Anhang Kap 4: Beweise der technischen Lemmata

Beweis von Lemma 4.1: Sei $(x^*, y^*, z^*) \in \mathcal{PD}^*$ und sei $(x,y,z) = (x^k, y^k, z^k)$. Wir setzen

$$(\bar{x}, \bar{y}, \bar{z}) := \nu_k (x^0, y^0, z^0) + (1-\nu_k)(x^*, y^*, z^*) - (x,y,z)$$

Es gilt:
$$A\bar{x} = \nu_k (Ax^0 - Ax^*) - (Ax - Ax^*)$$
$$= \nu_k (Ax^0 - b) - (Ax - b) = \nu_k r_p(x^0) - \nu_k r_p(x^k) = 0$$

Ebenso: $A^T\bar{y} + \bar{z} = 0$. Also gilt: $\bar{x} \in \mathcal{N}(A)$, $\bar{z} \in \mathcal{R}(A^T)$.

$$\Rightarrow 0 = \bar{x}^T\bar{z} = (\nu_k x^0 + (1-\nu_k)x^* - x)^T (\nu_k z^0 + (1-\nu_k) z^* - z)$$
$$= \nu_k^2 (x^0)^T z_0 + (1-\nu_k)^2 (x^*)^T z^* + \nu_k (1-\nu_k)\left((x^0)^T z^* + (z^0)^T x^*\right)$$
$$+ x^T z - \nu_k (z^T x^0 + x^T z^0) - (1-\nu_k)(z^T x^* + x^T z^*) \qquad (4.16)$$

Wir beachten nun $(x^*)^T z^* = 0$ und $z^T x^* + x^T z^* \geq 0$ und stellen um:

$$\nu_k (z^T x^0 + x^T z^0) \leq \nu_k^2 (x^0)^T z_0 + x^T z + \nu_k (1-\nu_k)\left((x_0)^T z^* + (z^0)^T x^*\right) \qquad (*)$$

Definiere nun die (positive!) Konstante

$$\eta := \min \{ x_i^0, z_i^0 \mid i = 1, \dots, n \} \qquad (4.17)$$

Man erhält damit die Beziehung

$$\eta \, \|(x,z)\|_1 \leq x^T z^0 + z^T x^0$$

Daraus bekommt man mit Hilfe von $(*)$ und wegen $\mu(x,z) = x^T z / n$, $\nu_k \in (0,1)$ folgende Abschätzung:

$$\eta \nu_k \|(x,z)\|_1 \leq \nu_k^2 n \mu(x^0, z^0) + n\mu(x,z) + \nu_k (1-\nu_k)\left(\|x^0\|_\infty \|z^*\|_1 + \|z_0\|_\infty \|x^*\|_1\right)$$
$$\leq \nu_k n \mu(x^0, z^0) + n\mu(x,z) + \nu_k \|(x^0, z^0)\|_\infty \cdot \|(x^*, z^*)\|_1 \qquad (4.18)$$

Indem wir nun ν_k gemäß der Ungleichung (4.10) ersetzen, folgt:

$$\eta \nu_k \|(x,z)\|_1 \leq \beta n \mu(x^k, z^k) + n\mu(x^k, z^k) + \beta \mu(x^k, z^k) \, \|(x^0, z^0)\|_\infty \cdot \|(x^*, z^*)\|_1 / \mu(x^0, z^0)$$

In dieser letzten Ungleichung kann auf der rechten Seite der Faktor $\mu(x^k, z^k)$ ausgeklammert werden; dividiert man diese Ungleichung durch η und setzt

$$C_1 := \frac{1}{\eta} \left[\beta n + n + \beta \, \|(x^0, z^0)\|_\infty \cdot \|(x^*, z^*)\|_1 / \mu(x^0, z^0) \right]$$

so folgt die im Lemma behauptete Ungleichung. ▨

Beweis von Lemma 4.2. Wir definieren

$$(\bar{x}, \bar{y}, \bar{z}) := (\Delta x, \Delta y, \Delta z) + \nu_k (x^0, y^0, z^0) - \nu_k (x^*, y^*, z^*) \qquad (4.19)$$

Es gilt: $A\bar{x} = A\Delta x + \nu_k (Ax^0 - \underbrace{Ax^*}_{=b}) = -r_p(\underbrace{x}_{=x^k}) + \nu_k r_p(x^0) = 0$

Also $\bar{x} \in \mathcal{N}(A)$; ebenso zeigt man: $\bar{z} \in \mathcal{R}(A^T)$.

$$\Rightarrow \quad (\Delta x + \nu_k (x^0 - x^*))^T (\Delta z + \nu_k (z^0 - z^*)) = 0 \qquad (4.20)$$

