Monografías Matemáticas *García de Galdeano* **35**, 155–162 (2010)

α-THEORY FOR NEWTON-MOSER METHOD

José M. Gutiérrez, Miguel A. Hernández and Natalia Romero

Abstract. We study the semilocal convergence of Newton-Moser method to solve nonlinear equations $F(x) = 0$ defined in Banach spaces. The method defines a sequence $\{x_n\}$ that under appropriate conditions converges to a solution of the aforesaid equation. In fact, by following the known as α-theory, we give conditions on the starting point x_0 and on the derivatives of the operator F in order to establish such convergence. Finally, as an application, we apply this theory to the study of a kind of integral equations.

Keywords: Newton's method, Moser's method, semilocal convergence.

AMS classification: 45G10, 47H17, 65J15.

§1. Introduction

Newton-Moser method is a method to numerically solve nonlinear equations. In order to consider the more general case, let us consider a nonlinear equation

$$F(x) = 0, \tag{1}$$

where F is an operator defined between two Banach spaces X and Y. Let us assume that x^* is a simple root of (1).

Newton-Moser method is an iterative method defined by

$$\begin{cases} x_{n+1} = x_n - B_n F(x_n), & n \geq 0, \\ B_{n+1} = 2B_n - B_n F'(x_{n+1}) B_n, & n \geq 0, \end{cases} \tag{2}$$

where x_0 is a given point in X and B_0 is a given linear operator from Y to X.

The method exhibits several attractive features. First, it avoids the calculus of inverse operators that appears in Newton's method, $x_{n+1} = x_n - F'(x_n)^{-1}F(x_n)$, $n \geq 0$. So it is not necessary to solve a linear equation at each iteration. Second, it has quadratic convergence, the same as Newton's method. Third, in addition to solve the nonlinear equation (1), the method produces successive approximations $\{B_n\}$ to the value of $F'(x^*)^{-1}$, being x^* a solution of (1). This property is very helpful when one investigates the sensitivity of the solution to small perturbations.

We find the origin of the method in a Moser's work [6] for investigating the stability of the N-body problem in Celestial Mechanics. The main difficulty in this, and similar problems involving small divisors, is the solution of a system of nonlinear partial differential equations. In fact, Moser proposed the following method

$$\begin{cases} x_{n+1} = x_n - A_n F(x_n), & n \geq 0, \\ A_{n+1} = A_n - A_n(F'(x_n)A_n - I), & n \geq 0, \end{cases} \tag{3}$$

for a given $x_0 \in X$, a given $A_0 \in \mathcal{L}(Y,X)$, the set of linear operators from Y to X, and where I is the identity operator in X.

Notice that the first equation is similar to Newton's method, but replacing the operator $F'(x_n)^{-1}$ by a linear operator A_n. The second equation is Newton's method applied to equation $g_n(A) = 0$ where $g_n : \mathcal{L}(Y,X) \to \mathcal{L}(X,Y)$ is defined by $g_n(A) = A^{-1} - F'(x_n)$. So $\{A_n\}$ gives us an approximation of $F'(x_n)^{-1}$.

Method (3), firstly proposed by Moser, has a rate of convergence of $(1+\sqrt{5})/2$ for simple roots. However, the variant (2) later introduced by Ulm [9] reaches quadratic convergence. Notice that in (2) $F'(x_{n+1})$ appears instead of $F'(x_n)$.

Since then, method (2) has been also considered by other authors. For instance, Hald [4] showed the quadratic convergence of the method. Later, Petzeltova [7] studied the convergence of the method under Kantorovich-type conditions.

