# Numerical Solution of Equations - Fixed Point Iteration

(16 pages; 27/3/20)

(1) Suppose that we wish to solve the equation $x = g(x)$ approximately.

Let our first estimate be $x_0$, and let $x_1 = g(x_0)$, $x_2 = g(x_1)$ etc.

If it happens that the sequence of $x_r$ converges on a particular value $\alpha$, then $\alpha = \mathrm{g}(\alpha)$ and $\alpha$ will be a solution of the equation.

(2) If we wish to solve the equation $x^3 - x = 1$ numerically, it can be rearranged into the following forms (for example):

(A) $x = x^3 - 1$

(B) $x = \sqrt[3]{x+1}$

(C) $x = \frac{1}{x} + \frac{1}{x^2}$

(by dividing the original equation by $x^2$ to give $x - \frac{1}{x} = \frac{1}{x^2}$ )

(3) The iterations can be carried out by calculator.

We need to find a suitable starting point $x_0$. To do this we can employ the Change of Sign method, for example; demonstrating (with a bit of trial and error) that for $f(x) = x^3 - x - 1$, $f(1) < 0$ and $f(2) > 0$, so that a solution lies between 1 and 2. Note that other solutions may exist elsewhere.

So let $x_0 = 1$, say.

For most Casio models, type in the following for (B):

1 = ANS DEL

$(ANS + 1)^{1/3}$

= [repeatedly]

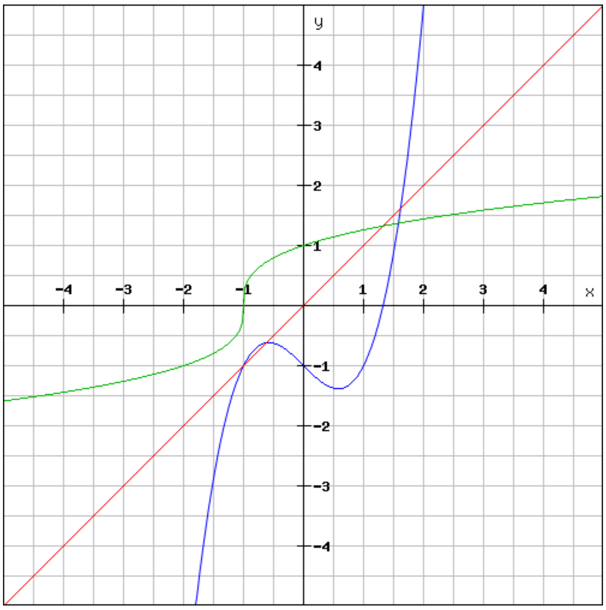
The iterations are as follows (to 5dp), with the results shown for (A) and (C) as well:

| (A) | (B) | (C) |
|---|---|---|
| 1 | 1 | 1 |
| 0 | 1.25992 | 2 |
| -1 | 1.31229 | 0.75 |
| -2 | 1.32235 | 3.11111 |
| -9 | 1.32427 | 0.42475 |
| -730 | 1.32463 | 7.89734 |
| | 1.32470 | 0.14266 |
| | 1.32472 | 56.14607 |
| | 1.32472 | |

Thus only (B) converges.

(4) The diagram below shows the graphs of $y = x^3 - x - 1$ (crossing the $x$-axis at 1.32472), $y = \sqrt[3]{x+1}$ (in green)

and $y = x$.

