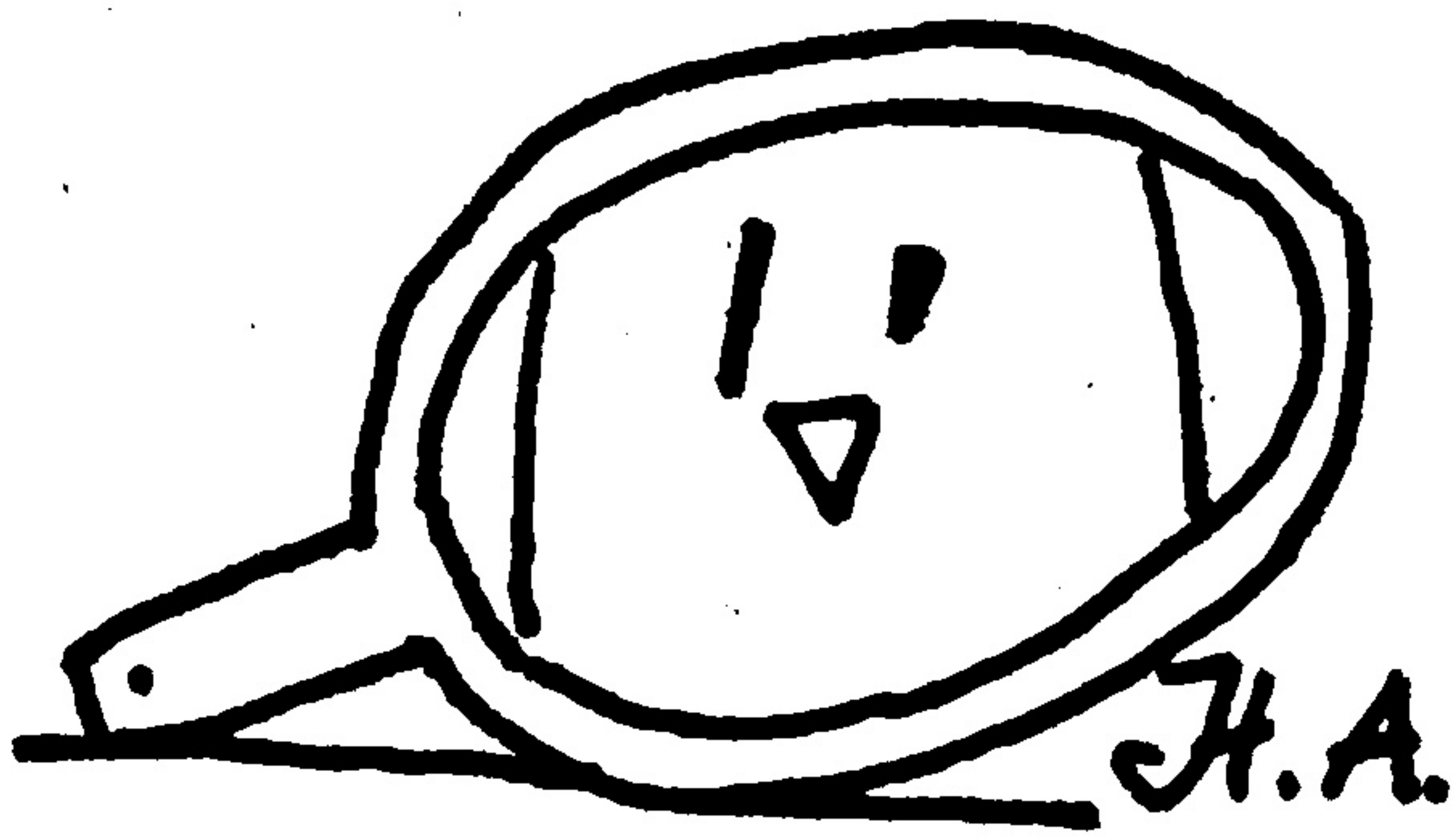
KIT Karlsruher Institut für Technologie

Fakultät für Mathematik
Institut für Angewandte und
Numerische Mathematik

Prof. Dr. Marlis Hochbruck

Modulklausur zur Vorlesung Numerik II

6.8.2014



- Es werden nur Lösungsvorschläge gewertet, bei denen der Lösungsweg klar zu erkennen ist. Die Angabe des Endresultats alleine gibt keine Punkte.
- Die Klausur dauert 90 Minuten. Bei vollständiger Lösung erhalten Sie für jede Aufgabe 8 Punkte. Zum Bestehen dieser Klausur sind **20** Punkte hinreichend.
- Als Hilfmittel ist eine DIN-A4-Seite, handgeschrieben und nicht kopiert mit beliebigem Inhalt erlaubt. Weitere Hilfsmittel sind nicht zugelassen, insbesondere keine elektronischen Geräte (z.B. Taschenrechner) und keine Zeichenschablonen (außer Lineal und Geodreieck).

Aufgabe 1: (weiß)

Gegeben sei die Funktion $f(x) = \dfrac{x^3+1}{4}$.

a) Zeigen Sie, dass f genau einen Fixpunkt $\widehat{x}$ im Intervall $[0,1]$ hat und die Fixpunktiteration für jeden Startwert $x_0 \in [0,1]$ gegen $\widehat{x}$ konvergiert.

b) Führen Sie mit dem Startwert $x_0 = 0$ so viele Schritte der Fixpunktiteration durch, dass $|x_k - \widehat{x}| \leq 0.05$ gilt.
Hinweis: Benutzen Sie die a-posteriori Fehlerschranke.

Aufgabe 2: (gelb)

a) Berechnen Sie mit Hilfe des Satzes von Gershgorin eine möglichst kleine Menge die alle Eigenwerte der Matrix

$$A = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 10 & 1 \\ 0 & 1 & 5 \end{bmatrix}$$

enthält.

