Pergamon

Computers Math. Applic. Vol. 31, No. 8, pp. 69–84, 1996

0898-1221/96 \$15.00 + 0.00

S0898-1221(96)00032-6

# Least Square Approximation of a Nonlinear Ordinary Differential Equation

T. BENOUAZ* AND O. ARINO
Laboratoire de Mathématiques Appliquées
URA 1204 CNRS, Avenue de l'Université
64000 Pau, France

*(Received and accepted September 1995)*

**Abstract**—The aim of this paper is to present an optimal approximation method for a nonlinear ordinary differential equation based on the minimization in the least square sense. The approximation is order two or higher in the vicinity of the origin. We provide a few examples.

**Keywords**—Nonlinear ordinary differential equations, Optimal approximation, Mean square convergence, Stability.

## 1. INTRODUCTION

Linearization methods play an important role in the analysis of ordinary differential equations. A classical linear approximation is obtained by the Fréchet derivative of a nonlinear equation. In [1], we presented a computational procedure which yields a linear map defined as the optimal linearization to a nonlinear ordinary differential equation.

The procedure is based on the minimization of a certain functional with respect to a curve starting from an initial value $x_0$ and going to 0 as $t$ goes to infinity. At each step, it gives a linear map, starting from the Jacobian matrix $DF(x)$ estimated at the initial value $x_0$. The optimal approximation of the nonlinear equation is obtained as a limit of the sequence of linear maps determined by the procedure. Our results are in the line of previous work by Vujanovic in 1973 [2], and Jordan *et al.* [3,4].

To the best of our knowledge, however, no theoretical evidence of the validity of the method introduced by Vujanovic [2] has been given up to now. It is our intention to make some progress in that direction, in order to later on apply the procedure to some problems. In [5], we have applied this procedure to a specific nonlinear ordinary differential equation for which we proved existence, uniqueness and convergence of the optimal approximation associate for this. The work presented in [1,6] is based on the applicability of the proposed method to the study of stability. In this paper, we present results concerning the proposed approximation. We give a necessary and sufficient condition for uniqueness of the elements of the sequence determined in the course of the optimal approximation and prove that the order of the approximation is two, or higher. In the scalar case, we give the analytic expression of the optimal approximation and prove that the limit as $x_0 \longrightarrow 0$ is the derivative of $f$ at 0. We give an example where the limit exists even if the derivative of $f$ does not exist. We have applied the procedure to some examples and computed the relative error for comparison.

*Permanent address: Universite de Tlemcen, Institut de Sciences Exactes, B.P 119 Tlemcen 13000 Algerie.

Typeset by $\mathcal{A}_{\mathcal{M}}\mathcal{S}$-TEX

# 2. THEORETICAL FRAMEWORK

## 2.1. Formulation of Problem

Consider the following system of nonlinear ordinary differential equations:

$$\frac{dx}{dt} = F(x(t)), \qquad x(0) = x_0, \tag{1}$$

where

$x = (x_1, \dots, x_n)$ is the unknown function,
$F = (f_1, \dots, f_n)$ is a given function on an open subset $\Omega$ of $\mathbb{R}^n$.

Our purpose is to elaborate a method of approximation, which will associate to system (1) a linear system of the form

$$\frac{dx}{dt} = A^* x(t), \qquad x(0) = x_0, \tag{2}$$

where $A^* \in \mathcal{M}_n(\mathbb{R})$ is to be determined. For this, we shall assume

(H1) $F(0) = 0$.
(H2) The spectrum $\sigma(DF(x))$ is contained in the set $\{z : \operatorname{Re} z < 0\}$ for every $x \neq 0$, in a neighborhood of 0, for which $DF(x)$ exists.
(H3) $F$ is $\gamma$-Lipschitz continuous.

System (2), corresponding to system (1), will give an optimal approximation with respect to curve, starting at the initial point $x_0$ and tending to 0 as $t$ goes to infinity.

## 2.2. Formalism

Consider the functional defined by

$$G(A) = \int_0^{+\infty} \|F(x(t)) - Ax(t)\|^2 \, dt, \tag{3}$$

where

$F(x)$ is as above.
$A \in \mathcal{M}_n(\mathbb{R})$ is to be determined successively.

For the time being, $x$ is just any function defined on $[0, +\infty[$, bounded, continuous and such that $x \in L^1(0, +\infty)$ and $F(x(\cdot)) \in L^1(0, +\infty)$.

Later on, we will consider functions $x(t)$ that are solutions of linear equations. The minimization of the functional $G(A)$ with respect to $A$ will allow us to get the optimal system (2). Differentiating (3) with respect to $A$ along a function $x$ yields

$$DG(A)\alpha = 2 \int_0^{+\infty} \langle Ax(t) - F(x(t)), \, \alpha x(t) \rangle \, dt, \tag{4}$$

for every matrix $\alpha$. In particular, for matrices $\alpha$ such that $\alpha_{l,m} = 1; \alpha_{i,j} = 0$, if $(i,j) \neq (l,m)$, we have

$$\int_0^{+\infty} \langle Ax(t) - F(x(t)), \, \alpha x(t) \rangle \, dt = \int_0^{+\infty} [Ax(t) - F(x(t))]_l \, x_m(t) \, dt, \tag{5}$$

and

$$\int_0^{+\infty} [Ax(t) - F(x(t))]_l \, x_m(t) \, dt, \qquad \forall 1 \leq l, \ m \leq n. \tag{6}$$

Assuming that $A$ minimizes (3) along a given function $x$, the above quantities are equal to zero, which leads to

$$\sum_{j=1}^{n} a_{l,j} \left( \int_0^{+\infty} x_j(t)\, x_m(t)\, dt \right)_{1 \le j,m \le n} = \left( \int_0^{+\infty} f_l(x(t))\, x_m(t)\, dt \right)_{1 \le l,m \le n}, \tag{7}$$

with obvious notations for the elements of matrix $A$.

Introducing valued function $\Gamma(x)$ defined by

$$\Gamma(x) = \int_0^{+\infty} [x(t)]\,[x(t)]^{\top}\, dt = \left( \int_0^{+\infty} x_j(t)\, x_m(t)\, dt \right)_{1 \le j,m \le n}, \tag{8}$$

and assuming $\Gamma(x)$ is nonsingular, we obtain

$$A = \left[ \int_0^{+\infty} [F(x(t))]\,[x(t)]^{\top}\, dt \right] [\Gamma(x)]^{-1}. \tag{9}$$

A necessary and sufficient condition for $\Gamma(x)$ to be invertible is given in the next lemma.

