Inverse Probleme

Vorlesung im Wintersemester 2015/2016

Andreas Kirsch

Institut für Angewandte und Numerische Mathematik

Karlsruher Institut für Technologie (KIT)

10. Februar 2016

5 Nichtlineare Probleme

In diesem Kapitel seien X und Y normierte Räume (meistens Hilberträume), $K : X \supset \mathcal{D}(K) \to Y$ eine (nichtlineare) Abbildung mit offenem Definitionsbereich $\mathcal{D}(K) \subset X$. Es sei $x^* \in \mathcal{D}(K)$ und $y \in Y$ mit $K(x^*) = y$. Es ist – wie im linearen Fall – die Aufgabe, aus gestörten Daten $y^\delta \in Y$ mit $\|y^\delta - y\|_Y \leq \delta$ eine Näherung an x^* zu bestimmen. Sei im Folgenden $B(x,r) = \{z : \|x - z\| < r\}$ die offene Kugel um x mit Radius r.

Definition 5.1 *Sei $x^* \in \mathcal{D}(K)$ und $y = K(x^*)$. Die Gleichung $K(x) = y$ heißt in x^** ***lokal schlecht gestellt****, wenn es zu jedem $r > 0$ eine Folge $x_n \in \mathcal{D}(K) \cap B(x^*, r)$ gibt mit $x_n \not\to x^*$ und $K(x_n) \to K(x^*)$ für $n \to \infty$.*

Beispiel 5.2 Sei $k \in C^1\big([0,1] \times [0,1] \times \mathbb{R}\big)$, $k = k(t,s,r)$, und es gebe $c_1 > 0$ mit $\big|\partial k(t,s,r)/\partial r\big| \leq c_1$ für alle $(t,s,r) \in [0,1] \times [0,1] \times \mathbb{R}$. Es sei

$$K(x)(t) \;=\; \int_0^1 k\big(t,s,x(s)\big)\,ds\,, \quad t \in [0,1]\,.$$

Dann ist K wohldefiniert von $L^2(0,1)$ nach $L^2(0,1)$, und die Gleichung $K(x) = y$ ist lokal schlecht gestellt in $x = 0$.

Beweis: Aus der Gleichung

$$k(t,s,r) \;=\; k(t,s,0) \;+\; \int_0^r \frac{\partial k}{\partial r}(t,s,r')\,dr'\,, \quad r \in \mathbb{R}\,, \tag{1}$$

folgt $\bigl|k(t,s,r)\bigr| \le \bigl|k(t,s,0)\bigr| + c_1|r|$, also $\bigl|k(t,s,r)\bigr|^2 \le 2\bigl|k(t,s,0)\bigr|^2 + 2c_1^2 r^2$, also mit der Cauchy-Schwarzschen Ungleichung

$$\bigl|K(x)(t)\bigr|^2 \;\le\; \int_0^1 k\bigl(t,s,x(s)\bigr)^2\,ds \;\le\; 2\int_0^1 \Bigl[\bigl|k(t,s,0)\bigr|^2 + c_1^2 x(s)^2\Bigr]\,ds \;=:\; c$$

für alle $t \in [0,1]$. Daher ist $K(x)$ messbar und $\bigl|K(x)(t)\bigr|^2$ beschränkt, was $K(x) \in L^2(0,1)$ impliziert.

Sei jetzt $r > 0$ beliebig und $x_n(t) = r\sqrt{2n+1}\,t^n$, $t \in [0,1]$. Dann ist $\|x_n\|^2_{L^2(0,1)} = r^2(2n+1)\int_0^1 t^{2n}dt = r^2$ und mit (1)

$$\bigl|K(x_n)(t) - K(0)(t)\bigr| \;\le\; c_1 \int_0^1 \bigl|x_n(s)\bigr|\,ds \;=\; c_1 r\sqrt{2n+1}\int_0^1 s^n\,ds \;=\; \frac{c_1 r\sqrt{2n+1}}{n+1} \;\to\; 0$$

für $n \to \infty$. Daher ist die Gleichung $K(x) = y$ lokal schlecht gestellt in $x = 0$.

Wie schon aus anderen Analysis-Vorlesungen bekannt, spielt bei nichtlinearen Problemen die Linearisierung eine wichtige Rolle. Diese wird mit Hilfe der Ableitung berechnet. Wir benötigen also den Begriff der Ableitung von (nichtlinearen) Abbildungen zwischen normierten Räumen.

5.1 Die Fréchet-Ableitung

Definition 5.3 *Es seien X und Y normierte Räume und $X \supset \mathcal{D}(K) \to Y$ eine Abbildung. K heißt in $x^* \in \mathcal{D}(K)$ Fréchet-differenzierbar, wenn es einen linearen und beschränkten Operator $A : X \to Y$ gibt mit*

$$\lim_{\|h\|_X \to 0} \frac{\|K(x^*+h) - K(x^*) - Ah\|_Y}{\|h\|_X} \;=\; 0\,. \tag{2}$$

Man schreibt $K'(x^*)$ für A.

Achtung: $K(x^*+h), K(x^*) \in Y$ sind Funktionswerte in Y, $K'(x^*) : X \to Y$ ist eine lineare, beschränkte Abbildung, also $K' : X \supset \mathcal{D}(K') \to L(X,Y) =$ Raum der linearen und beschränkten Operatoren von X nach Y.

Beispiele 5.4 (a) $X = \mathbb{R}^n$, $Y = \mathbb{R}^m$, $K : \mathbb{R}^n \to \mathbb{R}^m$ sei stetig partiell differenzierbar mit Funktionalmatrix (Jacobimatrix) $J_K(x^*) = \left(\frac{\partial K_i}{\partial x_j}(x^*)\right)_{i,j} \in \mathbb{R}^{m\times n}$. Dann ist bezüglich der kanonischen Orthonormalbasis von $\mathbb{R}^n$ und $\mathbb{R}^m$ die Fréchet-Ableitung von K gegeben durch $K'(x^*)h = J_K(x^*)h$ (links: Operator, angewandt auf h, rechts: Matrix mal Vektor).

(b) (vgl. Beispiel 5.2) Sei $k \in C^1\big([a,b]\times[c,d]\times\mathbb{R}\big)$, $k = k(t,s,r)$, und $K : C[c,d] \to C[a,b]$ sei gegeben durch

$$K(x)(t) \;=\; \int_c^d k\big(t,s,x(s)\big)\,ds\,, \quad t \in [a,b]\,.$$

Hier betrachten also jetzt die normierten Räume $C[c,d]$ und $C[a,b]$ mit der Maximumnorm. In diesen Räumen ist K wohldefiniert (klar) und sogar Fréchet-differenzierbar mit

$$K'(x^*)h(t) \;=\; \int_c^d \frac{\partial k}{\partial r}\big(t,s,x^*(s)\big)\,h(s)\,ds\,, \quad t \in [a,b]\,.$$

Beweis: Sei Ah der Term auf der rechten Seite. Dieser definiert einen linearen, beschränkten Operator $A : C[c,d] \to C[a,b]$ (Beginn der Vorlesung, A ist sogar kompakt). $\partial k/\partial r$ ist nach Voraussetzung stetig auf $[a,b]\times[c,d]\times\mathbb{R}$, also gleichmäßig stetig auf der kompakten Menge $M = [a,b]\times[c,d]\times[-\|x^*\|_\infty - 1, \|x^*\|_\infty + 1]$. Sei $\varepsilon > 0$. Wähle $\delta \in (0,1)$ mit $\left|\frac{\partial k}{\partial r}(t,s,r) - \frac{\partial k}{\partial r}(t,s,\tilde r)\right| \le \frac{\varepsilon}{d-c}$ für alle $(t,s,r),(t,s,\tilde r) \in M$ mit $|r-\tilde r| \le \delta$. Wir schätzen jetzt ab für $h \in C[c,d]$ mit $\|h\|_\infty \le \delta$:

$$\begin{aligned}
&\big|K(x^*+h)(t) - K(x^*)(t) - Ah(t)\big| \\
&= \left|\int_c^d \left[k\big(t,s,x^*(s)+h(s)\big) - k\big(t,s,x^*(s)\big) - \frac{\partial k}{\partial r}\big(t,s,x^*(s)\big)\,h(s)\right] ds\right| \\
&= \left|\int_c^d \int_{x^*(s)}^{x^*(s)+h(s)} \left(\frac{\partial k}{\partial r}(t,s,r) - \frac{\partial k}{\partial r}\big(t,s,x^*(s)\big)\right) dr\,ds\right| \\
&\le \int_c^d \int_{x^*(s)}^{x^*(s)+h(s)} \left|\frac{\partial k}{\partial r}(t,s,r) - \frac{\partial k}{\partial r}\big(t,s,x^*(s)\big)\right| |dr|\,ds \\
&\le \int_c^d \|h\|_\infty \frac{\varepsilon}{d-c}\,ds \;=\; \varepsilon\|h\|_\infty\,.
\end{aligned}$$

Dies gilt für alle $t \in [a,b]$, also auch für das Maximum. Daher gilt:

$$\frac{\|K(x^*+h) - K(x^*) - Ah\|_\infty}{\|h\|_\infty} \;\le\; \varepsilon \quad \text{für alle } h \in C[c,d] \text{ mit } \|h\|_\infty \le \delta\,.$$

Dies beweist die Behauptung.

(c) Sei $k \in C\big([a,b] \times [c,d] \times \mathbb{R}\big)$ und bzgl. r zweimal stetig differenzierbar mit beschränkter zweiter Ableitung, d.h. es existiere $\kappa > 0$ mit $\big|\partial^2 k(t,s,r)/\partial r^2\big| \le \kappa$ für alle $(t,s,r) \in [a,b] \times [c,d] \times \mathbb{R}$. Dann ist der Operator K von Teil (b) auch Fréchet-differenzierbar als Operator von $L^2(c,d)$ nach $L^2(a,b)$ mit der gleichen Darstellung der Ableitung.

Beweis: Definiere wieder den Operator A wie in Teil (b). Wir benutzen die Abschätzung $\big|\frac{\partial k}{\partial r}k(t,s,r) - \frac{\partial k}{\partial r}k(t,s,\tilde{r})\big| \le \kappa|r-\tilde{r}|$. Dann ist A beschänkt in den L^2-Räumen wegen $\big|Ah(t)\big| \le \int_c^d \big|\frac{\partial k}{\partial r}k\big(t,s,x^*(s)\big)\big|\,\big|h(s)\big|\,ds \le \int_c^d \big[\kappa_1 + \kappa|x^*(s)|\big]|h(s)|\,ds \le \big[\kappa_1\sqrt{d-c} + \kappa\|x^*\|_{L^2(c,d)}\big]\,\|h\|_{L^2(c,d)}$ für alle $t \in [a,b]$. Hier ist κ_1 eine Schranke für $\frac{\partial k}{\partial r}k(t,s,0)$.

Für die Berechnung der Ableitung können wir fast genauso vorgehen wie in Teil (b):

$$
\begin{aligned}
&\big|K(x^*+h)(t) - K(x^*)(t) - Ah(t)\big| \\
&\quad\le \int_c^d \int_{x^*(s)}^{x^*(s)+h(s)} \left|\frac{\partial k}{\partial r}(t,s,r) - \frac{\partial k}{\partial r}\big(t,s,x^*(s)\big)\right| |dr|\,ds \\
&\quad\le \kappa \int_c^d \int_{x^*(s)}^{x^*(s)+h(s)} \big|r - x^*(s)\big|\,|dr|\,ds \;\le\; \kappa \int_c^d h(s)^2\,ds \;=\; \kappa\,\|h\|^2_{L^2(c,d)}\,.
\end{aligned}
$$

Dies gilt für alle $t \in [a,b]$. Quadrieren und Integration liefern die Behauptung. □

Lemma 5.5 *Ist K in x^* differenzierbar, so auch stetig in x^*.*

Beweis: Sei $x_n \to x^*$. Aus der Differenzierbarkeit folgt

$$
\frac{\|K(x_n) - K(x^*) - K'(x^*)(x_n - x^*)\|_Y}{\|x_n - x^*\|_X} \;\longrightarrow\; 0\,, \quad n \to \infty\,,
$$

also insbesondere $\|K(x_n) - K(x^*) - K'(x^*)(x_n - x^*)\|_Y \to 0$ für $n \to \infty$ und daher $K(x_n) \to K(x^*)$. □

Definition 5.6 *K heißt in x^* stetig differenzierbar, wenn K in einer Umgebung U von x^* differenzierbar ist und $K' : X \supset U \to L(X,Y)$ in x^* stetig ist.*

Lemma 5.7 *Sei K differenzierbar in der Kugel $B(x^*,\rho)$ um $x^* \in \mathcal{D}(K)$ mit Radius $\rho > 0$, und es gebe $\gamma > 0$ mit $\|K'(x) - K'(x^*)\|_{L(X,Y)} \le \gamma \|x - x^*\|_X$ für alle $x \in B(x^*,\rho)$. Dann gilt:*

$$\big\|K(x) - K(x^*) - K'(x^*)(x - x^*)\big\|_Y \;\le\; \frac{\gamma}{2}\,\|x - x^*\|_X^2 \quad \textit{für alle } x \in B(x^*,\rho)\,.$$

Beweis: Sei $z \in Y$ und $x \in B(x^*,\rho)$ festgehalten. Setze $h = x - x^*$ und definiere die skalare Funktion $f(t) = \big(K(x^* + th), z\big)_Y$ für $|t| \le \rho/\|h\|_X$. Dann ist $f'(t) = \big(K'(x^* + th)h, z\big)_Y$ (verbleibt als Übung) und daher

$$\begin{aligned}
&\big|\big(K(x) - K(x^*) - K'(x^*)h\,,z\big)_Y\big| \\
&= \big|\big(K(x), z\big)_Y - \big(K(x^*), z\big)_Y - \big(K'(x^*)h, z\big)_Y\big| \\
&= \big|f(1) - f(0) - K'(x^*)h\,,z\big)_Y\big| \;=\; \left|\int_0^1 \left[\frac{d}{dt}f(t) - \big(K'(x^*)h, z\big)_Y\right] dt\right| \\
&= \left|\int_0^1 \big[\big(K'(x^* + th)h, z\big)_Y - \big(K'(x^*)h, z\big)_Y\big]\, dt\right| \\
&= \left|\int_0^1 \big(\big[K'(x^* + th) - K'(x^*)\big]h, z\big)_Y\, dt\right| \\
&\le \|z\|_Y \|h\|_X \int_0^1 \big\|K'(x^* + th) - K'(x^*)\big\|_{L(X,Y)}\, dt \\
&\le \gamma \|z\|_Y \|h\|_X^2 \int_0^1 t\, dt \;=\; \frac{\gamma}{2}\|z\|_Y \|h\|_X^2\,.
\end{aligned}$$

Setzt man jetzt $z = K(x) - K(x^*) - K'(x^*)h$, so erhält man die Behauptung. $\square$

Unter der Bedingung dieses Lemmas folgt aus der lokalen Schlechtgestelltheit die Schlechtgestelltheit der Linearisierung.

