Some Notes on Fixed-point Iterations and Newton's Method

This note is meant to clarify two points that are not fully explained in Kreyszig's book. The first concerns the following:

Theorem 1 (from Section 19.2 of Kreyszig, 9th or 10 edition) *Let $g : \mathbb{R} \longrightarrow \mathbb{R}$ and let $x = s$ be a solution of $g(x) = x$. Suppose g has a continuous derivative in some interval J containing s. If, there is a number $K > 0$ such that $|g'(x)| \leq K < 1$ for each $x \in J$, then the iteration process*

$$x_{n+1} = g(x_n), \qquad n = 0, 1, 2, \ldots,$$

converges for any $x_0 \in J$. The limit of the sequence $\{x_n\}$ is s.

Many students have raised this question: if the interval J is as in Theorem 1 and if s_1 and s_2 are two points in J that are solutions of $g(x) = x$, then which of the solutions $\{s_1, s_2\}$ does the s above refer to? The answer to this question, which is not clear from the text of Section 19.1, is that if g satisfies the conditions of Theorem 1 on J, and J contains a solution of the equation $g(x) = x$, then *there is exactly one solution in J*. Let us state this precisely:

Theorem 2 *Let $g : \mathbb{R} \longrightarrow \mathbb{R}$ and let $x = s$ be a solution of $g(x) = x$. Suppose g is differentiable in some interval J containing s. Suppose there is a number $K > 0$ such that $|g'(x)| \leq K < 1$ for each $x \in J$. Then, s is the only solution of $g(x) = x$ in J.*

Proof. Suppose there are two different points s_1 and s_2 in J that are solutions of $g(x) = x$. We may assume that $s_1 < s_2$. Then, g satisfies the conditions for applying the Mean Value Theorem on the interval $[s_1, s_2]$. Thus, there exists a $t \in (s_1, s_2)$ such that

$$g'(t) = \frac{g(s_2) - g(s_1)}{s_2 - s_1} = 1 > K.$$

The second equality occurs because $g(s_1) = s_1$ and $g(s_2) = s_2$, by assumption. Hence, we have found a $t \in J$ such that $|g'(t)| > K$. This contradicts the conditions on g. Thus, our assumption must be false, and there is only one solution. $\square$

The second point we would like to clarify is the issue of whether or not Newton's Method converges. The theorem on page 802 in the 10th edition of Kreyszig (page 793 in the 9th edition of Kreyszig) implies that if s is a solution of $f(x) = 0$ such that $f'(s) \neq 0$ and $f''(s) \neq 0$, then Newton's method converges when we choose x_0 sufficiently close to s. This is a bit imprecise; how close is "sufficiently close" ?

The key to answering this question is to realise that Newton's method is a special type of fixed-point iteration. Remember that the iteration rule in Newton's method is

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \qquad n = 0, 1, 2, \dots. \tag{1}$$

Let us call a solution s of $f(x) = 0$ for which $f'(s) \neq 0$ a "simple root". Let us call the right-hand side of equation (1) $g(x_n)$ Then, we see that s is a simple root of $f(x) = 0$ if $x = s$ is a solution of $g(x) = x$, and vice versa. This is why the search for the simple roots of f by *Newton's method is a special type of fixed-point iteration.*

Assume that for a function $f : \mathbb{R} \longrightarrow \mathbb{R}$, f' and f'' exist on an interval J. Then, with $g(x) = x - \frac{f(x)}{f'(x)}$, we compute

$$g'(x) = \frac{f(x)\, f''(x)}{\cancel{f'(x)} f'(x)^2}, \qquad x \in J.$$

Since Newton's method is a fixed-point iteration given by the rule (1), Theorem 1 tells us:

Theorem 3 *Let $f : \mathbb{R} \longrightarrow \mathbb{R}$ and suppose f' and f'' exist and are continuous. Let s be a root of $f(x) = 0$ such that $f'(s) \neq 0$. Fix a number $K \in (0, 1)$. Suppose J is an interval around s such that $|f(x)\, f''(x)/f'(x)^2| \leq K$ for each $x \in J$. Then, for any $x_0 \in J$, the sequence $\{x_n\}$ given by Newton's method, i.e. by (1) above, converges to s.*

Proof. We have discussed that Newton's method is a fixed-point iteration with

$$g(x) = x - \frac{f(x)}{f'(x)}, \qquad g'(x) = \frac{f(x)\, f''(x)}{\cancel{f'(x)} f'(x)^2}.$$

Let us choose and *fix* a $K \in (0, 1)$. Notice that $g'(s) = 0$. By our assumptions on f, g' is continuous wherever $f'(x) \neq 0$. So, we can certainly find an interval around s on which $|g'(x)| \leq K < 1$. Call this interval J. We see that all the assumptions of Theorem 1 are satisfied by the above choice of g for each x in J. Thus, the present theorem follows from Theorem 1. $\square$