LEMMA 1. *The matrix $\Gamma(x)$ is invertible if only and if the set $x(\mathbb{R}^+)$ is dense in $\mathbb{R}^n$.*

PROOF. $\Gamma(x)$ being a nonnegative symmetric matrix, a necessary and sufficient condition for $\Gamma(x)$ to be invertible is that

$$v^{\top}\Gamma(x)\, v > 0, \qquad \text{for each } v \in \mathbb{R}^n, \quad v \neq 0; \tag{10}$$

that is

$$v^{\top}\Gamma(x)\, v = \int_0^{+\infty} (\langle x(t), v\rangle)^2\, dt > 0. \tag{11}$$

This happens if and only if $\forall v \neq 0$, $\langle x(\cdot), v\rangle \neq 0$; that is, $x(\mathbb{R}^+)$ is dense in $\mathbb{R}^n$. ∎

Suppose now that $x$ is the solution of a linear equation

$$\frac{dx}{dt} = A_0 x(t), \qquad x(0) = x_0, \tag{12}$$

in which $A_0 = DF(x_0)$.

From Assumption H2 in Section 2.1, we know that $x(t)$ goes to zero exponentially, as $t$ goes to $+\infty$. In this case, Lemma 1 leads to the following condition for $\Gamma(x)$ to be invertible.

LEMMA 2. *Suppose $x(t) = \exp(tA_0)x_0$, where $A_0$ satisfies H2 in Section 2.1. Then, a necessary and sufficient condition for the matrix $\Gamma(x)$ to be invertible is that the rank of the family $x_0, A_0x_0, \dots, A_0^{n-1}x_0$ be equal to $n$,*

$$\operatorname{Rank}\left[x_0, A_0x_0, \dots, A_0^{n-1}x_0\right] = n. \tag{13}$$

PROOF. This result is an immediate consequence of the Cayley-Hamilton theorem [7]. ∎

REMARK 3. *The condition precludes notably the possibility for $x_0$ to be an eigenvector of $A_0$, or to belong to an invariant subspace of dimension less than $n$. In the case when $A_0$ has only simple eigenvalues, it holds if and only if $x_0$ has a nonzero projection on each of the eigenspaces.*

### 2.3. Algorithm

The computation presented in Section 2.2 will be used iteratively. We shall assume that the successive matrices are stable; that is, their spectrum lies in $\{z : \operatorname{Re} z < 0\}$.

The initial matrix is the Jacobian matrix of $F$ at $x_0$, where $x_0$ is an arbitrary point in a neighborhood of 0, such that $DF(x_0)$ exists.

Consider system (1)

$$\frac{dx}{dt} = F(x(t)), \qquad x(0) = x_0.$$

FIRST STEP. Compute $A_0 = DF(x_0)$.

SECOND STEP. Compute $A_1$ from the solution of the equation

$$\frac{dy}{dt} = A_0 y(t), \qquad y(0) = x_0 \tag{14}$$

by minimizing the functional

$$G(A) = \int_0^{+\infty} \|F(y(t)) - Ay(t)\|^2 \, dt, \tag{15}$$

$y$ being the solution of equation (14).

$A_1$ is uniquely determined by formula (9), where we let $x$ be the solution of equation (14). From this point on, the matrices determined by the procedure are no longer Jacobian matrices for $F$ at a given point. They are obtained as a sort of mean value of the derivatives of $F$ along trajectories linking $x_0$ to the origin. In order to continue, it is necessary that the above conditions be satisfied at each step.

Let us first assume that this holds. Then the procedure works as follows.

THIRD STEP. Assuming that $A_1, \dots, A_{j-1}$ have been computed, to compute $A_j$ from $A_{j-1}$, we first solve

$$\frac{dy}{dt} = [A_{j-1}]\, y(t), \qquad y(0) = x_0. \tag{16}$$

Let $y_j$ be the solution of equation (16). The minimization of the functional

$$G_j(A) = \int_0^{+\infty} \|F(y_j(t)) - Ay_j(t)\|^2 \, dt \tag{17}$$

yields $A_j$.

In fact, we have the following relationship between $A_{j-1}$ and $A_j$:

$$A_j \Gamma(y_j) = \int_0^{+\infty} [F(y_j(t))] [y_j(t)]^\top \, dt; \tag{18}$$

that is, assuming that $\Gamma(y_j)$ is invertible, $A_j$ can be written as

$$\begin{aligned} A_j &= \left[\int_0^{+\infty} [F(y_j(t))] [y_j(t)]^\top \, dt\right] [\Gamma(y_j)]^{-1} \\ &= \left[\int_0^{+\infty} \left[F\left(e^{tA_{(j-1)}} x_0\right)\right] \left[e^{tA_{(j-1)}} x_0\right]^\top dt\right] \left[\Gamma\left(e^{tA_{(j-1)}} x_0\right)\right]^{-1}. \end{aligned} \tag{19}$$

If the sequence $A_j$ converges, then the limit $A^*$ is by definition the optimal approximation of $F$ at $x_0$.

## 3. PROPERTIES OF THE PROCEDURE

We will now consider situations where the procedure converges.

### 3.1. Case When the Application is Linear

If $F$ is linear with $\sigma(F)$ in the negative part of the complex plane, then the procedure gives $F$ at the first iteration. Indeed, in this case, equation (7) reads

$$A\Gamma(x) = F\Gamma(x)\,; \tag{20}$$

it is clear that $A = F$ is a solution. It is unique if $\Gamma(x)$ is invertible. This happens if and only if condition (10) is satisfied.

Therefore, the optimal approximation of a linear system is the system itself.

### 3.2. General Case, When the System is the Sum of Linear and Nonlinear Terms

Consider the more general system of nonlinear equations with a nonlinearity of the form

$$F(x) = Mx + \tilde{F}(x)\,, \qquad x(0) = x_0, \tag{21}$$

where $M$ is linear.

The computation of the matrix $A_1$ gives

$$A_1 = \left[\int_0^{+\infty} [F(x(t))]\,[x(t)]^\top\,dt\right][\Gamma(x)]^{-1}, \tag{22}$$

which can be written as

$$A_1 = \left[M\Gamma(x) + \left(\int_0^{+\infty}\left[\tilde{F}(x(t))\right][x(t)]^\top\,dt\right)\right][\Gamma(x)]^{-1}, \tag{23}$$

and finally

$$A_1 = M + \left[\int_0^{+\infty}\left[\tilde{F}(x(t))\right][x(t)]^\top\,dt\right][\Gamma(x)]^{-1}. \tag{24}$$

Hence, $A_1 = M + \tilde{A}_1$ with

$$\tilde{A}_1 = \left[\int_0^{+\infty}\left[\tilde{F}(x(t))\right][x(t)]^\top\,dt\right][\Gamma(x)]^{-1}. \tag{25}$$

Then, for all $j$, we have

$$A_j = M + \tilde{A}_j, \tag{26}$$

with

$$\tilde{A}_j = \left[\int_0^{+\infty}\left[\tilde{F}(x_j(t))\right][x_j(t)]^\top\,dt\right][\Gamma(x_j(t))]^{-1}. \tag{27}$$

If, in particular, some components of $F$ are linear, then the corresponding components of $\tilde{F}$ are zero, and the corresponding components of $A_j$ are those of $F$.

