

**The Numerical Effect of Discretisation of Continuous
Systems and Bilinear Mappings of Non-linear Systems**

K. Shin

ISVR Technical Report No 275

February 1998



SCIENTIFIC PUBLICATIONS BY THE ISVR

Technical Reports are published to promote timely dissemination of research results by ISVR personnel. This medium permits more detailed presentation than is usually acceptable for scientific journals. Responsibility for both the content and any opinions expressed rests entirely with the author(s).

Technical Memoranda are produced to enable the early or preliminary release of information by ISVR personnel where such release is deemed to be appropriate. Information contained in these memoranda may be incomplete, or form part of a continuing programme; this should be borne in mind when using or quoting from these documents.

Contract Reports are produced to record the results of scientific work carried out for sponsors, under contract. The ISVR treats these reports as confidential to sponsors and does not make them available for general circulation. Individual sponsors may, however, authorize subsequent release of the material.

COPYRIGHT NOTICE

(c) ISVR University of Southampton All rights reserved.

ISVR authorises you to view and download the Materials at this Web site ("Site") only for your personal, non-commercial use. This authorization is not a transfer of title in the Materials and copies of the Materials and is subject to the following restrictions: 1) you must retain, on all copies of the Materials downloaded, all copyright and other proprietary notices contained in the Materials; 2) you may not modify the Materials in any way or reproduce or publicly display, perform, or distribute or otherwise use them for any public or commercial purpose; and 3) you must not transfer the Materials to any other person unless you give them notice of, and they agree to accept, the obligations arising under these terms and conditions of use. You agree to abide by all additional restrictions displayed on the Site as it may be updated from time to time. This Site, including all Materials, is protected by worldwide copyright laws and treaty provisions. You agree to comply with all copyright laws worldwide in your use of this Site and to prevent any unauthorised copying of the Materials.

UNIVERSITY OF SOUTHAMPTON
INSTITUTE OF SOUND OF VIBRATION RESEARCH
SIGNAL PROCESSING & CONTROL GROUP

**The Numerical Effect of Discretisation of Continuous Systems and Bilinear
Mapping of Non-linear Systems**

by

K Shin

ISVR Technical Report No. 275

February 1998

Authorized for issue by
Professor S J Elliott
Group Chairman

ABSTRACT

When a continuous system is mapped to an equivalent discrete system, one may use various numerical methods such as the Euler method, the Runge-Kutta method, the bilinear method, etc. In signal processing, the bilinear mapping method is widely used for linear systems, because it provides a very simple conversion from a continuous system to a discrete system without introducing aliasing problems. Thus, the bilinear mapping method is very important in designing digital filters from analogue filters for linear systems. In this report, the use of the bilinear mapping method for non-linear systems is investigated by applying it to simple non-linear ordinary differential systems. A brief review of the 'dynamics of numerics' is given and the bilinear mapping method is compared to other numerical methods for both linear and non-linear systems. The results show that the bilinear mapping method is very stable (even when the step size of integration is very large) and indicates promise for non-linear digital filter design.

CONTENTS

1. Introduction
2. Some Results of the ‘Dynamics of Numerics’
3. Linear Systems
 - 3.1 Homogeneous System
 - 3.2 The Effect of Forcing
4. Non-linear Systems
5. Discussion

