Newtonian Approximations to a Zero of a Function.

Autor(en): **Hartmann, Philip**

Objekttyp: **Article**

Zeitschrift: **Commentarii Mathematici Helvetici**

Band (Jahr): **21 (1948)**

PDF erstellt am: **11.09.2024**

Persistenter Link: https://doi.org/10.5169/seals-18613

Newtonian Approximations to a Zero of a Function

By PHILIP HARTMAN, Baltimore (U. S. A.)

This note is concerned with conditions required for the convergence of the Newtonian approximations to a zero of a function, $f(x)$. These approximations are given by the recursion formula

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

where x_0 is arbitrary and $f'(x)$ denotes the derivative of $f(x)$. The convergence of the numbers (1) to a zero of $f(x)$ is usually proved under the assumptions that $f(x)$ has a continuous second derivative, that the derivative $f'(x)$ is bounded away from zero,

$$|f'(x)| \geqslant m > 0\,, \tag{2}$$

and that x_0 is chosen sufficiently near a zero of $f(x)$. For such a proof, see *Runge* [3]. The "sufficiently near to a zero of $f(x)$" is usually defined in terms of upper and lower bounds of $|f'(x)|$ and the least upper bound of $|f''(x)|$.

For reasons of local convexity, there is considerable simplification in case $f(x)$ is a polynomial, cf., e.g., *Fricke* [1]. It may be noted that, in this case, the standard requirement that $f(x)$ have no multiple roots is superfluous.

It is known (cf., e.g., *Ostrowski* [2]) that the convergence statement can be so formulated that the existence of a zero of $f(x)$ is not presupposed and that the condition of the proximity of x_0 to a zero can be replaced by a condition to the effect that $f(x_0)$ be sufficiently small. The "sufficiently small" in this case is defined in terms of bounds of the first and second derivatives of $f(x)$ and the greatest lower bound of the absolute value of $f'(x)$, which is required to be different from zero.

It will be shown below that the sequence of numbers (1) converges to a zero of $f(x)$ even if the standard condition involving the *existence of* *a second derivative is dropped*, that is, if it is only assumed that $f(x)$ has a continuous first derivative satisfying (2), and that x_0 is chosen sufficiently near a zero of $f(x)$. The "sufficiently near" is defined in terms of the "modulus of continuity of $f'(x)$". Also, this condition on x_0 can be transcribed in terms of the smallness of $|f(x_0)|$. Needless to say, the results obtained for the problem in one dimension can be extended to higher dimensions without additional effort, so that only first order partial derivatives are needed for convergence statements.

This raises the question as to the possibility of improving the assumptions still further. The existence of a derivative $f'(x)$ must, of course, be assumed (at least for every x near to, but distinct from, the root) in order to define the recursion formula (1) at all.

As to the other conditions imposed on $f(x)$, a negative result, (ii) below, will show that condition (2) cannot be omitted. In other words, if $f(x)$ is defined and has a continuous derivative on $|x| \leqq 1$ and if $f(0) = 0$, and $f'(x) > 0$ for $x \neq 0$, then the sequence need not converge for all x_0 sufficiently near $x = 0$.

On the other hand, (i) below will show that the continuity of the derivative $f'(x)$ at the zero of $f(x)$ cannot be omitted; that is, if $f(x)$ is defined and possesses a derivative on $|x| \leqq 1$, and if $f(0) = 0$, and $f'(x)$ is continuous for all $x \neq 0$ and satisfies $0 < m < f'(x) < M$ on $|x| \leqq 1$, then the sequence of numbers (1) need not converge for all x_0 sufficiently near $x = 0$.

Thus, in a certain sense, the following theorem is the "best":

(I) *Let* $f(x)$ *be defined on* $|x| \leqq 1$ *and possess a derivative* $f'(x)$ *satisfying* (2). *Let* $f(0) = 0$ *and* $f'(x)$ *be continuous at* $x = 0$. *Then the sequence of numbers defined by* (1) *converges to zero whenever* $|x_0|$ *is sufficiently small.*

Let θ be arbitrarily chosen in the interval $0 < \theta < 1$. In virtue of the continuity[1]) of $f'(x)$ at $x = 0$, there exists a number $\delta > 0$ such that

$$|f'(x^1) - f'(x^2)| \leqq \theta m, \quad \text{whenever} \quad |x^1| \leqq \delta \ \text{and} \ |x^2| \leqq \delta, \qquad (3)$$

[1]) Professor *Ostrowski* has pointed out to me that the proof of the theorem (I) does not use fully the assumption that $f'(x)$ is continuous at $x = 0$ but merely the fact that the oscillation of $f'(x)$ at $x = 0$ is less than m, that is, that (3) holds for some pair of numbers θ and δ, where $0 < \theta < 1$ and $\delta > 0$. On the other hand, in the example constructed to prove (i), the oscillation of $f'(x)$ at $x = 0$ is exactly m, but the Newtonian approximations do not converge for all sufficiently small $|x_0|$.