Satz 5.8 *Sei K differenzierbar in der Kugel $B(x^*,\rho)$ um $x^* \in \mathcal{D}(K)$ mit Radius $\rho > 0$, und es gebe $\gamma > 0$ mit $\|K'(x) - K'(x^*)\|_{L(X,Y)} \le \gamma \|x - x^*\|_X$ für alle $x \in B(x^*,\rho)$. Sei die Gleichung $K(x) = y$ lokal schlecht gestellt in x^*. Dann ist $K'(x^*)$ nicht beschränkt invertierbar, d.h. die lineare Gleichung $K'(x^*)h = z$ ist ebenfalls schlecht gestellt.*

Beweis: Wir nehmen an, dass $K'(x^*)$ beschränkt invertierbar und wählen $r \in (0,\rho)$ so, dass $\frac{r\gamma}{2}\,\|K'(x^*)^{-1}\|_{L(Y,X)} =: q < 1$. Zu diesem r wählen wir eine Folge $x_n \in$

$B(x^*, r)$ gemäß Definition 5.1. Wegen des letzten Lemmas gilt die Darstellung

$$K(x_n) - K(x^*) \;=\; K'(x^*)(x_n - x^*) + R_n \quad \text{mit } \|R_n\|_Y \le \frac{\gamma}{2}\,\|x_n - x^*)\|_X^2 \quad \text{für alle } n \in \mathbb{N}\,,$$

d.h.

$$K'(x^*)^{-1}\big[K(x_n) - K(x^*)\big] \;=\; x_n - x^* \;+\; K'(x^*)^{-1}R_n \quad \text{für alle } n \in \mathbb{N}\,,$$

also

$$\begin{aligned}
\|x_n - x^*\|_X &\le \big\|K'(x^*)^{-1}\big[K(x_n) - K(x^*)\big]\big\|_X + \|K'(x^*)^{-1}R_n\|_X \\
&\le \|K'(x^*)^{-1}\|_{L(Y,X)}\,\big\|K(x_n) - K(x^*)\big\|_Y \\
&\quad + \frac{\gamma}{2}\,\|K'(x^*)^{-1}\|_{L(Y,X)}\,\|x_n - x^*\|_X^2 \\
&\le \|K'(x^*)^{-1}\|_{L(Y,X)}\,\big\|[K(x_n) - K(x^*)\big\|_Y \;+\; q\,\|x_n - x^*\|_X\,,
\end{aligned}$$

also

$$(1-q)\,\|x_n - x^*\|_X \;\le\; \|K'(x^*)^{-1}\|_{L(Y,X)}\,\big\|[K(x_n) - K(x^*)\big\|_Y\,,$$

und dieser Ausdruck konvergiert gegen Null wegen $K(x_n) \to K(x^*)$. Dies widerspricht $x_n \not\to x^*$. $\square$

5.2 Die nichtlineare Tikhonovregularisierung

Sei wieder $x^* \in \mathcal{D}(K)$ und $y \in Y$ mit $K(x^*) = y$ und $y^\delta \in Y$ mit $\|y^\delta - y\|_Y \le \delta$. Zusätzlich sei $\hat{x} \in X$ (gedacht als Näherung von x^*) gegeben. Definiere das Tikhonov-Funktional

$$J_{\alpha,\delta}(x) \;=\; \|K(x) - y^\delta\|_Y^2 \;+\; \alpha\|x - \hat{x}\|_X^2\,, \quad x \in \mathcal{D}(A)\,.$$

Wir machen die folgende Voraussetzung:

Voraussetzung (A): Für alle $\alpha, \delta > 0$ besitzt $J_{\alpha,\delta}$ globale Minimalstellen, d.h. die Menge $M_{\alpha,\delta}^* := \big\{\tilde{x} \in \mathcal{D}(K) : J_{\alpha,\delta}(\tilde{x}) \le J_{\alpha,\delta}(x)\ \forall\, x \in \mathcal{D}(K)\big\} \ne \emptyset$. Ferner existiere $x^{\alpha,\delta} \in M_{\alpha,\delta}^*$ mit $\|x^{\alpha,\delta} - \hat{x}\|_X \le \|\tilde{x} - \hat{x}\|_X$ für alle $\tilde{x} \in M_{\alpha,\delta}^*$, d.h. $x^{\alpha,\delta}$ ist beste Approximation an $\hat{x}$ in $M_{\alpha,\delta}^*$. Wir nennen $x^{\alpha,\delta}$ die **Minimumnorm Lösung bzgl**

$\hat{\mathbf{x}}$. Für $\alpha = \delta = 0$ ist $M^*_{0,0}$ die Menge aller Lösungen von $K(x) = y$, und wir nehmen an, dass x^* die Minimumnorm Lösung bzgl $\hat{x}$ ist.

Es lassen sich Bedingungen an K angeben, die diese Voraussetzungen garantieren. Die Formulierung (und der Nachweis) dieser Voraussetzungen erfordern aber weitergehende Kenntnisse aus der Funktionalanalysis, insbesondere über schwache Topologien. Darauf verzichten wir.

Jetzt formulieren wir die zur „Quellenbedingung“ (source condition) analogen Voraussetzung.

Voraussetzung (B): Sei $\mathcal{D}(K)$ konvex, x^* Minimumnorm Lösung bzgl $\hat{x}$ von $K(x) = y$ und K differenzierbar auf $\mathcal{D}(K)$.

(i) Es gebe $w \in Y$ mit $x^* - \hat{x} = K'(x^*)^* w$.

Ferner existiere für jedes $\rho > 2\|x^* - \hat{x}\|_X$ ein $\gamma > 0$ mit

(ii) $\gamma \, \|w\|_Y < 1$ und

(iii) $\|K'(x) - K'(x^*)\|_{L(X,Y)} \leq \gamma \|x - x^*\|_X$ für alle $x \in B(x^*, \rho) \cap \mathcal{D}(K)$.

Unter diesen Voraussetzungen können wir Konvergenz zeigen und die aus der linearen Theorie zu erwartende Konvergenzordnung.

Satz 5.9 *Es seien die Voraussetzungen (A) und (B) erfüllt. Setze $\alpha(\delta) = c\delta$ für ein beliebiges $c > 0$. Dann gilt:*

$$\big\|x^{\alpha(\delta),\delta} - x^*\big\|_X \;=\; \mathcal{O}\big(\sqrt{\delta}\big) \quad \textit{und} \quad \big\|K(x^{\alpha(\delta),\delta}) - y\big\|_Y \;=\; \mathcal{O}(\delta)\,, \quad \delta \to 0\,.$$

Beweis: Nach Definition von $x^{\alpha,\delta}$ ist

$$\big\|K(x^{\alpha,\delta}) - y^\delta\big\|_Y^2 + \alpha\big\|x^{\alpha,\delta} - \hat{x}\big\|_X^2 \;\leq\; J_{\alpha,\delta}(x^*) \;\leq\; \delta^2 + \alpha\big\|x^* - \hat{x}\big\|_X^2\,,$$

also

$$\begin{aligned} \big\|K(x^{\alpha,\delta}) - y^\delta\big\|_Y^2 + \alpha\big\|x^{\alpha,\delta} - x^*\big\|_X^2 \;&\leq\; \delta^2 + \alpha\Big\{\big\|x^* - \hat{x}\big\|_X^2 + \big\|x^{\alpha,\delta} - x^*\big\|_X^2 \\ &\qquad - \big\|(x^{\alpha,\delta} - x^*) + (x^* - \hat{x})\big\|_X^2\Big\} \\ &\leq\; \delta^2 \;+\; 2\alpha\big(x^* - \hat{x}\,,\, x^* - x^{\alpha,\delta}\big)_X\,. \end{aligned} \tag{3}$$

Sei jetzt $\alpha = c\delta$. Wir wollen zunächst $x^{\alpha,\delta} \in B(x^*, \rho)$ zeigen für ein geeignetes $\rho > 0$. Dazu benutzen wir (3), dividieren durch $\alpha = c\delta$ und benutzen die Cauchy-Schwarzsche Ungleichung (wir schreiben $x^{\alpha,\delta}$ statt $x^{\alpha(\delta),\delta}$):

$$\left\|x^{\alpha,\delta} - x^*\right\|_X^2 \;\le\; \frac{\delta}{c} \;+\; 2\left\|x^* - \hat{x}\right\|_X \left\|x^{\alpha,\delta} - x^*\right\|_X$$

und daher mit quadratischer Ergänzung

$$\left[\left\|x^{\alpha,\delta} - x^*\right\|_X - \left\|x^* - \hat{x}\right\|_X\right]^2 \;\le\; \frac{\delta}{c} \;+\; \left\|x^* - \hat{x}\right\|_X^2,$$

also

$$\left\|x^{\alpha,\delta} - x^*\right\|_X \;\le\; \left\|x^* - \hat{x}\right\|_X \;+\; \sqrt{\frac{\delta}{c} + \left\|x^* - \hat{x}\right\|_X^2}\,.$$

Für beliebiges $\rho > 2\|x^* - \hat{x}\|_X$ existiert $\delta_0 > 0$ mit $\left\|x^{\alpha(\delta),\delta} - x^*\right\|_X \le \rho$ für all $\delta \le \delta_0$, d.h. $x^{\alpha(\delta),\delta} \in B(x^*, \rho)$. Für diese ρ wählen wir γ nach Voraussetzung (B).

Jetzt benutzen wir (i) und wieder (3):

$$\left\|K(x^{\alpha,\delta}) - y^\delta\right\|_Y^2 + \alpha\left\|x^{\alpha,\delta} - x^*\right\|_X^2 \;\le\; \delta^2 \;+\; 2\alpha\big(w, K'(x^*)(x^* - x^{\alpha,\delta})\big)_Y\,.$$

Jetzt schreiben wir

$$K(x^{\alpha,\delta}) \;=\; K(x^*) \;+\; K'(x^*)(x^{\alpha,\delta} - x^*) \;+\; r^{\alpha,\delta}$$

und wissen nach Lemma 5.7, dass $\|r^{\alpha,\delta}\|_Y \le \frac{\gamma}{2}\|x^{\alpha,\delta} - x^*\|_X^2$. Daher ist

$$\begin{aligned}
&\left\|K(x^{\alpha,\delta}) - y^\delta\right\|_Y^2 + \alpha\left\|x^{\alpha,\delta} - x^*\right\|_X^2 \\
&\quad\le\; \delta^2 \;+\; 2\alpha\big(w, y - K(x^{\alpha,\delta}) + r^{\alpha,\delta}\big)_Y \\
&\quad=\; \delta^2 \;+\; 2\alpha\big(w, y - y^\delta\big)_Y \;+\; 2\alpha\big(w, y^\delta - K(x^{\alpha,\delta})\big)_Y \;+\; 2\alpha\big(w, r^{\alpha,\delta}\big)_Y \\
&\quad\le\; \delta^2 \;+\; 2\alpha\|w\|_Y\delta \;+\; 2\alpha\|w\|_Y\left\|y^\delta - K(x^{\alpha,\delta})\right\|_Y \;+\; \gamma\alpha\|w\|_Y\|x^{\alpha,\delta} - x^*\|_X^2\,,
\end{aligned}$$

also mit quadratischer Ergänzung

$$\begin{aligned}
&\left[\left\|K(x^{\alpha,\delta}) - y^\delta\right\|_Y - \alpha\|w\|_Y\right]^2 + \alpha\left[1 - \gamma\|w\|_Y\right]\left\|x^{\alpha,\delta} - x^*\right\|_X^2 \\
&\quad\le\; \alpha^2\|w\|_Y^2 + \delta^2 + 2\alpha\delta\|w\|_Y \;=\; \big(\alpha\|w\|_Y + \delta\big)^2\,.
\end{aligned}$$

Lassen wir jeweils einen der beiden (nicht-negativen!) Terme auf der linken Seite weg, so erhalten wir

$$\big\|K(x^{\alpha,\delta})-y^\delta\big\|_Y-\alpha\|w\|_Y \le \alpha\|w\|_Y+\delta \quad \text{und} \quad \alpha\big[1-\gamma\|w\|_Y\big]\,\big\|x^{\alpha,\delta}-x^*\big\|_X^2 \le \big(\alpha\|w\|_Y+\delta\big)^2 .$$

Für $\alpha = c\delta$ folgt hieraus $\big\|K(x^{\alpha,\delta}) - y^\delta\big\|_Y \le 2c\delta\|w\|_Y + \delta$ und $c\delta\big[1-\gamma\|w\|_Y\big]\,\big\|x^{\alpha,\delta} - x^*\big\|_X^2 \le \delta^2\big(c\|w\|_Y+1\big)^2$, d.h.

$$\big\|K(x^{\alpha,\delta}) - y\big\|_Y \;\le\; \big\|K(x^{\alpha,\delta}) - y^\delta\big\|_Y \;+\; \|y-y^\delta\|_Y \;\le\; 2\delta(1+c\|w\|_Y) \quad \text{und}$$

$$\big\|x^{\alpha,\delta} - x^*\big\|_X^2 \;\le\; \delta\,\frac{\big(c\,\|w\|_Y+1\big)^2}{c\,\big(1-\gamma\|w\|_Y\big)}$$

Damit ist der Satz bewiesen. □

Bemerkungen: Man kann zeigen:

(a) Unter der Voraussetzung $x^* - \hat{x} \in \mathcal{R}\big(K'(x^*)^* K'(x^*)\big)$ erhält man für $\alpha(\delta) = c\,\delta^{2/3}$ die Konvergenzrate $\big\|x^{\alpha(\delta),\delta} - x^*\big\|_X = \mathcal{O}\big(\delta^{2/3}\big)$.