References

Appendix. Bilinear Mapping from s-plane to z-plane

1. INTRODUCTION

When a continuous system is mapped to an equivalent discrete system, one may use various numerical methods such as the Euler method, the Runge-Kutta method, the bilinear method, etc. Recently, 'spurious' solutions introduced by the time discretisation of ordinary differential equations have been investigated. In order to describe an equivalent difference equation for a non-linear differential equation one can apply a variety of methods, and solutions from these methods may be completely different from that of the original differential system depending on the numerical approximation (e.g. step size). One question arising from this is just how small a step size is 'reliable'. One may use sophisticated and relatively reliable methods which use variable step sizes. However, in general, these methods are time consuming and in some cases (when the uniformly sampled values are needed) may not be easily incorporated with an algorithm which requires the numerical integration as a subroutine. For example, calculation of Lyapunov exponents (from a differential equation) requires the solution of multiple sets of differential equations at the same time and requires that the values of each set of differential equations are sampled at the same time. The numerical methods using variable step sizes will produce values at different times for each set of differential equations, and this makes it difficult to use such a method. Thus, in this report, we only consider numerical methods with a uniform step size. In this case, different step sizes are applied to different numerical methods according to the inherent accuracy of the applied numerical method. Iserles *et al* [1 - 5] have studied the problem, and some of their results are briefly discussed in the following section. Here, however, the problem is considered from a slightly different point of view, i.e., a view of the signal processing context. In signal processing, the bilinear mapping method is widely used for mapping continuous linear systems to discrete systems, because it is a simple conversion which avoids aliasing problems. Thus, the bilinear mapping method is very important in designing digital filters from analogue filters for linear systems. In essence it is the trapezoidal integration rule. The effects of bilinear mapping method for non-linear systems has not been reported.. Thus, in this report, the bilinear mapping method is extended to non-linear systems.

We consider the bilinear method for a simple linear system and compare it to the Euler method and the Runge-Kutta method in section 3. The Euler method and the Runge-Kutta method give erroneous results even for linear systems when the step size is large. In section 4, the bilinear method is applied to a simple known non-linear system and compared to the Runge-Kutta method. The numerical results show the fundamental property of the bilinear method for linear systems carries over to non-linear systems. In section 5, the use of the bilinear mapping method is discussed in terms of a signal processing context and related to the solutions of non-linear differential equations.

2. SOME RESULTS OF THE ‘DYNAMICS OF NUMERICS’

This section is a very brief review of the results obtained by Iserles *et al*, and shows the main problems of the Euler method and the Runge-Kutta method. The details of the discussion are given in references [1 - 5]. Consider a dynamical system described by an ordinary differential system (ODS) given by

$$\frac{dy}{dt} = f(y), \quad y(t_0) = y_0, \quad y(t) \in \mathfrak{R}^d \quad (2.1)$$

where \mathfrak{R}^d represents d-dimensional space and y_0 is the initial condition. When the solution of the system (2.1) is unique and exists for all $t \in \mathfrak{R}$, solutions define a ‘flow’ in the phase space. Consider a discrete map which approximates (2.1) by

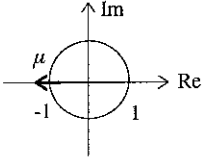
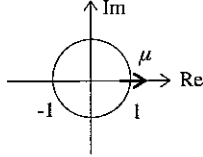
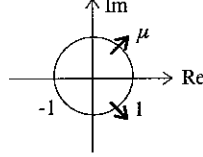
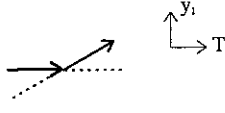
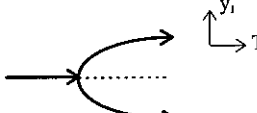
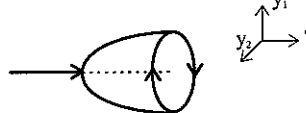
$$y_{n+1} = F_T(y_n), \quad y_0 \in \mathfrak{R}^d \quad (2.2)$$

where the ‘T’ is the fixed step size. In order to understand the behaviour of equation (2.2) consider the simplest case, namely that the solution of the continuous system (2.1) which tends to a fixed point. If y^* is an attractive fixed point of the flow, i.e., $y(t) \rightarrow y^*$ for all y_0 in an open neighbourhood of y^* , and both f and F_T are smoothly differentiable, then, for sufficiently small ‘T’, all the eigenvalues of the linearised system evaluated at y^* ($\frac{\partial F_T(y^*)}{\partial y}$) are within the complex unit disk and $y^* = F_T(y^*)$. In other words, for sufficiently small ‘T’, the numerical solutions approach the true fixed point. Since the eigenvalues of $\frac{\partial F_T(y^*)}{\partial y}$ are continuous in ‘T’, when increasing ‘T’ the following results hold [1].