b) Schätzen Sie mit den Resultaten aus a) möglichst genau die Konvergenzgeschwindigkeit der Potenzenmethode zur Approximation des betragsgrößten Eigenwerts λ_1 der Matrix A aus a) ab, d.h. bestimmen Sie $\widehat{\eta}$ und α so, dass für den Rayleigh–Quotienten ρ_k im k-ten Schritt

$$|\rho_k - \lambda_1| \leq C\widehat{\eta}^{\alpha}$$

für eine Konstante $C > 0$ gilt.

c) Es sei $A \in \mathbb{C}^{n,n}$ eine beliebige Matrix und $\alpha \in \mathcal{F}(A)$. Zeigen Sie, dass eine unitäre Matrix $U \in \mathbb{C}^{n,n}$ existiert so, dass α das $(1,1)$-Element von $U^H AU$ ist.

Hinweis: $\mathcal{F}(A)$ bezeichnet den Wertebereich von A.

Aufgabe 3: (rosa)

Gegeben sei eine nicht singuläre, diagonalisierbare Matrix $A \in \mathbb{R}^{n,n}$ und ein Vektor $b \in \mathbb{R}^n$, $b \neq 0$, $n \geq 10$. Wir verwenden die Notation:

$$A = X\Lambda X^{-1}, \qquad \Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n) \quad \text{mit} \quad \lambda(A) = \{\lambda_1, \ldots, \lambda_n\}.$$

Wir verwenden das GMRES Verfahren mit $x_0 = 0$ um das lineare Gleichungssystem $Ax = b$ zu lösen.

a) Zeigen Sie, dass die Residuennorm im m-ten Schritt die folgende Abschätzung erfüllt:

$$\|r_m\|_2 \leq \kappa_2(X) \min_{\substack{p \in \mathcal{P}_m \\ p(0)=1}} \max_{i=1,\ldots,n} |p(\lambda_i)| \, \|r_0\|_2 .$$

Hinweis: Zeigen Sie zunächst, dass $r_m = p(A)b$ für ein geeignetes Polynom p gilt.

b) Sei nun $\lambda_i = (-1)^i$ für $i = 1, \ldots, n$. Zeigen Sie, dass das GMRES Verfahren nach maximal 2 Schritten die exakte Lösung $A^{-1}b$ liefert.