Nun gilt unter Berücksichtigung der letzten Zeile von (4.1):

$$Z(\Delta x + \nu_k (x^0 - x^*)) + X(\Delta z + \nu_k (z^0 - z^*))$$
$$= -XZe + \sigma \mu(x^k, z^k) e + Z(x^0 - x^*) + \nu_k X (z^0 - z^*)$$

Multipliziert man dieses System mit $(XZ)^{-1/2}$ und beachtet man, dass $(XZ)^{-1/2} Z = D^{-1}$, $(XZ)^{-1/2} X = D$ ist, so erhält man:

$$D^{-1}(\Delta x + \nu_k (x^0 - x^*)) + D(\Delta z + \nu_k (z^0 - z^*)) \qquad (4.21)$$
$$= -(XZ)^{-1/2}\left(XZe - \sigma \mu(x^k, z^k) e\right) + \nu_k D^{-1}(x^0 - x^*) + \nu_k D (z^0 - z^*)$$

Nun gilt wegen (4.20):

$$\| D^{-1}(\Delta x + \nu_k(x^0 - x^*)) \|^2 + \| D(\Delta z + \nu_k(z^0 - z^*)) \|^2 \tag{4.22}$$

$$\underset{(4.20)}{=} \| D^{-1}(\Delta x + \nu_k(x^0 - x^*)) + D(\Delta z + \nu_k(z^0 - z^*)) \|^2$$

$$\underset{(4.21)}{\leq} \left\{ \|(XZ)^{-1/2}\| \cdot \| XZe - \sigma\mu(x^k, z^k)e\| + \nu_k \|D^{-1}(x^0 - x^*)\| + \nu_k \|D(z^0 - z^*)\| \right\}^2$$

Wir benützen diese Abschätzung jetzt für den 1. Summanden aus (4.22):

$$\|D^{-1}\Delta x\| = \|D^{-1}(\Delta x + \nu_k(x^0 - x^*)) - D^{-1}\nu_k(x^0 - x^*)\|$$

$$\leq \|D^{-1}(\Delta x + \nu_k(x^0 - x^*))\| + \nu_k \|D^{-1}(x^0 - x^*)\| \tag{4.23}$$

$$\underset{(4.22)}{\leq} \|(XZ)^{-1/2}\| \cdot \|Xz - \sigma\mu(x^k, z^k)e\| + 2\nu_k \|D^{-1}(x^0 - x^*)\| + 2\nu_k \|D(z^0 - z^*)\|$$

Wir vervollständigen den Beweis, indem wir zeigen, dass jeder Summand in (4.23) von der Größenordnung $O(\mu(x^k, z^k)^{1/2})$ ist. Zunächst gilt:

$$\| XZe - \sigma\mu(x^k, z^k)e \|^2 = \|XZe\|^2 - 2\sigma\mu(x^k, z^k)x^T z + \sigma^2 \mu(x^k, z^k)^2 n$$

$$\leq (n\mu(x^k, z^k))^2 - 2\sigma n \mu(x^k, z^k)^2 + \sigma^2 n \mu(x^k, z^k)$$

$$\leq n^2 \mu(x^k, z^k)^2$$

Außerdem: $$\|(XZ)^{-1/2}\| = \max_{i=1,\dots,n} \frac{1}{(x_i z_i)^{1/2}} \overset{\text{Def. } \mathcal{U}_{-\infty}(\gamma, \beta)}{\leq} \frac{1}{\gamma^{1/2}\mu(x^k, z^k)^{1/2}} \quad (*)$$

Wenn man die beiden letzten Abschätzungen kombiniert, so folgt:

$$\|(XZ)^{-1/2}\| \, \|XZe - \sigma\mu(x^k, z^k)e\| \leq \frac{n}{\gamma^{1/2}} \mu(x^k, z^k)^{1/2} \tag{4.24}$$

Wir schätzen nun die beiden letzten Summanden aus (4.23) ab:

$$\nu_k \|D^{-1}(x^0 - x^*)\| + \nu_k \|D(z^0 - z^*)\|$$

$$\leq \nu_k (\|D^{-1}\| + \|D\|) \max(\|x^0 - x^*\|, \|z^0 - z^*\|) \tag{4.25}$$