Recently, in [2] a system of recurrence relations is given in order to analyze the convergence of Newton-Moser method (2) under estimations at one point. This theory, introduced by Smale [8], is an alternative to Kantorovich theory [5] to study the semilocal convergence of iterative processes to solve nonlinear equations. Roughly speaking, if x_0 is an initial value such that the sequence $\{x_n\}$ satisfies

$$\|x_n - x^*\| \le \left(\frac{1}{2}\right)^{2^n-1} \|x_0 - x^*\|,$$

then x_0 is said to be an approximate zero of F. The following conditions were introduced by Smale [8] in order to prove that x_0 is an approximated zero

$$\left\|F'(x_0)^{-1}F(x_0)\right\| \le \beta, \tag{4a}$$

$$\sup_{k\ge 2}\left(\frac{1}{k!}\left\|F'(x_0)^{-1}F^{(k)}(x_0)\right\|\right)^{1/(k-1)} \le \gamma, \tag{4b}$$

$$\alpha = \beta\gamma \le 3 - 2\sqrt{2}. \tag{4c}$$

Wang and Zhao [10] pointed that condition (4) is too restrictive. Instead of (4) they assume

$$\left\|F'(x_0)^{-1}F(x_0)\right\| \le \beta, \tag{5a}$$

$$\frac{1}{k!}\left\|F'(x_0)^{-1}F^{(k)}(x_0)\right\| \le \gamma_k,\ k \ge 2, \tag{5b}$$

$$\begin{cases} \text{the equation } \phi(t) = 0 \text{ has at least a positive} \\ \text{solution, where } \phi(t) = \beta - t + \sum_{k\ge 2} \gamma_k t^k. \end{cases} \tag{5c}$$

In [2] the semilocal convergence of Newton-Moser method is established from a system of recurrence relations. However, a majorizing function, as the given in (5c), is not provided. In this paper we present a majorizing function for Newton-Moser method and we give an analysis of its convergence by following the patterns of the α-theory introduced by Smale. The semilocal convergence hypothesis and the main theorem are shown in section 2.

§2. Semilocal convergence results (α-theory)

In this section we study the semilocal convergence of Newton-Moser method (2) to solve the nonlinear equation (1). Let us assume that F is a nonlinear operator defined from an open subset Ω in a Banach space X to another Banach space Y. Let $x_0 \in \Omega$ be a given point and $B_0 \in \mathcal{L}(Y, X)$ a given linear operator defined from Y to X.

Instead the aforesaid conditions (4) or (5), we consider the following ones:

$$\|B_0 F(x_0)\| \le \gamma_0, \tag{6a}$$

$$\|I - B_0 F'(x_0)\| \le \beta < 1, \tag{6b}$$

$$\|B_0 F^{(j)}(x_0)\| \le \gamma_j, \text{ for } j \ge 2, \tag{6c}$$

$$\begin{cases} \text{there exists } R > 0 \text{ such that the series} \\ \sum_{j\ge 2} \gamma_j t^j / j! \text{ is convergent for } t \in [0, R), \end{cases} \tag{6d}$$

$$f(\hat{t}) < 0, \tag{6e}$$

where $\hat{t}$ is the absolute minimum of the function

$$f(t) = \gamma_0 + (\beta - 1)t + \sum_{j\ge 2} \frac{1}{j!}\gamma_j t^j, \quad t \ge 0. \tag{7}$$

In addition, we consider the following scalar sequence

$$\begin{cases} t_0 = 0, \quad b_0 = -1, \\ t_{n+1} = t_n - b_n f(t_n), \\ b_{n+1} = 2b_n - b_n f'(t_{n+1}) b_n. \end{cases} \tag{8}$$

Condition (6e) allows us to say that function $f(t)$ defined in (7) has at least one positive root. Let us denote t^* the smallest positive solution of $f(t) = 0$. With the rest of conditions in (6), (7), (8), we can show that $\{t_n\}$ is an increasing monotone sequence to t^* and

$$\|x_{n+1} - x_n\| \le t_{n+1} - t_n, \ n \ge 0. \tag{9}$$

Consequently, as $\{t_n\}$ is a convergent sequence and $\{x_n\}$ is a sequence defined in a Banach space, $\{x_n\}$ converges to a limit x^*, that can be shown it is a solution of the nonlinear equation (1).