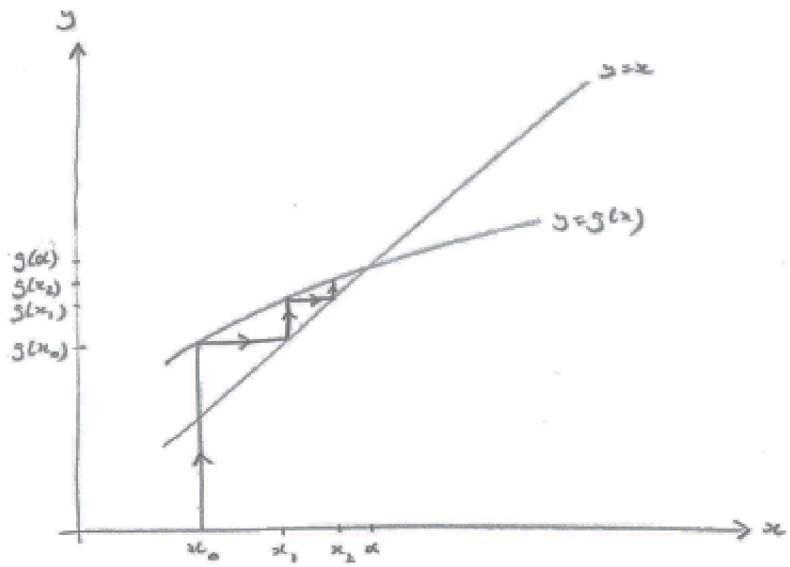



Note that the graphs of $y = \sqrt[3]{x+1}$ and $y = x$ intersect at $x = 1.32472$, corresponding to the solution of $x = g(x)$ for (B).

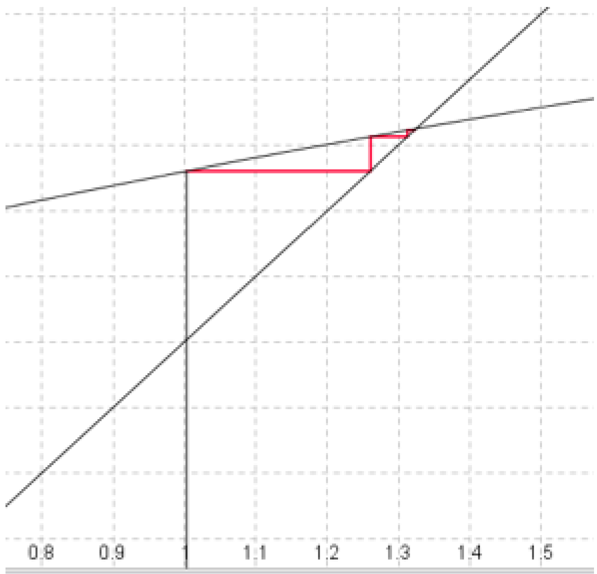
(5) The diagram below shows how the iterations for (B) approach $\alpha = 1.32472$

Starting with $x = x_0$, we obtain $x_1 = g(x_0)$ by following the vertical line up to the curve $y = g(x)$; $x_1$ is then the $y$-coordinate of the point reached; this is then turned into a point on the $x$-axis by following the horizontal line along to the line $y = x$; the

$x$-coordinate of the point reached is $x_1 = g(x_0)$; the process is then repeated to find $x_2$ etc (though the vertical lines don't need to start on the $x$-axis).




For this particular series of iterations, the $x_r$ approach $\alpha$ from below in a 'staircase' fashion.




$$g(x) = \sqrt[3]{x+1}$$

(6) We will now show that the convergence (or otherwise) of iterations of $x = g(x)$ depends on the value of $g'(\alpha)$.

From the diagram in (5), provided that $x_r$ is reasonably close to $\alpha$,

$g'(\alpha) \approx \frac{g(\alpha)-g(x_r)}{\alpha-x_r} = \frac{\alpha-g(x_r)}{\alpha-x_r} = \frac{g(x_r)-\alpha}{x_r-\alpha}$ (1),

using the fact that $\alpha$ is a solution of $x = g(x)$, so that $\alpha = g(\alpha)$.

Let $e_r = x_r - \alpha$ be the 'error' associated with $x_r$.

[Note: some textbooks define $e_r$ as $\alpha - x_r$]

Then $e_{r+1} = x_{r+1} - \alpha = g(x_r) - \alpha \approx g'(\alpha)\,(x_r - \alpha)$, from (1)

Thus $e_{r+1} \approx g'(\alpha)\,.\,e_r$

This is an example of what is termed **1st order convergence**,

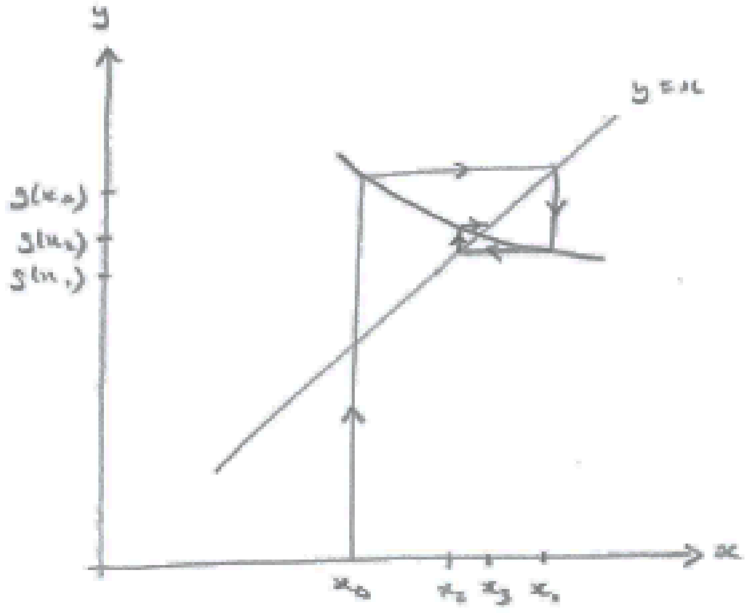
where each error term is proportional to the previous error term.