(b) Auch die Wahl von $\alpha = \alpha(\delta)$ über das Diskrepanzprinzip ist möglich und liefert die Konvergenzrate $\big\|x^{\alpha(\delta),\delta} - x^*\big\|_X = \mathcal{O}\big(\sqrt{\delta}\big)$.

Als generelle Kritik gegenüber der Tikhonovmethode im nichtlinearen Fall wird häufig genannt, dass man *globale* Minimalstellen der nicht-konvexen Funktion $J_{\alpha,\delta}$ bestimmen muss. Dies kann nur durch iterative Verfahren geschehen, die aber im besten Fall nur *lokale* Minima berechnen. Daher scheint es besser zu sein, gleich iterative Verfahren zur Lösung der nichtlinearen Gleichung $K(x) = y^\delta$ zu benutzen. Darauf werden wir im übernächsten Abschnitt eingehen.

5.3 Ein Parameter-Identifizierungsproblem

Sei $f \in L^2(0,1)$ gegeben. Es ist die Aufgabe, $c \in L^2(0,1)$, $c \ge 0$ auf $(0,1)$, in dem Randwertproblem

$$-u''(t) \;+\; c(t)\,u(t) \;=\; f(t)\,,\; 0<t<1\,, \quad u(0)=u(1)=0\,, \tag{4}$$

aus fehlerhaften Messungen $u^\delta(t)$ zu bestimmen. Wir erinnern an den Sobolevraum $H^1(0,0)$ als Vervollständigung von $C^1[0,1]$ bzgl. der Norm

$$\|x\|_{H^1(0,1)} \;=\; \sqrt{\|x\|^2_{L^2(0,1)} + \|x'\|^2_{L^2(0,1)}}\,.$$

Dann haben wir gezeigt, dass $H^1(0,1)$ beschränkt eingebettet ist in $C[0,1]$ (d.h. Funktionen aus $H^1(0,1)$ sind stetig). Außerdem besitzen Funktionen $u \in H^1(0,1)$ Ableitungen $u' \in L^2(0,1)$ in dem Sinn, dass der Hauptsatz der Differential- und Integralrechnung $\int_a^t x'(s)\,ds = x(t) - x(a)$ gilt für alle $a, t \in [0,1]$. Analog wird der Sobolevraum $H^2(0,1)$ definiert als Vervollständigung von $C^2[0,1]$ bzgl der Norm

$$\|x\|_{H^2(0,1)} \;=\; \sqrt{\|x\|^2_{H^1(0,1)} + \|x''\|^2_{L^2(0,1)}}\,.$$

Dann ist $C^2[0,1]$ beschränkt eingebettet in $C^1[0,1]$. Wir wollen zunächst Existenz und Eindeutigkeit der AWA betrachten und suchen eine Lösung in $H^2(0,1)$.

Lemma 5.10 *Es seien $f, c \in L^2(0,1)$. Ist $u \in H^2(0,1)$ Lösung von (4), so löst u die Integralgleichung*

$$u(t) \;+\; \int_0^1 G(t,s)\,c(s)\,u(s)\,ds \;=\; \int_0^1 G(t,s)\,f(s)\,ds\,, \quad t \in [0,1]\,, \tag{5}$$

wobei $G(t,s) = \begin{cases} s(1-t), & 0 \le s \le t \le 1, \\ t(1-s), & 0 \le t \le s \le 1. \end{cases}$

Sei umgekehrt $u \in C[0,1]$ eine Lösung von (5). Dann ist $u \in H^2(0,1)$ und Lösung von (4). Wir bemerken hierzu, dass die rechte Seite der Integralgleichung stetig ist.

Beweis: Sei $u \in H^2(0,1)$ Lösung von (4) und setze $g = f - cu$. Dann ist $g \in L^2(0,1)$ (da u stetig) und $-u'' = g$. Zweimalige Integration und Benutzung der Randbedingungen $u(0) = u(1) = 0$ liefert die Behauptung (Übung).

Sei umgekehrt $u \in C[0,1]$ eine Lösung von (5) und setze wieder $g = f - cu$. Dann ist wieder $g \in L^2(0,1)$, und u ist Lösung von $u(t) = \int_0^1 G(t,s)\,g(s)\,ds$ für $t \in (0,1)$. Wähle eine Folge $g_n \in C[0,1]$ mit $g_n \to g$ in $L^2(0,1)$ und setze $u_n(t) = \int_0^1 G(t,s)\,g_n(s)\,ds$ für $t \in (0,1)$. Dann konvergiert u_n gegen u in $L^2(0,1)$ und es ist

$$u_n(t) \;=\; (1-t)\int_0^t s\,g_n(s)\,ds \;+\; t\int_t^1 (1-s)\,g_n(s)\,ds\,, \quad t \in [0,1]\,.$$

Also ist u_n zweimal stetig differenzierbar, $u_n(0) = u_n(1) = 0$ und (nachrechnen!)

$$u_n'(t) = -\int_0^t s\, g_n(s)\, ds + \int_t^1 (1-s)\, g_n(s)\, ds\,, \quad u_n''(t) = -g_n(t)\,, \quad t \in [0,1]\,.$$

Da g_n gegen g konvergiert in $L^2(0,1)$, so u_n' gegen $-\int_0^t s\, g(s)\, ds + \int_t^1 (1-s)\, g(s)\, ds$ und u_n'' gegen $-g$ in $L^2(0,1)$. Also ist $u \in H^2(0,1)$ und Lösung von $-u'' = g = f - cu$ fast überall. $\square$

Satz 5.11 *Die Integralgleichung (5) und das Randwertproblem (4) haben eindeutige Lösungen für alle $f, c \in L^2(0,1)$, $c \geq 0$ fast überall auf $(0,1)$. Ferner existiert $\alpha > 0$ mit $\|u\|_{H^2(0,1)} \leq \alpha \|f\|_{L^2(0,1)}$, wobei α nicht von f abhängt (sondern nur von c).*

Beweis: Wir definieren den Integraloperator $A : L^2(0,1) \to L^2(0,1)$ durch

$$Au(t) \;=\; \int_0^1 G(t,s)\, u(s)\, ds \;=\; (1-t) \int_0^t s\, u(s)\, ds \;+\; t \int_t^1 (1-s)\, u(s)\, ds\,, \tag{6}$$

$t \in (0,1)$, $u \in L^2(0,1)$. Im Beweis vom vorigen Lemma haben wir die Beschränktheit von A von $L^2(0,1)$ nach $H^2(0,1)$ gezeigt. Wir schreiben die Integralgleichung in der Form

$$u \;+\; A(cu) \;=\; Af \tag{7}$$

Es ist $Af \in H^2(0,1) \subset C[0,1]$, und der Operator $T : u \mapsto A(cu)$ ist kompakt sowohl von $L^2(0,1)$ in sich als auch von $C[0,1]$ in sich (weshalb?). Wir benutzen jetzt folgenden allgemeinen Satz der Funktionalanalysis (Folgerung aus dem 3. Rieszschen Satz, bzw. erster Teil der Fredholmschen Alternative): Hat die Integralgleichung $u + Tu = h$ mit dem kompakten Operator $T : C[0,1] \to C[0,1]$ für $h = 0$ *nur* die triviale Lösung $u = 0$, so ist sie für alle $h \in C[0,1]$ eindeutig lösbar, und die Lösung ist beschränkt invertierbar. In anderen Worten: Ist $I + T$ injektiv, so bijektiv und die Inverse $(I+T)^{-1}$ ist beschränkt. Wir haben also die Injektivität von $I + T$ in $C[0,1]$ zu zeigen, tun dies sogar in $L^2(0,1)$. Sei also $u \in L^2(0,1)$ Lösung von (7) für $h = 0$. Dann ist u Lösung von (5) für $f = 0$, also $u \in H^2(0,1)$ Lösung von (4) für $f = 0$ nach dem letzten Lemma. Multiplikation von (4) mit $u(t)$ und Integration liefert

$$0 \;=\; \int_0^1 \big[-u''(t) + c(t)\, u(t)\big]\, u(t)\, dt \;=\; \int_0^1 \big[u'(t)^2 + c(t)\, u(t)^2\big]\, dt\,,$$

wobei wir hier partiell integriert und ausgenutzt haben, dass u am Rand verschwindet. Da $c \geq 0$ ist, so muss $u'(t) = 0$ sein für alle $t \in [0,1]$. Also ist u konstant und daher Null wegen der Randbedingungen. Daher ist der 3. Rieszsche Satz anwendbar und liefert, dass $I + T$ beschränkt invertierbar ist sowohl als Operator in $C[0,1]$ als auch in $L^2(0,1)$. Wir benutzen diese Eigenschaft hier nur in $C[0,1]$ an und erhalten auch die Existenz von $\alpha_1 > 0$ mit $\|u\|_\infty \leq \alpha_1 \|Af\|_\infty \leq \frac{\alpha_1}{4} \|f\|_{L^2(0,1)}$. Die letzte Ungleichung folgt, da $|G(t,s)| \leq \frac{1}{4}$ für alle $s, t \in [0,1]$. Schließlich folgt aus (7):

$$\begin{aligned}
\|u\|_{H^2(0,1)} &= \|A(f - cu)\|_{H^2(0,1)} \leq \|A\|_{L(L^2(0,1),H^2(0,1))} \|f - cu\|_{L^2(0,1)} \\
&\leq \|A\|_{L(L^2(0,1),H^2(0,1))} \left[\|f\|_{L^2(0,1)} + \|c\|_{L^2(0,1)} \|u\|_\infty\right] \\
&\leq \|A\|_{L(L^2(0,1),H^2(0,1))} \left[\|f\|_{L^2(0,1)} + \frac{\alpha_1}{4} \|c\|_{L^2(0,1)} \|f\|_{L^2(0,1)}\right] \\
&\leq \left(1 + \frac{\alpha_1}{4} \|c\|_{L^2(0,1)}\right) \|A\|_{L(L^2(0,1),H^2(0,1))} \|f\|_{L^2(0,1)} \,.
\end{aligned}$$

□

Jetzt zeigen wir, dass die Abbildung $K : c \mapsto u$ stetig und differenzierbar ist.

Satz 5.12 *Sei $\mathcal{D}(K) = \{c \in L^2(0,1) : c \geq 0$ fast überall$\}$, und $K : \mathcal{D}(K) \to L^2(0,1)$ sei definiert durch $K(c) = u$, wobei $u \in H^2(0,1)$ das Randwertpoblem (4) löst. (Nach dem letzten Satz ist K wohldefiniert.) K ist stetig und sogar differenzierbar in $c^* \in \mathcal{D}(K)$ und $K'(c^*)c = v$, wobei $v \in H^2(0,1)$ das Randwertproblem löst:*

$$-v''(t) \ + \ c^*(t)\, v(t) \ = \ -c(t)\, u^*(t)\,, \ 0 < t < 1\,, \quad v(0) = v(1) = 0\,, \qquad (8)$$

wobei $u^* \in H^2(0,1)$ die Lösung von (4) für c^* ist.

