RICE UNIVERSITY

A LINEAR TIME-VARYING MODEL FOR NONLINEAR SYSTEMS

BY

Tien-Lin Chang

A THESIS SUBMITTED IN
PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

Thesis Director's Signature:

C. S. Burrus

Houston, Texas

May 1969

ABSTRACT

A LINEAR TIME-VARYING MODEL FOR NONLINEAR SYSTEMS

Tien-Lin Chang

In this thesis we present an approach to the approximate solution of a class of nonlinear autonomous systems. We use a linear time-varying model, $\dot{x} + a(t)x = 0$ to approximate the nonlinear system $\dot{x} + f(x) = 0$, i.e., the solution of the former system is a good approximation to the solution of the latter one.

The function $a(t)$ will depend on the nature of $f(x)$ and on the initial condition $x(0)$. If we choose $a(t)$ to be a constant, then this method reduces to the conventional iterative method which uses the linear time-invariant model $\dot{x} + ax = 0$ as an approximation to the nonlinear system.

We may choose $a(t)$ in several ways. Here we assume a form with an undetermined parameter for $a(t)$ and then we determine the parameter by matching the trajectories of both systems in the phase plane.

This method can also be applied to second order system with some modification. Excellent results are obtained when applied to specific examples. Also it is a promising idea to extend this method to driven nonlinear systems.

ACKNOWLEDGMENT

The author would like to express his deepest gratitude to Dr. C.S. Burrus, research director, for his constant inspiration, valuable ideas, helpful suggestions, and patient guidance of this research. None of this research would have been possible without his assistance.

Special thanks are due to Drs. R.J.P. de Figueiredo and T.W. Parks for many suggestions and constructive criticisms.

Partial support for this study was provided by NSF Grant GK-807.

TABLE OF CONTENTS

I. INTRODUCTION

1-1. Nonlinear Systems

A physical system can sometimes be represented by a differential equation of the form $\dot{X} + F(X,t) = 0$, where X is a vector, $F(X,t)$ is a vector function of X and t. In this thesis we develop an approximate method to solve a class of nonlinear differential equations. This method uses a linear time-varying model $\dot{X} + A(t)X = 0$ to approximate the original nonlinear differential equation.

Typical examples of nonlinear systems are a mass on a nonlinear spring without friction, and a series RL circuit with an iron core inductor. The differential equations for these systems are

$$\ddot{x} + f(x) = 0 \quad , x(0), \dot{x}(0) \text{ are known}$$

$$\dot{x} + f(x) = 0 \quad , x(0) \text{ is known}$$

We shall consider only first and second order differential equations of the above type in this thesis.

1-2. Preliminary Considerations

We have a system of the form

$$\dot{X} + F(X) = 0 \quad , \quad X(t_o) = X_o \tag{1-1}$$

where $F(X)$ is defined, continuous and satisfies a Lipschitz condition in a certain region of (X,t) space. Then there exists a unique solution satisfying any initial condition in this region for the system. [1].

We now propose an auxiliary linear time-varying model for the original system

$$\dot{X} + A(t)X = 0 \quad , \quad X(t_o) = X_o \tag{1-2}$$

where $A(t)$ is defined and continuous in this region of interest, and then X always has a unique solution satisfying the same initial condition in this same region. [1]

If these two systems are exactly the same, then we have

$$A(t) = \frac{F(X(t))}{X(t)} \triangleq \begin{bmatrix} \frac{f_1(X(t))}{x_1(t)} & 0 \\ 0 & \frac{f_2(X(t))}{x_2(t)} \end{bmatrix}$$

where

$$X(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} \quad \text{and} \quad F(X(t)) = \begin{bmatrix} f_1(X(t)) \\ f_2(X(t)) \end{bmatrix}$$

Then the solution of (1-2) is

$$X(t) = \Phi(t,t_o)\, X(t_o)$$

where $\Phi(t,t_o) = \mathrm{EXP}\left[\int_{t_o}^{t} A(s)ds\right]$. Also, this is the solution to (1-1).

If we know $X(t)$, then it is a trivial problem to find $A(t)$. We now want to do the inverse problem. We want to find $A(t)$ from what we know about the differential equation. We hope we can get some information from $\frac{F(X)}{X}$. Here we introduce an approximation, we assume a functional form for $A(t)$ with some undetermined parameters in it. This function $A(t;\alpha) = A'(t; X_o,t_o; \alpha)$, where X_o is the initial value of $X(t)$ at initial time t_o and α is a vector, should more or less look like the composite function $\frac{F(X(t))}{X(t)}$, if $X(t)$ is not far away from linear or is fairly smooth. After assuming $A(t;\alpha)$, we solve (1-2), the solution is an approximate solution to (1-1)

$$\tilde{X} = \tilde{X}(t;X_o,t_o;\alpha) = \mathrm{EXP}\left[-\int_{t_o}^{t} A(s;X_o,t_o;\alpha)\,ds\right] \cdot X_o$$

Substituting into (1-1) we have

$$\dot{\tilde{X}} + F(\tilde{X}) = \epsilon(t;X_o,t_o;\alpha) \tag{1-3}$$

We are to find α by minimizing ϵ in some sense, then we have an approximate solution $\tilde{X}(t)$.

This method differs from usual averaging method. [2,3] In the averaging method, we assume a solution form with some undetermined parameters from our prior knowledge of the solution. Substituting into the differential equation and minimizing the equation error $\epsilon(t;a)$ in some sense, we find the a's for the assumed solution. While in this method we assume a function form for $\frac{\dot{X}}{X} = \frac{F(X)}{X}$, that means we approximate the trajectory of the nonlinear system in the $(X,\dot{X})$ phase plane by another curve that is the solution of a linear time-varying system. In general, getting a qualitative description of trajectory is easier than finding a qualitative description of the solution, but both methods have the drawback that the assumed function should not be too complicated to manipulate.

This method provides excellent results if we have the appropriate function form for $A(t)$. If $A(t)$ is a constant, then this method reduces to the ordinary iterative method. [4]

We shall illustrate the method in the following sections and solve two examples in detail. The function form we assume for $A(t;\alpha)$ in the first order system problem is $A(t;a) = 1-x_0^2\, e^{-at}$, where x_o is the initial value for $x(t)$, a is a positive undetermined parameter. The approximate solution obtained agrees extremely well with the true solution and is much better than either that obtained by the perturbation method or by the iteration method as shown in Fig. 1. In the second order system problem $A(t;X_o,t_o;\alpha)$ is assumed to be

$$\begin{bmatrix} a_2 \tan a_1 t & 0 \\ 0 & -a_3 \operatorname{ctn} a_1 t \end{bmatrix}$$ where X_o is implicit in the a's. We solve the famous differential equation that gives rise to elliptic function.[5] We again have the approximate solution very close to the true solution as shown in Fig. 2.

In the last section we shall give a simple error analysis in the phase plane that will tell how good our approximate solution is. Also we find an error bound for the approximate solution.

