Newton method  As we are almost always looking for critical points (i.e. solutions to $\nabla f = 0$), it makes sense to look for a general tool of finding roots, i.e. $x^*: F(x^*) = 0$, $F: \mathbb{R}^n \to \mathbb{R}^n$

Easy case: $F$ is linear: $F(x) = Ax + b$, then

$0 = F(x^*) = F(x) + A(x^* - x)$, and

$x^* = -A^{-1}F(x) + x$.

If $F$ is nonlinear, just pretend it is, using $\frac{\partial F}{\partial x}$ as approx for $A$:

So, Newton method just iterates

$$x_{k+1} = x_k - \left( \frac{\partial F}{\partial x} \right)^{-1} F(x_k)$$

From pictures, Newton method should be very fast.

**Example:** $n = 1$, $f = x^3 - x$

\[ N(x) \to x - \frac{x^3 - x}{3x^2 - 1} = \frac{2x^3}{3x^2 - 1} \]

$x_0 = 2$

$x_1 = \frac{16}{17}$

$x_2 = \frac{2084266232445329}{20717993079584776} \approx 1.00605$

However, there might be some problems, too:

$x_0 = \frac{1}{\sqrt{5}}$, $x_1 = \frac{2}{5\sqrt{5}} = -\frac{1}{\sqrt{5}}$

$x_2 = \frac{1}{\sqrt{5}}$, etc...

Thus, if $F: \mathbb{R}^n \to \mathbb{R}^n$ is continuously differentiable, $F(x^*) = 0$ and $\frac{\partial F}{\partial x}$ is invertible, then:
a) There is an open ball $B_0(\alpha_k) = B$ for any starting $\alpha_0$.

1. $x_k \in B$, $x_k \rightarrow x_k$ &
   \lim_{k \to \infty} \frac{1}{k} \log \|x_k - x_0\| \rightarrow -\infty$

b) If $F$ is $L$-Lipschitz, then

\[ \|x_k - x_0\| \leq K \|x_k - x_0\|^2 \]

(i.e. $\|x_k - x_0\| = O(e^{-c2^k})$).

Proof: Compute $x_{k+1} - x_k = x_k - x_k - (DF)^{-1}DF(x_k) = (DF)^{-1} \left[ DF(x_k) - F(x_k) \right] = (DF)^{-1} \left[ DF(x_k) - \int_0^1 D_{x_k}F \cdot (x_k - x_t) dt \right]$

Here $x_t = x_k + t(x_k - x_0)$

we use standard formula $F(x_0) - F(x_t) = \int_0^1 D_{x_t}F \cdot (x_t - x_t) dt$

$= (DF)^{-1} \left[ \int_0^1 (D_{x_k}F - D_{x_k}F)(x_k - x_t) dt \right]$

$\Rightarrow \|x_k - x_0\| \leq \|(DF)^{-1}\| \cdot \|x_k - x_0\| \cdot \int_0^1 \|D_{x_k}F - D_{x_k}F\| dt$ \((x)\)

As $D_{x_k}F$ is continuous in $x_k$, we can be assumed small for $\|x_k - x_0\|$ small, i.e. for $x_k \in B_0(\alpha_0)$, if $\|D_{x_k}F\| \leq M$ on $B_0(\alpha_0)$ we can shrink $p$ if needed so that $\|D_{x_k}F - D_{x_k}F\| \leq \frac{1}{2} M$ on $B$, and therefore $(x)$ holds.

This is actually going done with $p$, so that $a)$ holds.

Now, if $DF$ is Lipschitz, then $(*) \leq M \|x_k - x_0\| \cdot \|x_k - x_0\|^2$

This implies $b)$. 
A useful (and usually interesting) class of mappings is given by $f: \mathbb{C} \to \mathbb{C}$, a polynomial. Julia sets emerging there are exactly of fractals.

What happens when $Df$ is not invertible? Quadratic convergence disappears. Say $f(x) = x^2$, $x_0 = 0$.

Then $x_{k+1} = x_k - \frac{2x_k}{2x_k} = x_k/2 \Rightarrow$ bland geometric convergence.

Question about the region for starting points to converge to $x_f$ is important one. One typical result:

(valid, actually, for $\mathbb{C}$-dimension Baudouy)

Thus (Kantrowich) $F: D \to \mathbb{R}^n$; $D$ -open convex.

If for $x_0 \in D$, $\| (D_{x0}F)^{-1}F(x_0) \| \leq 1$;

$\| (D_{x0}F)^{-1}(D_{x0}F - D_{y0}F) \| \leq k \| x_0 - y_0 \|$ for $x, y \in D$

$0 < k < 1$,

$B_{2k}(x_0) \subseteq D$, then $B_{2k}(x_0) \ni x_f$, unique root $F(x_f) = 0$; and $\| x_{k+1} - x_f \| \leq \frac{d}{2^{n+1}} q^{2^{n-1}}$, $q = 2dkc4$

No proof here.

Use of Newton method in optimization requires finding $DF$, where $F(x) = \nabla f(x)$, i.e. $DF = H_f$. NM is therefore a second order method. Conceptually, one approximates $f$ with its Taylor expansion to 2nd order, and finds minimum.

$\nabla f(x_k) + \frac{1}{2} H_f(x_k)\left(x - x_k\right) + \frac{1}{2} (H_f(x_k), x - x_k)$