Beweis: Es seien A und $Tu = A(c^*u)$ die Operatoren von Satz 5.11. Zuerst zeigen wir die Stetigkeit: Es sei $c_n \in \mathcal{D}(K)$ eine Folge mit $c_n \to c^*$ in $L^2(0,1)$ und $T_n u = A(c_n u)$ für $u \in C[0,1]$ der zugehörige Operator. Dann gilt wegen $|G(t,s)| \leq 1/4$ für alle $s, t \in [0,1]$:

$$\|(T_n - T)u\|_\infty \ = \ \big\|A\big((c_n - c^*)u\big)\big\|_\infty \ \leq \ \frac{1}{4} \|c_n - c^*\|_{L^2(0,1)} \|u\|_\infty \,,$$

also $\|(T_n - T)u\|_{L(C[0,1],C[0,1])} \le \frac{1}{4}\,\|c_n - c^*\|_{L^2(0,1)}$. Daher konvergieren die Operatoren T_n in der Norm gegen T. Da $I+T$ beschränkt invertierbar ist in $C[0,1]$ (Satz 5.11), so auch $I + T_n$ für hinreichend große n nach einem bekannten Satz der Funktionalanalysis[1]. Ferner gibt es $n_0 \in \mathbb{N}$ mit $\|(I+T_n)^{-1}\|_{L(C[0,1],C[0,1])} \le 2\|(I+T)^{-1}\|_{L(C[0,1],C[0,1])}$ für $n \ge n_0$. Seien jetzt $u^*, u_n \in C[0,1]$ die zugehörigen Lösungen der Integralgleichungen, d.h. $(I+T)u^* = Af$ und $(I+T_n)u_n = Af$. Dann ist $(I+T_n)(u_n - u^*) = (T - T_n)u^*$, also

$$\begin{aligned} \|u_n - u^*\|_\infty &\le \|(I+T_n)^{-1}\|_{L(C[0,1],C[0,1])}\|T_n - T\|_{L(C[0,1],C[0,1])}\,\|u^*\|_\infty \\ &\le 2\|(I+T)^{-1}\|^2_{L(C[0,1],C[0,1])}\|T_n - T\|_{L(C[0,1],C[0,1])}\,\|Af\|_\infty \\ &\le \frac{1}{2}\,\|(I+T)^{-1}\|^2_{L(C[0,1],C[0,1])}\|c_n - c^*\|_{L^2(0,1)}\,\|Af\|_\infty\,, \end{aligned}$$

was die Stetigkeit von K sogar von $\mathcal{D}(K) \subset L^2(0,1)$ nach $C[0,1]$ beweist.

Sei nun $c^*, c \in L^2(0,1)$ mit $c^*, c^* + c \ge 0$ fast überall und $u^*, \tilde{u}$ Lösungen von (4) für c^* bzw. $c^* + c$, d.h. $u^* + A(c^* u^*) = Af$ sowie $\tilde{u} + A\big((c^* + c)\tilde{u}\big) = Af$. Sei v die Lösung von (8). Die Differenz $w = \tilde{u} - u^* - v$ löst das Randwertproblem

$$-w''(t) \;+\; c^*(t)\,w(t) \;=\; c(t)\,\big(u^*(t) - \tilde{u}(t)\big)\,,\ 0 < t < 1\,, \quad w(0) = w(1) = 0\,.$$

Dies ist äquivalent zur Integralgleichung $(I+T)w = A\big(c(u^* - \tilde{u})\big)$ und liefert die Abschätzung

$$\begin{aligned} \|K(c^*+c) - K(c^*) - v\|_\infty &= \|w\|_\infty \;\le\; \|(I+T)^{-1}\|_{L(C[0,1],C[0,1])}\,\|A\big(c(u^* - \tilde{u})\big)\|_\infty \\ &\le \|(I+T)^{-1}\|_{L(C[0,1],C[0,1])}\,\|c\|_{L^2(0,1)}\,\|u^* - \tilde{u}\|_{L^2(0,1)}\,, \end{aligned}$$

also

$$\frac{\|K(c^*+c) - K(c^*) - v\|_{L^2(0,1)}}{\|c\|_{L^2(0,1)}} \;\le\; \alpha\,\|K(c^*) - K(c^*+c)\|_{L^2(0,1)} \tag{9}$$

mit $\alpha = \|(I+T)^{-1}\|_{L(C[0,1],C[0,1])}$, was wegen der gerade bewiesenen Stetigkeit von K die Differenzierbarkeit beweist. $\square$

Wir benötigen für die Voraussetzung (B) den adjungierten Operator.

[1] im Zusammenhang mit der Neumannschen Reihe

Lemma 5.13 *Der adjungierte Operator* $K'(c^*)^* : L^2(0,1) \to L^2(0,1)$ *ist gegeben durch* $K'(c^*)^* w = -u^* y$*, wobei* $u^* = K(c^*)$ *und* $y \in H^2(0,1)$ *das folgende Randwertproblem löst:*

$$-y''(t) \; + \; c^*(t)\, y(t) \; = \; w(t)\,, \; 0 < t < 1\,, \quad y(0) = y(1) = 0\,, \tag{10}$$

Beweis: Es seien $c, w \in L^2(0,1)$ und v die Lösung die Lösung von (8) und y die Lösung von (10). Dann ist mit partieller Integration

$$\begin{aligned} \big(K'(c^*)c, w\big)_{L^2(0,1)} &= \int_0^1 v(t)\,[-y''(t) + c^*(t)\,y(t)]\,dt \\ &= \int_0^1 [-v''(t) + c^*(t)\,v(t)]\,y(t)\,dt \;=\; -\int_0^1 c(t)\,u^*(t)\,y(t)\,dt \\ &= -(c, u^* y)_{L^2(0,1)}\,. \end{aligned}$$

Dies beweist die Behauptung. □

Jetzt können wir die Bedingung (i) von Voraussetzung (B) formulieren:

Die Existenz von $w \in L^2(0,1)$ mit $c^* - \hat{c} = K'(c^*)^* w$ ist äquivalent zur Existenz von $y \in H^2(0,1)$ mit $y(0) = y(1) = 0$ und $c^* - \hat{c} = -u^* y$. Also ist die Bedingung äquivalent zu

$$\frac{c^* - \hat{c}}{u^*} \;\in\; H^2(0,1) \cap H_0^1(0,1)\,.$$

Dies beinhaltet sowohl eine Glattheit von $c^* - \hat{c}$ als auch eine hinreichend starke Randbedingung, da auch u^* am Rand verschwindet.

5.4 Die Nichtlineare Landweberiteration

Wir halten uns hier an die Arbeit

- M. Hanke, A. Neubauer, O. Scherzer: A Convergence Analysis of the Landweber Iteration for Nonlinear Ill-Posed Problems. Numer. Math. 72 (1995), 21–37.

sowie an das Buch

- B. Kaltenbacher, A. Neubauer, O. Scherzer: Iterative Regularization Methods for Nonlinear Ill-Posed Problems. De Gruyter, 2008.

Es seien in diesem Abschnitt wieder X und Y Hilberträume und $K : X \supset \mathcal{D}(K) \to Y$ eine stetig Fréchet-differenzierbare Abbildung mit offenem Definitionsbereich $\mathcal{D}(K)$ sowie $\hat{x} \in X$.

Zunächst erinnern wir noch einmal an das nichtlineare Tikhonovfunktional

$$J_{\alpha,\delta}(x) \;=\; \|K(x) - y^\delta\|_Y^2 \;+\; \alpha\|x - \hat{x}\|_X^2\,, \quad x \in \mathcal{D}(K)\,,$$

für $\alpha, \delta \geq 0$. Dieses Funktional ist differenzierbar:

Lemma 5.14 *$J_{\alpha,\delta}$ ist differenzierbar in jedem $x^* \in \mathcal{D}(K)$ und*

$$J'_{\alpha,\delta}(x^*)h \;=\; 2\big(K(x^*) - y^\delta, K'(x^*)h\big)_Y \;+\; 2\alpha\,(x^* - \hat{x}, h)_X\,, \quad h \in X\,.$$

Beweis: Es ist $K(x^* + h) = K(x^*) + K'(x^*)h + r(h)$ mit $\|r(h)\|_Y / \|h\|_X \to 0$ für $h \to 0$. Also ist mit der binomischen Formel $\|a + b\|^2 = \|a\|^2 + 2(a, b) + \|b\|^2$:

$$\begin{aligned}
J_{\alpha,\delta}(x^* + h) \;&=\; \|K(x^* + h) - y^\delta\|_Y^2 \;+\; \alpha\|x^* + h - \hat{x}\|_X^2 \\
&=\; \|(K(x^*) - y^\delta) + (K'(x^*)h + r(h))\|_Y^2 \;+\; \alpha\|(x^* - \hat{x}) + h\|_X^2 \\
&=\; \|K(x^*) - y^\delta\|_Y^2 + 2\big(K(x^*) - y^\delta, K'(x^*)h + r(h)\big)_Y + \|K'(x^*)h + r(h)\|_Y^2 \\
&\qquad + \alpha\|x^* - \hat{x}\|_X^2 + 2\alpha\,(x^* - \hat{x}, h)_X + \alpha\|h\|_X^2 \\
&=\; J_{\alpha,\delta}(x^*) + 2\big(K(x^*) - y^\delta, K'(x^*)h\big)_Y + 2\alpha\,(x^* - \hat{x}, h)_X + \tilde{r}(h)
\end{aligned}$$

mit

$$\tilde{r}(h) \;=\; 2\big(K(x^*) - y^\delta, r(h)\big)_Y \;+\; \|K'(x^*)h + r(h)\|_Y^2 \;+\; \alpha\|h\|_X^2\,,$$

also $\lim_{h\to 0} \tilde{r}(h)/\|h\|_X = 0$. □

Lemma 5.15 *Ist $x^* \in \mathcal{D}(K)$ lokale Minimalstelle von $J_{\alpha,\delta}$, so ist $J'_{\alpha,\delta}(x^*)h = 0$ für alle $h \in X$, also $K'(x^*)^*\big(K(x^*) - y^\delta\big) + \alpha\,(x^* - \hat{x}) = 0$.*

Beweis: Sei $h \neq 0$ fest. Für hinreichend kleines $t > 0$ ist $x^* + th \in \mathcal{D}(K)$, da $\mathcal{D}(K)$ offen ist. Also gilt für hinreichend kleines $t > 0$ wegen der Optimalität von x^* die Ungleichung $\big[J_{\alpha,\delta}(x^* + th) - J_{\alpha,\delta}(x^*)\big] / \|th\|_X \geq 0$, also

$$0 \;\leq\; \frac{J_{\alpha,\delta}(x^* + th) - J_{\alpha,\delta}(x^*) - J'_{\alpha,\delta}(x^*)(th)}{\|th\|_X} \;+\; \frac{J'_{\alpha,\delta}(x^*)h}{\|h\|_X}$$

und daher für $t \to 0$ die Ungleichung $J'_{\alpha,\delta}(x^*)h \geq 0$. Da dies für alle h gilt, so folgt die Behauptung. $\square$

Die Gleichung

$$K'(x^*)^*\big(K(x^*) - y^\delta\big) + \alpha\,(x^* - \hat{x}) \;=\; 0$$

reduziert sich im Fall, dass K linear (und $\hat{x} = 0$) ist, auf die schon bekannte Normalgleichung der Tikhonovregularisierung, denn im linearen Fall ist $K'(x^*) = K$. Im allgemeinen ist diese Gleichung aber nichtlinear, und man muss iterative Verharen anwenden. Dann kann man das gleich auf die unregularisierte Gleichung, d.h. für $\alpha = 0$, tun. Dies ergibt die Gleichung $K'(x^*)^*\big(K(x^*) - y^\delta\big) = 0$, also

$$x^* \;=\; x^* \;-\; a\,K'(x^*)^*\big(K(x^*) - y^\delta\big)$$

mit einer geeigneten Zahl $a > 0$. Dies ist eine Fixpunktgleichung, und die erste Idee ist es, sie iterativ zu lösen. Sei dazu $\hat{x} \in \mathcal{D}(K)$. Setze $x_0^\delta = \hat{x}$ und

$$x_{k+1}^\delta \;=\; x_k^\delta \;-\; a\,K'(x_k^\delta)^*\big(K(x_k^\delta) - y^\delta\big)\,,\ k = 0,1,2,\ldots$$

Dies ist die nichtlineare Landweberiteration. Noch ist nicht klar, dass sie wohldefiniert ist, d.h. dass die Iterierten im Definitionsbereich $\mathcal{D}(K)$ liegen. Jedenfalls stimmt sie im linearen Fall mit der bekannten Landweberiteration überein, da dann $K'(x_k^\delta) = K$ ist für alle k.

Wir machen die Voraussetzung, dass es $\rho > 0$ gibt mit $a\|K'(x)\|^2_{L(X,Y)} \leq 1$ für alle $x \in B(\hat{x},\rho)$. Skalieren wir die Gleichung $K(x) = y^\delta$, d.h. ersetzen wir sie durch die äquivalente Gleichung $\tilde{K}(x) = \tilde{y}^\delta$ mit $\tilde{K} = \sqrt{a}\,K$ und $\tilde{y}^\delta = \sqrt{a}\,y^\delta$, so ist $\|\tilde{K}'(x)\|^2_{L(X,Y)} \leq 1$ für alle $x \in B(\hat{x},\rho)$, und die Landweberiteration hat die Form

$$x_{k+1}^\delta \;=\; x_k^\delta \;-\; \tilde{K}'(x_k^\delta)^*\big(\tilde{K}(x_k^\delta) - \tilde{y}^\delta\big)\,,\ k = 0,1,2,\ldots$$

Daher können wir ohne Einschränkung annehmen, dass $a = 1$ ist. Die Landweberiteration hat also die Form:

$$x_0^\delta = \hat{x}\,,\quad x_{k+1}^\delta \;=\; x_k^\delta \;-\; K'(x_k^\delta)^*\big(K(x_k^\delta) - y^\delta\big)\,,\ k = 0,1,2,\ldots \tag{11}$$

Voraussetzung (V1):

Es gebe $\rho > 0$ und $\eta < \frac{1}{2}$ mit $B(\hat{x}, \rho) \subset \mathcal{D}(K)$ und $\|K'(x)\|_{L(X,Y)} \leq 1$ und

$$\big\|K(x) - K(\tilde{x}) - K'(\tilde{x})(x - \tilde{x})\big\|_Y \;\leq\; \eta \,\big\|K(\tilde{x}) - K(x)\big\|_Y \tag{12}$$

für alle $x, \tilde{x} \in \mathcal{D}(K)$.