II. DESCRIPTION OF THE METHOD

2-1. Convergence

We want first to establish the convergence property of this method. This method is in some sense an iterative procedure. [4]

We consider a differential equation with the initial conditions

$$\dot{X} + F(X) = 0 \quad , \quad X(0) = B \tag{2-1}$$

and the auxiliary linear time-varying model is

$$\dot{\tilde{X}} + A(t)\tilde{X} = 0 \ , \quad \tilde{X}(0) = B \quad . \tag{2-2}$$

Defining a function to represent the difference between nonlinear and linear equations by

$$G(X,t) = F(X) - A(t)X \tag{2-3}$$

We write the iterative procedure in the form

$$\tilde{X}^{(1)}(t) = e^{-K(t)}\tilde{X}(0)$$

where $K(t) = \int_o^t A(t)\ dt$

$$\tilde{X}^{(n+1)}(t) = \tilde{X}^{(n)}(t) - \int_o^t e^{-K(t)+K(\tau)}\epsilon^{(n)}(\tau)\ d\tau \tag{2-4}$$

where

$$\epsilon^{(n)} = \tilde{X}^{(n)} + F(\tilde{X}^{(n)}) \tag{2-5}$$

Substituting (2-3), (2-5) into (2-4) we have

$$\tilde{X}^{(n+1)} = \tilde{X}^{(n)} - \int_o^\tau e^{-K(t)+K(\tau)}\left[\dot{\tilde{X}}^{(n)} + A(\tau)\tilde{X}^{(n)} + G(\tilde{X}^{(n)},\tau)\right] d\tau \tag{2-6}$$

Consider the first term of the integral and integrate by parts to obtain

$$\int_o^t e^{-K(t)+K(\tau)}\left[\dot{\tilde{X}}^{(n)}(\tau)\right] d\tau = \tilde{X}^{(n)} - \tilde{X}^{(n)}(0)e^{-K(t)}$$

$$-\int_o^t \tilde{X}^{(n)}A(\tau)\left[e^{-K(t)+K(\tau)}\right] d\tau$$

substitute this result back into (2-6)

$$\tilde{X}^{(n+1)} = \tilde{X}^{(n)}(0)\ e^{-K(t)} - \int_0^t e^{-K(t)+K(\tau)}\ G(\tilde{X}^{(n)},\tau)\,d\tau$$

We will assume that X and t are restricted to a region R defined by*

$$|t| \leq \alpha \quad , \quad |\tilde{X} - \tilde{X}(0)| \leq \beta$$

the function G is assumed continuous, bounded and satisfies a Lipschitz condition in R , i.e.

$$|G(X,t)| \leq M \quad , \quad (X,t)\in R$$

and

$$|G(X,t) - G(Y,t)| \leq L|X - Y| \quad , (X,t) \text{ and } (Y,t)\in R$$

where L is a Lipschitz constant. We also assume

$$|e^{-K(t)+K(\tau)}| \leq H \quad , \quad t\in R \quad .$$

To show our iterative procedure converges in R we shall show if the initial approximation is in R , all succeeding approximations will be also, and that $|\tilde{X}^{(n+1)} - \tilde{X}^{(n)}| \to 0$ as $n \to \infty$.

---

* $|X|$ means a norm of X , here we have $|X| = \Sigma|x_i|$, $|A| = \sum_{i,j} |a_{ij}|$.

$$|\tilde{X}^{(2)} - \tilde{X}^{(1)}| = |\int_o^t e^{-K(t)+K(\tau)} G(\tilde{X}^{(1)},\tau)\, d\tau|$$

$$\leq HM \int_o^t d\tau$$

$$\leq HM|t|$$

$$\leq HM\alpha$$

If we restrict the time interval α such that

$$\alpha \leq \beta/HM$$

then

$$|\tilde{X}^{(2)} - \tilde{X}^{(1)}| \leq \beta$$

by induction we find all $\tilde{X}^{(n)}$ in R . Also we can find the difference, $|\tilde{X}^{(n+1)} - \tilde{X}^{(n)}| \to 0$ as $n \to \infty$, i.e.,

$$|\tilde{X}^{(n+1)} - \tilde{X}^{(n)}| \leq \frac{ML^n H^{n+1} t^{n+1}}{(n+1)!} \to 0 \quad \text{as} \quad n \to 0 \quad .$$

This gives a convergence of the solution in a restricted region. If in practical problems we cannot iterate analytically due to mathematical complexities, then we get only the first approximation. If we can iterate analytically, the approximate solution will converge to the true solution.

2-2. Method for First Order Systems and an Example

Consider a first order system

$$\dot{x} + f(x) = 0 \qquad x(0) = A \tag{2-7}$$

We propose that it is almost equivalent to another system

$$\dot{\tilde{x}} + K(t)\tilde{x} = 0 \qquad \tilde{x}(0) = A \tag{2-8}$$

The solution for (2-8) is easily obtained

$$\tilde{x}(t) = A \text{ EXP } \left[-\int_0^t K(t)dt \right] \tag{2-9}$$

substituting into (2-7) we have

$$\dot{\tilde{x}} + f(\tilde{x}) = \epsilon(t) \tag{2-10}$$

So $\tilde{x}$ is an approximate solution for the original system. Now we are to find the conditions on $K(t)$, if the time-varying linear system represents in some sense the original nonlinear system.

We know from (2-8) and (2-10)

$$K(t) = \frac{f(\tilde{x}(t)) - \epsilon(t)}{\tilde{x}(t)}$$

A first order system will start at an initial point and end at a singular point or go to infinite, so we match the initial point and the final point. At these two points $\epsilon(t) = 0$, $K(t)$ has known values. Assuming that $\epsilon(t)$ is small, we have

$$K(t) \approx \frac{f(\tilde{x}(t))}{\tilde{x}(t)}$$

or

$$\bar{K}(\tilde{x}) \approx \frac{f(\tilde{x})}{\tilde{x}}$$

and $\bar{K}(\tilde{x}(t)) = K(t)$.

We know of the graph of $\frac{f(x)}{x}$ v.s. x , so we know $\bar{K}(x)$. Now if the true solution $x(t)$ is linear or fairly smooth, then the composite function $\bar{K}(x(t))$ is approximately equal to $\bar{K}(t)$ and to $K(t)$. So we can assume a function form for $K(t)$ and leave some parameter undetermined, i.e.

$$K(t) = g(t,a)$$

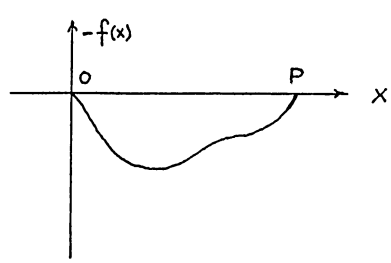
substituting into (2-9) and (2-10) we have $\epsilon(t,a)$. Minimizing ϵ in some sense we find a and $\tilde{x}(t)$.

We shall now restrict $f(x)$ to have some properties. First we consider two special cases.

Case I.

i) $f(x) : \begin{cases} = 0 , & x = 0 \text{ and } x = p > 0 \\ > 0 , & p > x > 0 \end{cases}$

ii) $\frac{d^2 f(x)}{dx^2} \leq 0 \quad \forall\, x \epsilon (0,P)$

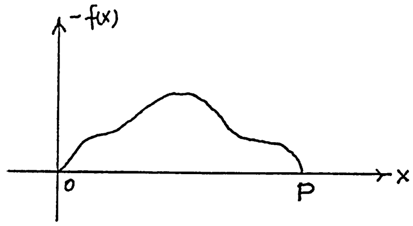


Then $\frac{f(x)}{x}$ is positive and monotonically increasing, so $K(t;a)$ is a positive monotonic increasing function in t .

Case II.

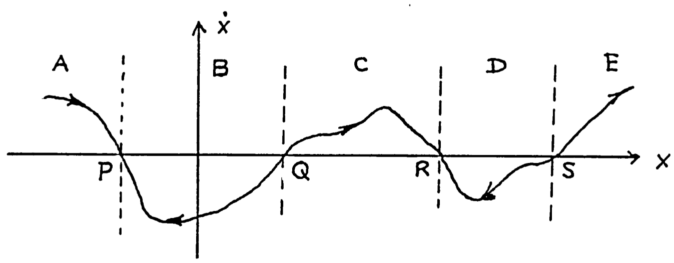
i) $f(x) : \begin{cases} = 0 & x = 0 \text{ and } x = P > 0 \\ < 0 & P > x > 0 \end{cases}$

ii) $\frac{d^2 f(x)}{dx^2} \geq 0 \qquad \forall\, x \in (0,P)$



Now $\frac{f(x)}{x}$ is negative and monotonically decreasing, so we can assume a negative monotonic decreasing function for $K(t;a)$.

We now consider a more general f(x) which has the graph



We divide it into five regions A, B, C, D and E. P and R are stable singular points, Q and S are unstable singular points. The direction of motion of representative points is also shown in the graph. We notice region D is the same as case I, if we shift the vertical axis to the point R. Let

$$x = R + x'$$

then $\dot{x} = \dot{x}'$ and $f(x) = f(R+x') = f'(x')$, so we have $\dot{x}' + f'(x') = 0$. We can solve this by using the techniques of case I. Similar procedures can be applied to other regions.

