Proof of THM:

PLAN: Show that the sequence of Newton iterates

$$\vec{x}_{k+1} = \vec{x}_k - D\vec{f}(\vec{x}_k)^{-1}\vec{f}(\vec{x}_k)$$

is a Cauchy sequence and therefore converges.

Step I: Show that $\forall \vec{x} \in D$, $D\vec{f}(\vec{x})$ is invertible, and

$$\| D\vec{f}(\vec{x})^{-1} \| \le 2 \| D\vec{f}(\vec{x}_0)^{-1} \|.$$

If $\vec{x} \in D$ then $\qquad D = \{x : \|x - x_1\| \le \delta\} \qquad \delta = \frac{r}{1-r}\|x_1 - x_0\|$

$$\|\vec{x} - \vec{x}_0\| \le \|x - x_1\| + \|x_1 - x_0\| \le \delta + \|x_1 - x_0\| = \left(\frac{r}{1-r} + 1\right)\|x_1 - x_0\|$$

$$= \left(\frac{1}{1-r}\right)\|x_1 - x_0\|$$

$$\frac{1}{1-r} = \frac{1}{r}\left[\frac{1}{\frac{1}{r} - 1}\right] \qquad r < \tfrac{1}{3} \Rightarrow \tfrac{1}{r} \ge 3 \Rightarrow \tfrac{1}{r} - 1 \ge 2$$

$$\Rightarrow \frac{1}{\frac{1}{r} - 1} \le \frac{1}{2}$$

$$\Rightarrow \|x - x_0\| \le \frac{1}{2r}\|x_1 - x_0\| \le \frac{1}{2r}\|Df(x_0)^{-1}\| \, \|f(x_0)\|$$

$$= \frac{1}{2L\|Df(x_0)^{-1}\|}$$

$r = L\|Df(x_0)^{-1}\| \cdot \|f(x_0)\|$

To make the notation more compact, let $A = Df(x_0)$ and $B = Df(x)$.

If $x \in D$, we know that $\|x - x_0\| \leq \frac{1}{2L\|A^{-1}\|}$.

$x \to Df(x)$ is Lip. cont $\Rightarrow$ For $x \in D$:

$$\|A - B\| \leq L\|x - x_0\| \leq \frac{1}{2\|A^{-1}\|}$$

$\Rightarrow \|A^{-1}\| \|A - B\| \leq \frac{1}{2}$.

By the theorem (11.2), B is invertible and

$$\|B^{-1}\| \leq \frac{\|A^{-1}\|}{1 - \underbrace{\|A^{-1}\| \|A - B\|}_{\leq \frac{1}{2}}} \leq 2\|A^{-1}\|$$

i.e. $\|Df(x)^{-1}\| \leq 2\|Df(x_0)^{-1}\|$.

✓

<u>Step II</u>: We show that the Newton iterates $x_k \in D$ $\forall k$ and

$$\|x_k - x_{k-1}\| \leq r^{k-1}\|x_1 - x_0\|. \quad (*)$$

The proof is by induction.

Assume that $x_1, \dots, x_k \in D$ and that $(*)$ holds $\forall i \leq k$. We will show that $(*)$ holds for $k+1$ and that $x_{k+1} \in D$.

$$\|x_{k+1} - x_k\| \leq \|Df(x_k)^{-1}\| \|f(x_k)\| \leq 2\|Df(x_0)^{-1}\| \|f(x_k)\|.$$

$\Rightarrow$ We need to estimate $\|f(x_k)\|$.