Diese Bedingung (12) heißt „tangentiale Kegelbedingung“, im Englischen „Tangential Cone Condition“. Definiert man nämlich die Menge

$$\begin{aligned} C_{\tilde{x}} \;&:=\; \Big\{(x, y) \in X \times Y : \|y - K(\tilde{x}) - K'(\tilde{x})(x - \tilde{x})\big\|_Y \;\leq\; \eta \,\big\|y - K(\tilde{x})\big\|_Y\Big\} \\ &=\; \big(\tilde{x}, K(\tilde{x})\big) \;+\; \Big\{(z, v) \in X \times Y : \|v - K'(\tilde{x})z\big\|_Y \;\leq\; \eta \,\|v\|_Y\Big\}\,, \end{aligned}$$

so lautet die Bedingung, dass $\Big\{\big(x, K(x)\big) : x \in B(\hat{x}, \rho)\Big\} \subset C_{\tilde{x}}$ für alle $\tilde{x} \in B(\hat{x}, \rho)$. Die Menge $C_{\tilde{x}}$ ist ein Kegel mit Spitze in $\big(\tilde{x}, K(\tilde{x})\big)$. Man skizziere die Situation für $X = Y = \mathbb{R}$.

Wir vergleichen diese Bedingung mit der Abschätzung

$$\big\|K(\tilde{x}) - K(x) - K'(x)(\tilde{x} - x)\big\|_Y \;\leq\; c\,\|\tilde{x} - x\|_X^2\,,$$

die für Lipschitzstetig differenzierbare Funktionen gilt (vgl. Lemma 5.7). Existiert also c' mit $\|\tilde{x} - x\|_X \leq c'\big\|K(\tilde{x}) - K(x)\big\|_Y$, so folgt (12). Solch eine Abschätzung gilt aber gerade nicht für schlecht gestellte Probleme. Die Bedingung (12) ist sehr stark, kann aber in einigen Anwendungen nachgewiesen werden (siehe Beispiel unten). Zunächst ziehen wir Folgerungen aus dieser Bedingung.

Lemma 5.16 *Unter der Voraussetzung (V1) gilt:*

$$\frac{1}{1+\eta}\,\big\|K'(x)(\tilde{x} - x)\big\|_Y \;\leq\; \big\|K(\tilde{x}) - K(x)\big\|_Y \;\leq\; \frac{1}{1-\eta}\,\big\|K'(x)(\tilde{x} - x)\big\|_Y \tag{13}$$

für alle $x, \tilde{x} \in B(\hat{x}, \rho)$.

Beweis: (a) Die rechte Ungleichung folgt aus

$$\begin{aligned} \big\|K(\tilde{x}) - K(x)\big\|_Y \;&\leq\; \big\|K(\tilde{x}) - K(x) - K'(x)(\tilde{x} - x)\big\|_Y \;+\; \big\|K'(x)(\tilde{x} - x)\big\|_Y \\ &\leq\; \eta\,\big\|K(\tilde{x}) - K(x)\big\|_Y \;+\; \big\|K'(x)(\tilde{x} - x)\big\|_Y\,, \end{aligned}$$

die linke aus

$$\begin{aligned}\big\|K(\tilde{x})-K(x)\big\|_Y &= \big\|K'(x)(\tilde{x}-x)-\big[K'(x)(\tilde{x}-x)-K(\tilde{x})+K(x)\big]\big\|_Y \\ &\geq \big\|K'(x)(\tilde{x}-x)\big\|_Y - \eta\big\|K(\tilde{x})-K(x)\big\|_Y .\end{aligned}$$

□

Aus diesem Lemma folgt sofort die Umkehrung von Satz 5.8.

Lemma 5.17 *Es gelte die Voraussetzung (V1) und es sei $\rho > 0$ und $x^* \in B(\hat{x},\rho)$ mit $K(x^*) = y$. Ferner sei die lineare Gleichung $K'(x^*)h = z$ schlecht gestellt. Dann ist die nichtlineare Gleichung $K(x) = y$ in x^* lokal schlecht gestellt.*

Beweis: Sei $r \in (0,\rho)$. Da die lineare Gleichung schlecht gestellt ist, existiert eine Folge $h_n \in X$ mit $\|h_n\|_X = 1$ und $\|K'(x^*)h_n\|_Y \to 0$ für $n \to \infty$. Wir setzen $x_n = x^* + rh_n$. Dann folgt für $\tilde{x} = x^*$ und $x = x_n$ in (12), dass $K(x_n) \to K(x^*)$ und $\|x_n - x^*\|_X = r$. □

Satz 5.18 *Es gelte die Voraussetzung (V1) und es sei $\rho > 0$ und $x^* \in B(\hat{x},\rho/2)$ mit $K(x^*) = y$. Wir setzen ferner voraus, dass es $\tau > \frac{2(1+\eta)}{1-2\eta}$ und $k_* \in \mathbb{N}$ gibt mit $x_k^\delta \in B(x^*,\rho/2)$ für alle $k = 0,\ldots,k_*-1$ und*

$$\|K(x_k^\delta)-y^\delta\|_Y \;\geq\; \tau\delta \quad \text{für alle } k = 0,\ldots,k_*-1\,. \tag{14}$$

Dann gilt:

(a) $\|x_{k+1}^\delta - x^*\|_X \leq \|x_k^\delta - x^*\|_X$ für alle $k = 0,\ldots,k_*-1$. Insbesondere gilt dann auch $\|x_{k_*}^\delta - x^*\|_X \leq \|x_0^\delta - x^*\|_X = \|\hat{x}-x^*\|_X < \rho/2$, d.h. $x_k^\delta \in B(x^*,\rho/2) \subset B(\hat{x},\rho)$ für alle $k = 0,\ldots,k_*$.

(b) Es gilt für alle $\ell \in \{0,\ldots,k_*-1\}$:

$$\frac{(1-2\eta)\tau-2(1+\eta)}{\tau}\sum_{k=\ell}^{k_*-1}\big\|K(x_k^\delta)-y^\delta\big\|_Y^2 \;\leq\; \|x_\ell^\delta-x^*\|_X^2 - \|x_{k_*}^\delta-x^*\|_X^2\,. \tag{15}$$

(c) Ist $\delta = 0$, so ist (14) für alle $k \in \mathbb{N}$ erfüllt, und es gilt (wir schreiben x_k an Stelle von x_k^0)

$$\sum_{k=0}^{\infty} \big\| K(x_k) - y \big\|_Y^2 \;\le\; \frac{\|\hat{x} - x^*\|_X^2}{1 - 2\eta}\,. \tag{16}$$

Insbesondere konvergiert $K(x_k)$ gegen y.

Beweis: Es ist wegen $\big\| K'(x_k^\delta)^* \big\|_{L(Y,X)} \le 1$

$$\begin{aligned}
&\|x_{k+1}^\delta - x^*\|_X^2 - \|x_k^\delta - x^*\|_X^2 \\
&\quad= \; 2\,(x_k^\delta - x^*, x_{k+1}^\delta - x_k^\delta)_X \;+\; \|x_{k+1}^\delta - x_k^\delta\|_X^2 \\
&\quad= \; 2\,\big(x_k^\delta - x^*, K'(x_k^\delta)^*(y^\delta - K(x_k^\delta))\big)_X \;+\; \big\|K'(x_k^\delta)^*(y^\delta - K(x_k^\delta))\big\|_X^2 \\
&\quad\le \; 2\,\big(K'(x_k^\delta)(x_k^\delta - x^*), y^\delta - K(x_k^\delta)\big)_X \;+\; \big\|y^\delta - K(x_k^\delta)\big\|_Y^2 \\
&\quad= \; 2\,\big(K'(x_k^\delta)(x_k^\delta - x^*) - K(x_k^\delta) + y^\delta, y^\delta - K(x_k^\delta)\big)_X \;-\; \big\|y^\delta - K(x_k^\delta)\big\|_Y^2 \\
&\quad\le \; \big[2\eta\,\big\|K(x_k^\delta) - y\big\|_Y + 2\delta - \big\|K(x_k^\delta) - y^\delta\big\|_Y\big]\,\big\|K(x_k^\delta) - y^\delta\big\|_Y \\
&\quad\le \; \big[(2\eta - 1)\,\big\|K(x_k^\delta) - y^\delta\big\|_Y + 2\delta + 2\eta\delta\big]\,\big\|K(x_k^\delta) - y^\delta\big\|_Y \\
&\quad= \; \big[2\delta(1+\eta) - (1 - 2\eta)\,\big\|K(x_k^\delta) - y^\delta\big\|_Y\big]\,\big\|K(x_k^\delta) - y^\delta\big\|_Y \qquad (*) \\
&\quad= \; \frac{1}{\tau}\,\big[2(1+\eta)\delta\tau - (1-2\eta)\tau\,\big\|K(x_k^\delta) - y^\delta\big\|_Y\big]\,\big\|K(x_k^\delta) - y^\delta\big\|_Y \\
&\quad\le \; \frac{2(1+\eta) - (1-2\eta)\tau}{\tau}\,\big\|K(x_k^\delta) - y^\delta\big\|_Y^2\,.
\end{aligned}$$

Hieraus folgt (a), da $2(1+\eta) - (1-2\eta)\tau < 0$ nach Wahl von τ.

Mit $\alpha = \big[(1-2\eta)\tau - 2(1+\eta)\big]/\tau > 0$ haben wir gerade die Ungleichung gezeigt:

$$\alpha\,\big\|K(x_k^\delta) - y^\delta\big\|_Y^2 \;\le\; \|x_k^\delta - x^*\|_X^2 \;-\; \|x_{k+1}^\delta - x^*\|_X^2\,, \quad k = 0, \dots, k_* - 1\,,$$

und daher für $\ell \in \{0, k_* - 1\}$:

$$\alpha \sum_{k=\ell}^{k_*-1} \big\|K(x_k^\delta) - y^\delta\big\|_Y^2 \;\le\; \|x_\ell^\delta - x^*\|_X^2 \;-\; \|x_{k_*}^\delta - x^*\|_X^2\,.$$

(c) Wir betrachten die obige Abschätzung bis $(*)$ für $\delta = 0$, d.h.

$$(1 - 2\eta)\,\big\|K(x_k) - y\big\|_Y^2 \;\le\; \|x_k - x^*\|_X^2 \;-\; \|x_{k+1} - x^*\|_X^2\,,$$

und daher für alle $m \in \mathbb{N}$

$$(1 - 2\eta) \sum_{k=0}^{m-1} \big\|K(x_k) - y\big\|_Y^2 \;\le\; \|x_0 - x^*\|_X^2 \;-\; \|x_m - x^*\|_X^2 \;\le\; \|x_0 - x^*\|_X^2\,.$$

Daher konvergiert die Reihe, und es gilt die behauptete Abschätzung. $\square$

Sei jetzt $\delta > 0$ und wieder $\tau > \frac{2(1+\eta)}{1-2\eta}$. Die Bedingung (14) kann nicht für alle k_* gelten. In der Tat, sonst wäre $\left(\|x_k - x^*\|_X^2\right)_{k=0}^\infty$ eine monoton fallende und nach unten beschränkte Folge, die also konvergieren würde. Aus (15) für $\ell = k_* - 1$ würde folgen, dass

$$\frac{(1-2\eta)\tau - 2(1+\eta)}{\tau} \left\|K(x_{k^*-1}^\delta) - y^\delta\right\|_Y^2 \;\leq\; \|x_{k^*-1}^\delta - x^*\|_X^2 \;-\; \|x_{k_*}^\delta - x^*\|_X^2 \;\longrightarrow\; 0$$

für $k_* \to \infty$, ein Widerspruch. Daher ist das folgende Abbruchkriterium wohldefiniert:

Stopregel: Sei $\delta > 0$ und $\tau > \frac{2(1+\eta)}{1-2\eta}$. Wir definieren $k_* = k_*(\delta)$ als die eindeutig bestimmte natürliche Zahl mit

$$\|K(x_k^\delta) - y^\delta\|_Y \;\geq\; \tau\delta \;>\; \|K(x_{k^*}^\delta) - y^\delta\|_Y \quad \text{für alle } k = 0, \ldots, k_* - 1\,. \tag{17}$$

Mit dieser Stopregel können wir Konvergenz zeigen für $\delta \to 0$.

Bemerkung: Sei $k \in \mathbb{N}$ festgehalten und es gelte $k \leq k_*(\delta)$ für alle $\delta > 0$. Dann ist die Abbildung $y^\delta \mapsto x_k^\delta$ wohldefiniert und stetig. Es werden nämlich nur k Schritte des Landweberverfahrens durchgeführt und in jedem Schritt stetige Operationen vorgenommen. Insbesondere ist $\lim_{\delta\to 0} x_k^\delta = x_k$, wobei x_k der $k-$te Iterationsschitt im Landweberverfahren für $\delta = 0$ ist.