If $f_k$ is linear, then the $k^{\text{th}}$ row of matrix $A_j$ is equal to $f_k$.

## 4. SCALAR CASE

### 4.1. Expression of the Optimal Approximation

Consider the following nonlinear scalar equation:

$$\frac{dx}{dt} = f(x(t))\,, \qquad x(0) = x_0, \tag{28}$$

where $f : \mathbb{R} \longrightarrow \mathbb{R}$, and satisfies the following:

(H1) $f(0) = 0$.

(H2) $f'(x) < 0$ at every point where $f'(x)$ exists in an interval $]-\alpha, +\alpha[, \alpha > 0$.

(H3) $f$ is absolutely continuous with respect to the Lebesgue measure.

Choose $x_0 \in ]-\alpha, +\alpha[$ such that $f'(x_0)$ exists. Set $a_0 = f'(x_0)$ and use the method presented in Section 2.3.

We solve the linear equation

$$\frac{dx}{dt} = a_0 x(t), \qquad x(0) = x_0 \tag{29}$$

to obtain

$$x(t) = \exp(a_0 t)\, x_0. \tag{30}$$

Substituting $f$ for $F$ in expression (9), we get

$$a_1 = \frac{\left(\int_0^{+\infty} f(e^{a_0 t} x_0)\, e^{a_0 t}\, dt\right)}{\left(\int_0^{+\infty} e^{2a_0 t}\, dt\right)} \frac{1}{x_0}. \tag{31}$$

For $x_0 \neq 0$, $f(x(t))$ is almost everywhere differentiable and

$$\frac{d}{dt}\left[f\left(e^{a_0 t} x_0\right)\right] = f'\left(e^{a_0 t} x_0\right) e^{a_0 t} x_0 a_0. \tag{32}$$

This gives

$$\int_0^{+\infty} f(x(t))\, e^{a_0 t}\, dt = \frac{1}{a_0}\left[f(x(t))\, e^{a_0 t}\right]_0^{+\infty} - \frac{1}{a_0}\int_0^{+\infty} \left(f'(x(t))\, e^{2a_0 t}\, dt\right) x_0 a_0, \tag{33}$$

from which we obtain $a_1$

$$a_1 = 2\left(\frac{f(x_0)}{x_0} + a_0 \int_0^{+\infty} f'(x(t))\, e^{2a_0 t}\, dt\right). \tag{34}$$

Changing the variable $t$ to $x(t)$ in the integral, we obtain

$$a_1 = \frac{2}{x_0^2}\int_0^{x_0} f(z)\, dz. \tag{35}$$

So, $a_1$ does not depend on $a_0$. Repeating the procedure, as indicated above will give the same result. In this case, the procedure leads to the optimal approximation in one step; i.e.,

$$a^*(x_0) = \frac{2}{x_0^2}\int_0^{x_0} f(z)\, dz. \tag{36}$$

### 4.2. Analysis of $a^*(x_0)$

LEMMA 4. *If the derivative of $f$ exists at 0 and $f$ is continuous, then* $\lim_{x_0 \longrightarrow 0} a^*(x_0) = f'(0)$.

PROOF. With $f(z) = zf'(0) + z\varepsilon(z)$, equation (36) can be written

$$a^*(x_0) = f'(0) + \frac{2}{x_0^2}\int_0^{x_0} z\varepsilon(z)\, dz. \tag{37}$$

The second term of equation (37)

$$\left|\frac{2}{x_0^2}\int_0^{x_0} z\varepsilon(z)\, dz\right| \leq \varepsilon(x_0)\frac{2}{x_0^2}\int_0^{x_0} z\, dz = \varepsilon(x_0)\left(\frac{2}{x_0^2}\right)\left(\frac{x_0^2}{2}\right) \tag{38}$$

converges to 0 as $x_0 \longrightarrow 0$. Hence, $\lim_{x_0 \longrightarrow 0} a^*(x_0) = f'(0)$. ∎

We can see that the optimal approximation defined by equation (37) depends on the initial value $x_0$ and converges to $f'(0)$ as $x_0 \longrightarrow 0$ if $f'(0)$ exist.

REMARK 5. It is possible to find a limit even if the derivative of $f$ at 0 does not exist.

EXAMPLE 6. Consider equation (36) and write $f(z)$ as follows:

$$f(z) = -zg(z). \tag{39}$$

This yields

$$a^*(x_0) = -\frac{2}{x_0^2}\int_0^{x_0} zg(z)\,dz. \tag{40}$$

Let us choose $g(z) = |\sin\log|z||$, for $z \neq 0$.

$$a^*(x_0) = -\frac{2}{x_0^2}\int_0^{x_0} z\,|\sin\log|z||\,dz. \tag{41}$$

Changing $z$ to $ux_0$,

$$a^*(x_0) = -2\int_0^1 u\,|\sin\log|ux_0||\,du, \tag{42}$$

and changing $-\log(ux_0)$ to $v$, we have

$$a^*(x_0) = -\frac{2}{x_0^2}\int_{\log(1/x_0)}^{+\infty} e^{-2v}\,|\sin(v)|\,dv, \tag{43}$$

and

$$\begin{aligned} a^*(x_0) &= -\frac{2}{x_0^2}\int_{k\pi}^{+\infty} e^{-2v}\,|\sin(v)|\,dv \\ &= -\frac{2}{x_0^2}\sum_{l=k}^{\infty}\int_0^{\pi} e^{-2(v+l\pi)}\sin(v)\,dv \\ &= -\frac{2}{x_0^2}\left(\sum_{l=k}^{\infty} e^{-2l\pi}\right)\int_0^{\pi} e^{-2v}\sin(v)\,dv \\ &= -\frac{2}{x_0^2}\frac{e^{-2k\pi}}{1-e^{-2\pi}}\int_0^{\pi} e^{-2v}\sin(v)\,dv. \end{aligned} \tag{44}$$

With $k\pi = \log(1/x_0) \Longrightarrow e^{-2k\pi} = x_0^2$, we have

$$a^*(x_0) = \frac{-2}{1-e^{-2\pi}}\int_0^{\pi} e^{-2v}\sin(v)\,dv. \tag{45}$$

Finally,

$$a^*(x_0) = -\frac{2}{5}\coth(\pi). \tag{46}$$

Hence, $\lim_{x_0\longrightarrow 0} a^*(x_0)$ exist.

