Lemma 14.9 f konveks, og $\lambda_{min}(\nabla^2 f(x)) \geq m$. Da er

($f^* = f(x^*)$, min. verdi) $f(x) - f^* \leq \frac{1}{2m} \|\nabla f(x)\|^2$

(er på linjestykket fra x til y) $x + th$

Bevis: 2. ordens Taylor: $y = x + h$

$$f(y) = f(x) + \nabla f(x)^T \overbrace{(y-x)}^{h} + \frac{1}{2}(y-x)^T \nabla^2 f(z)(y-x)$$

$$\geq f(x) + \nabla f(x)^T (y-x) + \frac{m}{2}\|y-x\|^2 \quad \text{(Lemma 14.2)}$$

minimerer denne med tanke på y

gradient: $\nabla f(x) + m(y-x)$

hvis min inntreffer for $y = y^*$

så må $y^* - x = -\frac{1}{m}\nabla f(x)$ $\left(\Rightarrow y^* = x - \frac{1}{m}\nabla f(x)\right)$

gradient til $\frac{1}{2}(y-x)^T A (y-x)$, I

er $A(y-x)$.

$$\geq f(x) + \nabla f(x)^T (y^* - x) + \frac{m}{2}\|y^* - x\|^2$$

$$= f(x) - \frac{1}{m}\nabla f(x)^T \nabla f(x) + \frac{1}{m^2}\frac{m}{2}\|\nabla f(x)\|^2$$

$$= f(x) - \frac{1}{m}\|\nabla f(x)\|^2 + \frac{1}{2m}\|\nabla f(x)\|^2$$

$$= f(x) - \frac{1}{2m}\|\nabla f(x)\|^2$$

$$f(y) \geq f(x) - \frac{1}{2m}\|\nabla f(x)\|^2$$

$y = x^*$

$$f^* \geq f(x) - \frac{1}{2m}\|\nabla f(x)\|^2$$

$$f(x) - f^* \leq \frac{1}{2m}\|\nabla f(x)\|^2.$$

Teorem 14.10 f konveks, to ganger kont. deriverbar og

1. $\lambda_{min}(D^2 f(x)) \geq m$ alle x
2. $\|D^2 f(x) - D^2 f(y)\|_2 \leq L \|x-y\|$ alle x, y

Anta x^* er min for f. Newtons metode med backtracking linjesøk gir en sekvens x_k som konvergerer kvadratisk til x^* $(\|x_{k+1} - x^*\| \leq C \|x_k - x^*\|^2)$

Bevis: Sett $f^* = f(x^*)$

Lemma 14.11 "Damped Newton phase"
For enhver η fins $\gamma > 0$ s.a. (for alle k) hvis $\|Df(x_k)\| \geq \eta$ så er $f(x_{k+1}) \leq f(x_k) - \gamma$.

Selv med stor gradient klarer vi altså å få ned funksjonsverdien noe.
Kan derfor ha kun endelig mange steg med Newton der $\|Df(x_k)\| \geq \eta$
antall steg $\leq \dfrac{f(x_k) - f_*}{\gamma}$

Lemma 14.11 brukes for å få gradienten $< \eta$

Lemma 14.12 Det finnes en η med $0 < \eta < \frac{m^2}{L}$, s.a. for hver k, hvis $\|Df(x_k)\| < \eta$ så tilfredsstiller $\alpha_k = 1$ stoppbetingelsen for backtracking linjesøk i Newton, og Newton gir
$$\frac{L}{2m^2}\|Df(x_{k+1})\| \leq \left(\frac{L}{2m^2}\|Df(x_k)\|\right)^2$$

Hopper over bevis.
Bevis for teorem 14.10 fra lemmaene: "Damped Newtonphase" har endelig mange iterasjoner.
Det er klart at
La η være som i lemma 14.12, og anta $\|Df(x_k)\| < \eta$

$$\|Df(x_{k+1})\| \leq \frac{2m^2}{L}\left(\frac{L}{2m^2}\|Df(x_k)\|\right)^2 = \frac{L}{2m^2}\|Df(x_k)\|^2 < \frac{1}{2}\frac{L}{m^2}\eta \cdot \eta < \frac{1}{2}\frac{L}{m^2}\frac{m^2}{L}\eta = \frac{\eta}{2} < \eta$$

Derfor vil også $\|Df(x_{k+1})\| < \eta$, $\|Df(x_{k+2})\| < \eta$, o.s.v.