Satz 5.19 *Es gelte die Voraussetzung (V1) und es sei $\rho > 0$ und $x^* \in B(\hat{x}, \rho/2)$ mit $K(x^*) = y$. Für $\delta = 0$ ist die Landweberiteration wohldefiniert (d.h. alle Iterierten x_k liegen in $\mathcal{D}(K)$) und die Folge $(x_m)_m$ konvergiert gegen eine Lösung $\tilde{x} \in B[x^*, \rho/2] \subset B(\hat{x}, \rho)$ von $K(x) = y$.*

Beweis: Wir haben schon in Satz 5.18 gesehen, dass alle Folgenglieder in $B(x^*, \rho/2)$ liegen. Wir werden jetzt zunächst zeigen, dass $(x_m)_m$ eine Cauchyfolge ist. Sei $\ell \leq m$ festgehalten. Bestimme dazu ein k mit $\ell \leq k \leq m$ und

$$\|K(x_k) - y\|_Y \;\leq\; \|K(x_i) - y\|_Y \quad \text{für alle } \ell \leq i \leq m\,.$$

Wegen $\|x_\ell - x_m\|^2 \leq 2\|x_\ell - x_k\|^2 + 2\|x_k - x_m\|^2$ schätzen wir beide Terme ab und benutzen dazu die Formel $\|u - v\|^2 + \|u\|^2 - \|v\|^2 = 2(u, u - v)$ sowie die linke Abschätzung von Lemma 5.16, (a).

$$\begin{aligned}
&\|x_k - x_m\|_X^2 + \|x_k - x^*\|_X^2 - \|x_m - x^*\|_X^2 \\
&= 2\big(x_k - x^*, x_k - x_m\big)_X \;=\; 2\sum_{i=k}^{m-1}\big(x_k - x^*, x_i - x_{i+1}\big)_X \\
&= 2\sum_{i=k}^{m-1}\big(x_k - x^*, K'(x_i)^*(K(x_i) - y)\big)_X \;=\; 2\sum_{i=k}^{m-1}\big(K'(x_i)(x_k - x^*), K(x_i) - y\big)_Y \\
&\leq 2\sum_{i=k}^{m-1} \|K(x_i) - y\|_Y \, \|K'(x_i)(x_k - x^*)\|_Y \\
&\leq 2\sum_{i=k}^{m-1} \|K(x_i) - y\|_Y \, \big[\|K'(x_i)(x_i - x^*)\|_Y + \|K'(x_i)(x_i - x_k)\|_Y\big] \\
&\leq 2(1+\eta)\sum_{i=k}^{m-1} \|K(x_i) - y\|_Y \, \big[\|K(x_i) - K(x^*)\|_Y + \|K(x_i) - K(x_k)\|_Y\big] \\
&\leq 2(1+\eta)\sum_{i=k}^{m-1} \|K(x_i) - y\|_Y \, \big[\|K(x_i) - y)\|_Y + \|K(x_i) - y\|_Y + \|y - K(x_k)\|_Y\big] \\
&\leq 2(1+\eta)\sum_{i=k}^{m-1} \|K(x_i) - y\|_Y \, \big[\|K(x_i) - y)\|_Y + \|K(x_i) - y\|_Y + \|y - K(x_i)\|_Y\big] \\
&= 6(1+\eta)\sum_{i=k}^{m-1} \|K(x_i) - y\|_Y^2 \,.
\end{aligned}$$

Ganz analog schätzt man ab:

$$\|x_k - x_\ell\|_X^2 + \|x_k - x^*\|_X^2 - \|x_\ell - x^*\|_X^2 \;\leq\; 6(1+\eta)\sum_{i=\ell}^{k-1} \|K(x_i) - y\|_Y^2$$

und daher

$$\begin{aligned}
\|x_\ell - x_m\|^2 &\leq 2\|x_\ell - x_k\|^2 + 2\|x_k - x_m\|^2 \\
&\leq 2\|x_\ell - x^*\|_X^2 - 2\|x_k - x^*\|_X^2 \;+\; 12(1+\eta)\sum_{i=\ell}^{k-1} \|K(x_i) - y\|_Y^2 \\
&\quad + 2\|x_m - x^*\|_X^2 - 2\|x_k - x^*\|_X^2 \;+\; 12(1+\eta)\sum_{i=k}^{m-1} \|K(x_i) - y\|_Y^2 \\
&= 2\|x_\ell - x^*\|_X^2 + 2\|x_m - x^*\|_X^2 - 4\|x_k - x^*\|_X^2 \\
&\quad + 12(1+\eta)\sum_{i=\ell}^{m-1} \|K(x_i) - y\|_Y^2 \,.
\end{aligned}$$

Da sowohl die Folge $\big(\|x_\ell - x^*\|_X\big)_\ell$ als auch die Reihe $\sum_{i=0}^\infty \|K(x_i) - y\|_Y^2$ nach Satz 5.18 konvergieren, so konvergiert $\|x_\ell - x_m\|^2$ gegen Null für $\ell \to \infty$. Daher ist (x_m) eine Cauchyfolge und daher konvergent: $x_m \to \tilde{x}$ mit $\tilde{x} \in B[x^*, \rho/2]$. Natürlich ist $K(\tilde{x}) = y$. □

Satz 5.20 *Es gelte die Voraussetzung (V1) und es sei $\rho > 0$ und $x^* \in B(\hat{x}, \rho/2)$ mit $K(x^*) = y$ sowie $\tau > \frac{2(1+\eta)}{1-2\eta}$. Die Folge breche mit der Stopregel ab und definiert dadurch $k_*(\delta)$ für alle $\delta > 0$. Dann gilt $\lim\limits_{\delta \to 0} x^\delta_{k_*(\delta)} = \tilde{x}$, wobei $\tilde{x} \in B[x^*, \rho/2]$ der Grenzwert der Folge für $\delta = 0$ ist. Dieser existiert ja nach dem letzten Satz und ist Lösung der Gleichung $K(\tilde{x}) = y$.*

Beweis: Es sei $(x_k)_k$ die von der Landweberiteration erzeugte Folge für $\delta = 0$ mit $\tilde{x}$ als Grenzwert und sei $(\delta_n)_n$ eine Nullfolge. Setze zur Abkürzung $k_n = k_*(\delta_n)$. Wir unterscheiden zwei Fälle.

1. Fall: Die Folge $(k_n)_n$ von natürlichen Zahlen habe einen endlichen Häufungspunkt k, d.h. $k_n = k$ für alle $n \in I$, wobei $I \subset \mathbb{N}$ unendliche Teilmenge ist. Nach der Definition von k_n gilt

$$\big\|K(x_k^{\delta_n}) - y^{\delta_n}\big\| \;<\; \tau\,\delta_n \quad \text{für alle } n \in I\,.$$

Da für dieses feste k die Iteration $x_k^{\delta_n}$ stetig von δ_n abhängt und y^{δ_n} gegen y konvergiert, so folgt für $n \in I$, $n \to \infty$, dass $x_k^{\delta_n} \to x_k$ (dies ist die k-te Iteration der Folge für $\delta = 0$) und $K(x_k) = y$. Die Landweberiteration für $\delta = 0$ impliziert aber, dass dann $x_m = x_k$ für alle $m \geq k$, also konstant, ist. Insbesondere ist $x_k = \tilde{x}$.

2. Fall: Die Folge $(k_n)_n$ konvergiert gegen unendlich. Ohne Beschränkung der Allgemeinheit nehmen wir an, dass k_n monoton gegen unendlich konvergiert. Sei $n > m$. Wir können Satz 5.18 für $\tilde{x}$ statt x^* anwenden und erhalten

$$\begin{aligned}\|x_{k_n}^{\delta_n} - \tilde{x}\|_X \;&\leq\; \|x_{k_{n-1}}^{\delta_n} - \tilde{x}\|_X \;\leq\; \cdots \;\leq\; \|x_{k_m}^{\delta_n} - \tilde{x}\|_X \\ &\leq\; \|x_{k_m}^{\delta_n} - x_{k_m}\|_X \;+\; \|x_{k_m} - \tilde{x}\|_X\,.\end{aligned}$$

Sei $\varepsilon > 0$. Wähle m so groß, dass $\|x_{k_m} - \tilde{x}\|_X \leq \varepsilon/2$. Für dieses feste m konvergiert $x_{k_m}^{\delta_n}$ gegen x_{k_m} für $n \to \infty$. Daher können wir n_0 finden mit $\|x_{k_m}^{\delta_n} - x_{k_m}\|_X \leq \varepsilon/2$ für alle $n \geq n_0$. Für diese $n \geq n_0$ gilt also $\|x_{k_n}^{\delta_n} - \tilde{x}\|_X \leq \varepsilon$. Damit haben wir Konvergenz jeder Folge $\delta_n \to 0$ gegen den selben Grenzwert $\tilde{x}$ gezeigt. $\square$

Bevor wir zu Konvergenzordnungen kommen, wollen wir noch etwas zur lokalen Eindeutigkeit sagen.

Lemma 5.21 *Es gelte die Voraussetzung (V1).*

(a) Die Menge $M^* := \{x \in B(\hat{x}, \rho) : K(x) = y\}$ aller Lösungen von $K(x) = y$ in der Kugel $B(\hat{x}, \rho)$ sei nicht leer. Dann gibt es ein eindeutig bestimmes $x^\dagger \in M^*$ mit minimalem Abstand zu $\hat{x}$.

(b) Es sei $\mathcal{N}\big(K'(x^\dagger)\big) \subset \mathcal{N}\big(K'(x)\big)$ für alle $x \in B(\hat{x}, \rho)$. Dann konvergieren die Folgen $(x_k)_k$ und $\big(x_{k(\delta_n)}^{\delta_n}\big)_n$ der Landweberiteration (siehe Sätze 5.19 und 5.20) gegen $x^\dagger$.

Beweis: (a) Sei $x^* \in M^*$. Aus (13) für $x = x^*$ folgt, dass ein $\tilde{x} \in B(\hat{x}, \rho)$ genau dann eine Lösung von $K(x) = y$ in $B(\hat{x}, \rho)$ ist, wenn $K'(x^*)(x^* - \tilde{x}) = 0$, d.h. wenn $x^* - \tilde{x} \in \mathcal{N}\big(K'(x^*)\big)$. Daher ist $M^* = B(\hat{x}, \rho) \cap \big[x^* + \mathcal{N}\big(K'(x^*)\big)\big]$. Der Unterraum $\mathcal{N}\big(K'(x^*)\big)$ ist abgeschlossen. Daher existiert eine e indeutig bestimmte beste Approximation p von $x^* - \hat{x}$ in $\mathcal{N}\big(K'(x^*)\big)$ (klar?). Dann ist $x^\dagger = x^* + p \in x^* + \mathcal{N}\big(K'(x^*)\big)$ beste Approximation an $\hat{x}$ in $x^* + \mathcal{N}\big(K'(x^*)\big)$, denn $x^\dagger - \hat{x} \perp \mathcal{N}\big(K'(x^*)\big)$. Es ist $\|x^\dagger - \hat{x}\|_X \leq \|x^* - \hat{x}\|_X < \rho$, also $x^\dagger \in M^*$.

(b) Wir zeigen induktiv, dass $x_k^\delta - x^\dagger \perp \mathcal{N}\big(K'(x^\dagger)\big)$ für alle $k = 0, \ldots, k_*(\delta)$ (wobei $k_*(\delta) = \infty$ gesetzt wird für $\delta = 0$). Für $k = 0$ gilt wegen $x_0^\delta = \hat{x}$ und der Definition von $x^\dagger$. Sei dies also für k richtig und sei $z \in \mathcal{N}\big(K'(x^\dagger)\big)$. Dann ist auch $z \in \mathcal{N}\big(K'(x_k)\big)$, also

$$(x_{k+1}^\delta - x^\dagger, z)_X \;=\; (x_k^\delta - x^\dagger, z)_X \;+\; \big(K(x_k^\delta) - y^\delta, K'(x_k)z\big)_Y \;=\; 0\,.$$

Sei nun $\tilde{x} = \lim_{k\to\infty} x_k$ (im Fall $\delta = 0$, sonst analog). Da das orthogonale Komplement $\mathcal{N}\big(K'(x^\dagger)\big)$ abgeschlossen ist, so gilt auch $\tilde{x} - x^\dagger \perp \mathcal{N}\big(K'(x^\dagger)\big)$. Außerdem ist natürlich $\tilde{x} - x^\dagger \in \mathcal{N}\big(K'(x^\dagger)\big)$, da beides Lösungen von $K(x) = y$ sind. Also ist $\tilde{x} = x^\dagger$. $\square$

Wir wollen jetzt eine Konvergenzrate beweisen und benötigen natürlich eine „source condition“ wie im linearen Fall oder der nichtlinearen Tikhonovregularisierung. Das reicht aber nicht aus, sondern wir benötigen noch eine Verschärfung von (V1).

Voraussetzung (V2):

Sei (V1) erfüllt, die Gleichung $K(x) = y$ sei in $B(\hat{x}, \rho)$ lösbar und $x^\dagger \in B(\hat{x}, \rho)$ sei die Minimumnormlösung an $\hat{x}$ von $K(x) = y$. (Diese existiert nach Lemma 5.21). Ferner sei $K'(x^\dagger)$ kompakt und es gebe $C > 0$ und eine Schar $\{R_x : x \in B(\hat{x}, \rho)\}$ von linearen beschränkten Operatoren $R_x : Y \to Y$ mit

$$K'(x) = R_x K'(x^\dagger) \quad \text{und} \quad \|R_x - I\|_{L(Y,Y)} \leq C \|x - x^*\|_X \quad \text{für } x \in B(\hat{x}, \rho)\,.$$

Im linearen Fall sind die beiden Bedingungen natürlich für $R_x = I$ erfüllt, da $K'(x) = K$ für alle x.

Lemma 5.22 *Unter der Voraussetzung (V2) gilt $\mathcal{N}\big(K'(x^\dagger)\big) \subset \mathcal{N}\big(K'(x)\big)$ für alle $x \in B(\hat{x}, \rho)$ und*

$$\big\|K(x) - K(x^\dagger) - K'(x^\dagger)(x - x^\dagger)\big\|_Y \leq \frac{C}{2} \|x - x^\dagger\|_X \big\|K'(x^\dagger)(x - x^\dagger)\big\|_Y\,, \tag{18}$$

$$\big\|K(x) - K(x^\dagger) - K'(x^\dagger)(x - x^\dagger)\big\|_Y \leq \frac{C}{2(1-\rho C)} \|x - x^\dagger\|_X \big\|K(x) - K(x^\dagger)\big\|_Y \tag{19}$$

für alle $x \in B(\hat{x}, \rho)$, wobei ρ so klein ist, dass $1 - \rho C > 0$. Wählen wir also ρ so klein, dass $\eta := \frac{C\rho}{2-\rho C} < 1/2$, so ist die Bedingung (12) für $x = x^\dagger$ erfüllt.