Für die Matrixnorm $\|D^{-1}\|$ erhalten wir:

$$\|D^{-1}\| = \max_{i=1,\dots,n} |D_{ii}^{-1}| = \|D^{-1}e\|_\infty = \|(XZ)^{-1/2} Ze\|_\infty$$

$$\leq \|(XZ)^{-1/2}\| \, \|z\|_1 \tag{4.26a}$$

Ebenso: $$\|D\| \leq \|(XZ)^{-1/2}\| \, \|x\|_1 \tag{4.26b}$$

Wir verwenden nun Lemma 4.1, (*) und (4.26a, b):

$$\nu_k \|D^{-1}(x^0 - x^*)\| + \nu_k \|D(z^0 - z^*)\|$$

$$\underset{(4.25/4.26)}{\leq} \nu_k \|(x, z)\|_1 \|(XZ)^{-1/2}\| \max(\|x^0 - x^*\|, \|z^0 - z^*\|)$$

$$\underset{\text{Le } 4.1}{\leq} \frac{C_1}{\gamma^{1/2}} \mu(x^k, z^k)^{1/2} \max(\|x^0 - x^*\|, \|z^0 - z^*\|) \tag{4.27}$$

Unter Beachtung von (4.23), (4.24) und (4.27) folgt, dass $\|D^{-1}\Delta x\| \leq C_2 \mu(x^k, y^k)^{1/2}$ ist, wobei C_2 gegeben ist durch

$$C_2 = \frac{n}{\gamma^{1/2}} + 2\frac{C_1}{\gamma^{1/2}} \max(\|x^0 - x^*\|, \|z^0 - z^*\|).$$

Dass man $\|D^{-1}\Delta z\|$ ebenso abschätzen kann, dürfte klar sein. ▨

Anhang zu Kap. 6: Beweis zu Satz 6.3

Vss.: $\{w^k\}$ besitze einen HP.
Beh.: Dann gilt: $\tau_k \to 0$.

(Beachte: $t \hat{=} \alpha$, $t_k \hat{=} \alpha_k$)

Beweis: Nach Konstruktion ist die Folge (τ_k) monoton fallend und durch 0 nach unten beschränkt; sie ist also konvergent. Wir setzen $\lim_{k\to\infty} \tau_k = \bar{\tau}$.

Annahme: Es sei $\bar{\tau} > 0$. Aus der Aufdatierungsvorschrift in Schritt 4) folgt dann:

$$\gamma_k \downarrow 0 \quad (\text{für } k \to \infty) \qquad (6.12)$$

Sei nun w^* ein Häufungspunkt von (w^k) und $(w^k)_K$ eine Teilfolge, die gegen w^* konvergiert. Da $\Phi'_{\bar{\tau}}(w^*)$ regulär ist (Lemma 6.2), existiert eine Konstante $c > 0$ mit

$$\| \Phi'_{\tau_k}(w^k)^{-1} \| \le c$$

für alle hinreichend großen $k \in K$. Wir betrachten nun die folgenden beiden Fälle:

1) $\exists\, \bar{\alpha} \ge 0$ mit $\alpha_k \ge \bar{\alpha} \quad \forall\, k \in K$ $\quad(\bar{\alpha} := \bar{t})$ $\quad(\alpha_k := t_k)$
2) $\exists$ Teilfolge (α_{k_ν}) von $(\alpha_k)_K$ mit $\alpha_{k_\nu} \to 0$

Fall 1: Sei also $\alpha_k \ge \bar{\alpha} > 0 \quad \forall\, \alpha \in K$. Wegen (6.10) folgt:

$$\psi_{\tau_k}(w^{k+1}) \le (1 - 2\alpha_k \sigma)\psi_{\tau_k}(w^k) \le (1 - 2\bar{\alpha}\sigma)\psi_{\tau_k}(w^k)$$

Aus der Definition von ψ_τ erhält man mit einem $\bar{c} \in (2\bar{\alpha}\sigma, 1)$:

$$\| \Phi_{\tau_k}(w^{k+1}) \| \le (1 - \bar{c}) \cdot \| \Phi_{\tau_k}(w^k) \| \le (1 - \bar{c})\beta \tau_k \qquad (6.13)$$

Da offenbar die Ungleichung

$$\| \Phi_\tau(w) - \Phi_{\tau'}(w) \| \le \varkappa |\tau - \tau'|$$

für alle $\tau, \tau' > 0$ und alle w gilt, wobei $\varkappa$ die Konstante aus Lemma 6.1 ist, erhalten wir mit (6.13):

$$\| \Phi_{(1-\gamma)\tau_k}(w^{k+1}) \| \le \| \Phi_{\tau_k}(w^{k+1}) \| + \| \Phi_{(1-\gamma)\tau_k}(w^{k+1}) - \Phi_{\tau_k}(w^{k+1}) \|$$
$$\le (1 - \bar{c})\beta\tau_k + \varkappa\gamma\tau_k$$
$$= (\beta - \bar{c}\beta + \varkappa\gamma)\tau_k$$

für alle $\gamma \in (0,1)$. Man kann nun leicht nachprüfen, dass

$$(\beta - \bar{c}\beta + \varkappa\gamma)\tau_k \le \beta(1-\gamma)\tau_k$$

für alle $\gamma > 0$ mit $\gamma \le \bar{c}\beta/(\varkappa + \beta)$ gilt. Aus der Berechnungsvorschrift des Schrittes 4) folgt damit:

$$\gamma_k / \varrho \ge \bar{c}\beta/(\varkappa + \beta)$$

für alle hinreichend großen $k \in K$. Dies widerspricht aber $\gamma_k \to 0$ (sh. (6.12)).