The linear approximation of f at x_{k-1} evaluated at x_k:

$$f(x_k) = f(x_{k-1}) + Df(x_{k-1})(x_k - x_{k-1}) + R(x_k, x_{k-1})$$

But x_k is chosen such that

$$f(x_{k-1}) + Df(x_{k-1})(x_k - x_{k-1}) = 0$$

$$\Rightarrow \|f(x_k)\| = \|R(x_k, x_{k-1})\| \underset{\text{Lemma}}{\leq} \frac{L}{2}\|x_k - x_{k-1}\|^2 \underset{\text{Induction hypothesis}}{\leq}$$

$$\leq \frac{L}{2}\left(r^{k-1}\|x_1 - x_0\|\right)^2 = \frac{Lr^{2k-2}}{2}\underbrace{\|x_1 - x_0\|}_{\text{(one iteration of Newton)}}\cdot\|x_1 - x_0\|$$

$$\leq \frac{r^{2k-2}L}{2}\|Df(x_0)^{-1}\|\,\|f(x_0)\|\,\|x_1 - x_0\|$$

Combining this result with the previous estimate:

$$\|x_{k+1} - x_k\| \leq 2\|Df(x_0)^{-1}\|\,\|f(x_0)\| \leq r^{2k-2}\underbrace{L\,\|Df(x_0)^{-1}\|^2\|f(x_0)\|}_{=r}\|x_1 - x_0\|$$

$$\leq r^{2k-1}\|x_1 - x_0\| \underset{2k-1 \geq k \text{ and } r \leq 1/3}{\leq} r^k\|x_1 - x_0\|.$$

So we are done with proving ⊛ for $k+1$.

To show that $x_{k+1} \in D$:

$$\|x_{k+1} - x_0\| \leq \|x_{k+1} - x_k\| + \ldots + \|x_2 - x_1\| \leq (r^k + \ldots + r)\|x_1 - x_0\|$$

$$= r\underbrace{(1 + \ldots + r^{k-1})}_{\leq \frac{1}{1-r}}\|x_1 - x_0\| \leq \frac{r}{1-r}\|x_1 - x_0\| = \delta. \quad \Rightarrow x_{k+1} \in D. \checkmark$$

<u>Step 3:</u> The sequence $\{x_k\}$ converges and the limit is a root of f.

- From the inequality $\|x_k - x_{k-1}\| \le r^{k-1} \|x_1 - x_0\|$ it can be immediately shown that $\{x_k\}$ is a Cauchy Sequence (just like in the proof of the contraction mapping THM) $\Longrightarrow$ x_k converges to a limit $x_* \in D$. (Since D is closed).

- Due to the estimate $\|f(x_k)\| \le \frac{r^{2k-2}L}{2} \|Df(x_0)^{-1}\| \, \|f(x_0)\| \, \|x_1 - x_0\|$ We know that $f(x_k) \xrightarrow{k \to \infty} 0$.

The continuity of f implies that

$$f(x_*) = f(\lim_{k \to \infty} x_k) = \lim_{k \to \infty} f(x_k) = 0 \quad \text{so } x_* \text{ is a root.}$$

<u>Step 4:</u> The convergence is quadratic, i.e. there exists a constant K such that for k sufficiently large

$$\|x_{k+1} - x_k\| \le K \|x_k - x_*\|^2.$$

Proof: The $k+1$ iterate is defined as

$$0 = f(x_k) + Df(x_k)(x_{k+1} - x_k)$$

With the linear approximation of f at x_k:

$$0 = f(x_*) = f(x_k) + Df(x_k)(x_* - x_k) + R(x_*, x_k)$$

Combining this with the previous line:

$$0 = Df(x_k)(x_* - x_{k+1}) + R(x_*, x_k)$$

$$\Rightarrow \quad x_{k+1} - x_* = Df(x_k)^{-1} R(x_*, x_k)$$

$$\Rightarrow \|x_{k+1} - x_*\| \leq \|Df(x_k)^{-1}\| \, \|R(x_*, x_k)\|$$

$$\leq 2 \|Df(x_0)^{-1}\| \frac{L}{2} \|x_k - x_*\|^2$$

Take $K = L \|Df(x_0)^{-1}\|$ and we are done with the proof of (at least) quadratic convergence.

<u>QED</u>

<u>Remark:</u> In the Thm we assumed that $x \to Df(x)$ is Lip cont. with Lip const. L. The following Thm will provide a practical way to estimate L.
We expect this L to be bounded by the max. of the 2nd-order derivatives.