Beweis: Die erste Aussage ist trivial. Für die zweite wählen wir ein beliebiges $z \in Y$ und schreiben

$$\begin{aligned}
&\big(K(x) - K(x^\dagger) - K'(x^\dagger)(x - x^\dagger), z\big)_Y \\
&= \int_0^1 \left[\frac{d}{dt}\big(K\big(tx + (1-t)x^\dagger\big), z\big)_Y - \big(K'(x^\dagger)(x - x^\dagger), z\big)_Y\right] dt \\
&= \int_0^1 \big(\big[K'\big(tx + (1-t)x^\dagger\big) - K'(x^\dagger)\big](x - x^\dagger), z\big)_Y\, dt \\
&= \int_0^1 \big((R_{tx+(1-t)x^\dagger} - I)K'(x^\dagger)(x - x^\dagger), z\big)_Y\, dt \\
&\leq C \int_0^1 t\, dt\, \|x - x^\dagger\|_X \big\|K'(x^\dagger)(x - x^\dagger)\big\|_Y \|z\|_Y \\
&= \frac{C}{2} \|x - x^\dagger\|_X \big\|K'(x^\dagger)(x - x^\dagger)\big\|_Y \|z\|_Y
\end{aligned}$$

und daher (setze $z = K(x) - K(x^\dagger) - K'(x^\dagger)(x - x^\dagger)$)

$$
\begin{aligned}
& \left\|K(x) - K(x^\dagger) - K'(x^\dagger)(x - x^\dagger)\right\|_Y \\
\leq\ & \frac{C}{2}\,\|x - x^\dagger\|_X \left\|K'(x^\dagger)(x - x^\dagger)\right\|_Y \quad \text{und weiter} \\
\leq\ & \frac{C}{2}\,\|x - x^\dagger\|_X \left\|K'(x^\dagger)(x - x^\dagger) + K(x^\dagger) - K(x)\right\|_Y \\
& + \frac{C}{2}\,\|x - x^\dagger\|_X \left\|K(x^\dagger) - K(x)\right\|_Y \\
\leq\ & C\rho \left\|K'(x^\dagger)(x - x^\dagger) + K(x^\dagger) - K(x)\right\|_Y \\
& + \frac{C}{2}\,\|x - x^\dagger\|_X \left\|K(x^\dagger) - K(x)\right\|_Y
\end{aligned}
$$

und daher

$$
(1 - C\rho) \left\|K(x) - K(x^\dagger) - K'(x^\dagger)(x - x^\dagger)\right\|_Y \;\leq\; \frac{C}{2}\,\|x - x^\dagger\|_X \left\|K(x^\dagger) - K(x)\right\|_Y .
$$

Dies beweist die Abschätzung (19). □

Unter der Voraussetzung (V2) wissen aus den letzten beiden Lemmas, dass $\left(x^{\delta_n}_{k(\delta_n)}\right)_n$ gegen $x^\dagger$ konvergiert. Für den Beweis der Fehlerabschätzung benötigen wir die folgenden Abschätzungen:

Lemma 5.23 *Sei $k \in \mathbb{N}$ und $x \in (0,1]$ und $\alpha \in (0,1]$. Dann gilt:*

$$
\begin{aligned}
\frac{1 - (1 - x^2)^k}{x} &\leq \sqrt{k}\,, && (20) \\
x\,(1 - x^2)^k &\leq \frac{1}{\sqrt{k+1}}\,, && (21) \\
x^2\,(1 - x^2)^k &\leq \frac{1}{k+1}\,, && (22) \\
\sum_{j=0}^{k-1} (j+1)^{-\alpha}(k-j)^{-3/2} &\leq \frac{c(\alpha)}{(k+1)^\alpha}\,, && (23)
\end{aligned}
$$

wobei die Konstante $c(\alpha)$ nur von α abhängt.

Beweis: (20) Sei $0 < x \leq 1/\sqrt{k}$. Die Bernoullische Ungleichung besagt $(1 - x^2)^k \geq 1 - kx^2$, also $\frac{1-(1-x^2)^k}{x} \leq \frac{1-(1-kx^2)}{x} = kx \leq \sqrt{k}$. Sei jetzt $x \geq 1/\sqrt{k}$. Dann ist $\frac{1-(1-x^2)^k}{x} \leq \frac{1}{x} \leq \sqrt{k}$.

(21) und (22) sind schon früher bewiesen worden (und verbleiben als Übung).

(23) Definiere die Funktion $f(x) = (x+1)^{-\alpha}(k-x)^{-3/2}$ für $-1 < x < k$. Dann rechnet man die zweite Ableitung aus (Übung) und erkennt, dass $f''(x) \geq 0$ für alle $-1 < x < k$.[2] Also ist mit der Taylorformel

$$f(x) \;=\; f(j) \;+\; f'(j)(x-j) \;+\; \frac{1}{2} f''(z_{j,x})(x-j)^2 \;\geq\; f(j) \;+\; f'(j)(x-j)\,,$$

für $j - 1/2 \leq x \leq j + 1/2$ mit einem Zwischenpunkt $z_{j,x}$, also

$$\int\limits_{j-1/2}^{j+1/2} f(x)\,dx \;\geq\; \int\limits_{j-1/2}^{j+1/2} \big[f(j) + f'(j)(x-j)\big]\,dx \;=\; f(j)\,,$$

da das Integral $\int\limits_{j-1/2}^{j+1/2} (x-j)\,dx$ wegen der Symmetrie des Intervalls verschwindet. Damit ist

$$\begin{aligned}\sum_{j=0}^{k-1}(j+1)^{-\alpha}(k-j)^{-3/2} \;&=\; \sum_{j=0}^{k-1} f(j) \;\leq\; \sum_{j=0}^{k-1}\int\limits_{j-1/2}^{j+1/2} f(x)\,dx \;=\; \int\limits_{-1/2}^{k-1/2} f(x)\,dx \\ &=\; \int\limits_{-1/2}^{k/2} f(x)\,dx \;+\; \int\limits_{k/2}^{k-1/2} f(x)\,dx\,.\end{aligned}$$

Es ist für das erste Integral:

$$\begin{aligned}\int\limits_{-1/2}^{k/2} f(x)\,dx \;&\leq\; \left(\frac{2}{k}\right)^{3/2} \int\limits_{-1/2}^{k/2} (x+1)^{-\alpha}\,dx \;\leq\; \left(\frac{2}{k}\right)^{3/2} \frac{1}{1-\alpha}\left(\frac{k}{2}+1\right)^{1-\alpha} \\ &\leq\; c(\alpha)\,k^{-\alpha-1/2}\end{aligned}$$

für $\alpha < 1$ und analog $\int\limits_{-1/2}^{k/2} f(x)\,dx \leq c(\alpha)\,k^{-1}$ für $\alpha = 1$. Für das andere Integral schätzen wir analog ab:

$$\begin{aligned}\int\limits_{k/2}^{k-1/2} f(x)\,dx \;&\leq\; \left(\frac{2}{k}\right)^{\alpha} \int\limits_{k/2}^{k-1/2} (k-x)^{-3/2}\,dx \;=\; 2\left(\frac{2}{k}\right)^{\alpha} (k-x)^{-1/2}\Big|_{k/2}^{k-1/2} \\ &\leq\; 2\left(\frac{2}{k}\right)^{\alpha}\sqrt{2}\,.\end{aligned}$$

Dies beendet den Beweis. □

Mit (20), (21) und (22) erhält man direkt:

[2] Also ist f konvex.

Korollar 5.24 *Es sei $A : X \to Y$ linear und kompakt mit Adjungiertem A^* und $\|A\|_{L(X,Y)} \le 1$. Dann gilt*

(a) $\big\|(I - A^*A)^k A^*\|_{L(Y,X)} \;\le\; (k+1)^{-1/2}$,

(b) $\big\|(I - A\,A^*)^k A\,A^*\|_{L(Y,Y)} \;\le\; (k+1)^{-1}$,

(c) $\left\|\sum_{j=0}^{k-1}(I - A^*A)^j A^*\right\|_{L(Y,X)} \;\le\; \sqrt{k}$.

Beweis: Sei $\{\sigma_i, x_i, y_i : i \in I\}$ eine singuläres System für A. Für $y = y_0 + \sum_{i\in I} \alpha_i\, y_i \in Y$ mit $A^* y_0 = 0$ ist

$$\big\|(I - A^*A)^k A^* y\|_X^2 \;=\; \sum_{i\in I} \big[(1-\sigma_i^2)^k\, \sigma_i\big]^2 \alpha_i^2 \;\le\; \frac{1}{k+1} \sum_{i\in I} \alpha_i^2 \;\le\; \frac{1}{k+1}\, \|y\|_Y^2$$

wegen (21). Teile (b) und (c) folgen genauso aus (22) und (20). □

Satz 5.25 *Es gelte Voraussetzung (V2), und $x^\dagger$ sei Minimumnormlösung von $K(x) = y$ bzgl. $\hat{x}$. Definiere die Konatante*

$$c_* \;=\; 2\left(1 \;+\; \frac{1-2\eta}{\eta\,(5-4\eta)}\right).$$

Außerdem gelte die ***Quellenbedingung (Q):*** *Es gebe $w \in Y$ mit*

$$x^\dagger - \hat{x} \;=\; K'(x^\dagger)^* w \tag{24}$$

und $\frac{9}{2}\, C\, c_*^2\, c(\alpha)\, \|w\|_Y \le 1$, wobei C die Konstante von (V2) ist und $c(\alpha)$ die größere der beiden Konstanten von (23) für $\alpha = 1/2$ und $\alpha = 1$. Dann existiert Konstante $c > 0$ mit

$$\|x^\delta_{k_*(\delta)} - x^\dagger\|_X \;\le\; c\,\|w\|_Y^{1/2}\, \delta^{1/2}\,, \qquad \|K(x^\delta_{k_*(\delta)}) - y\|_Y \;\le\; (1+\tau)\,\delta\,. \tag{25}$$

Hier ist wieder $k_*(\delta)$ der Abbruchindex von (17).

Beweis: Die zweite Abschätzung folgt sofort aus der Eigenschaft des Abbruchindex $k_*(\delta)$ und der Dreiecksungleichung. Setze zur Abkürzung $e_k = x_k^\delta - x^\dagger$ und $A =$

$K'(x^\dagger)$. Dann ist $K'(x_k^\delta) = R_{x_k^\delta} A$ wegen (V2) und daher

$$\begin{aligned}
e_{k+1} &= e_k - K'(x_k^\delta)^*\big(K(x_k^\delta) - y^\delta\big) \\
&= e_k - A^*\big(K(x_k^\delta) - y^\delta\big) + A^*(I - R^*_{x_k^\delta})\big(K(x_k^\delta) - y^\delta\big) \\
&= (I - A^*A)e_k - A^*\big(K(x_k^\delta) - y^\delta - K'(x^\dagger)e_k\big) + A^*(I - R^*_{x_k^\delta})\big(K(x_k^\delta) - y^\delta\big) \\
&= (I - A^*A)e_k + A^*(y^\delta - y) - A^*\big[K(x_k^\delta) - K(x^\dagger) - K'(x^\dagger)(x_k^\delta - x^\dagger)\big] \\
&\quad + A^*(I - R^*_{x_k^\delta})\big(K(x_k^\delta) - y^\delta\big) \\
&= (I - A^*A)e_k + A^*(y^\delta - y) + A^*z_k\,, \quad k = 0,\dots,k_*(\delta) - 1\,, \qquad (26)
\end{aligned}$$

mit

$$z_k = (I - R^*_{x_k^\delta})\big(K(x_k^\delta) - y^\delta\big) - \big[K(x_k^\delta) - K(x^\dagger) - K'(x^\dagger)(x_k^\delta - x^\dagger)\big]\,.$$

Die Lösung von (26) ist gegeben durch

$$\begin{aligned}
e_k &= (I - A^*A)^k e_0 + \sum_{j=0}^{k-1}(I - A^*A)^j A^*\big[(y^\delta - y) + z_{k-j-1}\big] \\
&= -(I - A^*A)^k A^* w + \left[\sum_{j=0}^{k-1}(I - A^*A)^j A^*\right](y^\delta - y) \\
&\quad + \sum_{j=0}^{k-1}(I - A^*A)^j A^* z_{k-j-1} \qquad (27)
\end{aligned}$$

für $k = 0,\dots,k_*(\delta)$. (Beweis mit vollständiger Induktion nach k.) Außerdem ist wegen $A(I - A^*A)^j = (I - A\,A^*)^j A$ auch

$$\begin{aligned}
Ae_k &= -(I - A\,A^*)^k A\,A^* w + \sum_{j=0}^{k-1}(I - A\,A^*)^j A\,A^*(y^\delta - y) \\
&\quad + \sum_{j=0}^{k-1}(I - A\,A^*)^j A\,A^* z_{k-j-1} \\
&= -(I - A\,A^*)^k A\,A^* w - \sum_{j=0}^{k-1}\big[(I - A\,A^*)^{j+1} - [(I - A\,A^*)^j\big](y^\delta - y) \\
&\quad + \sum_{j=0}^{k-1}(I - A\,A^*)^j A\,A^* z_{k-j-1} \\
&= -(I - A\,A^*)^k A\,A^* w + \big[I - (I - A\,A^*)^k\big](y^\delta - y) \\
&\quad + \sum_{j=0}^{k-1}(I - A\,A^*)^j A\,A^* z_{k-j-1} \qquad (28)
\end{aligned}$$

für $k = 0, \ldots, k_*(\delta)$. Für $k = 0, \ldots, k_*(\delta) - 1$ ist wegen der Stopregel und (13)

$$\tau\,\delta \;\le\; \|K(x_k^\delta) - y^\delta\|_Y \;\le\; \|K(x_k^\delta) - K(x^\dagger)\|_Y \;+\; \delta \;\le\; \frac{1}{1-\eta}\,\|Ae_k\|_Y \;+\; \delta\,,$$

also wegen $\tau - 1 > \frac{2(1+\eta)}{1-2\eta} - 1 = \frac{1+4\eta}{1-2\eta}$

$$\delta \;\le\; \frac{1-2\eta}{(1+4\eta)(1-\eta)}\,\|Ae_k\|_Y\,, \quad k = 0, \ldots, k_*(\delta) - 1\,. \tag{29}$$

