Error Behaviour of Newton's Method

Newton's method is a procedure for finding approximate solutions to equations of the form $f(x) = 0$. The procedure is to

1) Make a preliminary guess x_1.
2) Define $x_2 = x_1 - \frac{f(x_1)}{f'(x_1)}$.
3) Iterate. That is, once you have computed x_n, define $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$.

Newton's method usually works spectacularly well, provided your initial guess is reasonably close to a solution of $f(x) = 0$. A good way to select this initial guess is to sketch the graph of $y = f(x)$. In these notes we shall see why "Newton's method usually works spectacularly well, provided your initial guess is reasonably close to a solution of $f(x) = 0$". We shall assume that there are two numbers $L, M > 0$ such that f obeys:

H1) $\big|f'(x)\big| \ge L$ for all x
H2) $\big|f''(x)\big| \le M$ for all x

Let r be any solution of $f(x) = 0$. Then $f(r) = 0$. Suppose that we have already computed x_n. The error in x_n is $\big|x_n - r\big|$. We now derive a formula that relates the error after the next step, $\big|x_{n+1} - r\big|$, to $\big|x_n - r\big|$. We have seen in class that

$$f(x) = f(x_n) + f'(x_n)(x - x_n) + \tfrac{1}{2} f'(c)(x - x_n)^2$$

for some c between x_n and x. In particular, choosing $x = r$,

$$0 = f(r) = f(x_n) + f'(x_n)(r - x_n) + \tfrac{1}{2} f'(c)(r - x_n)^2 \tag{1}$$

By the definition of x_{n+1},

$$0 = f(x_n) + f'(x_n)(x_{n+1} - x_n) \tag{2}$$

(In fact, we defined x_{n+1} as the solution of $0 = f(x_n) + f'(x_n)(x - x_n)$.) Subtracting (2) from (1).

$$\begin{aligned} 0 = f'(x_n)(r - x_n+1) + \tfrac{1}{2} f''(c)(r - x_n)^2 \ &\Rightarrow\ x_{n+1} - r = \tfrac{f''(c)}{2f'(x_n)}(x_n - r)^2 \\ &\Rightarrow\ \big|x_{n+1} - r\big| = \tfrac{|f''(c)|}{2|f'(x_n)|}|x_n - r|^2 \end{aligned}$$

If the guess x_n is close to r, then c, which must be between x_n and r, is also close to r and $\big|x_{n+1} - r\big| \approx \frac{|f''(r)|}{2|f'(r)|}|x_n - r|^2$. Even if x_n is not close to r, by the hypotheses (H1) and (H2) on the behaviour of f

$$\big|x_{n+1} - r\big| \le \tfrac{M}{2L}|x_n - r|^2 \tag{3}$$

Let's denote by ε_1 the error $|x_1 - r|$ of our initial guess. In fact, let's denote by ε_n the error $|x_n - r|$ in x_n. Then (3) says

$$\varepsilon_{n+1} \le \tfrac{M}{2L}\varepsilon_n^2$$

In particular

$$\begin{aligned}
\varepsilon_2 &\le \tfrac{M}{2L}\varepsilon_1^2 \\
\varepsilon_3 &\le \tfrac{M}{2L}\varepsilon_2^2 \le \tfrac{M}{2L}\left(\tfrac{M}{2L}\varepsilon_1^2\right)^2 &&= \left(\tfrac{M}{2L}\right)^3\varepsilon_1^4 \\
\varepsilon_4 &\le \tfrac{M}{2L}\varepsilon_3^2 \le \tfrac{M}{2L}\left(\left(\tfrac{M}{2L}\right)^3\varepsilon_1^4\right)^2 &&= \left(\tfrac{M}{2L}\right)^7\varepsilon_1^8 \\
\varepsilon_5 &\le \tfrac{M}{2L}\varepsilon_4^2 \le \tfrac{M}{2L}\left(\left(\tfrac{M}{2L}\right)^7\varepsilon_1^8\right)^2 &&= \left(\tfrac{M}{2L}\right)^{15}\varepsilon_1^{16}
\end{aligned}$$

By now we can see a pattern forming, that is easily verified by induction

$$\varepsilon_n \le \left(\tfrac{M}{2L}\right)^{2^{n-1}-1}\varepsilon_1^{2^{n-1}} = \tfrac{2L}{M}\left(\tfrac{M}{2L}\varepsilon_1\right)^{2^{n-1}}$$

As long as $\frac{M}{2L}\varepsilon_1 < 1$ (which tells us quantitatively how good our first guess has to be in order for Newton's method to converge), this goes to zero extremely quickly as n increases. For example, suppose that $\frac{M}{2L}\varepsilon_1 \le \frac{1}{2}$. Then

$$\varepsilon_n \le \tfrac{2L}{M}\left(\tfrac{1}{2}\right)^{2^{n-1}} \le \tfrac{2L}{M}\begin{cases} 0.25 & \text{if } n = 2 \\ 0.0625 & \text{if } n = 3 \\ 0.0039 & \text{if } n = 4 \\ 0.000015 & \text{if } n = 5 \\ 0.00000000023 & \text{if } n = 6 \\ 0.000000000000000000054 & \text{if } n = 7 \end{cases}$$

Each time you increase n by one, the number of zeroes after the decimal place roughly doubles.