2.5 A General Theory for One-Point Iteration Methods

AN INTRODUCTION TO NUMERICAL ANALYSIS

Second Edition

Kendall E. Atkinson

University of Iowa

WILEY

John Wiley & Sons

New York Chichester Brisbane Toronto Singapore

Table 2.5 Muller's method, example 3

IT	Root	f(root)
41	2.9987526 −6.98E − 4i	−3.33E − 11 +6.70E − 11i
17	2.9997591 −2.68E − 4i	5.68E − 14 −6.48E − 14i
31	3.0003095 −3.17E − 4i	−3.41E − 13 +3.22E − 14i
10	3.0003046 +3.14E − 4i	3.98E − 13 −3.83E − 14i
6	5.97E − 15 +3.000000000i	4.38E − 11 −1.19E − 11i
3	5.97E − 15 −3.000000000i	4.38E − 11 +1.19E − 11i

***Example* 1.** $f(x) = x^{20} - 1$. All 20 roots were found with an accuracy of 10 or more significant digits. In all cases, the approximate root z satisfied $|f(z)| < 10^{-10}$, generally much less. The number of iterates ranged from 1 to 18, with an average of 8.5.

2. $f(x)$ = Laguerre polynomial of degree 12. The real parts of the roots as shown are correct, rounded to the number of places shown, but the imaginary parts should all be zero. The numerical results are given in Table 2.4. Note that $f(x)$ is quite large for many of the approximate roots.

3.

$$f(x) = x^6 - 12x^5 + 63x^4 - 216x^3 + 567x^2 - 972x + 729$$

$$= (x^2 + 9)(x - 3)^4$$

The numerical results are given in Table 2.5. Note the inaccuracy in the first four roots, which is inherent due to the noise in $f(x)$ associated with $\alpha = 3$ being a repeated root. See Section 2.7 for a complete discussion of the problems in calculating repeated roots.

The last two examples demonstrate why two error tests are necessary, and they indicate why the routine requests a maximum on the number of iterations to be allowed per root. The form (2.4.2) of Muller's method is due to Traub (1964, pp. 210–213). For a computational discussion, see Whitley (1968).

2.5 A General Theory for One-Point Iteration Methods

We now consider solving an equation $x = g(x)$ for a root α by the iteration

$$x_{n+1} = g(x_n) \qquad n \geq 0 \tag{2.5.1}$$

with x_0 an initial guess to α. The Newton method fits in this pattern with

$$g(x) \equiv x - \frac{f(x)}{f'(x)} \tag{2.5.2}$$

Table 2.6 Iteration examples for $x^2 - 3 = 0$

	Case (i)	Case (ii)	Case (iii)
n	x_n	x_n	x_n
0	2.0	2.0	2.0
1	3.0	1.5	1.75
2	9.0	2.0	1.732143
3	87.0	1.5	1.732051

Each solution of $x = g(x)$ is called a *fixed point* of g. Although we are usually interested in solving an equation $f(x) = 0$, there are many ways this can be reformulated as a fixed-point problem. At this point, we just illustrate this reformulation process with some examples.

Example Consider solving $x^2 - a = 0$ for $a > 0$.

(i) $x = x^2 + x - a$, or more generally, $x = x + c(x^2 - a)$ for some $c \neq 0$

(ii) $x = \dfrac{a}{x}$

(iii) $x = \dfrac{1}{2}\left(x + \dfrac{a}{x}\right)$ (2.5.3)

We give a numerical example with $a = 3$, $x_0 = 2$, and $\alpha = \sqrt{3} = 1.732051$. With $x_0 = 2$, the numerical results for (2.5.1) in these cases are given in Table 2.6.

It is natural to ask what makes the various iterative schemes behave in the way they do in this example. We will develop a general theory to explain this behavior and to aid in analyzing new iterative methods.

