Newton's method is an iterative process that converges to the root of a function. Newton did not use this method in it's modern form. Instead he approximated the root using a sequence of polynomials. We choose an x_0 from an appropriate interval around a single root. Then, the $n+1$ term is given by the $x-intercept$ of the tangent line at $(x_n, f(x_n))$. Under the right conditions this converges to the root.

Draw picture and derive recursive sequence.

Let f, g be real-valued functions because $f(x) = g(x) \Leftrightarrow f(x) - g(x) = 0$ finding the roots of $f - g$ is the same as finding the intersection of f and g. So Newton's method can be used to find the intersection of two functions if the conditions of the theorem are met. The fixed points of a function f is the special case $f(x) - x = 0$.

Example 1: Consider $f(x) = \cos(x) - x$. As already noted the roots of this function are also the fixed points of $\cos(x)$. First, note that $\cos\left(\frac{\pi}{6}\right) = \frac{1}{2} > \frac{\pi}{6}$ and that $\cos\left(\frac{\pi}{3}\right) = \frac{\sqrt{3}}{2} < \frac{\pi}{3}$. Thus $f\left(\frac{\pi}{6}\right) > 0 > f\left(\frac{\pi}{3}\right)$. So we know there is a root between $\frac{\pi}{6}$ and $\frac{\pi}{3}$. Therefore, it is a good idea to let $x_0 = \frac{\pi}{4}$.

To apply Newton's method we need $f'(x) = -\sin(x) - 1$.

$$\begin{aligned} x_1 &= x_0 - \frac{f(x_0)}{f'(x_0)} \\ &= \frac{\pi}{4} - \frac{\cos\left(\frac{\pi}{4}\right) - \frac{\pi}{4}}{-\sin\left(\frac{\pi}{4}\right) - 1} \\ &= \frac{\pi}{4} - \frac{\frac{\sqrt{2}}{2} - \frac{\pi}{4}}{-\frac{\sqrt{2}}{2} - 1} \\ &\approx 0.7395361337 \end{aligned}$$

$$\begin{aligned} x_2 &= x_1 - \frac{f(x_1)}{f'(x_1)} \\ &\approx 0.7390851781 \\ x_3 &= x_2 - \frac{f(x_2)}{f'(x_2)} \\ &\approx 0.7390851332 \\ x_4 &= x_3 - \frac{f(x_3)}{f'(x_3)} \\ x_4 &\approx 0.7390851332 \end{aligned}$$

This is a very good approximation to our root and it got there very quickly. Because x_4 and x_3 agree it is reasonable to think that the error is less than 10^{-10}.

There are certain instances where Newton's method may not converge. To guarantee convergence the derivative of the function needs to be continuous in a neighborhood around the root, the derivative at any of the estimations cannot be zero (to avoid division by zero), and the initial estimate has to be close enough to the root. What follows is an example of the last case not being met.

Example 2: Consider the function $f(x) = \frac{x}{x^2+\pi^2}$. Clearly, $f(x) > 0$ for all $x > 0$ and $f(x) < 0$ for all $x < 0$. So, the only root of f is at $x = 0$. This function has a derivative, $f'(x) = \frac{\pi^2 - x^2}{(\pi^2+x^2)^2}$, that is continuous on $\mathbb{R}$. Also, $f'(x) > 0$ for all $x \in [-3, 3]$. However, if we let $x_0 = 3$ then terrible things happen. Note that

$$\frac{f(x)}{f'(x)} = \frac{x}{\pi^2 + x^2} \frac{\left(\pi^2 + x^2\right)^2}{\pi^2 - x^2} = \frac{\pi^2 x + x^3}{\pi^2 - x^2}$$

$$
\begin{aligned}
x_1 &= x_0 - \frac{f\left(x_0\right)}{f'\left(x_0\right)} \\
&= 3 - \frac{\pi^2 3 + 3^3}{\pi^2 - 3^2} \approx -62.0972018223 \\
x_2 &= x_1 - \frac{f\left(x_1\right)}{f'\left(x_1\right)} \approx -124.513095316 \\
x_3 &= x_2 - \frac{f\left(x_2\right)}{f'\left(x_2\right)} \approx -249.184822805
\end{aligned}
$$