### 4.3. Another Expression of $a^*(x_0)$

Assuming $f$ is analytic

$$f(x) = \sum_{n=1}^{\infty}\frac{f^{(n)}(0)}{n!}x^n, \tag{47}$$

we can give another expression of $a^*(x_0)$

$$a^*(x_0) = \frac{\sum_{n=1}^{\infty}\left(f^{(n)}(0)/n!\right)x_0^{n+1}\int_0^{+\infty} e^{(n+1)sa_0}\,ds}{x_0^2\int_0^{+\infty} e^{2sa_0}\,ds}, \tag{48}$$

and finally,

$$a^*(x_0) = 2\sum_{n=1}^{\infty} \frac{f^{(n)}(0)}{(n+1)!} x_0^{n-1}, \tag{49}$$

and

$$a^*(x_0) = f'(0) + \frac{1}{3} x_0 f''(0) + \cdots + \frac{2}{(n+1)!} x_0^{n-1} f^{(n)}(0) + \cdots. \tag{50}$$

REMARK 7. The role of the optimal approximation in the study of stability is evidenced in the scalar case [1,8] by the fact that in this case, the function $x \longrightarrow v(x) = x^2 a^*(x)$ is a Lyapunov function for the equation. Indeed, if $x(t)$ is a solution of equation (28), differentiating $[(x(t))^2 a^*(x(t))]$ with respect to $t$, we obtain

$$\frac{d}{dt}\left[(x(t))^2 a^*(x(t))\right] = (f(x(t)))^2. \tag{51}$$

Since, on the other hand, $v(x)x < 0$ (in view of assumption H1 and H2 in Section 4.1), we obtain that $v(x(t)) \longrightarrow 0$ as $t \longrightarrow +\infty$, therefore, $x(t) \longrightarrow 0$ as $t \longrightarrow +\infty$.

## 5. ORDER OF THE APPROXIMATION

In order to estimate the order of the approximation, we will evaluate the functional defined by relation (3):

$$\int_0^{+\infty} \|F(y(t)) - Ay(t)\|^2 \, dt, \tag{52}$$

where $A$ is any matrix. Starting from an arbitrary matrix $A_0$, the first matrix obtained in the optimal procedure minimizes the functional

$$\int_0^{+\infty} \|F(y_0(t)) - A_1 y_0(t)\|^2 \, dt, \tag{53}$$

where $y_0(t)$ is the solution of equation (14). We have the following relationship between $A_1$ and $A$:

$$\int_0^{+\infty} \|F(y_0(t)) - A_1 y_0(t)\|^2 \, dt \le \int_0^{+\infty} \|F(y_0(t)) - Ay_0(t)\|^2 \, dt, \tag{54}$$

and between $A_j$ and $A$,

$$\int_0^{+\infty} \|F(y_j(t)) - (A_{j+1}) y_j(t)\|^2 \, dt \le \int_0^{+\infty} \|F(y_j(t)) - Ay_j(t)\|^2 \, dt, \tag{55}$$

where $y_j(t)$ is the solution of equation (16).

In the limit ($j \longrightarrow +\infty$), we obtain

$$\int_0^{+\infty} \|F(y^*(t)) - A^* y^*(t)\|^2 \, dt \le \int_0^{+\infty} \|F(y^*(t)) - Ay^*(t)\|^2 \, dt, \tag{56}$$

where $y^*(t)$ is the solution of equation

$$\frac{dy}{dt} = A^* y(t), \qquad y(0) = x_0. \tag{57}$$

So

$$\int_0^{+\infty} \|F(y^*(t)) - A^* y^*(t)\|^2 \, dt = \inf_{\substack{\forall A \in \mathcal{M}_n \mathbb{R} \\ \operatorname{Re}\sigma(A) \subset ]-\infty, 0[}} \int_0^{+\infty} \|F(y^*(t)) - Ay^*(t)\|^2 \, dt. \tag{58}$$

In particular, for $A = DF(0)$, we have

$$\int_0^{+\infty} \|F(y^*(t)) - A^*y^*(t)\|^2 \, dt \le \int_0^{+\infty} \|F(y^*(t)) - DF(0)\, y^*(t)\|^2 \, dt, \tag{59}$$

and

$$\int_0^{+\infty} \|F(y^*(t)) - A^*y^*(t)\|^2 \, dt \le 0\left(\|x_0\|^2\right)^2. \tag{60}$$

We will now evaluate the difference $\|x(t) - y^*(t)\|$ where $x$ is a solution of equation (1) and $y^*$ is the solution of the optimal approximation, both having the same initial value. We have

$$\frac{dx}{dt} - \frac{dy^*}{dt} = F(x(t)) - A^*y^*(t) = F(x(t)) - F(y^*(t)) + F(y^*(t)) - A^*y^*(t). \tag{61}$$

From Assumption H3 in Section 2.1, we have

$$\frac{d}{dt}\|x(t) - y^*(t)\| \le \gamma \|x(t) - y^*(t)\| + \|F(y^*(t)) - A^*y^*(t)\|, \tag{62}$$

and using Gronwall's lemma, we obtain

$$\begin{aligned} \|x(t) - y^*(t)\| &\le \int_0^t e^{\gamma t} \|F(y^*(s)) - A^*y^*(s)\| \, ds \\ &\le \left(\int_0^t e^{2\gamma(t-s)} \, ds\right)^{1/2} \left(\int_0^t \|F(y^*(s)) - A^*y^*(s)\|^2 \, ds\right)^{1/2}. \end{aligned} \tag{63}$$

For every $T > 0$, there exists $M \ge 0$ such that

$$\|x(t) - y^*(t)\| \le M \|x_0\|^2, \qquad \text{for } 0 \le t \le T. \tag{64}$$

This approximation can be extended to $\mathbb{R}^+$ if we assume that $F$ is dissipative, namely, if for some $\alpha > 0$, we have $\langle F(x) - F(y), x - y\rangle \le -\alpha\|x - y\|^2$ for every $x, y$. With $F(0) = 0$, we obtain

$$\|x(t) - y^*(t)\| \le \left(\int_0^t e^{-2\alpha(t-s)} \, ds\right)^{1/2} \left(\int_0^t \|F(y^*(s)) - A^*y^*(s)\|^2 \, ds\right)^{1/2}. \tag{65}$$

Finally, there exists $M \ge 0$

$$\|x(t) - y^*(t)\| \le \left[\frac{1}{(2\alpha)^{1/2}}\right] \left[M\left(\|x_0\|^2\right)\right]. \tag{66}$$

The proposed approximation is of order two or higher. More generally, it has the same order as the nonlinearity.

EXAMPLE 8. Consider the following equation:

$$\frac{dx}{dt} = F(x) = -x + x^3, \qquad x(0) = x_0. \tag{67}$$

In this case,

$$\|x(t) - y^*(t)\| = 0\left(|x_0|^3\right), \tag{68}$$

where $y^*$ is the solution of the equation

$$\frac{dy}{dt} = a^*y(t), \qquad y(0) = x_0. \tag{69}$$

Hence, the approximation is of order three.

# 6. APPLICATIONS

Prior to the study of examples of nonlinear systems, we present the computational procedure.

## 6.1. Computational Procedure

The computational procedure is based on the algorithm presented in Section 2.3, and written in Fortran language. The differential equations have been solved using the fourth-order Runge-Kutta method [7].

Input $x_0, A_0$.