We shall now work out an example in detail. The system is

$$\dot{x} + x - x^3 = 0 \qquad x(0) = A = 0.9 \tag{2-11}$$

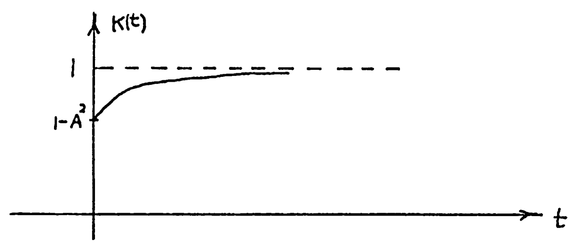
The auxiliary linear time-varying model is

$$\dot{\tilde{X}} + K(t)\,\tilde{x} = 0 \qquad \tilde{x}(0) = A = 0.09 \tag{2-12}$$

We know then

$$\frac{f(x)}{x} = \frac{x - x^3}{x} = 1 - x^2$$

and $K(t) \approx 1 - x^2(t)$. $x(t)$ is a monotonic decreasing function, so $K(t)$ is a monotonic increasing function. $K(0)$ is known and equals $1 - A^2$, $K(\infty) = 1$, we can propose the following graph for $K(t)$



Here we can have various function forms for $K(t)$, we want the method to be carried out analytically so we search a tractable elementary function for $K(t)$. $K(t)$ is assumed to be

$$K(t;b) = 1 - A^2 e^{-bt} \tag{2-13}$$

where b is a positive undetermined parameter. Then our approximate solution is

$$\tilde{x}(t) = A\ \mathrm{EXP}\left[-\int_0^t K(t;b)\,dt\right] \tag{2-14}$$

Substituting (2-13), (2-14) into (2-11), we have

$$\epsilon(t;b) = -A^3\left(e^{-3(t + \frac{A^2}{b}(e^{-bt} - 1))} - e^{-(t + \frac{A^2}{b}(e^{-bt} - 1))} \cdot e^{-bt}\right)$$

We can minimize ϵ in various ways. For instance, we let

$$\int_0^\infty \epsilon(t;b)dt = 0$$

or

$$\int_0^\infty \epsilon \frac{\partial\epsilon}{\partial b} dt = 0$$

to find the best value for b . Since $\epsilon(t;b)$ is too complicated, we can hardly integrate it analytically. We then examine the plot of $\dot{\tilde{x}}$ vs. $\tilde{x}$. We find that the second derivative of $\dot{\tilde{x}}$ with respect to $\tilde{x}$ is always greater than zero, so we know the graph of $\dot{\tilde{x}}$ vs. $\tilde{x}$ is a concave upward curve. This curve indeed must look like the phase plot of our original system. Now our criterion is to let $\epsilon(T;b) = 0$ at a certain fixed time T . If we have $\epsilon(t;b) = \bar{\epsilon}(\tilde{x};b)$, then the above criterion is similar to letting $\int_0^A \bar{\epsilon}\ (\tilde{x};b)\ d\tilde{x} = 0$. We also know that when time t is large, the system behaves more or less like a linear system, so we choose the T in $\epsilon(T;b) = 0$ to be small. Then we solve a nonlinear algebraic equation.

$$\epsilon(T;b) = -A^3\left[e^{-3(T + \frac{A^2}{b}(e^{-bT} - 1))} - e^{-(T + \frac{A^2}{b}(e^{-bT} - 1))} \cdot e^{-bT}\right] = 0$$

For simplicity, choose $T = 1$, we have the value for b , $b = 0.98$, if $A = 0.9$. The approximate solution is

$$\tilde{x} = 0.9 \text{ EXP}\left[-t - 0.826(e^{-0.98t} - 1)\right]$$

We compare this solution with the true solution and the solutions obtained by using perturbation and iteration methods in Fig. 1.* We can have a better approximate solution if we let $\epsilon(t;b) = 0$ at $T = 0.8$, "better" is in the sense of smaller maximum solution error; if T is chosen smaller than 0.8, we get again larger maximum error.

This method is not difficult to carry out, the solution will in general be "better" than that of perturbation and iteration methods. The critical parts are the assumption of the function form for $K(t;b)$ and the choice at T at which $\epsilon(T;b)$ is zero. There is always a trade-off between the "accuracy" of the approximate solution and the complexity of the assumed function.

2-3. Method for Second Order Systems and an Example

We have a second order system of the form

$$\ddot{X} + f(x) = 0$$

We shall consider only conservative systems here, $f(x)$ is assumed to be continuous and satisfy a Lipschitz condition. We are not to consider singular solutions; we are given two initial conditions, $x(0) = A$, $\dot{x}(0) = B$. Then we find a trajectory corresponding to the solution of the system;it must be a closed curve. Now let

$$x_1 = x$$

$$x_2 = \dot{x}$$

* Figures shown at the end.

so the original equation becomes

$$\begin{cases} \dot{x}_1 = x_2 & x_1(0) = A \\ \dot{x}_2 = -f(x_1) & x_2(0) = B \end{cases} \tag{2-15}$$

We assume as before an auxiliary linear time-varying model

$$\begin{cases} \dot{\tilde{x}}_1 = -a_{11}(t)\tilde{x}_1 & \tilde{x}_1(0) = Z \\ \dot{\tilde{x}}_2 = -a_{22}(t)\tilde{x}_2 & \tilde{x}_2(0) = B \end{cases} \tag{2-16}$$

Now if these two systems are more or less equivalent, then from (2-15) and (2-16) we have

$$a_{11}(t) = \frac{-[\tilde{x}_2(t) + \epsilon_1(t)]}{\tilde{x}_1(t)} \tag{2-17a}$$

$$a_{22}(t) = \frac{f[\tilde{x}_1(t)] - \epsilon_2(t)}{\tilde{x}_2(t)} \tag{2-17b}$$

We apply the same techniques, we assume function forms for $a_{11}(t;b_1,b_2,\ldots b_n)$ and $a_{22}(t;b_1,b_2,\ldots b_n)$, substitute into the original equation (2-15) we find $\epsilon_1(t;b_1,b_2,\ldots b_n)$ and $\epsilon_2(t;b_1,b_2,\ldots b_n)$, and minimize ϵ_1,ϵ_2 in some sense we get $a_{11}(t)$ and $a_{22}(t)$. The approximate solution then is

$$\tilde{x}_1 = Ae^{-\int_o^t a_{11}(t)dt} \tag{2-18a}$$

$$\tilde{x}_2 = Be^{-\int_o^t a_{22}(t)dt} \tag{2-18b}$$

Because f(x) is too general to give any insight into the method, we shall consider an example in detail.