Aufgabe 4: (grün)

Gegeben sei die folgende Variante des BiCG-Verfahrens:

Wähle $x_0 \in \mathbb{C}^n$, setze $r_0 = b - Ax_0$; $\rho_0 = \|r_0\|$,
Wähle einen linken Startvektor $w_1 \neq 0$;
Initialisiere $p_0 = q_0 = 0$, $\delta_0 = 1$, $\mu_1 = 0$, $m = 1$;
while $\rho_{m-1} \geq$ tol **do**
    Berechne $\delta_m = w_m^H r_{m-1}$, $\mu'_m = \delta_m / \delta_{m-1}$;
    Setze $p_m = r_{m-1} + \mu'_m p_{m-1}$, $q_m = w_m + \overline{\mu}'_m q_{m-1}$;
    Berechne $\delta'_m = q_m^H A p_m$;
    if $\delta'_m \neq 0$ **then**
        Berechne $\omega_m = \delta_m / \delta'_m$
    else
        Setze $L = m$, Stopp;
    end if
    Setze $r_m = r_{m-1} - \omega_m A p_m$, $w_{m+1} = w_m - \overline{\omega}_m A^H q_m$;
    Setze $\rho_m = \|r_m\|$;
    Update: $x_m = x_{m-1} + \omega_m p_m$;
    if $\delta_m = 0$ **then**
        Setze $L = m + 1$, Stopp;
    end if
    $m = m + 1$;
end while

---

a) Geben Sie den minimalen Rechenaufwand dieses Verfahrens für den m-ten Iterationsschritt, $m > 1$, gemessen an der Anzahl der Matrix-Vektormultiplikationen, saxpy's und Skalarprodukte an.

b) Geben Sie die minimale Anzahl von Vektoren an, die für die Implementierung des Verfahrens gespeichert werden müssen (der Speicheraufwand für die Matrix A soll unberücksichtigt bleiben).

c) Zeigen Sie, dass für $m = 1, 2, \ldots, L$ gilt:

$$\mathrm{span}\{r_0, \ldots, r_{m-1}\} = \mathrm{span}\{p_1, \ldots, p_m\} = \mathcal{K}_m(A, r_0).$$

Hinweis zu a) und b): Begründen Sie Ihre Aussagen. Sie können hierfür die angegebenen Zeilennummern verwenden.

Aufgabe 5: (blau)

Geben Sie bei maximal 8 der folgenden 10 Aussagen an, ob sie jeweils wahr oder falsch sind, indem Sie die Nummer der Aussage in den entsprechenden Kasten auf dem blauen Blatt eintragen.

Es sei eine nicht singuläre Matrix $A \in \mathbb{R}^{n,n}$ und ein beliebiger Vektor $b \in \mathbb{R}^n \setminus \{0\}$ gegeben. Ferner sei $\widehat{x}$ eine Lösung von $Ax = b$ und der Startwert $x_0 = 0$ gewählt.

1) Es gilt $5b + 2A^2b \in \mathcal{K}_5(A, b)$.

2) Es gilt $\dim \mathcal{K}_n(A, b) = \mathrm{Rank}(A)$.

3) Die m-te Iterierte des GMRES-Verfahrens minimiert die Maximumsnorm des Residuums im Raum $\mathcal{K}_m(A, b)$.

4) Ist $\mathcal{K}_m(A, b) = \mathcal{K}_{m+1}(A, b)$ und x_m die m-te Iterierte des QMR-Verfahrens, dann gilt $x_m = \widehat{x}$.