LEVEL (I). Computation of $A_1$ in terms of $A_0$

$$A_1 = \left[\int_0^{+\infty} \left[F\left(e^{A_0 t}x_0\right)\right]\left[e^{A_0 t}x_0\right]^{\top} dt\right]\left[\int_0^{+\infty} \left[e^{A_0 t}x_0\right]\left[e^{A_0 t}x_0\right]^{\top} dt\right]^{-1}. \tag{70}$$

LEVEL (II). Computation of $A_{(j)}$ in terms of $A_{(j-1)}$

$$A_{(j)} = \left[\int_0^{+\infty} \left[F\left(e^{A_{(j-1)} t}x_0\right)\right]\left[e^{A_{(j-1)} t}x_0\right]^{\top} dt\right]\left[\int_0^{+\infty} \left[e^{A_{(j-1)} t}x_0\right]\left[e^{A_{(j-1)} t}x_0\right]^{\top} dt\right]^{-1}. \tag{71}$$

LEVEL (III). Computation of

$$\left\|A_{(j)} - A_{(j-1)}\right\|. \tag{72}$$

LEVEL (IV). If

$$\left\|A_{(j)} - A_{(j-1)}\right\| < \varepsilon, \tag{73}$$

where $\varepsilon$ is the desired level of approximation, then set $A^* = A_{(j)}$. $A^*$ is the optimal approximation of $F$ at $x_0$. Else set $A_{(j-1)} = A_{(j)}$ and go to Level (II).

## 6.2. Case of a System which Cannot be Linearized at 0 Using the Fréchet Derivative

EXAMPLE 9. Consider a system with a function of the absolute value type, that is, nondifferentiable at 0.

$$\begin{aligned} \frac{dx}{dt} &= -x + \alpha \sin(|y|), \\ \frac{dy}{dt} &= -y + \alpha \sin(|x|), \end{aligned} \qquad (x_0, y_0) = (1, 0.5), \quad |\alpha| < 1. \tag{74}$$

Then we have, for $\alpha = 0.5$,

$$DF(x_0, y_0) = \begin{bmatrix} -1 & 0.4387 \\ 0.2701 & -1 \end{bmatrix}, \qquad (x_0, y_0) = (1, 0.5). \tag{75}$$

After five iterations, the computational procedure gives ($\varepsilon = 10^{-6}$)

$$A^* = \begin{bmatrix} -1.0207 & 0.5172 \\ 0.3502 & -0.8336 \end{bmatrix}, \qquad (x_0, y_0) = (1, 0.5). \tag{76}$$

Table 1 shows the values of the solutions of systems (74) and (76) and the relative error. The formula of relative error is given by

$$E_r = \frac{\|y(t) - y^*(t)\|}{\|y(t)\|}, \tag{77}$$

where $y$ is the solution of equation (1) and $y^*$ is the solution of equation (2).

The curves in Figures 1 and 2 represent the graphs of the respective components $(x(t), y(t))$ of solutions of systems (74) and (76) as a function of time.

Note that the method enables us to associate a linear system (optimal approximation) to a nonlinear system in the neighborhood of 0.

Table 1.

| $t$ | $X_{nl}(t)$ | $X_{lin}(t)$ | $Y_{nl}(t)$ | $Y_{lin}(t)$ | $E_r$ |
|---|---|---|---|---|---|
| 0 | .1000000E+01 | .1000000E+01 | .5000000E+00 | .5000000E+00 | 0 |
| 1 | .5008458E+00 | .5012365E+00 | .3794262E+00 | .379014E+00 | .0009 |
| 2 | .2781739E+00 | .2783531E+00 | .2502546E+00 | .2505841E+00 | .001 |
| 3 | .1628160E+00 | .1630481E+00 | .1565178E+00 | .1580888E+00 | .007 |
| 4 | .9741862E−01 | .9789484E−01 | .9600762E−01 | .9793934E−01 | .014 |
| 5 | .5878825E−01 | .5941472E−01 | .5847295E−01 | .6022978E−01 | .022 |
| 6 | .3559002E−01 | .3622664E−01 | .3551962E−01 | .3692688E−01 | .030 |
| 7 | .2157151E−01 | .2213122E−01 | .2155580E−01 | .2261112E−01 | .039 |
| 8 | .1308045E−01 | .1353121E−01 | .1307694E−01 | .1383805E−01 | .047 |
| 9 | .7932950E−02 | .8275914E−02 | .7932167E−02 | .8467011E−02 | .056 |
| 10 | .4811411E−02 | .5062412E−02 | .4811237E−02 | .5180184E−02 | .065 |
| 11 | .2918231E−02 | .3096885E−02 | .2918192E−02 | .3169155E−02 | .074 |
| 12 | .1769989E−02 | .1894538E−02 | .1769980E−02 | .1938808E−02 | .083 |
| 13 | .1073550E−02 | .1159008E−02 | .1073549E−02 | .1186105E−02 | .093 |
| 14 | .6511409E−03 | .7090407E−03 | .6511404E−03 | .7256214E−03 | .10 |
| 15 | .3949368E−03 | .4337673E−03 | .3949367E−03 | .4439117E−03 | .11 |

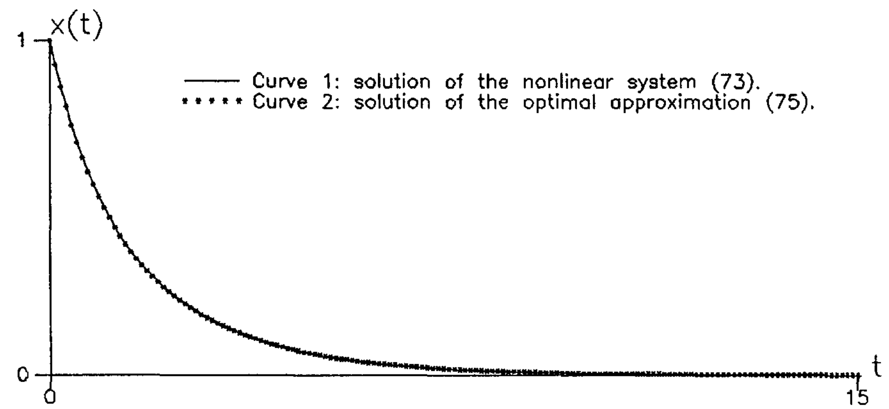


Figure 1. Represents the variation of the solution $x(t)$ as a function of time for the initial conditions $(x(0), y(0)) = (1, 0.5)$.