Let $f(x) = x - x^3$, so the system is the differential equation that gives rise to elliptical function

$$\ddot{x} + x - x^3 = 0 \qquad x(0) = \alpha \quad \dot{x}(0) = \beta$$

Write it as two first order differential equations

$$\begin{cases} \dot{x}_1 = x_2 \quad , \quad x_1(0) = \alpha \\ \dot{x}_2 = x_1^3 - x_1 \, , \, x_2(0) = \beta \end{cases} \tag{2-19}$$

The trajectories in the phase plane are

$$x_1^2 - \tfrac{1}{2}x_1^4 + x_2^2 = C \tag{2-20}$$

where C is determined by initial conditions. Let the auxiliary linear time-varying model be

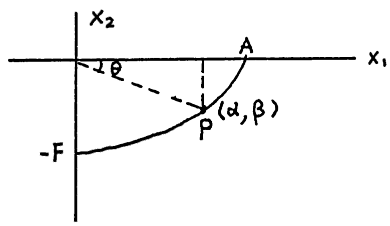
$$\begin{cases} \dot{\tilde{x}}_1 = -a_{11}\tilde{x}_1 \quad , \quad \tilde{x}_1(0) = \alpha \\ \dot{\tilde{x}}_2 = -a_{22}\tilde{x}_2 \quad , \quad \tilde{x}_2(0) = \beta \end{cases} \tag{2-21}$$

From (2-19) and (2-21) we have

$$a_{11}(t) = \frac{-\tilde{x}_2(t) - \epsilon_1(t)}{\tilde{x}_1(t)} \tag{2-22a}$$

$$a_{22}(t) = \frac{-\tilde{x}_1^3(t) + \tilde{x}_1(t) - \epsilon_2(t)}{\tilde{x}_2(t)} \qquad (2\text{-}22b)$$

Let us illustrate a_{11} and a_{22} in the phase plane. If ϵ_1 and ϵ_2 are small, then $\tilde{x}_1 \approx x_1$ and $\tilde{x}_2 \approx x_2$, we have then $a_{11} \approx -x_2/x_1$ and $a_{22} \approx (-x_1^3 + x_1)/x_2$. We show here a quadrant of the trajectory.



a_{11} is just the tangent of the angle θ , and a_{22} is just the slope at point P . Both a_{11} and a_{22} are functions of time.

Then we assume a function for $a_{11}(t)$

$$a_{11}(t) \approx c_2 \tan \theta_1(t) \ , \ 0 \leq \theta_1(t) \leq \frac{\pi}{2} \qquad (2\text{-}23)$$

where $\theta_1(t)$ is an unknown function. We know that $\theta_1(t)$ is an increasing function and its value at initial time $\theta_1(0) = 0$ and at quarter period $\theta_1(\frac{\pi}{4}) = \frac{\pi}{2}$. The simplest way to approximate $\theta_1(t)$ is to let

$$\theta_1(t) = c_1 t \qquad (2\text{-}24)$$

where c_1 is a positive parameter, we introduce error here. The same arguments apply to a_{22} , we have

$$a_{22}(t) = -c_3 \operatorname{ctn} \theta_2(t) \quad , \quad 0 \leq \theta \leq \frac{\pi}{2} \qquad (2\text{-}25)$$

and $\theta_2(t)$ is also assumed to be a linear function of t, i.e.,

$$\theta_2(t) = c_1{}'t \tag{2-26}$$

We let $c_1{}' = c_1$ for simplicity.

a_{11} and a_{22} should be related; let us examine the relationship between them. From (2-22a) and (2-22b) by neglecting ϵ_1 and ϵ_2 we have

$$a_{11}a_{22} \approx \frac{\tilde{x}_1\tilde{x}_2(\tilde{x}_1{}^2 - 1)}{\tilde{x}_1\tilde{x}_2} = \tilde{x}_1{}^2 - 1 \tag{2-27}$$

also from (2-23) and (2-25) we have

$$a_{11}a_{22} \approx -c_2c_3 \tan \theta_1(t) \operatorname{ctn} \theta_2(t)$$

If $\theta_1(t)$ and $\theta_2(t)$ were exact known functions we should have $a_{11}a_{22} = x^2 - 1$. Here we let $\theta_1(t) = c_1t = \theta_2(t)$ and c_2 and c_3 be constant, we commit errors to obtain an approximate analytic solution. The approximate solution after proper choice of time base looks like

$$\begin{aligned} \tilde{x}_1(t) &= Ae^{-\int_o^t c_2 \tan c_1 t dt} \\ &= Ae^{(c_2/c_1) \log \cos c_1 t} \\ &= A(\cos c_1 t)^{c_2/c_1} \qquad , \ 0 \le c_1 t \le \frac{\pi}{2} \end{aligned} \tag{2-28}$$

We have the solution in the time interval $0 \le t \le \frac{\pi}{2c_1}$, because when $c_1t > \frac{\pi}{2}$, $\cos(c_1t)$ is negative and in general c_2/c_1 is not an integer, we get no real solution. We know $\tilde{x}_1(t)$ is a periodic function, so the solution for a cycle shall be

$$\left.\begin{aligned}\tilde{x}_1(t) &= A[\cos(c_1t)]^{c_2/c_1} , \quad 0 \le c_1t \le \frac{\pi}{2} \\ &= -A[\cos\,(\pi - c_1t)]^{c_2/c_1} , \quad \frac{\pi}{2} \le c_1t \le \pi \\ &= -A[\cos\,(c_1t - \pi)]^{c_2/c_1} , \quad \pi \le c_1t \le \frac{3}{2}\pi \\ &= A[\cos\,(2\pi - c_1t)]^{c_2/c_1} , \quad \frac{3}{2}\pi \le c_1t < 2\pi \end{aligned}\right\} \tag{2-29}$$

We can get $\tilde{x}_2(t)$ in the same way, but need a little modification. If we proceed just the same as before, we have

$$\tilde{x}_2(t) = Ee^{-\int_0^t c_3 \operatorname{ctn}(c_1t)\,dt} . \tag{2-30}$$

This function is not defined at origin $t = 0$. We can evaluate $\tilde{x}_2(t)$ in another quadrant, because $\tilde{x}_2(t)$ is periodic and symmetric with respect to both vertical axis and origin. So we start at another point $\tilde{x}_2(0) = \max \tilde{x}_2(t) = F$, then we have

$$\tilde{x}_2(t) = Fe^{-\int_0^t c_3 \tan(c_1t)dt}$$

$$\tilde{x}_2(t) = F(\cos c_1t)^{c_3/c_1} , \quad 0 \le t \le \frac{\pi}{2c_1} \tag{2-31a}$$

$$= -F[\cos(\pi - c_1t)]^{c_3/c_1} , \quad \frac{\pi}{2} < c_1t \le \pi \tag{2-31b}$$

Now we change the time base, let

$$c_1t = \frac{\pi}{2} + c_1t'$$

substituting into (2-31b) we have

$$\tilde{x}_2(t') = -F[\cos(\pi - \frac{\pi}{2} - c_1 t')]^{c_3/c_1}$$
$$= -F[\sin c_1 t']^{c_3/c_1} \quad , \; 0 \le c_1 t' \le \frac{\pi}{2} \tag{2-32}$$

Now change t' back to t, we have $\tilde{x}_1(t)$ and $\tilde{x}_2(t)$ in the same time interval, i.e.,

$$\left.\begin{aligned} \tilde{x}_1(t) &= A[\cos(c_1 t)]^{c_2/c_1} \quad , \quad 0 \le c_1 t \le \frac{\pi}{2} \\ \tilde{x}_2(t) &= -F[\sin(c_1 t)]^{c_3/c_1} \quad , \quad 0 \le c_1 t \le \frac{\pi}{2} \end{aligned}\right\} \tag{2-33}$$

where A and F can be obtained from (2-20) if α, β are given. These two functions look nice, we have c_1 to approximate the frequency of the true solution $x_1(t)$, and the power of cos function to approximate the amplitude of the true solution $x_1(t)$. Let us substitute (2-33) into (2-22a) and (2-22b), we find

$$\epsilon_1(t;c_1,c_2,c_3) = -c_2 \tan(c_1 t)\, A\, (\cos c_1 t)^{c_3/c_1} + F(\sin c_1 t) \tag{2-34}$$

$$\epsilon_2(t;c_1,c_2,c_3) = -c_3\, \mathrm{ctn}(c_1 t)\, F(\sin c_1 t)^{c_3/c_1} - A^3 (\cos c_1 t)^{3c_2/c_1} + A(\cos c_1 t)^{c_2/c_1} \tag{2-35}$$

Before we find c_1, c_2, c_3 by minimizing ϵ_1, ϵ_2 in some sense, let us examine the plot of $\tilde{x}_2$ vs. $\tilde{x}_1$. We anticipate this graph resembles the plot of x_2 vs. x_1 , because this method is really just to fit the trajectory of the original nonlinear system in the phase plane by another trajectory of an auxiliary linear time-varying system which has a known analytic solution. Let us see the plot of $\tilde{x}_2$ vs. $\tilde{x}_1$, we have

$$\frac{d\tilde{x}_2}{d\tilde{x}_1} = \frac{-Ac_2(\cos c_1 t)^{c_2/c_1} (\sin c_1 t)^2}{-Fc_3(\sin c_1 t)^{c_3/c_1} (\cos c_1 t)^2}$$