In a more explicit way, the aforementioned comments are shown in the following results.

Theorem 1. *Let us consider the scalar sequences $\{t_n\}$ and $\{b_n\}$ defined in (8). Then the following relations hold:*

1. $b_n < 0$.
2. $b_n f'(t_n) < 1$.
3. $t_n < t_{n+1} < t^*$, *where t^* is the smallest positive root of (7).*

Proof. Firstly we notice that $f''(t) > 0$ for $t > 0$. Then, as $f'(0) = \beta - 1 < 0$ and $\lim_{t\to\infty} f(t) = \infty$, there exists a only value $\hat{t} \in (0, \infty)$ such that $f(\hat{t}) = 0$. Then, condition (6d) guarantees the existence of positive roots of function $f(t)$ defined in (7).

Now we prove the aforementioned are true for $n \geq 0$ by following an inductive reasoning. For $n = 0$ these relations are obviously true. If we suppose they are true for a given value of n, then $b_{n+1} = b_n(2 - b_n f'(t_{n+1})) < 0$, since $b_n f'(t_{n+1}) < b_n f'(t_n) < 1$.

In addition, as $(1 - b_n f'(t_{n+1}))^2 > 0$, then $b_{n+1} f'(t_{n+1}) = 2b_n f'(t_{n+1}) - b_n^2 f'(t_{n+1})^2 < 1$. Now we have $t_{n+2} - t_{n+1} = -b_{n+1} f(t_{n+1}) > 0$ and finally,

$$t^* - t_{n+2} = (1 - b_{n+1} f'(\eta_{n+1}))(t^* - t_{n+1}),$$

for $\eta_{n+1} \in (t_{n+1}, t^*)$. As $b_{n+1} f'(\eta_{n+1}) < b_{n+1} f'(t_{n+1}) < 1$, we conclude $t^* - t_{n+2} > 0$ and the induction is completed. □

Theorem 2. *Under conditions (6), the scalar sequence* $\{t_n\}$ *defined in (8) is a majorizing function for* $\{x_n\}$ *defined in (2), that is,*

$$\|x_{n+1} - x_n\| \leq t_{n+1} - t_n, \ n \geq 0. \tag{10}$$

Consequently, $\{x_n\}$ *converges to a limit* x^*.

Proof. Formula (10) can be proved by following an inductive reasoning. In fact, we can prove that the following inequalities hold for $n \geq 0$:

(I) $\|I - B_n F'(x_n)\| \leq 1 - b_n f'(t_n)$.

(II) $\|B_n F(x_n)\| \leq -b_n f(t_n)$.

(III) $\|B_n F^{(j)}(x_n)\| \leq -b_n f^{(j)}(t_n)$, $j \geq 2$.

Notice that (II) is equivalent to (10).

The aforesaid inequalities are clear for $n = 0$, just by taking into account (6). Now, if we assume they are true for $0, 1, \ldots, n$, then we can prove they are also true for $n + 1$.

Firstly, by (2), we have the following relationships:

$$\begin{gathered} I - B_{n+1}F'(x_{n+1}) = (I - B_n F'(x_{n+1}))^2, \\ I - B_n F'(x_{n+1}) = I - B_n F'(x_n) - \sum_{j\geq 1} \frac{1}{j!} B_n F^{(j+1)}(x_n)(x_{n+1} - x_n)^j, \\ \|I - B_n F'(x_{n+1})\| \leq 1 - b_n f'(t_{n+1}), \\ \|I - B_{n+1}F'(x_{n+1})\| \leq (1 - b_n f'(t_{n+1}))^2 = 1 - b_{n+1} f'(t_{n+1}). \end{gathered} \tag{11}$$

Then, (I) happens for $n + 1$.