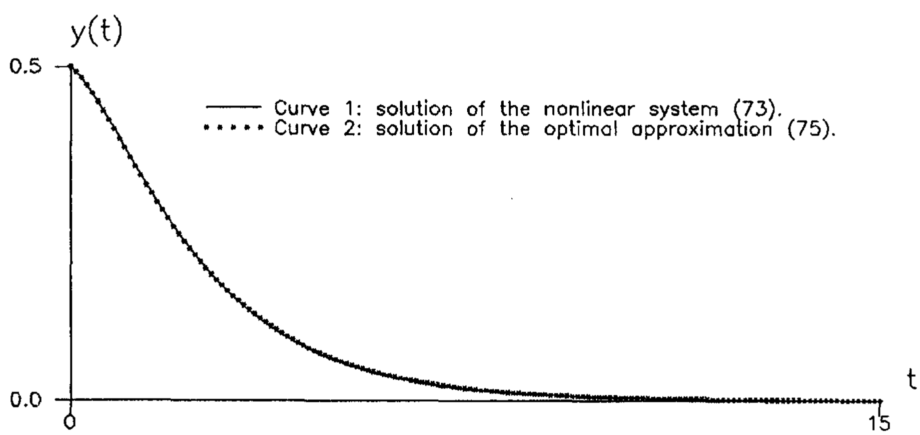


Figure 2. Represents the variation of the solution $y(t)$ as a function of time for the initial conditions $(x(0), y(0)) = (1, 0.5)$.

### 6.3. Examples of a Nonlinear Ordinary Differential Equation

EXAMPLE 10. Consider the following system:

$$\begin{aligned}\frac{dx}{dt} &= -\frac{x^3}{x^2+y^2},\\ \frac{dy}{dt} &= -\frac{y^3}{x^2+y^2},\end{aligned}\qquad (x_0, y_0) = (1, 0.5). \tag{78}$$

Here, the first and second derivatives of $F$, $DF(0)$ and $D^2F(0)$ are equal to zero. In this case, the proposed approximation is of order three

$$\|x(t) - y^*(t)\| = 0\left(\|x_0\|^3\right). \tag{79}$$

With the Jacobian matrix $DF(x)$ computed at $(x_0, y_0)$

$$DF(x_0, y_0) = \begin{bmatrix} -1.12 & 0.64 \\ 0.16 & -0.52 \end{bmatrix}, \qquad (x_0, y_0) = (1, 0.5), \tag{80}$$

and after the seven iterations, the computational procedure gives ($\varepsilon = 10^{-6}$)

$$A^* = \begin{bmatrix} -1.1030 & 0.6127 \\ 0.3085 & -0.7844 \end{bmatrix}, \qquad (x_0, y_0) = (1, 0.5). \tag{81}$$

Table 2 shows the values of the solutions of systems (78) and (81) and the relative error (77). The curves in Figures 3 and 4 represent the graphs of the respective components $(x(t), y(t))$ of solutions of systems (78) and (81) as a function of time.

Table 2.

| $t$ | $X_{\text{nl}}(t)$ | $X_{\text{lin}}(t)$ | $Y_{\text{nl}}(t)$ | $Y_{\text{lin}}(t)$ | $E_r$ |
|---|---|---|---|---|---|
| 0 | .1000000E+01 | .1000000E+01 | .5000000E+00 | .5000000E+00 | 0 |
| 1 | .4914913E+00 | .4909525E+00 | .3742482E+00 | .3723457E+00 | .003 |
| 2 | .2736490E+00 | .2723794E+00 | .2472790E+00 | .2457177E−02 | .005 |
| 3 | .1607496E+00 | .1607859E+00 | .1548591E+00 | .1557334E+00 | .003 |
| 4 | .9635604E−01 | .9755226E−01 | .9504148E−01 | .9722868E−01 | .018 |
| 5 | .5818971E−01 | .5986447E−01 | .5789638E−01 | .6035041E−01 | .036 |
| 6 | .3523752E−01 | .3690589E−01 | .3517207E−01 | .3737442E−01 | .055 |
| 7 | .2136008E−01 | .2279388E−01 | .2134548E−01 | .2312476E−01 | .075 |
| 8 | .1295274E−01 | .1408825E−01 | .1294948E−01 | .1430294E−01 | .096 |
| 9 | .7855611E−02 | .8710063E−02 | .7854884E−02 | .8845290E−02 | .11 |
| 10 | .4764529E−02 | .5385614E−02 | .4764367E−02 | .5469839E−02 | .13 |
| 11 | .2889802E−02 | .3330189E−02 | .2889766E−02 | .3382419E−02 | .16 |
| 12 | .1752747E−02 | .2059256E−02 | .1752739E−02 | .2091590E−02 | .18 |
| 13 | .1063093E−02 | .1273371E−02 | .1063091E−02 | .1293374E−02 | .20 |
| 14 | .6447982E−03 | .7874095E−03 | .6447978E−03 | .7997808E−03 | .23 |
| 15 | .3910898E−03 | .4869080E−03 | .3910897E−03 | .4945586E−03 | .25 |

EXAMPLE 11. Consider the following system [9]:

$$\begin{aligned}\frac{dx}{dt} &= -x - \frac{2y}{\ln(x^2+y^2)},\\ \frac{dy}{dt} &= -y + \frac{2x}{\ln(x^2+y^2)},\end{aligned}\qquad (x_0, y_0) = (0, 0.5) \tag{82}$$

in the open unit disk $\{(x, y) \in \mathbb{R}^2;\ x^2 + y^2 < 1\}$.

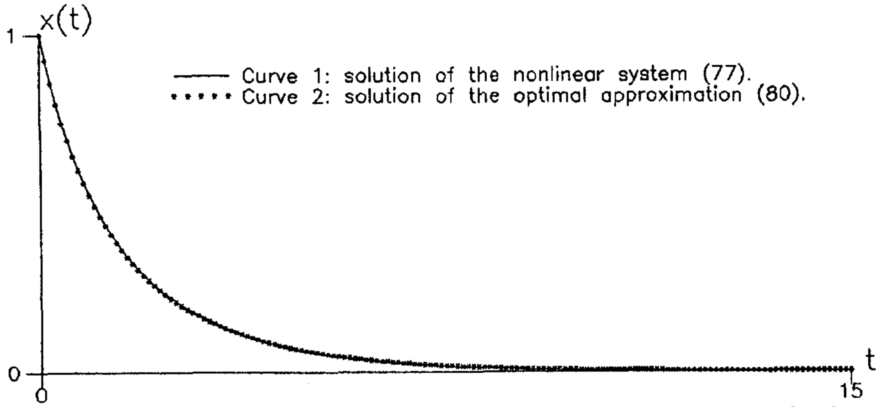


Figure 3. Represents the variation of the solution $x(t)$ as a function of time for the intitial conditions $(x(0), y(0)) = (1, 0.5)$.

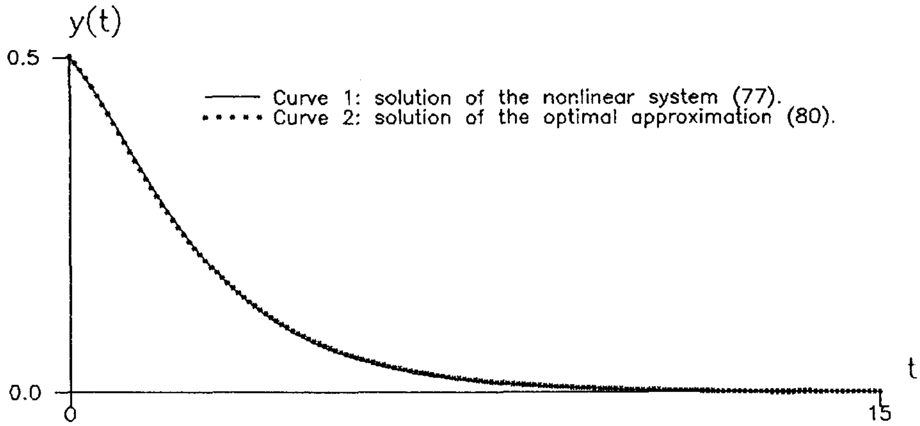


Figure 4. Represents the variation of the solution $y(t)$ as a function of time for the initial conditions $(x(0),\ y(0)) = (1, 0.5)$.