In order to match the slopes of two plots at $c_1 t = 0$, or $x_2 = 0$, i.e., to let $\frac{d\tilde{x}_2}{d\tilde{x}_1} = \frac{dx_2}{dx_1} = \infty$ at origin, we must have

$$\frac{c_3}{c_1} < 2 \tag{2-36}$$

To match the slopes at $c_1 t = \frac{\pi}{2}$ or $x_1 = 0$, i.e., to let $\frac{d\tilde{x}_2}{d\tilde{x}_1} = \frac{dx_2}{dx_1} = 0$ at quarter period, we must have

$$\frac{c_2}{c_1} < 2 \tag{2-37}$$

Also we have

$$\frac{d^2\tilde{x}_2}{d\tilde{x}_1{}^2} = \frac{\frac{d\tilde{x}_1}{dt} \cdot \frac{d^2\tilde{x}_2}{dt^2} - \frac{d^2\tilde{x}_1}{dt^2} \cdot \frac{d\tilde{x}_2}{dt}}{\left(\frac{d\tilde{x}_1}{dt}\right)^3} \geq 0 \qquad \forall\ 0 \leq c_1 t \leq \frac{\pi}{2}$$

if $c_3 \leq 2c_1$ and $c_2 \leq 2c_1$. Because of the construction of the assumed function for $\tilde{x}_1(t)$ and $\tilde{x}_2(t)$, we have the above restrictions on c_1, c_2, c_3 .

Now we are to find c_1, c_2, c_3 by minimizing ϵ_1 and ϵ_2 in some sense, for instance we want to minimize the following quantity PI

$$\min PI = \min_{c_1, c_2, c_3} \{\int_0^t [\epsilon_1^2 + \epsilon_2^2]^{\frac{1}{2}} dt\}$$

or

$$\min PI = \min_{c_1, c_2, c_3} \{\int_0^t [|\epsilon_1(t)| + |\epsilon_2(t)|] dt\}$$

However, we can't perform the minimization procedures, the equations are really too messy. We therefore re-examine the trajectory in the phase plane; by a very crude pencil and paper work we can get a qualitative description of the solution. We find it looks like a flattened cosine function and its period can be estimated by the methods in References [6,9]. We then use the old techniques, let $\epsilon_1(t)$ and $\epsilon_2(t)$ be zero at certain instances. We also know c_2 is smaller than c_1 to ensure $\tilde{x}_1(t)$ is a flattened cosine function. We have 3 parameters to be determined, we need 3 equations. If we let $\epsilon_1(t) = 0$ at $t = t_1, t_2, t_3$, then we shall have very accurate quarter period of the solution [6], but the amplitude form is still different from the true solution as it should. It is rather easy to solve for c_1, c_2, c_3 by letting $\epsilon_1(t_1) = \epsilon_1(t_2) = \epsilon_1(t_3) = 0$, but we don't get any useful results. So we let

$$\epsilon_1(t_1; c_1, c_2, c_3) = 0$$
$$\epsilon_1(t_2; c_1, c_2, c_3) = 0$$
$$\epsilon_2(t_3; c_1, c_2, c_3) = 0$$

and solve for c_1, c_2, c_3 . We try to find some suitable value for t_1, t_2, and t_3 in the phase plane. The equation becomes

$$
\begin{aligned}
c_2 \cdot \tan(c_1t_1) \cdot A \cdot [\cos(c_1t_1)]^{c_2/c_1} &= F(\sin c_1t_1)^{c_3/c_1} \\
c_2 \cdot \tan(c_1t_2) \cdot A \cdot [\cos(c_1t_2)]^{c_2/c_1} &= F(\sin c_1t_2)^{c_3/c_1} \qquad (2\text{-}38) \\
-c_3 \cdot \tan(c_1t_3) \cdot F \cdot [\sin(c_1t_3)]^{c_3/c_1} &= A^3 \cdot [\cos(c_1t_3)]^{3c_2/c_1} \\
&\quad -A[\cos(c_1t_3)]^{c_2/c_1}
\end{aligned}
$$

These are very nonlinear algrbraic equations, we can hardly get exact solutions, so we use a simple method to estimate the quarter period [getting an approximate value for c_1 , we solve the first two equations in (2-38), then adjust c_1, c_2, c_3 to satisfy almost the last equation in in (2-38). These c_1, c_2 and c_3 have also to satisfy the construction restriction (2-36), (2-37) and (2-27), i.e.,

$$
\begin{aligned}
c_3 &\le 2c_1 \\
c_2 &\le 2c_1 \qquad (2\text{-}39) \\
0 \le c_2c_3 &\le 1
\end{aligned}
$$

We find in the phase plane the suitable values for c_1t_1, c_1t_2 and c_2t_3 will be

$$
\begin{aligned}
c_1t_1 &= \theta_1 = 15° \\
c_1t_2 &= \theta_2 = 80° \\
c_2t_3 &= \theta_3 = 45°
\end{aligned}
$$

From the above reasoning and choice, we finally find the values for c_1, c_2 and c_3

$$c_1 = 0.6$$

$$c_2 = 0.34$$

$$c_3 = 0.934$$

and the approximate solution is

$$\tilde{x}_1 = 0.9\ [\cos(c_1 t)]^{c_2/c_1}$$

$$\tilde{x}_2 = -0.69432\ [\sin(c_1 t)]^{c_3/c_1}$$

The above is just one of the many possible approximate solutions. We then compare the approximate solution with the true solution in the phase plane, we can have some rough idea of how c_1, c_2 and c_3 effect the approximate solution.

Time and labor is saved by using the above hybrid method to evaluate c_1, c_2 and c_3. To solve (2-38) then is not necessary.

We compare our approximate solution with the solution obtained by perturbation method and the true solution. The time-varying model method solution is much better than that of perturbation method as shown in Fig. 2.

III. A SIMPLE ERROR ANALYSIS

3-1. Errors in the Phase Plane for 1st Order Systems

Here we consider two kinds of errors, we define,

Solution error: $\delta(t) = x(t) - \tilde{x}(t)$

Equation error: $\epsilon(t) = \dot{\tilde{x}}(t) - f(\tilde{x}, t)$

where $\hat{x}(t)$ is an approximate solution, $x(t)$ is the true solution to the differential equation $\dot{x} = f(x,t)$. In general we cannot find $\delta(t)$. If we could, then we could get the true solution by adding $\delta(t)$ to $\tilde{x}(t)$. What we would like to do is to find some relation between $\delta(t)$ and $\epsilon(t)$, and hopefully we can deduce some information about $\delta(t)$ from our knowledge of $\varepsilon(t)$ [4]. Our ultimate goal is to find out how good our approximation is. We shall consider first order systems in this section. We have

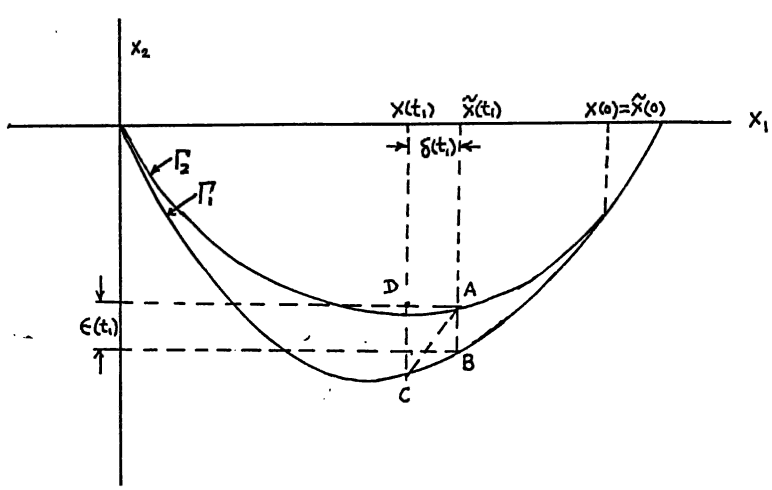
$$\begin{aligned}\varepsilon(t) &= \dot{\tilde{x}}(t) - f(\tilde{x},t) \\ &= \dot{x}(t) + \dot{\delta}(t) - f(x(t) + \delta(t),t) \\ &= \dot{x}(t) + \dot{\delta}(t) - f(x(t),t) - \frac{\partial f}{\partial x}\,\delta(t) + O(\delta(t))^2\end{aligned}$$