So, if the error is supposed to be getting smaller, we want $|g'(\alpha)| < 1$ (and the smaller the better).

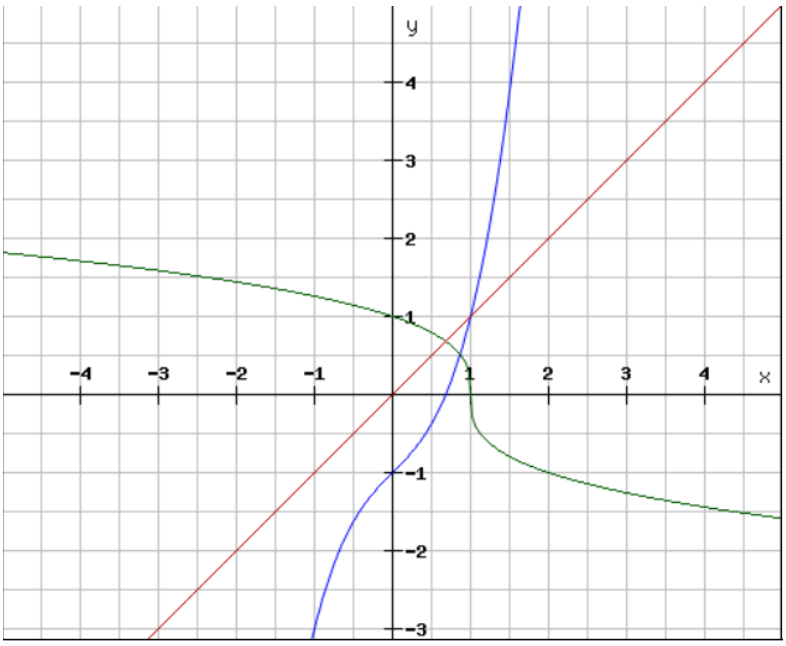
Referring to the diagram in (5), we can see whether $g'(\alpha) < 1$, by comparing the slopes of $y = g(x)$ and $y = x$ at $x = \alpha$. By imagining the line perpendicular to $y = x$ at $x = \alpha$, we can establish whether $g'(\alpha) > -1$.

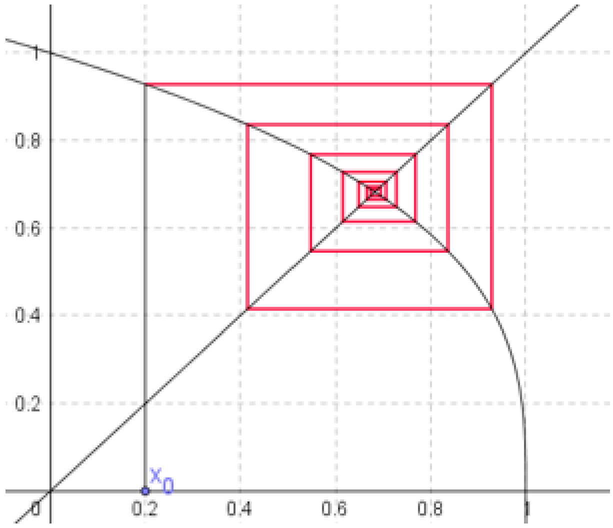
In the case of (B) above, $0 < g'(\alpha) < 1$, so that each error $e_r$ is smaller than the previous one, and of the same sign. This gives rise to the staircase pattern.

In cases where $-1 < g'(\alpha) < 0$, the signs of the errors alternate, and a 'cobweb' pattern is obtained, as shown below.



An example of this occurs for the solution of $x^3 + x - 1 = 0$, when $g(x) = \sqrt[3]{1-x}$, as shown below:






$$g(x) = \sqrt[3]{1-x}$$

(7) The above condition for convergence can be applied without having to draw any graphs, if we have a rough estimate of the solution.