The linearization of $F$ at $(x_0, y_0) = (0, 0.5)$ gives

$$DF(x_0, y_0) = \begin{bmatrix} -1 & 3.524 \\ -1.4426 & -1 \end{bmatrix}, \qquad (x_0, y_0) = (0, 0.5). \tag{83}$$

After the ten iterations, the computational procedure gives ($\varepsilon = 10^{-6}$)

$$A^* = \begin{bmatrix} -1.4934 & 1.2489 \\ -0.5213 & -1.1254 \end{bmatrix}, \qquad (x_0, y_0) = (0, 0.5). \tag{84}$$

Table 3 shows the values of the solutions of systems (82) and (84) and the relative error (77). In both Figures 5 and 6, curve 1 corresponds to the solution of the nonlinear system (82) and curve 2 corresponds to the solution of the optimal approximation (84).

### 6.4. Comparison between the Linearization by the Fréchet Derivative and the Optimal Approximation

EXAMPLE 12. Consider the system

$$\begin{aligned} \frac{dx}{dt} &= -7.10^2 x - 2.10^3 x^2 - 2.10^5 y, \\ \frac{dy}{dt} &= 2.10^3 x - 2.10^5 y, \end{aligned} \qquad (x_0, y_0) = (5, 0). \tag{85}$$

The Fréchet derivative at 0 can be written

$$\begin{bmatrix} -7.10^2 & -2.10^5 \\ 2.10^3 & -2.10^5 \end{bmatrix}, \qquad (x_0, y_0) = (5, 0), \tag{86}$$

Table 3.

| $t$ | $X_{\text{nl}}(t)$ | $X_{\text{lin}}(t)$ | $Y_{\text{nl}}(t)$ | $Y_{\text{lin}}(t)$ | $E_r$ |
|---|---|---|---|---|---|
| 0 | .0000000 | .0000000E+00 | .5000000E+00 | .5000000E+00 | 0 |
| 1 | .1432933E+00 | .1517739E+00 | .1153286E+00 | .1177924E+00 | 0.048 |
| 2 | .6613008E−01 | .5793614E−01 | .1434192E−01 | .8519892E−02 | 0.148 |
| 3 | .2476355E−01 | .1105308E−01 | −.2539680E−02 | −.5333531E−02 | 0.562 |
| 4 | .8627967E−02 | −.3666260E−05 | −.3069739E−02 | −.2656959E−02 | 0.946 |
| 5 | .2898245E−02 | −.8070500E−03 | −.1717583E−02 | −.6254748E−03 | 1.14 |
| 6 | .9504736E−03 | −.3078052E−03 | −.7953907E−03 | −.4509668E−04 | 1.18 |
| 7 | .3056629E−03 | −.5867217E−04 | −.3383072E−03 | .2837578E−04 | 1.06 |
| 8 | .9643728E−04 | .3895649E−07 | −.1372356E−03 | .1411885E−04 | 0.932 |
| 9 | .2978096E−04 | .4291440E−05 | −.5404247E−04 | .3321252E−05 | 1.01 |
| 10 | .8954184E−05 | .1635317E−05 | −.2085933E−04 | .2386980E−06 | 0.983 |
| 11 | .2595794E−05 | .3114442E−06 | −.7937176E−05 | −.1509661E−06 | 0.971 |
| 12 | .7123896E−06 | −.3104541E−09 | −.2988374E−05 | −.7502629E−07 | 0.971 |
| 13 | .1780326E−06 | −.2281944E−07 | −.1116058E−05 | −.1763572E−07 | 0.988 |
| 14 | .3641637E−07 | −.8688157E−08 | −.4141681E−06 | −.1263416E−08 | 0.999 |
| 15 | .3343012E−08 | −.1653209E−08 | −.1529153E−06 | .8031747E−09 | 1.00 |

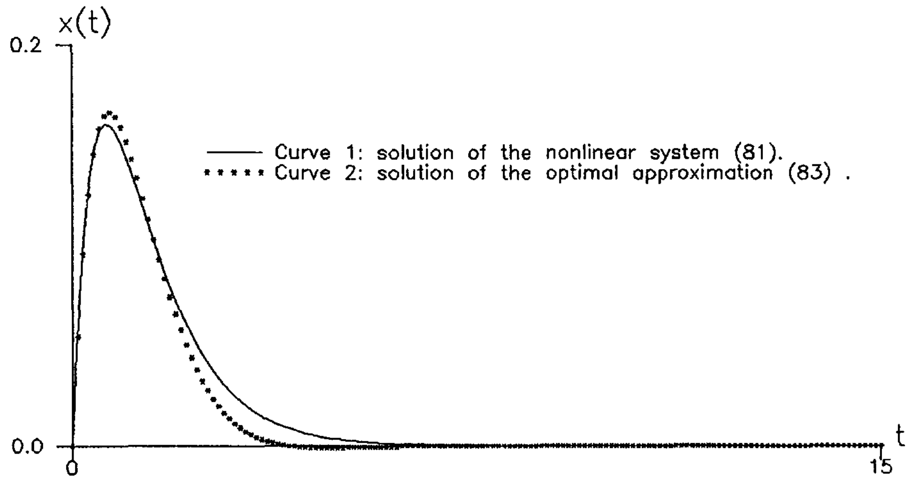


Figure 5. Represents the variation of the solution $x(t)$ as a function of time for the initial conditions $(x(0), y(0)) = (0, 0.5)$.