Kvadratisk konv.: Sett $\mu_l = \frac{L}{2m^2}\|Df(x_l)\|$, $l \geq k$

$$\mu_l = \frac{1}{2}\frac{L}{m^2}\|Df(x_l)\| < \frac{1}{2}\frac{L}{m^2}\eta < \frac{1}{2}\frac{L}{m^2}\frac{m^2}{L} = \frac{1}{2}$$

fra lemma 14.12 ser vi at $\mu_{l+1} \leq \mu_l^2 \leq ((\mu_{l-1})^2)^2$

Induksjon: $\mu_l \leq (\mu_k)^{2^{l-k}} < \left(\frac{1}{2}\right)^{2^{l-k}}$

Lemma 14.9: $f(x_l) - f^* \leq \frac{1}{2m}\|Df(x_l)\|^2 = \frac{1}{2m}\frac{4m^4}{L^2}\left(\frac{L}{2m^2}\|Df(x_l)\|\right)^2 = \frac{2m^3}{L^2}\mu_l^2$
$$\leq \frac{m^3}{L^2}\left(\frac{1}{2}\right)^{2^{l-k+1}}$$

derfor $f(x_l) \searrow f^*$, min. er da nådd.

Kap. 5 min. problemer med betingelser

(15.1) $\min f(x)$ } kont. deriverbare fra $\mathbb{R}^n$ til $\mathbb{R}$

$h_i(x) = 0 \quad 1 \le i \le m$

$g_j(x) \le 0 \quad 1 \le j \le r$

x tillatt ("feasible") betyr at alle $m+r$ likheter/ulikheter er oppfylt.

(15.2) $\min f(x)$

$h_i(x) = 0 \quad 1 \le i \le m$

Vi setter $\vec{H}(\vec{x}) = (h_1(\vec{x}), \ldots, h_m(\vec{x}))$ $\quad G(\vec{x}) = (g_1(\vec{x}), \ldots, g_r(\vec{x}))$

x^* kaltes regulært for (15.2) hvis $\{Dh_i(x^*)\}_{i=1}^m$ er lineært uavhengige.

Teorem 15.1 Anta x^* lokalt min. i (15.2), og at x^* er regulær

Finnes da en unik $\lambda^* = (\lambda_1^*, \ldots, \lambda_m^*) \in \mathbb{R}^m$ s.a.

$Df(x^*) + \sum \lambda_i Dh_i(x^*) = 0$

Hvis f, h_i er to ganger kont. deriverbare gjelder at

for $h \in T(x^*) = \{h \in \mathbb{R}^n \mid Dh_i(x^*)^T \vec{h} = 0 \;\; \forall i\}$

har vi at $h^T \left(D^2 f(x^*) + \sum_{i=1}^m \lambda_i D^2 h_i(x^*)\right) h \ge 0$

λ^* kalles Lagrangemultiplikatorer. $\rightarrow$ i $m+n$ variable.

Hvis vi setter $L(x,\lambda) = f(x) + \sum_{i=1}^n \lambda_i h_i(x)$ (kalles Lagrangefunksjonen) $(f(x) + \lambda^T H(x))$

så er $D_x L(x,\lambda) = Df(x) + \sum_{i=1}^n \lambda_i Dh_i(x)$

$D_\lambda L(x,\lambda) = (h_1(x), \ldots, h_m(x)) = H(x)$

Med andre ord, $Df(x) + \sum_{i=1}^n \lambda_i Dh_i(x) = 0$ og $H(x) = 0$ er det samme som

$D_x L(x,\lambda) = 0 \;,\; D_\lambda L(x,\lambda) = 0$

$\Updownarrow$

$DL(x,\lambda) = 0$

Vi trenger altså bare å vise at (x^*, λ^*) er et stasjonært punkt for $L(x,\lambda)$

$T(x^*)$: inneholder $\vec{h}$ som gjør at betingelsene ikke blir brutt.