Thus, for the rearrangements (A), (B) and (C), we can obtain $g'(1)$, for example. The accurate figure of $g'(1.32472)$ is also shown, for comparison.

(A): $g'(x) = 3x^2 \Rightarrow g'(1) = 3 \;\; \& \;\; g'(1.32472) = 5.26$

(B): $g'(x) = \frac{1}{3}\,(x+1)^{-\frac{2}{3}} \Rightarrow g'(1) = 0.21 \;\; \& \;\; g'(1.32472) = 0.19$

(C): $g'(x) = -x^{-2} - 2x^{-3} \Rightarrow g'(1) = -3 \;\; \& \;\; g'(1.32472) = -1.43$

Thus, in the case of (B), the fact that $-1 < 0.21 < 1$ strongly suggests convergence, whilst convergence is very unlikely in the cases of (A) and (C).

(8) The fixed point method provides an automatic interval for the solution where $-1 < g'(\alpha) < 0$; namely $(x_{r-1}, x_r)$ or $(x_r, x_{r-1})$, depending on whether $x_{r-1} < x_r$.

In order to obtain an interval for cases where $0 < g'(\alpha) < 1$, we take $x_r$ as one of the bounds (lower or upper, depending on whether $x_r < \alpha$) and then estimate a suitable value for the other bound, checking it by the Change of Sign method.

Thus for (B) above, we could take 1.324 as the lower bound, and 1.325, as $f(1.324) = 1.324^3 - 1.324 - 1 = -0.0031 < 0$ and $f(1.325) = 1.325^3 - 1.325 - 1 = 0.0012 > 0$

(9) Number of iterations needed

For (B), with $x_0 = 1$, it is clear after a few iterations that

$\alpha \approx 1.32$, so that $e_0 \approx 1 - 1.32 = -0.32$

Also $g'(1.32) = 0.19\ (2dp)$ and $e_{r+1} \approx g'(\alpha)\,.\,e_r$ ,

so that $e_r \approx -0.32(0.19)^r$

Thus, if $r = 8$, $e_r \approx -5 \times 10^{-7}$

The actual value of $e_r$ is $1.324715 - 1.324717957$

$= -3 \times 10^{-6}$ (6dp), so the approximation is not especially good.

If we want to find the approximate number of iterations needed in order to achieve a particular level of error, then we can take logs.

Thus, to achieve an error of approximately $-5 \times 10^{-6}$ (giving an answer approximately correct to 5dp):

$-5 \times 10^{-6} = -0.32(0.19)^r$

$\Rightarrow 1.5625 \times 10^{-5} = (0.19)^r$

$\Rightarrow log_{10}(1.5625) - 5 = r log_{10}(0.19)$

$\Rightarrow r = 6.7$

ie 7 iterations will be required.

(10) Quick way of estimating $\alpha$

As $e_{r+1} \approx g'(\alpha) \,.\, e_r$,

$x_{r+1} - \alpha \approx g'(\alpha)(x_r - \alpha)$

Then , given $x_r$ & $x_{r+1}$, $\alpha$ can be found approximately if an estimate of $g'(\alpha)$ is available.

To find $g'(\alpha)$:

(a) base on an estimate of $\alpha$ (as in (9))

or (b) find an approximation to the gradient at $\alpha$:

$g'(\alpha) \approx \frac{g(x_{r+1})-g(x_r)}{x_{r+1}-x_r} = \frac{x_{r+2}-x_{r+1}}{x_{r+1}-x_r}$

This is referred to as the 'ratio of differences', and a more accurate value for $g'(\alpha)$ can be obtained by increasing $r$.

| $x_r$ for (B) | $x_{r+1} - x_r$ | $\frac{x_{r+2} - x_{r+1}}{x_{r+1} - x_r}$ |
|---|---|---|
| 1 | 0.259921 | |
| 1.25992 | 0.052373 | 0.201495 |
| 1.31229 | 0.010060 | 0.192084 |
| 1.32235 | 0.001915 | 0.190351 |
| 1.32427 | 0.000364 | 0.190023 |
| 1.32463 | 0.000069 | 0.189961 |
| 1.32470 | 0.000013 | 0.189950 |
| 1.32472 | 0.000002 | |
| 1.32472 | | |

The ratio of differences can of course be presented in the form

$\frac{x_{r+1}-x_r}{x_r-x_{r-1}}$ instead (where the value of $r$ has been increased by 1).