where m is the positive number occurring in (2). It will be shown that sequence (1) converges to zero whenever

$$|x_0| \leqq \delta. \tag{4}$$

By the mean value theorem,

$$f(0) = f(x_n) - x_n f'(x_n) + x_n(f'(x_n) - f'(\xi_n)), \tag{5}$$

$n = 0, 1, 2, \ldots,$ where ξ_n is a number between 0 and x_n; so that, in particular,

$$|\xi_n| < |x_n|. \tag{6}$$

Dividing (5) by $f'(x_n)$, using (1) and the fact that $f(0) = 0$, one obtains

$$x_{n+1} = x_n(f'(x_n) - f'(\xi_n))/f'(x_n), \quad n = 0, 1, 2, \ldots. \tag{7}$$

Placing $n = 0$ in (7), the inequalities (2), (3), (4) and (6) imply $|x_1| \leqq \theta |x_0|$; in particular, $|x_1| \leqq \delta$, and so, by induction,

$$|x_{n+1}| \leqq \theta |x_n| \leqq \theta^{n+1} |x_0|, \qquad n = 0, 1, \ldots. \tag{8}$$

This completes the proof of (I).

It may be remarked that, in the standard proofs requiring the existence of a continuous second derivative, the estimate (8) can be replaced by

$$K |x_n| / 2m \leqq (K |x_0| / 2m)^{2^n}, \qquad n = 1, 2, \ldots, \tag{9}$$

where m, as above, is a lower bound for $|f'(x)|$ and K is an upper bound for $|f''(x)|$. Actually, the above proof of (8) implies

$$|x_n| \leqq |x_{n-1}| \, \omega(|x_{n-1}|) / m, \qquad n = 1, 2, \ldots, \tag{10}$$

where $\omega(\delta)$ is the modulus of continuity of $f'(x)$ at $x = 0$, that is,

$$\omega(\delta) = \text{l. u. b.} \, |f'(x^1) - f'(x^2)| \quad \text{for} \quad |x^1| \leqq \delta, \quad |x^2| \leqq \delta.$$

Thus, if $f'(x)$ satisfies a Lipschitz condition at $x = 0$ (for instance, if it is differentiable at $x = 0$), then there exists a constant $C > 0$ such that

$$\omega(\delta) \leqq C\,\delta.$$

In this case, (10) becomes

$$| x_n | \leqq C \, | x_{n-1} |^2 / m \,, \qquad n = 1, 2, \ldots,$$

or

$$C \, | x_n | / m \leqq (C \, | x_0 | / m)^{2^n} \,, \qquad n = 1, 2, \ldots \,. \tag{9 bis}$$

In order to avoid a condition involving the proximity of x_0 to a zero of $f(x)$, one can restate (I) as follows:

(II) *Let* $f(x)$ *be defined for* $| x | \leqq a$ *and possess a continuous derivative* $f'(x)$ *satisfying* (2). *Let* $\delta > 0$ *be such that, for some* $\theta < 1$,

$$| f'(x^1) - f'(x^2) | \leqq \theta m \quad \text{whenever} \quad | x^1 - x^2 | \leqq \delta \,. \tag{11}$$

Let x_0 *be any number such that*

$$| f(x_0) | < \delta m \quad \text{and} \quad | x_0 | < a - 2 \delta \,. \tag{12}$$

Then the sequence of numbers (1) *determined by this* x_0 *converges to a zero of* $f(x)$.

It is clear that conditions (12), (2) and the continuity of $f'(x)$ imply the existence of a zero within a distance δ of x_0. The proof is now an immediate consequence of (I).

The two negative results mentioned above will now be set forth.

(i) *If* $f(x)$, *where* $| x | \leqq 1$, *has the properties that* $f(0) = 0$, *that* $f(x)$ *is differentiable for* $| x | \leqq 1$, *that the derivative* $f'(x)$ *is continuous for every* $x \neq 0$ *and is such that there exist constants* m, M *satisfying*

$$M > f'(x) \geqq m > 0 \,, \qquad | x | \leqq 1 \,,$$

then the sequence of numbers (1) *need not converge for all sufficiently small* $| x_0 |$.

Let $\frac{1}{2} < \alpha_0 < \alpha_1 < \cdots$ be a sequence of increasing numbers such that $\alpha_n \to 1$ as $n \to \infty$. Let $a_n = 2^{-2n} \alpha_n^2$, $b_n = 2^{-2n}$, $n = 0, 1, 2, \ldots$. For $| x | \leqq 1$, define the function $f(x)$ as follows:

$$f(0) = 0$$
$$f(x) = 2^{-n+1} x^{\frac{1}{2}} \,, \quad \text{if} \quad a_n \leqq x \leqq b_n$$
$$f(x) = - f(-x) \,.$$

For the moment, $f(x)$ remains undefined if $b_{n+1} < | x | < a_n$.

Consider the sequence of numbers (1) if x_0 is chosen on the interval $a_n \leqq x \leqq b_n$. From the case $n = 0$ of (1),

$$x_1 = x_0 - 2^{-n+1} x_0^{\frac{1}{2}} / 2^{-n} x_0^{-\frac{1}{2}} = -x_0 .$$

But $x_2 = -x_1 = x_0$, since $f(x) = -f(-x)$, so that the sequence defined by (1) is $x_0, -x_0, x_0, \ldots$, which does not converge.