Fall 2: Es gebe eine Teilfolge von $(\alpha_k)_{k\in K}$, die gegen Null strebt; o.E. gelte: $\alpha_k \to 0$ (für $k \in K$). Die Armijo-Regel in Schritt 4) ist dann für $\tilde{\alpha}_k := \alpha_k/\varrho$ und alle hinreichend großen $k \in K$ nicht erfüllt. Es gilt also:

$$\psi_{\tau_k}(w^k + \tilde{\alpha}_k \Delta w) \ge \psi_{\tau_k}(w^k) + \tilde{\alpha}_k \sigma \nabla\psi_{\tau_k}(w^k)^T \Delta w$$

$$\Longleftrightarrow \quad \frac{\psi_{\tau_k}(w^k + \tilde{\alpha}_k \Delta w) - \psi_{\tau_k}(w^k)}{\tilde{\alpha}_k} \ge \sigma \nabla\psi_{\tau_k}(w^k)^T \Delta w$$

Nach dem MWS der Differentialrechnung gibt es ein ξ^k auf der Verbindungsstrecke von w^k und $w^k + \tilde{\alpha}_k \Delta$ mit

$$\nabla\psi_{\tau_k}(\xi^k)^T \Delta w \ge \sigma \nabla\psi_{\tau_k}(w^k)^T \Delta w \qquad (6.14)$$

Man beachte im Folgenden, dass auch Δw von w^k abhängt, schreibe deshalb: Δw^k an Stelle von Δw.

Man beachte nun, dass wegen $(\tau_k)_K \to \bar{\tau}$, $(w^k)_K \to w^*$ gilt: $(\Delta w^k)_K \to \Delta w^*$ und dass Δw^* das folgende LGS löst:

$$\Phi'_{\bar{\tau}}(w^*)\Delta w = -\Phi_{\bar{\tau}}(w^*)$$

Betrachtet man den Grenzwert $k \to \infty$ $(k \in K)$ für die Folgen $(\tilde{\alpha}_k)$, (τ_k), so folgt aus (6.14) wegen $\tilde{\alpha}_k \to 0$, $\tau_k \to \bar{\tau}$:

$$\nabla \psi_{\bar{\tau}}(w^*)^T \Delta w^* \geq \sigma \nabla \psi_{\bar{\tau}}(w^*)^T \Delta w^*$$

Wegen $\sigma \in (0,1)$ folgt daraus:

$$\nabla \psi_{\bar{\tau}}(w^*)^T \Delta w^* \geq 0$$

Andererseits folgt aus (6.9) durch Grenzübergang:

$$\nabla \psi_{\bar{\tau}}(w^*)^T \Delta w^* \leq 0$$

Also ergibt sich weiter: $\nabla \psi_{\bar{\tau}}(w^*)^T \Delta w^* = 0$.

Mit $\nabla \psi_{\bar{\tau}}(w^*) = \Phi'_{\bar{\tau}}(w^*) \Phi_{\bar{\tau}}(w^*)$ und $\Delta w^* = -\Phi'_{\bar{\tau}}(w^*) \Phi_{\bar{\tau}}(w^*)$ erhält man weiter:

$$\Phi_{\bar{\tau}}(w^*) = 0 \qquad (6.15)$$

Andererseits erfüllt $\gamma_k/8$ den Test (6.7) in Schritt 4) nicht, so dass gilt:

$$\| \Phi_{(1-\gamma_k/8)\tau_k}(w^{k+1}) \| > \beta (1 - \gamma_k/8) \tau_k$$

Für $k \to \infty$ $(k \in K)$ folgt daraus wegen $\gamma_k \downarrow 0$:

$$\| \Phi_{\bar{\tau}}(w^*) \| \geq \beta \bar{\tau} > 0$$

Dies ist ein Widerspruch zu (6.15).

Also ist die Annahme $\bar{\tau} > 0$ falsch und $\lim_{k \to \infty} \tau_k = 0$. ▨

Unter Verwendung von Satz 6.3 sind wir nun in der Lage, einen globalen Konvergenzsatz für Algorithmus GV zu beweisen.