It is clear that Newton's method is sending us away from the root very quickly. Because we know the root, we can fairly easily determine the interval from which we can draw a suitable x_0. We want the distance from x_0 to 0 to be greater than the distance from x_1 to 0. In other words,

$$
\begin{aligned}
|x| > \left| x - \frac{\pi^2 x + x^3}{\pi^2 - x^2} \right| &\Rightarrow 1 > \left| 1 - \frac{\pi^2 + x^2}{\pi^2 - x^2} \right| \\
&\Rightarrow 1 > \left| \frac{\pi^2 - x^2}{\pi^2 - x^2} - \frac{\pi^2 + x^2}{\pi^2 - x^2} \right| \\
&\Rightarrow 1 > \left| \frac{-2x^2}{\pi^2 - x^2} \right| = \frac{2\left|x^2\right|}{\left|\pi^2 - x^2\right|} = \frac{2x^2}{\pi^2 - x^2} \\
&\Rightarrow \pi^2 - x^2 > 2x^2 \\
&\Rightarrow \pi^2 > 3x^2 \\
&\Rightarrow |x| < \frac{\pi}{\sqrt{3}} \approx 1.81379936
\end{aligned}
$$

I want to emphasize that we can only do this because we know the root of our function and thus can find the distance to the root. In a real world situation that would not be the case and we would have to use other means to find a small enough interval on which to apply Newton's method. Let's try Newton's method on this function again but with $x_0 = 1.81$.

$$
\begin{aligned}
x_1 &= x_0 - \frac{f\left(x_0\right)}{f'\left(x_0\right)} \\
&= 1.81 - \frac{1.81\pi^2 + 1.81^3}{\pi^2 - 1.81^2} \approx -1.7986614217 \\
x_2 &= x_1 - \frac{f\left(x_1\right)}{f'\left(x_1\right)} \approx 1.7541842043 \\
x_3 &= x_2 - \frac{f\left(x_2\right)}{f'\left(x_2\right)} \approx -1.5893869155 \\
&\dots \\
x_8 &= x_7 - \frac{f\left(x_7\right)}{f'\left(x_7\right)} \approx 0
\end{aligned}
$$

I computed x_8 using the attached MatLab code. So, it converges but very slowly because we were so close to the boundary of the allowable interval. The closer to the root we get the faster it will converge.

Example 3: We can use Newton's method to approximate square roots. We need a function with $\sqrt{a}$ as one of its solutions. The obvious choice is $f(x) = x^2 - a$. Then

$$x_{n+1} = x_n - \frac{x_n^2 - a}{2x_n} = \frac{1}{2}\left(x_n + \frac{a}{x_n}\right).$$

Let's approximate $\sqrt{471}$ using Newton's method.

$$22^2 = 484 > 471 > 441 = 21^2$$

So we will start with $x_0 = 22$.

$$\begin{aligned} x_1 &= \frac{1}{2}\left(x_0 + \frac{a}{x_0}\right) \\ &= \frac{1}{2}\left(22 + \frac{471}{22}\right) \approx 21.70454545 \\ x_2 &= \frac{1}{2}\left(x_1 + \frac{a}{x_1}\right) \approx 21.70253451 \\ x_3 &= \frac{1}{2}\left(x_2 + \frac{a}{x_2}\right) \approx 21.70253441 \\ x_4 &= \frac{1}{2}\left(x_3 + \frac{a}{x_3}\right) \approx 21.70253442 \end{aligned}$$

This is the same formula that Heron, a roman mathematician from around 60 AD, used to calculate the square root of a real number 16 centuries before Newton.

Definition: A mapping $\phi : [a, b] \to [a, b]$ is called a *contraction mapping* if there is some $0 < k < 1$ such that

$$|\phi(x) - \phi(y)| \leq k\,|x - y|$$

for all $x, y \in [a, b]$.

Contraction Mapping Theorem Let $\phi\,[a, b] \to [a, b]$ be a contraction mapping. Then ϕ has a unique fixed point x_*. Moreover, given $x_0 \in [a, b]$, the sequence $\{x_n\}_n^\infty$ defined recursively by

$$x_{n+1} = \phi(x_n)$$

converges to x_*. In particular,

$$|x_n - x_*| \leq \frac{k^n\,|x_0 - x_1|}{1 - k}$$

for each n.