It remains to be shown that the definition of $f(x)$ can be extended for all x in $|x| \leqq 1$, so that $f(x)$ will have the stated properties. First, if $a_n \leqq x \leqq b_n$,

$$f'(x) = 2^{-n} x^{-\frac{1}{2}} ,$$

so that

$$1 \leqq f'(x) \leqq \alpha_n^{-1} ;$$

in addition

$$2 \leqq f(x) / x \leqq 2\alpha_n^{-1} .$$

The definition of $f(x)$ will be extended in such a way that

$$f'(0) = \lim_{x \to 0} f(x) / x = 2 , \tag{11}$$

and that, for $x \neq 0, f'(x)$ exists, is continuous and is not less than 1.

The mean value of $f'(x)$ on the interval $b_{n+1} \leqq x \leqq a_n$,

$$(a_n - b_{n+1})^{-1} \int_{b_{n+1}}^{a_n} f'(x)\, dx = \big(f(a_n) - f(b_{n+1})\big) / (a_n - b_{n+1}) ,$$

is seen to be

$$2(4\alpha_n - 1) / (4\alpha_n^2 - 1) ;$$

and has, therefore, the limit 2 as $n \to \infty$. Let $\delta_1, \delta_2, \ldots$ be a sequence of numbers which tends to zero so rapidly that $\delta_n / b_n \to 0$ as $n \to \infty$. Then a sequence of numbers $m_1, m_2, \ldots$ can be determined in such a way that $m_n \to 2 + 0$ as $n \to \infty$, and that, if $f'(x)$ is defined to be the constant m_n on the interval $b_{n+1} + \delta_n \leqq x \leqq a_n - \delta_n$ and linear on the intervals $b_{n+1} \leqq x \leqq b_{n+1} + \delta_n$, $a_n - \delta_n \leqq x \leqq a_n$, then the function $f(x)$, obtained by integration, has the required smoothness. In the two intervals of length δ_n, the derivative $f'(x)$ will increase from 1 to m_n and decrease from m_n to α_n^{-1}, respectively; while

$$m_n > 2(4\alpha_n - 1) / (4\alpha_n^2 - 1) .$$

The existence and continuity of $f'(x)$ for all $x \neq 0$ is clear, as is, also, the inequality $f'(x) \geqq 1$. The limit relation (11) follows from $m_n \to 2$ and $\delta_n / b_n \to 0$ as $n \to \infty$.

This completes the proof of (i).

(ii) *If* $f(x)$, *where* $|x| \leqq 1$, *has the properties that* $f(0) = 0$, *that* $f'(x)$ *exists and is continuous for* $|x| \leqq 1$, *and that* $f'(x) > 0$ *if* $x \neq 0$, *then the sequence of numbers* (1) *need not converge for all sufficiently small* $|x_0|$.

Let $a_n = 2^{-4n-2}$ and $b_n = 2^{-4n}$ for $n = 0, 1, \ldots$, and, for $|x| \leqq 1$, let

$$\begin{aligned} &f(0) = 0, \\ &f(x) = 2^{-4n-2} x^{\frac{1}{2}}, \quad \text{if } a_n \leqq x \leqq b_n, \ n = 0, 1, \ldots, \\ &f(x) = -f(-x). \end{aligned}$$

For a moment, $f(x)$ remains undefined for $b_{n+1} < |x| < a_n$.

It is clear from the proof in the last example that, if x_0 is chosen on the interval $a_n \leqq x \leqq b_n$, then the sequence of numbers (1) becomes $x_0, -x_0, x_0, \ldots$ and is, therefore, divergent.

In order to complete the definition of $f(x)$ with the stated properties, note that, if $a_n \leqq x \leqq b_n$,

$$0 < f'(x) = 2^{-4n-3} x^{-\frac{1}{2}} \leqq 2^{-2n-2}$$

and

$$0 < f(x) / x = 2^{-4n-2} x^{-\frac{1}{2}} \leqq 2^{-2n-1}.$$

The mean value of $f'(x)$ on the interval $b_{n+1} \leqq x \leqq a_n$ is seen to be

$$(2^{-6n-3} - 2^{-6n-8}) / (2^{-4n-2} - 2^{-4n-4}) = 2^{-2n-4}\, 31/3$$

which has the limit 0 as $n \to \infty$. The construction may now be carried out as before; in this case, $m_n \to +0$ as $n \to \infty$.

(Reçu le 13 novembre 1947.)

[1] *R. Fricke*, Lehrbuch der Algebra, Braunschweig (1924), vol. I, pp. 242—243.

[2] *A. Ostrowski*, Über die Konvergenz und die Abrundungsfestigkeit des Newtonschen Verfahrens, Recueil Mathématique, vol. 2 (1937), pp. 1073—1095.
—, Über einen Fall der Konvergenz des Newtonschen Näherungsverfahrens, Recueil Mathématique, vol. 3 (1938), pp. 254—258.

[3] *C. E. Runge*, Separation und Approximation der Wurzeln, Enzyklopädie der Mathematischen Wissenschaften, I B 3a, pp. 404—448 (1899).