If we consider $\delta(t)$ is small, and f is fairly smooth, then we neglect all higher order terms except the linear term in the Taylor expansion of $f(x,t)$ about the true solution x. $\varepsilon(t)$ and $\delta(t)$ is related by

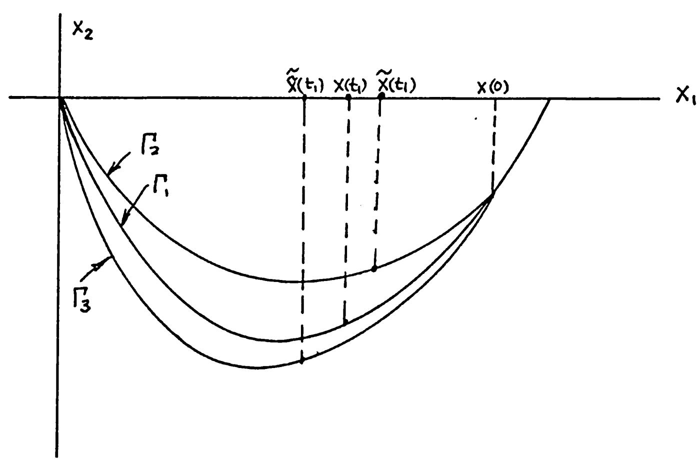
$$\varepsilon(t) = \dot{\delta}(t) - \frac{\partial f}{\partial X}\,\delta(t)$$

where $\frac{\partial f}{\partial X}$ is evaluated at $\tilde{x}(t)$; this makes us assume that $\tilde{x}$ is fairly close to x , i.e., $\delta(t)$ should be very small. But when we don't know in advance what $\delta(t)$ will be, we restore to an error analysis in phase plane.

We have in the phase plane two trajectories, one is the exact solution Γ_1 , the other is an approximate solution Γ_2 .

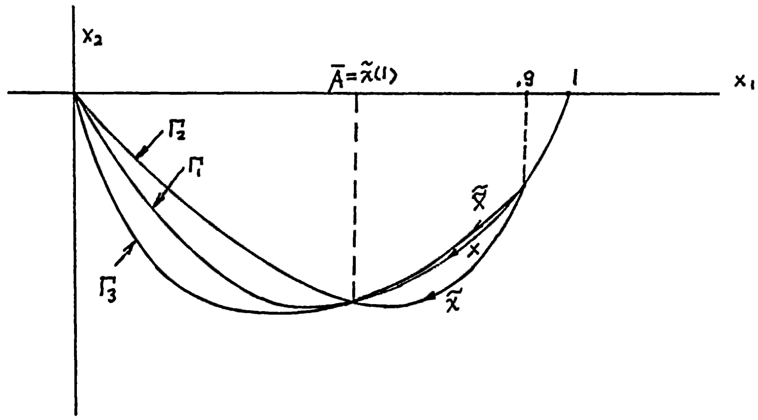


We know for first order system, $\dot{x} = f(x)$, $t = \int \frac{1}{\dot{x}}\, dx = \int_{x(0)}^{x} \frac{dx}{f(x)}$, so we have $\tilde{x}(t_1) > x(t_1)$ as shown in the figure. At a certain time t_1 , $\overline{AB}$ is our equation error, $\overline{AD}$ is our solution error, $\overline{AC}$ is a line joining two representative points on Γ_1 and Γ_2 . Unfortunately we can hardly find any relationship between $\overline{AB}$, $\overline{AC}$, and $\overline{AD}$, so in general $\epsilon(t)$ small is not necessarily to imply that $\delta(t)$ is small. We see that when $\overline{AC}$ is small, $\delta(t)$ is small, but there is no way to locate point C in phase plane. We present here a method to find a bound for $\delta(t)$. Let Γ_2 , Γ_3 be two approximate solution trajectories as shown in the figure



We know $\tilde{x}(t_1) > x(t_1) > \tilde{\tilde{x}}(t_1)$ for all t_1 in the interval of our interest, so the true solution is bounded in the range from $\tilde{x}(t)$ to $\tilde{\tilde{x}}(t)$, and $\delta(t) < |\tilde{x}(t) - \tilde{\tilde{x}}(t)|$. When we have these two approximate solutions, it is not difficult to evaluate $\tilde{x}(t) - \tilde{\tilde{x}}(t)$, so the bound for $\delta(t)$ is easily obtained, now $\delta(t)$ is either $(\tilde{x}(t) - x(t))$ or $(x(t) - \tilde{\tilde{x}}(t))$. We may take $x(t) \approx \bar{x}(t) = \frac{1}{2}\,(\tilde{x}(t) + \tilde{\tilde{x}}(t))$ to be a better approximation.

We have for our example in Chap. II-2, the following phase plane plot.



Γ_1 is the true solution, Γ_2 is an approximate solution,

$$\tilde{x}(t) = 0.9 \text{ EXP}[-t + 0.826 (\exp(-0.98 t) -1)]$$

Γ_3 is another approximate solution which we shall find later. Before the point A , we know from previous analysis that the representative point $\tilde{x}$ moves faster than x , and x faster than $\tilde{\tilde{x}}$, that is $\tilde{\tilde{x}}(t) > x(t) > \tilde{x}(t)$ for $t < t_o$, $\tilde{x}(t_o) = \bar{A}$. After crossing point A , $\tilde{\tilde{x}}$ moves faster than x , x faster than $\tilde{x}$, but we don't know whether $\tilde{\tilde{x}}(t)$ is still greater than x(t) . However, Γ_1 is still between Γ_2 and Γ_3. We can "fairly" assume either $\tilde{\tilde{x}}(t) \geq x(t) \geq \tilde{x}(t)$ or $\tilde{x}(t) \geq x(t) \geq \tilde{\tilde{x}}(t)$, i.e., x(t) is still between $\tilde{x}(t)$ and $\tilde{\tilde{x}}(t)$.

We are now to find $\tilde{\tilde{x}}(t)$, we assume

$$K'(t) = 1 - A^2(e^{-bt^2/c})$$

so

$$K'(t) > K(t) \quad \text{for} \quad t < T$$

$$K'(t) < K(t) \quad \text{for} \quad t \geq T$$

where K(t) is the same K(t) in (2-13) of Chap. II, b is the same as b in K(t), c is a parameter to be determined. By proper choice of c and the point A, we can get Γ_3 as just what we show in the phase plane plot. We let

$$\tilde{\tilde{x}}(t') = \tilde{x}(t=1) \tag{3-1}$$

$$\epsilon'(t') = 0 \ , \ \epsilon'(t) = \dot{\tilde{\tilde{x}}} + f(\tilde{\tilde{x}}) \tag{3-2}$$

from the two equations above we find the values for t' and c, then we get $\tilde{\tilde{x}}(t)$.

$$\tilde{\tilde{x}}(t) = A \cdot \text{EXP}\ [-\int K'(t)dt] \ , \quad A = 0.9 \tag{3-3}$$

$$\int_o^t (1 - A^2(e^{-bt^2/c}))dt = t - A^2\sqrt{\frac{c}{b}}\frac{\sqrt{\pi}}{2}\,\text{erf}(\sqrt{\frac{b}{c}}\,t) \tag{3-4}$$

substituting (3-3), (3-4) into (3-1) and (3-2), after rearranging we have

$$2(t' - A^2\sqrt{\frac{c}{b}}\frac{\sqrt{\pi}}{2}\,\text{erf}(\sqrt{\frac{b}{c}}\,t')) = \frac{b(t')^2}{c}$$

$$t' - A^2\sqrt{\frac{c}{b}}\frac{\sqrt{\pi}}{2}\,\text{erf}(\sqrt{\frac{b}{c}}\,t') = \log\frac{A}{\tilde{x}(t=1)}$$