Lemma 2.4 Let $g(x)$ be continuous on the interval $a \leq x \leq b$, and assume that $a \leq g(x) \leq b$ for every $a \leq x \leq b$. (We say g sends $[a, b]$ into $[a, b]$, and denote it by $g([a, b]) \subset [a, b]$.) Then $x = g(x)$ has at least one solution in $[a, b]$.

Proof Consider the continuous function $g(x) - x$. At $x = a$, it is positive, and at $x = b$ it is negative. Thus by the intermediate value theorem, it must have a root in the interval $[a, b]$. In Figure 2.5, the roots are the intersection points of $y = x$ and $y = g(x)$. ∎

Lemma 2.5 Let $g(x)$ be continuous on $[a, b]$, and assume $g([a, b]) \subset [a, b]$. Furthermore, assume there is a constant $0 < \lambda < 1$, with

$$|g(x) - g(y)| \leq \lambda |x - y| \qquad \text{for all} \quad x, y \in [a, b] \tag{2.5.4}$$

Then $x = g(x)$ has a unique solution α in $[a, b]$. Also, the iterates

$$x_n = g(x_{n-1}) \qquad n \geq 1$$

will converge to α for any choice of x_0 in $[a, b]$, and

$$|\alpha - x_n| \leq \frac{\lambda^n}{1-\lambda}|x_1 - x_0| \tag{2.5.5}$$

Proof Suppose $x = g(x)$ has two solutions α and β in $[a, b]$. Then

$$|\alpha - \beta| = |g(\alpha) - g(\beta)| \leq \lambda|\alpha - \beta|$$

$$(1 - \lambda)|\alpha - \beta| \leq 0$$

Since $0 < \lambda < 1$, this implies $\alpha = \beta$. Also we know by the earlier lemma that there is at least one root α in $[a, b]$.

To examine the convergence of the iterates x_n, first note that they all remain in $[a, b]$. To see this, note that the result

$$x_n \in [a, b] \qquad \text{implies} \qquad x_{n+1} = g(x_n) \in [a, b]$$

can be used with mathematical induction to prove $x_n \in [a, b]$ for all n. For the convergence,

$$|\alpha - x_{n+1}| = |g(\alpha) - g(x_n)| \leq \lambda|\alpha - x_n| \tag{2.5.6}$$

and by induction,

$$|\alpha - x_n| \leq \lambda^n|\alpha - x_0| \qquad n \geq 0 \tag{2.5.7}$$

As $n \to \infty$, $\lambda^n \to 0$; thus, $x_n \to \alpha$.

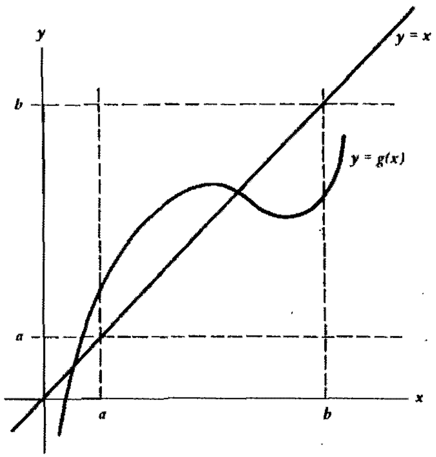

Figure 2.5 Example of Lemma 2.4.

To prove the bound (2.5.5), begin with

$$|\alpha - x_0| \leq |\alpha - x_1| + |x_1 - x_0| \leq \lambda|\alpha - x_0| + |x_1 - x_0|$$

where the last step used (2.5.6). Then solving for $|\alpha - x_0|$, we have

$$|\alpha - x_0| \leq \frac{1}{1-\lambda}|x_1 - x_0| \tag{2.5.8}$$

Combining this with (2.5.7) will complete the proof. ■

The bound (2.5.6) shows that the sequence $\{x_n\}$ is linearly convergent, with the rate of convergence bounded by λ, based on the definition (2.0.13). Also from the proof, we can devise a possibly more accurate error bound than (2.5.5). Repeating the argument that led to (2.5.8), we obtain

$$|\alpha - x_n| \leq \frac{1}{1-\lambda}|x_{n+1} - x_n|$$

Further, applying (2.5.6) yields the bound

$$|\alpha - x_{n+1}| \leq \frac{\lambda}{1-\lambda}|x_{n+1} - x_n| \tag{2.5.9}$$

When λ is computable, this furnishes a practical bound in most situations. Other error bounds and estimates are discussed in the following section.