Secondly,

$$B_n F(x_{n+1}) = (I - B_n F'(x_n)) B_n F(x_n) + \sum_{j\geq 2} \frac{1}{j!} B_n F^{(j)}(x_n)(x_{n+1} - x_n)^j.$$

Consequently,

$$\begin{aligned}\|B_nF(x_{n+1})\| &\le (1-b_nf'(t_n))(-b_nf(t_n)) + \sum_{j\ge 2}\frac{1}{j!}(-b_nf^{(j)}(t_n))(t_{n+1}-t_n)^j\\ &= -b_nf(t_n)) - b_nf'(t_n)(t_{n+1}-t_n) + \sum_{j\ge 2}\frac{1}{j!}(-b_nf^{(j)}(t_n))(t_{n+1}-t_n)^j = -b_nf(t_{n+1}).\end{aligned}$$

Then, by taking norms in $B_{n+1}F(x_{n+1}) = (2I - B_nF'(x_{n+1}))B_nF(x_{n+1})$, we show that (II) also holds for $n+1$. In fact,

$$\|B_{n+1}F(x_{n+1})\| \le -(2-b_nf'(t_{n+1})(b_nf(t_{n+1})) = -b_{n+1}f(t_{n+1}).$$

Finally,

$$\begin{aligned}\|B_{n+1}F^{(j)}(x_{n+1})\| &\le (2-b_nf'(t_{n+1}))\sum_{k\ge 0}\frac{1}{k!}(-b_nf^{(k+j)}(t_n))(t_{n+1}-t_n)^k\\ &= -(2-b_nf'(t_{n+1}))(b_nf^{(j)}(t_{n+1})) = -b_{n+1}f^{(j)}(t_{n+1}).\end{aligned}$$

Then (III) also holds and the induction is complete.

Now, as $\{t_n\}$ is a increasing sequence that converges to t^*, and the sequence $\{x_n\}$ is defined in a Banach space, $\{x_n\}$ converges to a limit x^*. □

Theorem 3. *Let x^* be the limit of the sequence $\{x_n\}$ defined in (2). Then, if $\|B_0\| \le 1$, x^* is a solution of (1), that is $F(x^*) = 0$.*

Proof. Notice that $\|B_0\| \le 1 = -b_0$. Then, taking into account (11) and the relationship $B_n = (I + (I - B_{n-1}F'(x_n))B_{n-1}$, we can show that $\|B_n\| \le -b_n$ for $n \ge 0$.

In addition, as $B_{n+1} - B_n = ((I - B_nF'(x_{n+1}))B_n$, we have $\|B_{n+1} - B_n\| \le b_{n+1} - b_n$ for $n \ge 0$ and then $\{B_n\}$ is a Cauchy sequence. Consequently, there exists a linear operator B^* such that $B^* = \lim_{n\to\infty} B_n$, $B^*F'(x^*) = I$. Then (see [5, Th. 2, p. 153]) there exists $F'(x^*)^{-1}$ and $\|F'(x^*)^{-1}\| \le -1/f'(t^*)$. This fact, together with (II) in the proof of Theorem 2 guarantees that $F(x^*) = 0$. □

§3. Application to Fredholm integral equations

In this section we consider the following integral equation:

$$x(t) = z(t) + \lambda\int_a^b k(t,s)H(x(s))\,ds, \qquad t \in [a,b],$$

where z is a given continuous function, H is an analytic function, k is a kernel continuous in its two variables and λ is a real parameter. This equation can be written as a equation $F(x) = 0$, where $F : X \to X$ is an operator defined on $X = C[a,b]$, the space of continuous functions in the interval $[a,b]$. The expression of such operator is the following:

$$F(x)(t) = x(t) - z(t) - \lambda\int_a^b k(t,s)H(\phi(s))\,ds, \qquad t \in [a,b]. \tag{12}$$