(11) Alternative derivation of the ratio of differences result:

$x_r = \alpha + e_r$ and $e_r \approx ke_{r-1}$

$\frac{x_{r+1}-x_r}{x_r-x_{r-1}} = \frac{(\alpha+e_{r+1})-(\alpha+e_r)}{(\alpha+e_r)-(\alpha+e_{r-1})}$

$= \frac{e_{r+1}-e_r}{e_r-e_{r-1}} = \frac{ke_r-e_r}{ke_{r-1}-e_{r-1}}$

$= \frac{(k-1)e_r}{(k-1)e_{r-1}} \approx k$

ie $\frac{x_{r+1}-x_r}{x_r-x_{r-1}} \approx k$

(12) What if $g'(\alpha) = 0$?

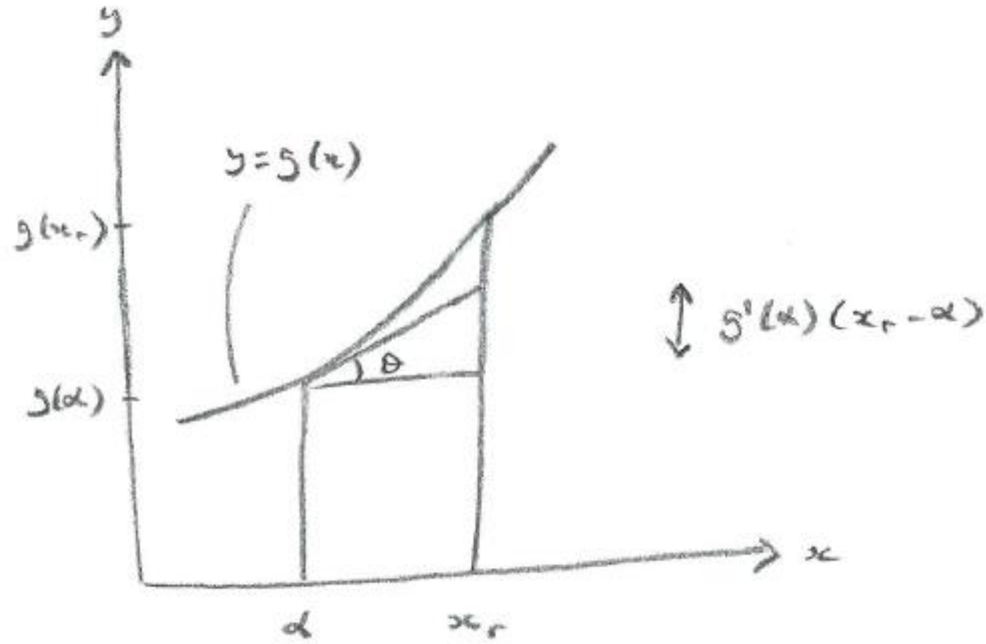
$e_{r+1} \approx g'(\alpha) . e_r$ becomes $e_{r+1} \approx 0$, which isn't very helpful.

$g'(\alpha) \approx \frac{g(\alpha)-g(x_r)}{\alpha-x_r}$ can be written as

$g(x_r) \approx g(\alpha) + g'(\alpha)(x_r - \alpha)$ (A)

Referring to the diagram below, an approximate value for $g(x_r)$ is obtained by adding $tan\theta(x_r - \alpha)$ to $g(\alpha)$, where $tan\theta = g'(\alpha)$

[in this case, it gives an underestimate]




(A) gives the first two terms of the Taylor expansion of $g(x_r)$ about $\alpha$

The 3rd term is $g''(\alpha)\frac{(x_r-\alpha)^2}{2!}$

Then if $g'(\alpha) = 0$,

$g(x_r) \approx g(\alpha) + g''(\alpha)\frac{(x_r-\alpha)^2}{2!}$

so that $e_{r+1} = x_{r+1} - \alpha = g(x_r) - g(\alpha) \approx \frac{1}{2}g''(\alpha){e_r}^2$ (B)

Thus each error term is proportional to the square of the previous error term. This is referred to as **quadratic or 2nd order convergence**.

From (B), if $g''(\alpha)>0$, the errors after $e_0$ will all be positive; ie giving a staircase pattern. If $g''(\alpha)<0$, the errors after $e_0$ will all be negative; again giving a staircase pattern.