If $g(x)$ is differentiable on $[a, b]$, then

$$g(x) - g(y) = g'(\xi)(x - y) \qquad \xi \text{ between } x \text{ and } y$$

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

$$\lambda = \underset{a \leq x \leq b}{\text{Max}} |g'(x)|$$

Then

$$|g(x) - g(y)| \leq \lambda|x - y| \qquad \text{all} \quad x, y \in [a, b]$$

Theorem 2.6 Assume that $g(x)$ is continuously differentiable on $[a, b]$, that $g([a, b]) \subset [a, b]$, and that

$$\lambda = \underset{a \leq x \leq b}{\text{Max}} |g'(x)| < 1 \tag{2.5.10}$$

Then

(i) $x = g(x)$ has a unique solution α in $[a, b]$

(ii) For any choice of x_0 in $[a, b]$, with $x_{n+1} = g(x_n)$, $n \geq 0$,

$$\underset{n \to \infty}{\text{Limit}}\, x_n = \alpha$$

(iii)

$$|\alpha - x_n| \le \lambda^n |\alpha - x_0| \le \frac{\lambda^n}{1-\lambda}|x_1 - x_0|$$

$$\operatorname*{Limit}_{n\to\infty} \frac{\alpha - x_{n+1}}{\alpha - x_n} = g'(\alpha) \qquad (2.5.11)$$

Proof Every result comes from the preceding lemmas, except for the rate of convergence (2.5.11). For it, use

$$\alpha - x_{n+1} = g(\alpha) - g(x_n) = g'(\xi_n)(\alpha - x_n) \qquad n \ge 0 \quad (2.5.12)$$

with ξ_n an unknown point between α and x_n. Since $x_n \to \alpha$, we must have $\xi_n \to \alpha$, and thus

$$\operatorname*{Limit}_{n\to\infty} \frac{\alpha - x_{n+1}}{\alpha - x_n} = \operatorname*{Limit}_{n\to\infty} g'(\xi_n) = g'(\alpha)$$

If $g'(\alpha) \ne 0$, then the sequence $\{x_n\}$ converges to α with order exactly $p = 1$, linear convergence. ■