5) Ist $A = A^T$ positiv definit, so stimmt die FOM-Iterierte mit der cg-Iterierten überein.

6) Ist $A = A^T$ positiv definit, so minimiert die m-te Iterierte des cg-Verfahrens die A-Norm des Residuums im Raum $\mathcal{K}_m(A, b)$.

7) Die FOM-Iterierten existieren für jedes $m = 1, 2, \ldots, \mathrm{Rank}(A)$.

8) Ist $A = A^T$ positiv definit, so konvergiert das cg-Verfahren linear.

9) Ist A normal, so kann man mit dem Satz von Bendixson das kleinste achsenparallele Rechteck finden, welches alle Eigenwerte von A enthält.

10) Mit Hilfe des Satzes von Gershgorin, kann man ein Intervall konstruieren, in dem alle Eigenwerte von A enthalten sind.

Hinweis zur Bewertung: (zu Aufgabe 5)

Wenn Sie mehr als 8 Aussagen bearbeiten, werden die 8 Aussagen mit den niedrigsten Nummern gewertet.

Sie erhalten für jede richtige Antwort 1 Punkt, für jede falsche -1 Punkt und 0 Punkte, wenn Sie keine Antwort geben. Die minimale Gesamtpunktzahl beträgt 0 Punkte.

Lösungsvorschlag: **Aufgabe 1**

a) Wir zeigen, dass f eine kontrahierende Selbstabbildung auf der abgeschlossenen Menge $[0,1]$ ist. Dann folgt die Behauptung direkt aus dem Banach'schen Fixpunktsatz.

Selbstabbildung Die Funktion f is monoton wachsend. Daraus folgt

$$f([0,1]) = [f(0), f(1)] = [\frac{1}{4}, \frac{1}{2}] \subset [0,1],$$

also ist f eine Selbstabbildung.

Kontraktion Aus

$$|f(x) - f(y)| = \frac{1}{4}\left|x^3 - y^3\right| = \frac{1}{4}\left|x^2 + xy + y^2\right| |x-y| \leq \frac{3}{4}|x-y|$$

folgt, dass f Lipschitz–stetig ist, mit Lipschitzkonstante $L = \frac{3}{4} < 1$. Also f ist kontraktiv. Alternativ genügt es

$$L = \max_{x\in[0,1]} |f'(x)| = \frac{3}{4} \max_{x\in[0,1]} \left|x^2\right| = \frac{3}{4}$$

abzuschätzen.

b) Die A-posteriori-Schranke ist:

$$|e_k| = \frac{L}{1-L}|x_k - x_{k-1}|$$

Aus Aufgabenteil a) wissen wir $L = 3/4$ also,

$$\frac{L}{1-L} = 3.$$

Die gewünschte Genauigkeit ist dann schon nach 2 Iterationen erreicht:

$x_1 = 1/4,$ $\qquad |e_1| \leq 3\,|1/4 - 0| = 3/4,$

$x_2 = ((1/4)^3 + 1)/4 = (65/64)/4 = 65/256,$ $\qquad |e_2| \leq 365/256 - 1/4 = 3/256 < 0.05.$

Lösungsvorschlag: **Aufgabe 2**

a) Da $A = A^T$ symmetrisch ist, betrachten wir nur die Gershgorin–Kreise von A und wissen vorher schon das $\lambda(A) \subset \mathbb{R}$. Dann erhalten wir

$$\lambda(A) \subset [-1,1] \cup [4,6] \cup [8,12].$$

b) Da A symmetrisch und damit auch normal ist, wissen wir aus Satz 6.24, dass die Potenzenmethode mit

$$\rho_A(y_k) = \lambda_1 + \mathcal{O}(\eta^{2k}) .$$

konvergiert. Da die Gershgorin–Kreise aus Teil a) disjunkt sind wissen wir (mit Satz 6.16), dass

$$\lambda_3 \in [-1, 1], \qquad \lambda_2 \in [4, 6], \qquad \lambda_1 \in [8, 12] .$$