(13) Relaxed iteration

Consider $h(x) = (1 - \lambda)x + \lambda g(x)$

$h(\alpha) = (1 - \lambda)\alpha + \lambda g(\alpha) = (1 - \lambda)\alpha + \lambda\alpha = \alpha$

Thus, $h(x)$ fulfills the same role as $g(x)$ and, with a suitable $\lambda$, it may be possible to convert an unfavourable $g(x)$ , where there isn't convergence, to a favourable $h(x)$, where there is convergence. Where there is already convergence for $g(x)$, it may be possible to obtain faster convergence with $h(x)$.

A spreadsheet can be used to do this, as shown below (with $k$ instead of $\lambda$).

B6 =(1-B$2)*A6+B$2*((A6^3)-1)

| | A | B | C | D | E | F | G | H |
|---|---|---|---|---|---|---|---|---|
| 1 | (A) g(x) = x^3 - 1 | | | (A) g(x) = x^3 - 1 | | | (A) g(x) = x^3 - 1 | |
| 2 | k | 0.2 | | k | -0.1 | | k | -0.3 |
| 3 | | | | | | | | |
| 4 | x | h(x) | | x | h(x) | | x | h(x) |
| 5 | 1 | 0.8 | | 1 | 1.1 | | 1 | 1.3 |
| 6 | 0.8 | 0.5424 | | 1.1 | 1.1769 | | 1.3 | 1.3309 |
| 7 | 0.5424 | 0.265835 | | 1.1769 | 1.231578 | | 1.3309 | 1.322945 |
| 8 | 0.265835 | 0.016425 | | 1.231578 | 1.267932 | | 1.322945 | 1.32521 |
| 9 | 0.016425 | -0.18686 | | 1.267932 | 1.290886 | | 1.32521 | 1.32458 |
| 10 | -0.18686 | -0.35079 | | 1.290886 | 1.304863 | | 1.32458 | 1.324756 |
| 11 | -0.35079 | -0.48927 | | 1.304863 | 1.313175 | | 1.324756 | 1.324707 |
| 12 | -0.48927 | -0.61484 | | 1.313175 | 1.318045 | | 1.324707 | 1.324721 |
| 13 | -0.61484 | -0.73836 | | 1.318045 | 1.320873 | | 1.324721 | 1.324717 |
| 14 | -0.73836 | -0.87119 | | 1.320873 | 1.322507 | | 1.324717 | 1.324718 |
| 15 | -0.87119 | -1.02919 | | 1.322507 | 1.323448 | | 1.324718 | 1.324718 |
| 16 | -1.02919 | -1.24139 | | 1.323448 | 1.323989 | | 1.324718 | 1.324718 |
| 17 | -1.24139 | -1.57572 | | 1.323989 | 1.3243 | | 1.324718 | 1.324718 |
| 18 | -1.57572 | -2.24304 | | 1.3243 | 1.324478 | | 1.324718 | 1.324718 |
| 19 | -2.24304 | -4.25147 | | 1.324478 | 1.32458 | | 1.324718 | 1.324718 |
| 20 | -4.25147 | -18.9702 | | 1.32458 | 1.324639 | | 1.324718 | 1.324718 |

Alternatively, suppose that we are able to estimate $\alpha$, and hence calculate $g'(\alpha)$ approximately.

Then $h'(\alpha) = 0 \Rightarrow (1 - \lambda) + \lambda g'(\alpha) = 0$

$\Rightarrow \lambda(g'(\alpha) - 1) = -1$

$\Rightarrow \lambda = \frac{1}{1-g'(\alpha)}$

For example, we can consider the three earlier rearrangements of $x^3 - x = 1$, and assume that it is known that a root lies between 1 and 2, with $x_0 = 1$.