$(Dh_i(x^*))^T \vec{h} = Dh_i(x^*) \cdot \vec{h} = 0 \rightarrow$ retningsderiverte til h_i i retning h.

Df er ortogonal på tillatte retninger h, dvs. $h \in T(x^*)$.

Bevis: Sett $F^k(x) = f(x) + \underbrace{\frac{k}{2}\|H(x)\|^2}_{\text{straffeterm}} + \underbrace{\frac{\alpha}{2}\|x-x^*\|^2}_{\text{teknisk term}}$

Anta x^* lokalt min. Finnes da $\varepsilon > 0$ s.a. $f(x^*) \le f(x)$ $\forall x \in \bar{B}(x^*, \varepsilon)$ $(H(x^*) = H(x) = 0)$

La x^k være min. for F^k på $\bar{B}(x^*, \varepsilon)$. (ubetinget min)

For alle k gjelder at

$$F^k(x^k) = f(x^k) + \boxed{\frac{k}{2}\|H(x^k)\|^2} + \frac{\alpha}{2}\|x^k - x^*\|^2$$
$$\le F^k(x^*) = f(x^*) + \frac{k}{2}\|H(x^*)\|^2 + \frac{\alpha}{2}\|x^* - x^*\|^2 = f(x^*)$$

Lar vi $k \to \infty$ ser vi at $\lim_{k\to\infty} \|H(x^k)\| = 0$

enhver grense $\bar{x}$ for følgen x^k vil oppfylle at $H(\bar{x}) = 0$ pga. kontinuitet.

Får også: $\boxed{f(x^k) + \frac{\alpha}{2}\|x^k - x^*\| \le f(x^*)}$ (droppet en term)

fordi x^* er min for $\bar{x}$ tillatt, siden $H(\bar{x}) = 0$)

La $k \to \infty$: $\boxed{f(\bar{x}) + \frac{\alpha}{2}\|\bar{x} - x^*\|^2 \le f(x^*) \le f(\bar{x})}$

Derfor er $\frac{\alpha}{2}\|\bar{x} - x^*\|^2 = 0 \Rightarrow \bar{x} = x^*$, slik at $x_k \to x^*$.

x^k er ubetinget min for F^k, og derfor

Teorem 14.1: $DF^k(x^k) = 0$

$$DF^k(x^k) = D\Big(f(x^k) + \underbrace{\frac{k}{2}\|H(x^k)\|^2}_{\text{neste gang}} + \frac{\alpha}{2}\|x^k - x^*\|^2\Big)$$
$$= Df(x^k) + kH'(x_k)^T H(x^k) + \alpha(x^k - x^*)$$

Vi vet at radene i $H'(x^*)$ er lin. uavh. $\Rightarrow H'(x^*)$ inverterbar

$\Rightarrow H'(x^*)(H'(x^*))^T$ også inverterbar.

Siden $x^k \to x^*$ så vil $\underbrace{H'(x^k)(H'(x^k))^T}_{\text{ganger med den inverse til denne på begge sider og får:}}$ også være inverterbar for store k. $H'(x_k)$

$$\underbrace{\big(H'(x^k)(H'(x^k))^T\big)^{-1} H'(x_k)(H'(x_k))^T}_{I} k H(x^k) = kH(x^k).$$

Får da: $-(H'(x^k)H'(x^k)^T)^{-1} H'(x_k)\big(Df(x^k) + \alpha\underbrace{(x^k - x^*)}_{0}\big) = kH(x^k)$

$k \to \infty$: $\boxed{-(H'(x^*)H'(x^*)^T)^{-1} H'(x^*) \quad Df(x^*)} = \lim_{k\to\infty} kH(x^k) = \lambda^*$

I grensen: $0 = Df(x^*) + H'(x^*)^T \lambda^*$