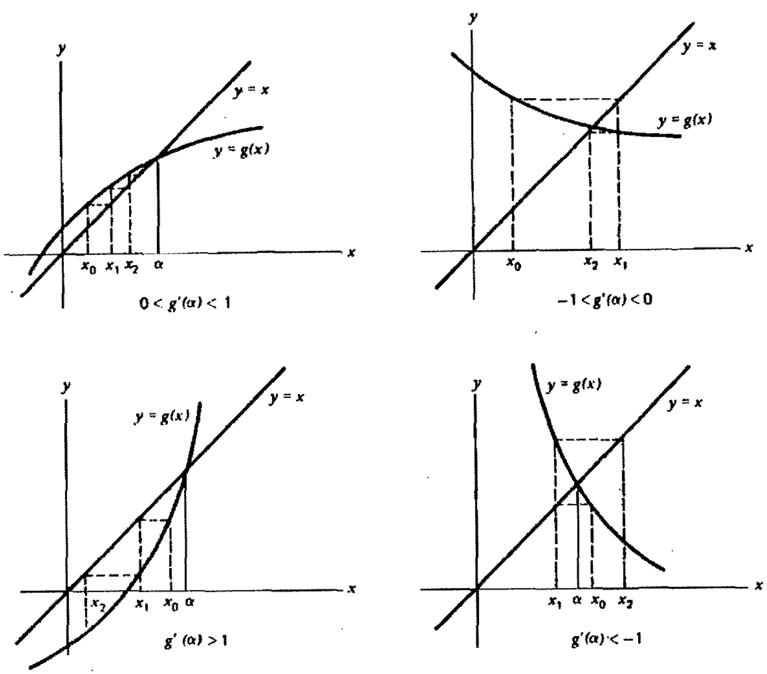

Figure 2.6 Examples of convergent and nonconvergent sequences $x_{n+1} = g(x_n)$.

This theorem generalizes to systems of m nonlinear equations in m unknowns. Just regard x as an element of $\mathbf{R}^m$, $g(x)$ as a function from $\mathbf{R}^m$ to $\mathbf{R}^m$, replace the absolute values by vector and matrix norms, and replace $g'(x)$ by the Jacobian matrix for $g(x)$. The assumption $g([a, b]) \subset [a, b]$ must be replaced with a stronger assumption, and care must be exercised in the choice of a region generalizing $[a, b]$. The lemmas generalize, but they are nontrivial to prove. This is discussed further in Section 2.10.

To see the importance of the assumption (2.5.10) on the size of $g'(x)$, suppose $|g'(\alpha)| > 1$. Then if we had a sequence of iterates $x_{n+1} = g(x_n)$ and a root $\alpha = g(\alpha)$, we have (2.5.12). If x_n becomes sufficiently close to α, then $|g'(\xi_n)| > 1$ and the error $|\alpha - x_{n+1}|$ will be greater than $|\alpha - x_n|$. Thus convergence is not possible if $|g'(\alpha)| > 1$. We graphically portray the computation of the iterates in four cases (see Figure 2.6).

To simplify the application of the previous theorem, we give the following result.

Theorem 2.7 Assume α is a solution of $x = g(x)$, and suppose that $g(x)$ is continuously differentiable in some neighboring interval about α with $|g'(\alpha)| < 1$. Then the results of Theorem 2.6 are still true, provided x_0 is chosen sufficiently close to α.

Proof Pick a number λ satisfying $|g'(\alpha)| < \lambda < 1$. Then pick an interval $I = [\alpha - \epsilon, \alpha + \epsilon]$ with

$$\operatorname*{Max}_{x\in I} |g'(x)| \le \lambda < 1$$

We have $g(I) \subset I$, since $|\alpha - x| \le \epsilon$ implies

$$|\alpha - g(x)| = |g(\alpha) - g(x)| = |g'(\xi)| \cdot |\alpha - x| \le \lambda|\alpha - x| \le \epsilon$$

Now apply the preceding theorem using $[a, b] = [\alpha - \epsilon, \alpha + \epsilon]$. ■

Example Referring back to the earlier example in this section, calculate $g'(\alpha)$.

(i) $g(x) = x^2 + x - 3 \qquad g'(\alpha) = g'(\sqrt{3}) = 2\sqrt{3} + 1 > 1$

(ii) $g(x) = \dfrac{3}{x} \qquad g'(\sqrt{3}) = \dfrac{-3}{(\sqrt{3})^2} = -1$

(iii) $g(x) = \dfrac{1}{2}\left(x + \dfrac{3}{x}\right) \qquad g'(x) = \dfrac{1}{2}\left(1 - \dfrac{3}{x^2}\right) \qquad g'(\sqrt{3}) = 0$