(A) $x = x^3 - 1$ ; $g'(1) = 3$; $\lambda = \frac{1}{1-g'(1)} = -0.5$

(B) $x = \sqrt[3]{x+1}$ ; $g'(1) = 0.21$; $\lambda = \frac{1}{1-g'(1)} = 1.26582$

(C) $x = \frac{1}{x} + \frac{1}{x^2}$ ; $g'(1) = -3$; $\lambda = \frac{1}{1-g'(1)} = 0.25$

Convergence, or otherwise, of $h(x)$ can now be investigated for the calculated value of $\lambda$, using a spreadsheet:

| (A) | | | (B) | | | (C) | |
|---|---|---|---|---|---|---|---|
| x | h(x) | | x | h(x) | | x | h(x) |
| 1 | 1.5 | | 1 | 1.329014 | | 1 | 1.25 |
| 1.5 | 1.0625 | | 1.329014 | 1.324608 | | 1.25 | 1.2975 |
| 1.0625 | 1.494019 | | 1.324608 | 1.324721 | | 1.2975 | 1.314303 |
| 1.494019 | 1.073635 | | 1.324721 | 1.324718 | | 1.314303 | 1.320669 |
| 1.073635 | 1.491667 | | 1.324718 | 1.324718 | | 1.320669 | 1.323135 |
| 1.491667 | 1.077968 | | | | | 1.323135 | 1.324097 |
| 1.077968 | 1.490645 | | | | | 1.324097 | 1.324475 |
| 1.490645 | 1.079845 | | | | | 1.324475 | 1.324622 |
| 1.079845 | 1.490183 | | | | | 1.324622 | 1.32468 |
| 1.490183 | 1.080691 | | | | | 1.32468 | 1.324703 |
| 1.080691 | 1.489971 | | | | | 1.324703 | 1.324712 |
| 1.489971 | 1.081079 | | | | | 1.324712 | 1.324716 |
| 1.081079 | 1.489873 | | | | | 1.324716 | 1.324717 |
| 1.489873 | 1.081258 | | | | | 1.324717 | 1.324718 |
| 1.081258 | 1.489827 | | | | | 1.324718 | 1.324718 |

As can be seen, the calculated values of $\lambda$ are effective for (B) and (C), but not for (A). In the case of (B), there was already convergence with $g(x)$, but $h(x)$ gives a faster convergence. In the case of (C), there previously wasn't any convergence.

For (A), the situation can be improved by taking the average of $g'(1) = 3$ and $g'(2) = 12$ (as the root lies between 1 and 2).

Then we can take $\lambda = \frac{1}{1-\frac{1}{2}(3+12)} = -0.28571$

This gives the following convergence:

| x | h(x) |
|---|---|
| 1 | 1.285714 |
| 1.285714 | 1.331529 |
| 1.331529 | 1.323177 |
| 1.323177 | 1.325052 |
| 1.325052 | 1.324645 |
| 1.324645 | 1.324734 |
| 1.324734 | 1.324714 |
| 1.324714 | 1.324719 |
| 1.324719 | 1.324718 |
| 1.324718 | 1.324718 |
| 1.324718 | 1.324718 |

(14) Exercise: Define $h(x) \equiv \lambda g(x) + (1-\lambda)x$

Show that $g(\alpha) = \alpha \Leftrightarrow h(\alpha) = \alpha$, provided one condition is met, and state that condition.

**Solution**

$g(\alpha) = \alpha \Rightarrow h(\alpha) = \lambda g(\alpha) + (1-\lambda)\alpha$

$= \lambda\alpha + (1-\lambda)\alpha = \alpha$

and $h(\alpha) = \alpha \Rightarrow \lambda g(\alpha) + (1-\lambda)\alpha = \alpha$

$\Rightarrow \lambda(g(\alpha) - \alpha) = 0$

$\Rightarrow g(\alpha) = \alpha$ , provided $\lambda \neq 0$

[If $\lambda = 0$, then $h(x) \equiv x$, so that $h(\alpha) = \alpha$ is always true.]

(15) Exercise: By employing the relaxed iteration

$h(x) = \lambda g(x) + (1-\lambda)x$ , where $g(x) = x^3 - 1$, with a suitable value of $\lambda$, find the root of the equation

$x^3 - x - 1 = 0$ that lies between 1.3 and 1.4, to 4 d.p.

**Solution**

Let $\alpha$ be the required root.

$g(x) = x^3 - 1 \Rightarrow g'(x) = 3x^2$

$\alpha \approx 1.3$ and $g'(1.3) = 5.07$

Let $\lambda = \frac{1}{1-5.07} = -0.2457$

Then $h(x) = -0.2457(x^3 - 1) + 1.2457x$

and with $x_{r+1} = -0.2457({x_r}^3 - 1) + 1.2457x_r$ and $x_0 = 1.3$,

$x_1 = 1.32531$,

$x_2 = 1.32469$,

$x_3 = 1.32472$,

$x_4 = 1.32472$

so that the root is 1.3247 (4 d.p.)