Wir schätzen jetzt $\|z_k\|_Y$ in den Darstellungen (27) und (28) ab. Wegen $\tau > 2$ und $\|K(x_k^\delta) - y^\delta\|_Y \ge \tau\delta > 2\delta$ ist

$$\|K(x_k^\delta) - y^\delta\|_Y \;\le\; \|K(x_k^\delta) - y\|_Y \;+\; \delta \;\le\; \|K(x_k^\delta) - y\|_Y \;+\; \frac{1}{2}\,\|K(x_k^\delta) - y^\delta\|_Y$$

und daher $\|K(x_k^\delta) - y^\delta\|_Y \le 2\,\|K(x_k^\delta) - y\|_Y$. Mit (13) und wegen $\eta < 1/2$ ist also

$$\big\|(I - R_{x_k^\delta}^*)\big(K(x_k^\delta) - y^\delta\big)\big\|_Y \;\le\; \|I - R_{x_k^\delta}^*\|_{L(Y,Y)}\,\|K(x_k^\delta) - y^\delta\|_Y \;\le\; 4C\,\|e_k\|_X\,\|Ae_k\|_Y\,.$$

Außerdem ist mit (18) für $x = x_k^\delta$:

$$\big\|K(x_k^\delta) - K(x^\dagger) - K'(x^\dagger)(x_k^\delta - x^\dagger)\big\|_Y \;\le\; \frac{C}{2}\,\|e_k\|_X\,\|Ae_k\|_Y$$

und daher $\|z_k\|_Y \le \frac{9}{2}C\,\|e_k\|_X\,\|Ae_k\|_Y$. Wir zeigen jetzt

$$\|e_j\|_X \;\le\; c_*\|w\|_Y\,\frac{1}{\sqrt{j+1}}\,, \quad \|Ae_j\|_Y \;\le\; c_*\|w\|_Y\,\frac{1}{j+1} \tag{30}$$

für $j = 0, \ldots, k$ und jedes $k \le k_*(\delta) - 1$. Wir zeigen dies durch vollständige Induktion nach $k < k_*$. Für $k = 0$ ist dies richtig, denn $\|e_0\|_X \le \|w\|_Y$ und $\|Ae_0\|_X \le \|w\|_Y$ (wegen $c_* \ge 1$). Sei jetzt $k \ge 1$ und (30) richtig für $j = 0, \ldots, k-1$. Aus obiger Darstellung von e_k, Teile (a) und (c) von Korollar 5.24, der Induktionsvoraussetzung und (23) für $\alpha = 1/2$ folgt:

$$\begin{aligned}
\|e_k\|_X \;&\le\; \frac{\|w\|_Y}{\sqrt{k+1}} \;+\; \sqrt{k}\,\delta \;+\; \frac{9}{2}\,C\sum_{j=0}^{k-1}\frac{1}{\sqrt{j+1}}\,\|e_{k-j-1}\|_X\|Ae_{k-j-1}k\|_Y \\
&\le\; \frac{\|w\|_Y}{\sqrt{k+1}} \;+\; \sqrt{k}\,\delta \;+\; \frac{9}{2}\,C\,c_*^2\,\|w\|_Y^2\sum_{j=0}^{k-1}\frac{1}{\sqrt{j+1}}\,\frac{1}{k-j}\,\frac{1}{k-j} \\
&\le\; \frac{\|w\|_Y}{\sqrt{k+1}} \;+\; \sqrt{k}\,\delta \;+\; \frac{9}{2}\,C\,c_*^2\,\|w\|_Y^2\,\frac{c(\alpha)}{\sqrt{k+1}} \\
&\le\; \sqrt{k+1}\,\delta \;+\; 2\,\frac{\|w\|_Y}{\sqrt{k+1}}
\end{aligned}$$

nach Voraussetzung für $\|w\|_Y$. Analog folgt für $\|Ae_k\|_Y$ mit Teil (b) von Korollar 5.24 und (23) für $\alpha = 1$:

$$\begin{aligned}
\|Ae_k\|_Y &\leq \frac{\|w\|_Y}{k+1} + \delta + \frac{9}{2} C \sum_{j=0}^{k-1} \frac{1}{j+1} \|e_{k-j-1}\|_X \|Ae_{k-j-1}k\|_Y \\
&\leq \frac{\|w\|_Y}{k+1} + \delta + \frac{9}{2} C\, c_*^2 \,\|w\|_Y^2 \sum_{j=0}^{k-1} \frac{1}{j+1} \, \frac{1}{k-j} \, \frac{1}{k-j} \\
&\leq \frac{\|w\|_Y}{k+1} + \delta + \frac{9}{2} C\, c_*^2 \,\|w\|_Y^2 \, \frac{c(\alpha)}{k+1} \\
&\leq \delta + 2\, \frac{\|w\|_Y}{k+1} \,.
\end{aligned}$$

Nun müssen wir noch δ nach oben abschätzen. Es ist mit (29) und der gerade bewiesenen Abschätzung von $\|Ae_k\|_Y$:

$$\delta \;\leq\; \frac{1-2\eta}{(1+4\eta)(1-\eta)} \|Ae_k\|_Y \;\leq\; \frac{1-2\eta}{(1+4\eta)(1-\eta)} \left[\delta + 2\, \frac{\|w\|_Y}{k+1} \right].$$

Wegen $1 - \frac{1-2\eta}{(1+4\eta)(1-\eta)} = \frac{\eta(5-4\eta)}{(1+4\eta)(1-\eta)}$ folgt

$$\delta \;\leq\; \frac{2(1-2\eta)}{\eta(5-4\eta)} \, \frac{\|w\|_Y}{k+1} \,. \tag{31}$$

Dies setzen wir nun oben in den Abschätzungen von $\|e_k\|_X$ und $\|Ae_k\|_Y$ ein:

$$\begin{aligned}
\|e_k\|_X &\leq 2 \left[1 + \frac{1-2\eta}{\eta(5-4\eta)} \right] \frac{\|w\|_Y}{\sqrt{k+1}} = c_* \frac{\|w\|_Y}{\sqrt{k+1}} \,, \\
\|Ae_k\|_Y &\leq 2 \left[1 + \frac{1-2\eta}{\eta(5-4\eta)} \right] \frac{\|w\|_Y}{k+1} = c_* \frac{\|w\|_Y}{k+1} \,.
\end{aligned}$$

Damit sind die Behauptungen (30) für k gezeigt. Damit ist (30) bewiesen für alle $0 \leq j < k_*(\delta)$. Auch die Abschätzung (31) gilt für alle $k = 0, \ldots, k_*(\delta) - 1$ (für $k = 0$ beweist man sie getrennt mit $\|Ae_0\|_Y \leq \|w\|_Y$). Hieraus folgt im Fall $k_*(\delta) \geq 1$ für $k = k_*(\delta) - 1$, dass

$$k_*(\delta) \;\leq\; \frac{2(1-2\eta)}{\eta(5-4\eta)} \, \frac{\|w\|_Y}{\delta} \,. \tag{32}$$

Aus der Abschätzung $\|K(x_k^\delta) - y^\delta\|_Y \leq 2\|K(x_k^\delta) - K(x^\dagger)\|_Y \leq \frac{2}{1-\eta} \|Ae_k\|_Y \leq 4\,\|Ae_k\|_Y$ für $k < k_*(\delta)$ folgt

$$\|K(x_k^\delta) - y^\delta\|_Y \;\leq\; 4\, c_* \|w\|_Y \, \frac{1}{k+1} \,, \quad k = 0, \ldots, k_*(\delta) - 1 \,. \tag{33}$$

Jetzt kommen wir zum letzten Teil des Beweises. Wir schreiben e_k von (27) in der Form

$$
\begin{aligned}
e_k &= -(I - A^*A)^k A^* w \;+\; \left[\sum_{j=0}^{k-1} (I - A^*A)^j A^*\right](y^\delta - y) \\
&\quad + \sum_{j=0}^{k-1} (I - A^*A)^j A^* z_{k-j-1} \\
&= -A^*(I - AA^*)^k w \;+\; \left[\sum_{j=0}^{k-1} (I - A^*A)^j A^*\right](y^\delta - y) \\
&\quad + A^* \sum_{j=0}^{k-1} (I - AA^*)^j z_{k-j-1} \\
&= A^* w_k \;+\; \left[\sum_{j=0}^{k-1} (I - A^*A)^j A^*\right](y^\delta - y) \qquad (34)
\end{aligned}
$$

mit

$$
w_k \;=\; -(I - A^*)^k w \;+\; \sum_{j=0}^{k-1} (I - AA^*)^j z_{k-j-1}\,.
$$

Es ist mit $\|(I - AA^*)^j\|_{L(Y,Y)} \le 1$ für alle j:

$$
\begin{aligned}
\|w_k\|_Y &\le \|w\|_Y \;+\; \sum_{j=0}^{k-1} \|z_{k-j-1}\|_Y \;\le\; \|w\|_Y \;+\; \frac{9}{2}\, C \sum_{j=0}^{k-1} \|e_{k-j-1}\|_X\, \|Ae_{k-j-1}\|_Y \\
&\le \|w\|_Y + \frac{9}{2}\, C\, c_*^2 \|w\|_Y^2 \sum_{j=0}^{k-1} \frac{1}{\sqrt{k-j}}\, \frac{1}{k-j} \;\le\; \|w\|_Y + \frac{9}{2}\, C\, c_*^2 \|w\|_Y^2 \sum_{j=1}^{\infty} \frac{1}{j^{3/2}} \\
&\le \|w\|_Y \left[1 \;+\; \frac{9}{2}\, C\, c_*^2 \sum_{j=1}^{\infty} \frac{1}{j^{3/2}}\right]
\end{aligned}
$$

für alle $k = 0, \ldots, k_*(\delta)$. Hier haben wir ohne Einschränkung $\|w\|_Y \le 1$ benutzt. Die Schranke auf der rechten Seite hängt nicht von δ ab. Genauso kann mit

$$
\begin{aligned}
A \sum_{j=0}^{k-1} (I - A^*A)^j A^* &= \sum_{j=0}^{k-1} (I - AA^*)^j AA^* \\
&= -\sum_{j=0}^{k-1} (I - AA^*)^{j+1} \;+\; \sum_{j=0}^{k-1} (I - AA^*)^j \;=\; I - (I - AA^*)^k
\end{aligned}
$$

der Term abgeschätzt werden:

$$\begin{aligned} \|AA^* w_k\|_Y &\le \|Ae_k\|_Y + \left\| A \sum_{j=0}^{k-1} (I - A^*A)^j A^* \right\|_{L(Y,Y)} \|y^\delta - y\|_Y \\ &\le \|Ae_k\|_Y + \left\| I - (I - AA^*)^k \right\|_Y \delta \le \|Ae_k\|_Y + \delta \end{aligned}$$

für $k = 0, \ldots, k_*(\delta)$. Für $k = k_*(\delta)$ folgt weiter wegen (13) und der Stopregel:

$$\begin{aligned} \|AA^* w_{k_*(\delta)}\|_Y &\le (1+\eta) \|K(x^\delta_{k_*(\delta)}) - y\|_Y + \delta \\ &\le (1+\eta) \big[\|K(x^\delta_{k_*(\delta)}) - y^\delta\|_Y + \delta \big] + \delta \le \big[(1+\eta)(1+\tau) + 1 \big] \delta \,. \end{aligned}$$

Hieraus folgt

$$\begin{aligned} \|A^* w_{k_*(\delta)}\|_Y^2 &= \big(A^* w_{k_*(\delta)}, A^* w_{k_*(\delta)} \big)_Y = \big(AA^* w_{k_*(\delta)}, w_{k_*(\delta)} \big)_Y \\ &\le \|AA^* w_{k_*(\delta)}\|_Y \, \|w_{k_*(\delta)}\|_Y \le c_1 \, \delta \, \|w\|_Y \end{aligned}$$

mit einer Konstanten c_1, die von δ und w unabhängig ist. Jetzt kommen wir zurück zu (34) und erhalten mit Teil (c) von Korollar 5.24 und (32) (im Fall $k_* \ge 1$, sonst direkt)

$$\begin{aligned} \|e_{k_*(\delta)}\|_X &\le \sqrt{c_1 \delta \, \|w\|_Y} + \sqrt{k_*(\delta)} \, \delta \\ &\le \sqrt{c_1 \delta \, \|w\|_Y} + \sqrt{\frac{2(1-2\eta)}{\eta(5-4\eta)}} \, \|w\|_Y^{1/2} \sqrt{\delta} \le c \, \|w\|_Y^{1/2} \sqrt{\delta} \,. \end{aligned}$$

Damit ist der Satz bewiesen. □