Solving the two equations above we get

$$t' = 1.22844$$

$$c = 1.52775$$

so

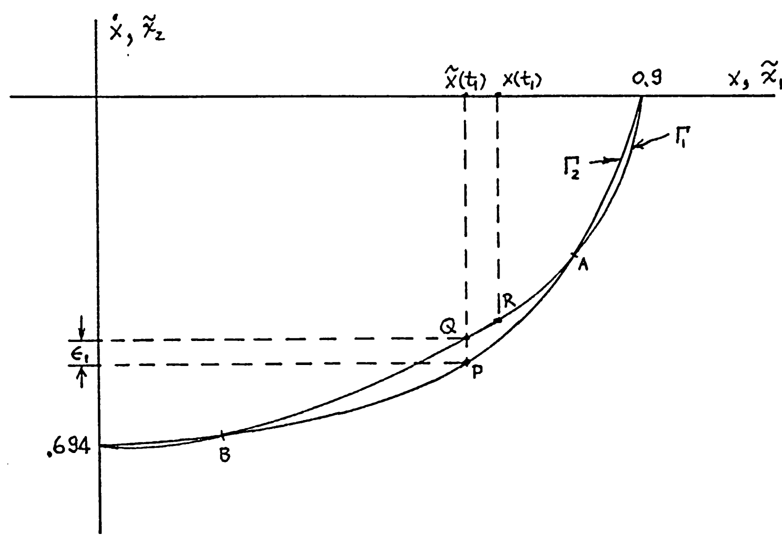
$$\tilde{\tilde{x}}(t) = 0.9\ \text{EXP}\ [-t + 0.88521\ \text{erf}(0.64147\ t)]$$

We can be fairly sure that the true solution will lie between $\tilde{x}(t)$ and $\tilde{\tilde{x}}(t)$.

3-2. Errors in the Phase Plane for 2nd Order Systems

A second order system is completely specified if we know the values for $x(t)$, $\dot{x}(t)$ and $\ddot{x}(t)$. We can plot a solution of a second order system in the $(x, \dot{x}, \ddot{x})$ state space. What we shall have is a curve in the three-dimensional space. We can hardly define the discrepancy between two curves in three-dimensional space, so we have no way to compare an approximation solution with the true solution in state space. Then we still restore our error analysis in phase plane. There are a number of works done on phase plane analysis,[8,7] but none of them seem to give an error analysis in phase plane.

We shall consider an error analysis for our specific example in Chap. II-3. We have in the phase plane the following plot



Γ_1 is the integral curve. Γ_2 is an approximation solution. They meet at points A and B . We have

$$\delta(t) = x_1(t) - \tilde{x}_1(t)$$

$$\epsilon_1(t) = \dot{\tilde{x}}_1 - \tilde{x}_2$$

$$\epsilon_2(t) = \dot{\tilde{x}}_2 + \tilde{x}_1 - \tilde{x}_1^3$$

At a certain instance, $t = t_1$, $x(t_1)$ is the true solution at point R , $\tilde{x}(t_1)$ is an approximate solution at point P as shown in the above figure, $\epsilon_1(t)$ is the projection of $\overline{PQ}$ on x_2-axis , ϵ_2 is the difference of slopes at point P and point Q . If $\epsilon_1 = \epsilon_2 = 0$, we have the true solution.

We can estimate the period of our example from the phase plane plot [6,9]. From [6] we know that if we match five points of the two trajectories, we get the exact period, but the amplitude form of the two solutions is still different. Now in the example we match four points.; hopefully we can get a close approximation. Because of the nature of the assumed function form for the approximate solution, we don't worry too much about $\epsilon_2(t)$, so we solve for c_1, c_2 and c_3 by letting $\epsilon_1(t)$ equal to zero at two instances, t_1 and t_2; $\epsilon_2(t)$ equals to zero at $t = t_3$.

We shall say we can use the same technique to find an $\tilde{\tilde{x}}$, such that the true solution will lie in between $\tilde{x}(t)$ and $\tilde{\tilde{x}}(t)$ as in the case for first order systems. As we do we find the complexity and difficulty in solving four simultaneous nonlinear algebraic equations, so we don't proceed. In general if the integral curve is bound between two approximation solutions, we have the true solution almost lying between the two approximate solutions.

IV. CONCLUSION AND REMARKS

The purpose of this study is to find an approximate method for solving a class of nonlinear differential equations of the type

$$\dot{x} + f(x) = 0$$

and

$$\ddot{x} + g(x) = 0$$

where f , g are of polynomial type functions.

The method used was to approximate the system by another linear time-varying system. The linear systems are

$$\dot{x} + K(t)x = 0$$

and

$$\dot{x}_1 = a_{11}(t)x_1$$
$$\dot{x}_2 = a_{22}(t)x_2$$

We want to approximate the trajectory by another curve which is the exact solution to the auxiliary linear time-varying system. We then examine the system behavior in the phase plane, and assume certain function forms for $K(t)$ and $a_{11}(t)$ and $a_{22}(t)$, with some undetermined parameters in them. We substitute the approximate solution into the original system equations, and find the undetermined parameters by minimizing equation errors in some sense. Here we demand equation errors to be zero at certain instances, i.e.,

$$\epsilon(t) = 0 \qquad \text{for} \quad t = t_1$$

and

$$\epsilon_1(t) = 0 \qquad \text{for} \quad t = t_1 \quad \text{and} \quad t = t_2$$

$$\epsilon_2(t) = 0 \qquad \qquad t = t_3$$

The results obtained are satisfactory as shown in the two examples.

As we stated in the introduction, we can use this method to solve a larger class of nonlinear differential equations than the two types we wrote down here. The difficulty of extension lies in the complexities of integral curves in phase plane. We cannot represent the $K(t)$, $a_{11}(t)$ and $a_{22}(t)$ by suitable functions. Maybe there is some modification that will allow some 2nd order driven systems to be solved by this method.

We also gave a simple error analysis in the phase plane and find an error bound by finding another approximate solution based on our first approximate solution.

Finally, the goal of finding an approximate solution better than that of using ordinary perturbation or iteration method has been achieved, and valuable insight into the system behavior has been obtained, and it is a promising idea to make some modification to allow this method to be suitably applicable to a larger class of driven nonlinear systems.