Proof By the definition of contraction mapping

$$\left|x_{n+1}-x_n\right| = \left|\phi\left(x_n\right)-\phi\left(x_{n-1}\right)\right| \leq k\left|x_n-x_{n-1}\right|.$$

It follows easily by induction that $\left|x_{n+1}-x_n\right| \leq k^n\left|x_1-x_0\right|$. Now let $\varepsilon > 0$ and choose N large enough that $k^N < \varepsilon\frac{1-k}{|x_0-x_1|}$. For $n, m \geq N$

$$\begin{aligned}
\left|x_n-x_m\right| &= \left|x_n-x_{n+1}+x_{n+1}-\cdots-x_{m-1}+x_{m-1}-x_m\right| \\
&\leq \left|x_n-x_{n+1}\right|+\cdots+\left|x_{m-1}-x_m\right| \\
&\leq \left|x_1-x_0\right|\sum_{i=n}^{m-1}k^i = k^n\left|x_1-x_0\right|\sum_{i=0}^{m-n-1}k^i \\
&\leq k^n\left|x_1-x_0\right|\sum_{i=0}^{\infty}k^i = k^n\left|x_1-x_0\right|\frac{1}{1-k} \\
&\leq k^N\frac{\left|x_0-x_1\right|}{1-k} < \varepsilon
\end{aligned}$$

Hence, $\left(x_n\right)_n$ is Cauchy, and therefore converges to a point $x_* \in [a,b]$. Note that letting m go to infinity in the above calculuation gives us the result that $\left|x_n-x_*\right| \leq k^n\frac{|x_1-x_0|}{1-k}$. Our function ϕ is a contraction mapping which implies that it is uniformly continuous. So, we can say

$$\phi\left(x_*\right) = \phi\left(\lim_{n\to\infty}x_n\right) = \lim_{n\to\infty}\phi\left(x_n\right) = \lim_{n\to\infty}x_{n+1} = x_*.$$

Giving us that x_* is a fixed point of ϕ.

Finally, if x_{**} is also a fixed point of ϕ then we would have that

$$\left|x_*-x_{**}\right| = \left|\phi\left(x_*\right)-\phi\left(x_{**}\right)\right| \leq k\left|x_*-x_**\right|.$$

This implies that $k \geq 1$ which is a contradiction.

□

The theorem given to us to present is a variation of Newton's method in which f' is replaced with an upper bound of f' in the recursively defined sequence.

Newton's Method Let $f : [a,b] \to \mathbb{R}$ be a differentiable function with $f(a) < 0 < f(b)$ and $0 < m < f'(x) \leq M$ for all $x \in [a,b]$. Given $x_0 \in [a,b]$, the sequence $\left(x_n\right)_n$ defined by

$$x_{n+1} = x_n - \frac{f\left(x_n\right)}{M}$$

converges to the unique root $x_* \in [a,b]$ of the function $f(x)$. In particular,

$$\left|x_n-x_*\right| \leq \frac{\left|f\left(x_0\right)\right|}{m}\left(1-\frac{m}{M}\right)^n$$

for each n.

Proof Define $\phi : [a, b] \to \mathbb{R}$ by $\phi(x) = x - \frac{f(x)}{M}$. We want to show that ϕ is a contraction mapping on $[a, b]$. Since $\phi'(x) = 1 - \frac{f'(x)}{M}$, we see that

$$0 \leq \phi'(x) \leq 1 - \frac{m}{M} = k < 1,$$

so ϕ is a nondecreasing function. Therefore $a < a - \frac{f(a)}{M} = \phi(a) \leq \phi(x) \leq \phi(b) = b - \frac{f(b)}{M}$ for all $x \in [a, b]$, because $f(a) < 0 < f(b)$. Thus $\phi([a, b]) = [a, b]$ and ϕ is a contraction mapping of $[a, b]$.

Therefore, by the Contraction Mapping Theorem, ϕ has a unique fixed point

$$x_* = \phi(x_*) = x_* - \frac{f(x_*)}{M},$$

which is clearly the unique root of $f(x)$ and the sequence (x_n) defined by

$$x_{n+1} = \phi(x_n).$$

converges to x_*. Moreover

$$|x_n - x_*| \leq \frac{k^n |x_0 - x_1|}{1 - k}$$

gives us that

$$|x_n - x_*| \leq \frac{k^n |x_0 - x_1|}{1 - k} = \frac{|f(x_0)|}{m} \left(1 - \frac{m}{M}\right)^n.$$

$\square$