In the space of continuous functions in $[a,b]$ we consider the max-norm:

$$\|g\| = \max_{t\in[a,b]} |g(t)|, \;\; g \in C[a,b].$$

For the kernel k we define

$$\|k\| = \max_{t\in[a,b]} \int_a^b |k(t,s)|\,ds.$$

In [3] Newton's method has been considered for studying the solution of (12). The two main problems of using Newton's method for solving a nonlinear equation is the choice of the initial approximation x_0 and the calculus of the inverses $F'(x_k)^{-1}$ (or the corresponding solution of a linear equation) at each step. In [3] the initial approximation is chosen as $x_0(t) = z(t)$ and then it is established a set of values for the parameter λ in order equation (12) has a solution. An estimate for the norm of $F'(x_0)^{-1}$ is also given.

Now, in this section we use Newton-Moser method (2) for studying the solution of (12). We consider the same choice for the initial approximation, that is $x_0(t) = z(t)$, but the calculus of $F'(x_0)^{-1}$ it is not required now.

To construct the majorizing function (7) we need to calculate the parameters γ_0, β and γ_j, $j \geq 2$, given in (6), by taking as starting point the function $x_0 = z$. The derivatives of order j of (12) are j-linear operators from the space X^j on X given by:

$$F'(x)[y_1](t) = y_1(t) - \lambda \int_a^b k(t,s)H'(x(s))y_1(s)\,ds,$$

$$F^{(j)}(x)[y_1,\ldots,y_j](t) = -\lambda \int_a^b k(t,s)H^{(j)}(x(s))y_1(t)\cdots y_j(t)\,ds, \;\; j \geq 2.$$

Now we consider a particular integral equation of type (12). We take $x_0(t) = z(t)$ and $B_0 = I$, the identity operator, as starting values for Newton-Moser method (2) and we study the existence of solutions for the corresponding majorizing equation $f(t) = 0$, with f defined in (7). Notice that different convergence results could be obtained under different choices for $x_0(t)$ and B_0.

Let us consider the nonlinear integral equation

$$F(x)(t) = x(t) - 1 - \lambda \int_0^1 \cos(\pi st)x(s)^m\,ds. \tag{13}$$

We take $x_0(t) = 1$ for all $t \in [0,1]$ and $B_0 = I$. Then, $\gamma_0 = |\lambda|, \beta = m|\lambda|$ and

$$\gamma_j = \begin{cases} |\lambda| m(m-1)\cdots(m-j+1), & \text{if } 2 \leq j \leq m, \\ 0 & \text{if } j > m. \end{cases}$$

Consequently the majorizing function (7) is given by

$$f(t) = |\lambda| + (m|\lambda| - 1)t + |\lambda| \sum_{j=2}^{m} \binom{m}{j} t^j = |\lambda|(1+t)^m - t.$$

n	Newton-Moser method (2)	ρ
1	1.180118×10^{-1}	1.78711
2	2.14224×10^{-3}	1.84597
3	3.57016×10^{-5}	1.92453
4	1.30970×10^{-8}	1.96938
5	2.24399×10^{-15}	1.98755

Table 1: Error estimates (10) and the computational order of convergence (14)

If $m|\lambda| < 1$ this function has an absolute minimum $\hat{t} = -1+(m\,|\lambda|)^{-1/(m-1)}$ and, in addition, $f(\hat{t}) < 0$.

Then, according with the results of the previous section, we have established a result on the existence of solution for equations (13). In fact, if $|\lambda| < 1/m$, the integral equation (13) has a solution. In addition, this solution can be approximated by using Newton-Moser method (2) starting with $x_0(t) = 1$ and $B_0 = I$.

For instance, if we consider $m = 5$ and $\lambda = \frac{1}{20}$ then, function

$$f(t) = \frac{1}{20}\left(1 - 15t + 10t^2 + 10t^3 + 5t^4 + t^5\right),$$

is the majorizing function of sequence $\{x_n\}$ and, $t^* = 0.0701898$ is the smallest positive root of f.

