorigen Abschnitte zu Nutze macht. Dieser „[...] *fast and stable algorithm for solving the symmetric eigenvalue problem* [...]“ [Pol09, Abstract] wurde seither weiterentwickelt und zahlreichen Analysen unterzogen.⁶ Wir wollen in diesem Abschnitt die wichtigen Punkte des oben zitierten Papers skizzieren und damit das vierte Kapitel abschließen.

⁶Siehe etwa [TP16],[JK15] und [ARLP12].

Polizzi beginnt seine Ausführungen mit einem Blick in die Physik und motiviert das Lösen von Eigenwertproblemen anhand der Schrödinger Eigenwertgleichung

$$\mathbf{H}\Psi = E\mathbf{S}\Psi$$

welche die Fragestellung modelliert, ob gewisse Quantenobjekte, die von einer Wellenfunktion Ψ beschrieben werden, kinetische Energie besitzen oder nicht. Hierbei ist \mathbf{H} eine hermitesche Matrix und \mathbf{S} eine symmetrisch positiv definite Matrix.

Der Autor merkt an, dass das Lösen solcher Systeme, insbesondere dann, wenn sie sehr groß sind, enorme Anforderungen an die Numeriker stellt. Die Frage ist stets, wie Eigenpaare effizient berechnet werden können und welche Genauigkeit man erwarten darf. Polizzi bewirbt seinen Algorithmus mit hoher Geschwindigkeit, Robustheit und guter Skalierbarkeit. Seinem Urteil nach, werden Methoden, wie „*Rayleigh-quotient multigrid*“ oder „*parallel Chebyshev subspace iteration*“ als weniger effizient eingeschätzt.

Ausgehend von einer Ausführung über das Konzept der Konturintegration zur Bestimmung von Eigenpaaren, baut Polizzi seinen Algorithmus auf, der auf ein N -dimensionales Eigenwertproblem (A, B) mit hermiteschem oder reell symmetrischem A und symmetrisch positiv definitem B ausgelegt ist und auf einem reellen Intervall $[\lambda_1, \lambda_2]$ die $m \in \mathbb{N}$ Eigenpaare finden soll.

Sind $(\omega_j, s_j)_{j=1:k}$ die für die Gauß'sche Quadraturformel benötigten Gewichte und Stützstellen, so lässt sich der vorgestellte FEAST-Algorithmus wie folgt notieren.

Algorithmus 7 FEAST-Algorithmus (Vgl. [Pol09, Abschnitt III])

- 1: Wähle $M > m$ linear unabhängige Vektoren $Y \leftarrow [y_i]_{i=1:M} \in \mathbb{C}^{n,M}$, setze $Q \leftarrow 0_{n,M}$ und $r \leftarrow (\lambda_1 - \lambda_2)/2$.
 - 2: **repeat**
 - 3: **for** $j=1:k$
 - 4: Berechne $\theta_{(j)} \leftarrow -(\pi/2)(s_j - 1)$.
 - 5: Berechne $t_{(j)} \leftarrow (\lambda_2 + \lambda_1)/2 + re^{i\theta_{(j)}}$.
 - 6: Löse $(t_{(j)}B - A)Q_{(j)} = Y$ in $Q_{(j)}$.
 - 7: Setze $Q \leftarrow Q - (\omega_j/2) \cdot \Re(re^{i\theta_{(j)}}Q_{(j)})$.
 - 8: **endfor**
 - 9: Reduziere die Dimension: $\tilde{A} \leftarrow Q^H A Q$, $\tilde{B} \leftarrow Q^H B Q$.
 - 10: Löse das transformierte Problem $\tilde{A}\tilde{X} = \tilde{B}\tilde{X}\tilde{\Lambda}$ in \tilde{X} und $\tilde{\Lambda} = \text{diag}(\tilde{\lambda}_1, \dots, \tilde{\lambda}_M)$.
 - 11: Setze $X \leftarrow [x_i]_{i=1:M} = Q\tilde{X}$.
 - 12: Gilt $\tilde{\lambda}_i \in [\lambda_1, \lambda_2]$, so gibt Eigenpaar $(\tilde{\lambda}_i, x_i)$ aus.
 - 13: **until** Abbruchkriterium ist erfüllt.
-

Dieser Algorithmus stellt eine Umsetzung der beschleunigten Rayleigh-Ritz Iteration dar. Die Zeilen drei bis acht erfüllen den Zweck des Berechnens von $\rho(B^{-1}A)Y$, wobei man sich hier einer ähnlichen Umformulierung bedient, wie in Gleichung (4.3).⁷

⁷Die entsprechende Anpassung von Algorithmus 6 ist in [TP16, S. 365] zu finden.

Es ist zu beachten, dass sich hier durch die Verwendung einer negativ orientierten Kurve – zu erkennen in der vierten und fünften Zeile – Vorzeichenwechsel ergeben. Da bei einer entsprechenden Parametrisierung auf der rechten Seite der Identität (3.9) der zusätzliche Faktor (-1) benötigt wird, muss dies entsprechend bei der Konstruktion der rationalen Funktion ρ berücksichtigt werden.

Die weiteren Abschnitte des Papers beschäftigen sich vornehmlich mit numerischen Experimenten. Polizzi demonstriert anhand elektronenstrukturell bezogenen Berechnungen auf Kohlenstoffnanoröhren die numerische Stabilität, Robustheit und Skalierbarkeit des FEAST-Algorithmus. Dabei vergleicht er FEAST mit ARPACK – einer Bibliothek auf FORTAN77 basierenden Methoden zum Lösen hochdimensionaler Eigenwertprobleme.

Abschließend hebt der Autor einige Merkmale seines Algorithmus’ heraus, von denen an dieser Stelle eine bescheidene Auswahl vorgestellt wird. Zunächst einmal kommt FEAST gänzlich ohne Orthogonalisierungsmethoden aus, wie es etwa bei der unbeschleunigten Rayleigh-Ritz Iteration nötig ist. Des Weiteren weist Polizzi auf das Potential der parallelen Implementierung hin.

In Papern aus den vergangenen zwei Jahren, wie etwa in [JK15, Abschnitt 4.1], wird auf dieses Potential genauer eingegangen. Inzwischen wurden auch Versuche unternommen, FEAST auf nicht hermitesche Eigenwertprobleme auszuweiten.⁸

⁸Siehe [JK15].

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

`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 λ_{\min} und λ_{\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 Φ_k wird im Quelltext mit vq bezeichnet.

Wir wollen annehmen, dass der Input des Users die Voraussetzungen des Problems (??) 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

$$I_n \mathbb{N}, \mathbb{N}_0 \quad A^H \gg \perp \quad \| \cdot \|_B$$

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 (Jordan'sche Normalform). blubb

Satz B.5 (Cauchy'scher Integralsatz). bla

Satz B.6 (Potenzmethode). 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 λ_1

Beweis.

□

Satz B.7 (QR-Zerlegung).

Anhang C

Nebenrechnungen

Rechnung C.1 (Zu Satz ??).

Literatur

- [ARLP12] D. Zhang A. R. Levin und E. Polizzi. “FEAST fundamental framework for electronic structure calculations: reformulation and solution of the muffin-tin problem”. In: (2012).
- [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.
- [loa00] golub loan. *Matrix Computations*. 1500.
- [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.
- [Pla58] Robert Plato. *Numerische Mathematik kompakt*. 1758.
- [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).