Damit lässt sich das Verhältnis zwischen zweit-größtem und größtem Eigenwert abschätzen durch

$$\eta = \frac{|\lambda_2|}{|\lambda_1|} \le \frac{6}{8} = \frac{3}{4} .$$

und wir erhalten

$$\rho_A(y_k) = \lambda_1 + \mathcal{O}\left((3/4)^{2k}\right) .$$

c) Ist $\alpha \in \mathcal{F}(A)$, dann gibt es $u_1 \in \mathbb{C}^n$ mit $\|u_1\| = 1$ so, dass $u_1^H A u_1 = \alpha$. Wir ergänzen u_1 durch $u_2, \ldots, u_n \in \mathbb{C}^n$ zu einer Orthonormalbasis des $\mathbb{C}^n$. Dann ist $U = [u_1\, u_2 \, \ldots \, u_n]$ unitär und das $(1, 1)$-Element ist

$$e_1^T U^H A U e_1 = u_1^H A u_1 = \alpha.$$

Lösungsvorschlag: **Aufgabe 3**

a) Das GMRES Verfahren verwendet die Bedingung (M), also

$$\|r_m\|_2 = \|b - Ax_m\|_2 = \min_{x \in \mathcal{K}_m(A,b)} \|b - Ax\|_2 .$$

Sei nun $x_m \in \mathcal{K}_m(A, b)$, dann gibt es ein $q \in \mathcal{P}_{m-1}$ mit $x_m = q(A)b$. Es gilt:

$$r_m = b - Ax_m = b - Aq(A)b = (I - q(A))b = p(A)b$$

mit

$$p(z) = 1 - zq(z), \; p \in \mathcal{P}_m, \; p(0) = 1.$$

Also ist

$$\begin{aligned}\|r_m\|_2 = \min_{\substack{p \in \mathcal{P}_m \\ p(0)=1}} \|p(A)b\|_2 &= \min_{\substack{p \in \mathcal{P}_m \\ p(0)=1}} \left\|X p(\Lambda) X^{-1} b\right\|_2 \\ &\le \min_{\substack{p \in \mathcal{P}_m \\ p(0)=1}} \|X\|_2 \, \|p(\Lambda)\|_2 \left\|X^{-1}\right\|_2 \|b\|_2\end{aligned}$$

Die Behauptung folgt nun mit

$$r_0 = b, \qquad \kappa_2(X) = \|X\|_2 \left\|X^{-1}\right\|_2, \qquad \|p(\Lambda)\|_2 = \max_{i=1,\ldots,n} |p(\lambda_i)| .$$

b) Wir wählen das Polynom $\hat{p}(\lambda) = 1 - \lambda^2$. Dann gilt:

$$\hat{p} \in \mathcal{P}_2, \qquad \hat{p}(0) = 1, \qquad \hat{p}(-1) = \hat{p}(1) = 0 .$$

Mit der Ungleichung aus (a) Teil folgt also

$$\|r_2\|_2 \le \kappa_2(X) \max_{i=1,\ldots n} |\hat{p}(\lambda_i)| \, \|r_0\|_2 = 0 .$$

Also ist x_2 schon die exakte die Lösung.

Alternativ: Da $\Lambda = \operatorname{diag}(-1,1,-1,1,\dots,(-1)^n)$ ist

$$A^2b = X\Lambda X^{-1}X\Lambda X^{-1}b = X\Lambda^2X^{-1}b = XX^{-1}b = b$$

und der dritte Krylov–Raum damit invariant, d.h.

$$\mathcal{K}_3(A,b) = \mathcal{K}_2(A,b)\,.$$

Die exakte Lösung liegt also schon in $\mathcal{K}_2(A,b)$ und ist durch x_2 gegeben.