Example For $x = x + c(x^2 - 3)$, pick c to ensure convergence. Since the solution is $\alpha = \sqrt{3}$, and since $g'(x) = 1 + 2cx$, pick c so that

$$-1 < 1 + 2c\sqrt{3} < 1$$

For a good rate of convergence, pick c so that

$$1 + 2c\sqrt{3} \doteq 0$$

Table 2.7 Numerical example of iteration (2.5.13)

n	x_n	$\alpha - x_n$	Ratio
0	2.0	−2.68E − 1	
1	1.75	−1.79E − 2	.0668
2	1.7343750	−2.32E − 3	.130
3	1.7323608	−3.10E − 4	.134
4	1.7320923	−4.15E − 5	.134
5	1.7320564	−5.56E − 6	.134
6	1.7320516	−7.45E − 7	.134
7	1.7320509	−1.00E − 7	.134

This gives

$$c \doteq \frac{-1}{2\sqrt{3}}$$

Use $c = -\frac{1}{4}$. Then $g'(\sqrt{3}) = 1 - (\sqrt{3}/2) \doteq .134$. This gives the iteration scheme

$$x_{n+1} = x_n - \frac{1}{4}(x_n^2 - 3) \qquad n \geq 0 \tag{2.5.13}$$

The numerical results are given in Table 2.7. The ratio column gives the values of

$$\frac{\alpha - x_n}{\alpha - x_{n-1}} \qquad n \geq 1$$

The results agree closely with the theoretical value of $g'(\sqrt{3})$.

Higher order one-point methods We complete the development of the theory for one-point iteration methods by considering methods with an order of convergence greater than one, for example, Newtons' method.

Theorem 2.8 Assume α is a root of $x = g(x)$, and that $g(x)$ is p times continuously differentiable for all x near to α, for some $p \geq 2$. Furthermore, assume

$$g'(\alpha) = \cdots = g^{(p-1)}(\alpha) = 0 \tag{2.5.14}$$

Then if the initial guess x_0 is chosen sufficiently close to α, the iteration

$$x_{n+1} = g(x_n) \qquad n \geq 0$$

will have order of convergence p, and

$$\operatorname*{Limit}_{n\to\infty} \frac{\alpha - x_{n+1}}{(\alpha - x_n)^p} = (-1)^{p-1} \cdot \frac{g^{(p)}(\alpha)}{p!}$$

Proof Expand $g(x_n)$ about α:

$$x_{n+1} = g(x_n) = g(\alpha) + (x_n - \alpha)g'(\alpha) + \cdots + \frac{(x_n - \alpha)^{p-1}}{(p-1)!}g^{(p-1)}(\alpha)$$
$$+ \frac{(x_n - \alpha)^p}{p!}g^{(p)}(\xi_n)$$

for some ξ_n between x_n and α. Using (2.5.14) and $\alpha = g(\alpha)$,

$$\alpha - x_{n+1} = -\frac{(x_n - \alpha)^p}{p!}g^{(p)}(\xi_n)$$

Use Theorem 2.7 and $x_n \to \alpha$ to complete the proof. ■

The Newton method can be analyzed by this result

$$g(x) = x - \frac{f(x)}{f'(x)} \qquad g'(x) = \frac{f(x)f''(x)}{[f'(x)]^2}$$

$$g'(\alpha) = 0 \qquad g''(\alpha) = \frac{f''(\alpha)}{f'(\alpha)}$$

This and (2.5.14) give the previously obtained convergence result (2.2.3) for Newton's method. For other examples of the application of Theorem 2.8, see the problems at the end of the chapter.

The theory of this section is only for one-point iteration methods, thus eliminating the secant method and Muller's method from consideration. There is a corresponding fixed-point theory for multistep fixed-point methods, which can be found in Traub (1964). We omit it here, principally because only the one-point fixed-point iteration theory will be needed in later chapters.

2.6 Aitken Extrapolation for Linearly Convergent Sequences

From (2.5.11) of Theorem 2.6,

$$\operatorname*{Limit}_{x\to\infty} \frac{\alpha - x_{n+1}}{\alpha - x_n} = g'(\alpha) \tag{2.6.1}$$

for a convergent iteration

$$x_{n+1} = g(x_n) \qquad n \geq 0$$

In this section, we concern ourselves only with the case of linear convergence.