Using the majorizing sequence $\{t_n\}$, we show in Table 1 a priori error estimates (10) and the computational order of convergence [1]:

$$\rho \approx \ln \frac{\|t_{n+1} - t^*\|}{\|t_n - t^*\|} \Big/ \ln \frac{\|t_n - t^*\|}{\|t_{n-1} - t^*\|}, \qquad n \in \mathbb{N}, \tag{14}$$

when Newton-Moser method (2) is applied to solve equation (13).

Now, from Theorem 3 the integral equation (13) has a solution x^* in $B(1, 0.0701898)$ which is the limit of the iterations of Newton-Moser method (2) starting with $x_0(t) = 1$ and $B_0 = I$:

$$\begin{aligned}
x_1(t) &= 1 + 0.015915493\, t^{-1} \sin(3.14159\, t),\\
x_2(t) &= 1 + 0.017615759\, t^{-1} \sin(3.14159\, t),\\
x_3(t) &= 1 + 0.017633935\, t^{-1} \sin(3.14159\, t),\\
x_4(t) &= 1 + 0.017633938\, t^{-1} \sin(3.14159\, t).
\end{aligned}$$

Considering iteration $x_4(t)$ as a numerical solution x^* of integral equation (13) and the computational order of convergence:

$$\rho_n \approx \ln \frac{\|x_{n+1}(t) - x^*\|}{\|x_n(t) - x^*\|} \Bigg/ \ln \frac{\|x_n(t) - x^*\|}{\|x_{n-1}(t) - x^*\|}, \qquad n \in \mathbb{N}, \tag{15}$$

Newton-Moser method reach computationally the R-order of convergence at least two. In fact, $\rho_1 = 1.95368$ and $\rho_2 = 1.97401$.

Acknowledgements

This paper has been supported by the Spanish Ministry of Science and Technology (Ref. MTM2008-01952/MTM).

References

[1] Grau-Sánchez, M., and Noguera, M. A variant of Cauchy's method with accelerated fifth-order convergence. *Appl. Math. Lett. 17* (2004), 509–517.

[2] Gutiérrez, J. M., Hernández, M. A., and Romero, N. A note on a modification of Moser's method. *Journal of Complexity 24* (2008), 185–197.

[3] Gutiérrez, J. M., Hernández, M. A., and Salanova, M. A. α-theory for nonlinear Fredholm integral equations. *Grazer Mathematische Berichte 346* (2004), 187–196.

[4] Hald, O. H. On a Newton-Moser type method. *Numer. Math. 23* (1975), 411–425.

[5] Kantorovich, L. V. *Functional analysis*. Pergamon Press, Oxford, 1982.

[6] Moser, J. *Stable and random motions in dynamical systems with special emphasis on celestian mechanics*, vol. 77 of *Herman Weil Lectures, Annals of Mathematics Studies*. Princeton Univ. Press, Princeton, New Jersey, 1973.

[7] Petzeltova, H. Remark on Newton-Moser type method. *Commentationes Mathematicae Universitatis Carolinae 21* (1980), 719–725.

[8] Smale, S. *Newton's method estimates from data at one point*. The Merging of Disciplines: New Directions in Pure, Applied and Computational Mathematics. Springer-Verlag, New York, 1986.

[9] Ulm, S. On iterative methods with successive aproximation of the inverse operator (Russian). *Izv. Akad Nauk Est. SSR 16* (1967), 403–411.

[10] Wang, D. R., and Zhao, F. G. The theory of Smale's point estimation and its applications. *J. Comput. Appl. Math. 60* (1995), 253–269.

José M. Gutiérrez, Miguel A. Hernández and Natalia Romero
Dep. Mathematics and Computation, University of La Rioja
C/ Luis de Ulloa s/n
26004 Logroño, Spain
{jmguti, mahernan, natalia.romero}@unirioja.es