Lösungsvorschlag: **Aufgabe 4**

a) Rechenaufwand:

- 2 Matrix-Vektormultiplikationen: Ap_m in Zeile 7, A^Hq_m in Zeile 13.
- 5 saxpy's: je 2 in Zeile 6 und 13, eins in Zeile 15
- 3 Skalarprodukte: δ_m in Zeile 5, δ'_m in Zeile 7, ρ_m in Zeile 14

b) Speicheraufwand 6 Vektoren: je ein Vektor für p_m, q_m, r_m, w_m, x_m sowie ein temporärer Vektor für Ap_m und A^Hq_m (wird in Zeile 7 und 13 benötigt).

c) Da $m \leq L$ gilt $\omega_m, \mu'_m \neq 0$.

Der Beweis wird durch Induktion nach m geführt. Sei also $m = 1$, dann ist

$$r_0 \in \mathcal{K}_1(A,r_0)$$
$$p_1 = r_0 - \mu'_1 \underbrace{p_0}_{=0} = r_0 \in \mathcal{K}_1(A,r_0)$$

und die Behauptung folgt, da $\dim \mathcal{K}_1(A,r_0) = 1$.

Wir nehmen nun an, dass die Behauptung für ein festes m (mit $m+1 \leq L$) gelte. Dann

$$r_m = \underbrace{r_{m-1}}_{\in \mathcal{K}_m(A,r_0)} - \omega_m \underbrace{Ap_m}_{\in \mathcal{K}_{m+1}(A,r_0)\setminus\mathcal{K}_m(A,r_0)} \in \mathcal{K}_{m+1}(A,r_0) \setminus \mathcal{K}_m(A,r_0)$$

also

$$\mathcal{K}_{m+1}(A,r_0) = \operatorname{span}\{r_0,\dots,r_m\}\,.$$

Analog gilt

$$p_{m+1} = \underbrace{r_m}_{\in \mathcal{K}_{m+1}(A,r_0)\setminus\mathcal{K}_m(A,r_0)} - \mu'_{m+1} \underbrace{p_m}_{\in \mathcal{K}_m(A,r_0)} \in \mathcal{K}_{m+1}(A,r_0) \setminus \mathcal{K}_m(A,r_0)$$

und damit auch $\operatorname{span}\{p_1,\dots,p_{m+1}\} = \mathcal{K}_{m+1}(A,r_0)$.

Lösungsvorschlag: **Aufgabe 5**

1) (wahr) Es ist sogar $5b + 2A^2b \in \mathcal{K}_2(A,b) \subsetneq \mathcal{K}_5(A,b)$.
2) (falsch) Z.B. gilt für einen Eigenvektor b, dass $\dim \mathcal{K}_n(A,b) = 1$.
3) (falsch) GMRES minimiert das Residuum in der Euklid-Norm.
4) (wahr) Satz 7.5

5) (wahr) Die Residuenvektoren des cg-Verfahrens sind orthogonal und spannen den Krylov-Raum auf.

6) (falsch) Die cg-Iterierten minimieren die A-Norm des Fehlers

7) (falsch) Die Iterierten existieren genau dann, wenn H_m nicht singulär ist.

8) (wahr) Konvergenzrate $2\Big((\sqrt{\kappa}-1)/(\sqrt{\kappa}+1)\Big)$

9) (wahr) Da A normal ist, ist der Wertebereich die konvexe Hülle der Eigenwerte von A, welche symmetrisch zur reellen Achse liegen.

R sei das Rechteck aus dem Satz von Bendixson, d.h. $\mathcal{F}(A) \subset R$. Es genügt zu zeigen, dass R das kleinste Rechteck ist. Dies folgt aus

$$\frac{1}{2}x^H(A+A^H)x = \operatorname{Re} x^H Ax, \qquad \frac{1}{2i}x^H(A-A^H)x = \operatorname{Im} x^H Ax,$$

10) (falsch) Die Eigenwerte von A müssen nicht reell sein.