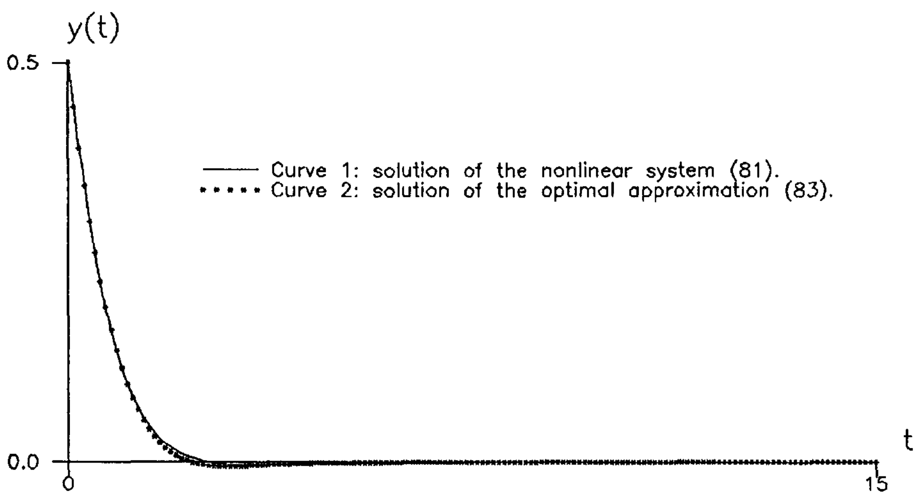


Figure 6. Represents the variation of the solution $y(t)$ as a function of time for the initial conditions $(x(0), y(0)) = (0, 0.5)$.

and the optimal approximation gives

$$\begin{bmatrix} -13.270^3 & 4.3110^5 \\ 2.10^3 & -2.10^5 \end{bmatrix}, \qquad (x_0, y_0) = (5, 0). \tag{87}$$

In Figures 7 and 8, curve 1 corresponds to the solution of nonlinear system, curve 2 the solution of the optimal approximation, and curve 3 the solution of the Fréchet derivative at 0.

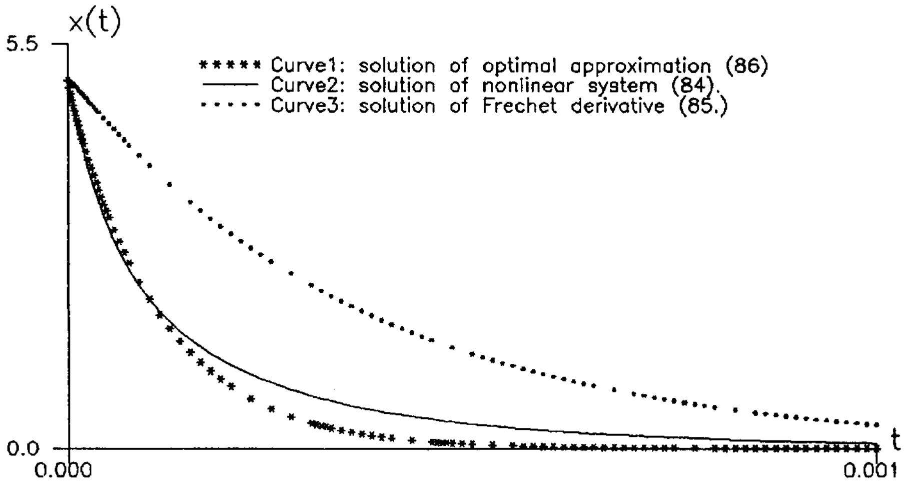


Figure 7. Represents the variation of the solution $x(t)$ as a function of time for the initial conditions $(x(0), y(0)) = (5, 0)$.

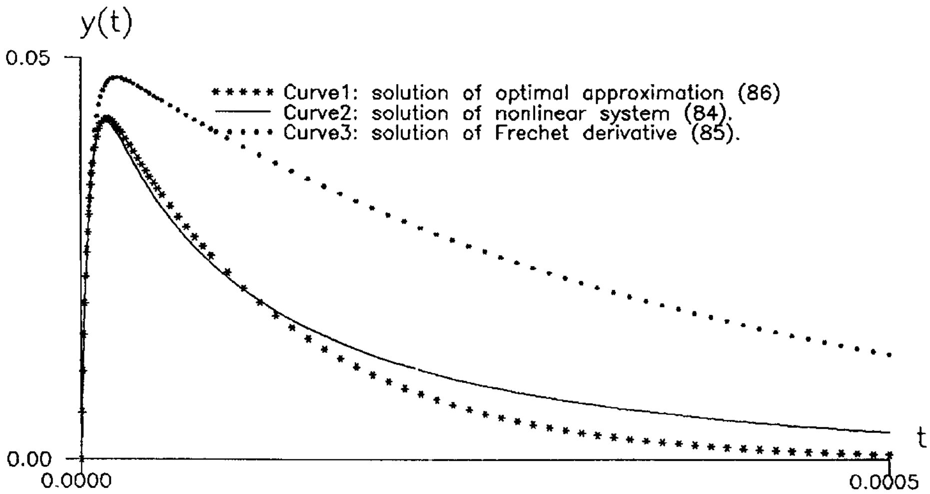


Figure 8. Represents the variation of the solution $y(t)$ as a function of time for the initial conditions $(x(0), y(0)) = (5, 0)$.

## 7. COMMENTS

As a continuation of earlier work [1], we have presented in this paper further developments regarding the optimal linearization method. The emphasis here was put on the use of the method as an approximation procedure. Our main results stipulate that the approximation is of order two with respect to the initial value, and is generally of the same order as the nonlinearity.

We included several examples showing satisfactory adequacy of approximate results compared to the exact ones. This is confirmed by the computation of the relative error which never exceeds 120%. This could be considered a high figure, but we point out that the highest rates correspond to portion of the solutions very close to the origin, where in fact the best approximation is certainly the one provided by the standard linearized equation.

Differences between the two approximations are well reflected in Figures 7 and 8, where one can see that the optimal approximation does much better than the Fréchet derivative as long as $x(t)$ stays far from zero.

## REFERENCES


1. T. Benouaz and O. Arino, Determination of the stability of a nonlinear ordinary differential equation by least square approximation: 1. Computational procedure, *Applied Mathematics and Computer Science* (to appear).
2. B. Vujanovic, Application of the optimal linearization method to the heat transfer problem, *International Journal Heat Mass Transfer* **16**, 1111–1117 (1973).
3. A. Jordan, M. Benmouna, A. Bensenane and A. Borucki, Optimal linearization method applied to the resolution of state equation, *RAIRO-APII* **21**, 175–185 (1987).
4. A. Jordan, M. Benmouna, A. Bensenane and A. Borucki, Optimal linearization of non-linear state equation, *RAIRO-APII* **21**, 263–271 (1987).
5. T. Benouaz and O. Arino, Existence, unicité et convergence de l'approximation au sens des moindres carrés d'une équation differentielle ordinaire non-linéaire, 94/14, Université de Pau, CNRS URA, p. 1204, (1994).
6. T. Benouaz, Application of the optimal linearization for study stability problem, *Third International Colloquium on Numerical Analysis,* Plovdiv, Bulgaria, 13–17 August 1994.
7. W. Ralston, *Mathematical Methods for Digital Computers*, pp. 110–120, Wiley, New York, (1960).
8. R.E. Kalman and E.J.E. Bertran, Control system analysis and design via second method of Lyapunov, *Jun. Journal of Basic Engineering,* 371–393 (1960).
9. J.P. Demailly, *Analyse numérique et équations différentielles*, Presses universitaires de Grenoble, (1991).
10. H. Reinhart, *Equations Différentielles, Fondements et Applications*, Gauthier-Villars.