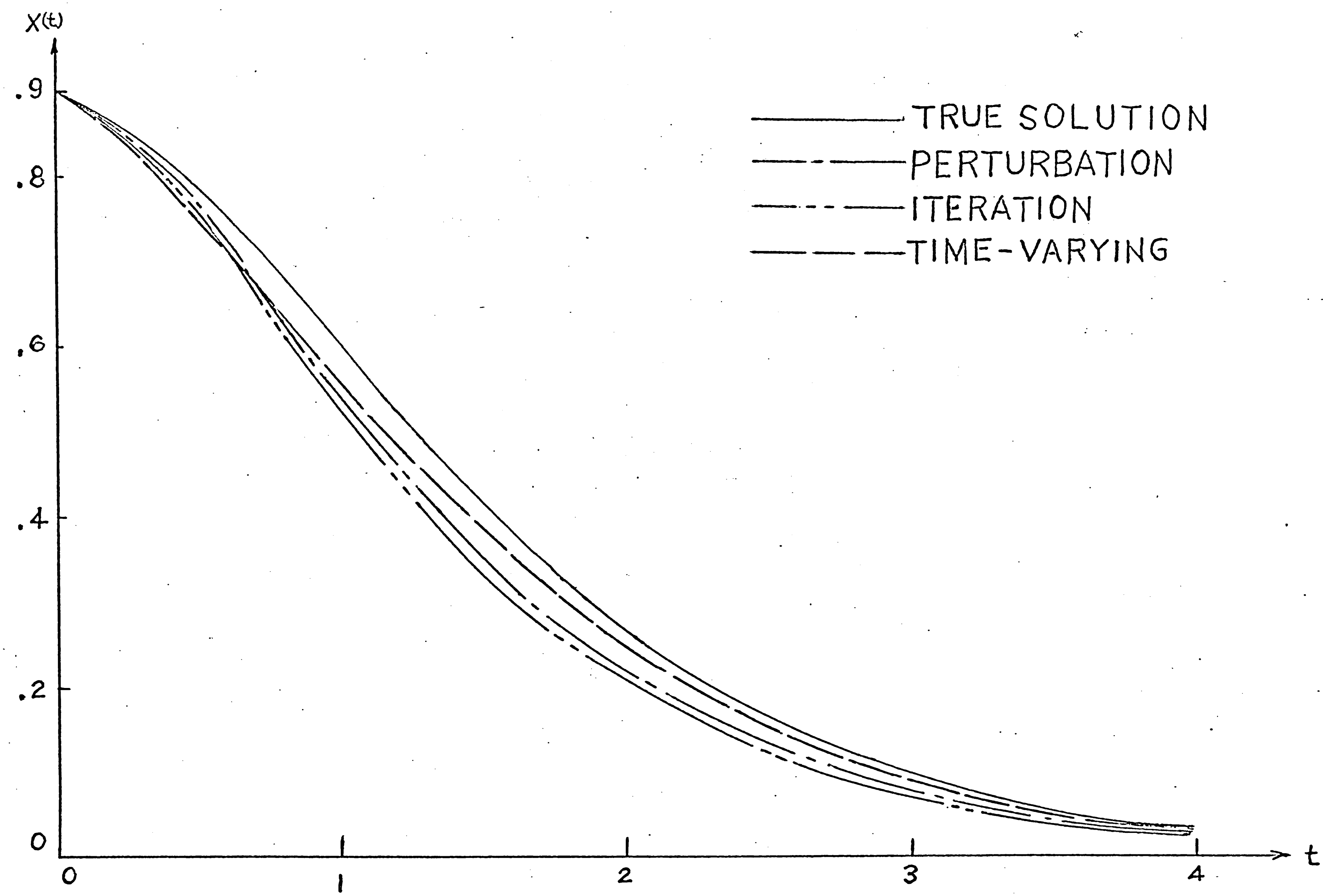

X(t)
.9
.8
.6
.4
.2
0
0
1
2
3
4
t
TRUE SOLUTION
PERTURBATION
ITERATION
TIME-VARYING

FIG. 1

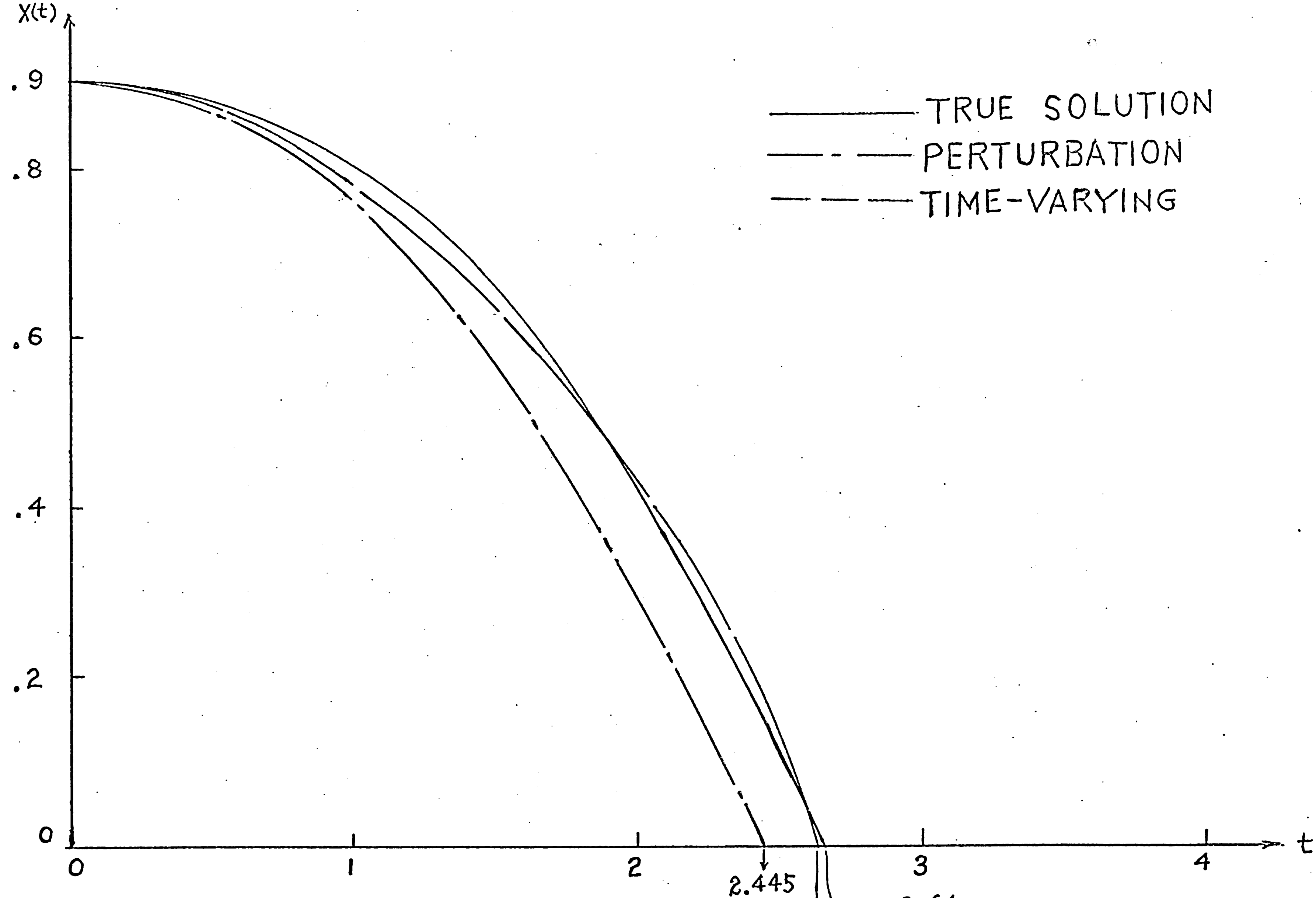

X(t)
.9
.8
.6
.4
.2
0
TRUE SOLUTION
PERTURBATION
TIME-VARYING
0
1
2
3
4
t
2.445
2.64
2.62

FIG. 2

REFERENCES

1. Birkhoff, G., and G. Rota, Ordinary Differential Equations, Ginn and Co., New York, 1962.

2. Cunningham, W.J., Introduction to Nonlinear Analysis, McGraw-Hill, New York, 1958.

3. Saaty, T.L., and J. Bram, Nonlinear Mathematics, McGraw-Hill, New York, 1964.

4. Burrus, C.S., "An Iterative Procedure for the Analysis of Nolinear Networks," Technical Report No. 6657-1, Stanford University, Stanford, 1965.

5. Bowman, F., Introduction to Elliptic Functions with Applications, Dover, New York, 1961.

6. Soudack, A.C., "Jacobean Elliptic and Other Functions as Approximate Solutions to a Class of Grossly Nonlinear Differential Equations," Technical Report No. 2054-1, Stanford University, Stanford, California, 1961.

7. Ku, Y.H., Analysis and Control of Nonlinear Systems, The Ronald Press Co., 1958.

8. Andronov, A.A., A.A. Vitt, and S.E. Khaikin, Theory of Oscillators, Pergamon Press Ltd., 1966.

9. Chen, C.F. and I.J. Haas, "Simple method for estimating time from phase trajectories," Automatic Control, Volume AC-13, No. 4, pp. 449-450, August 1968.