- (i) An eigenvalue may hit the unit disk at ‘1’ and this corresponds to a steady bifurcation (or transcritical bifurcation: from a stable equilibrium point to a new branch of stable equilibrium point and the original branch of equilibrium point becomes unstable).
- (ii) An eigenvalue may hit the unit disk at ‘-1’ and this corresponds to a flip bifurcation (or period-doubling bifurcation).

- (iii) A pair of complex eigenvalues may cross the unit disk and this corresponds to a Hopf bifurcation.

The graphical illustration of the above results are

Case (i)	Case (ii)	Case (iii)
<p>Eigenvalue (μ)</p> 	<p>Eigenvalue (μ)</p> 	<p>Eigenvalue (μ)</p> 
<p>Steady Bifurcation</p> 	<p>Flip Bifurcation</p> 	<p>Hopf Bifurcation</p> 

Note that these results are for the asymptotic behaviour (steady state) of the system, so we may see different transient behaviour of the system even when ‘T’ is small.

Consider the logistic equation as an example.

$$\frac{dy}{dt} = y(1 - y), \quad y(0) = y_0 \in (0,1) \quad (2.3)$$

The analytic solution is obtained easily and $y(t) \rightarrow 1$ (a stable fixed point) as $t \rightarrow \infty$, and there is no chaotic solution. Using the Euler method, i.e., $\frac{dy}{dt} = \frac{y(t+T) - y(t)}{T}$, the above equation is mapped by

$$y_{n+1} = y_n + Ty_n(1 - y_n) \quad (2.4)$$

Where $y_n \approx y(t_0 + nT)$. Since the eigenvalue of $\frac{\partial F_T(1)}{\partial y}$ is ‘ $1 - T$ ’, as long as ‘ $T < 2$ ’, the fixed point of (2.3) at $y^* = 1$ remains the same, but when ‘ $T = 2$ ’ the eigenvalue hits the unit disk at ‘ -1 ’ and this yields period doubling, and for larger step sizes this results in the Feigenbaum cascade. Thus, period doubling in this case is a numerical artefact.

Next, equation (2.3) may be solved by the Runge-Kutta method (second order) which results in

$$y_{n+1} = y_n + T \left[y_n + \frac{T}{2} y_n (1 - y_n) \right] \left[1 - y_n - \frac{T}{2} y_n (1 - y_n) \right] \quad (2.5)$$

The eigenvalue of $\frac{\partial F_T(1)}{\partial y}$ is $'1 - T + \frac{1}{2}T^2'$, and when at $'T = 2'$, it crosses the unit disk at $'1'$ and this yields a steady bifurcation and for increased step sizes this also follows the period doubling route to chaos.

Iserles [1] suggested a “Good Method Guide” to obtain the solutions exactly in step with the original flow (2.1) for the steady state of the system, and suggested the following definitions.

- (i) A numerical method is $R^{[1]}$ if when applied to an ODS, $y_n \rightarrow y^*$. This means that y^* is a genuine fixed point of the original flow. So the Runge-Kutta method is not $R^{[1]}$ because the Runge-Kutta method can yield a steady bifurcation which is a numerical artifact.
- (ii) A numerical method is $R^{[2]}$ if it never produces a flip bifurcation.
- (iii) A numerical method is $R^{[H]}$ if it has a Hopf bifurcation when (and only when) the original ODS also does. Thus the Euler method is neither $R^{[2]}$ nor $R^{[H]}$.

For the “Good Method Guide”, a multistep method is suggested in the form

$$\sum_{i=0}^k \alpha_i y_{n+i} = T \sum_{i=0}^k \beta_i f(y_{n+i}) \quad (2.6)$$

where α_i, β_i are coefficients. When β_k is not zero, the method is *implicit*. One can intuitively anticipate that the multistep methods are likely to model dynamics more faithfully since increasing the order in equation (2.6) may result in increasing the accuracy, i.e., an m -step method uses the ‘ m ’ previous points. Typically an implicit method has a larger region of stability than the corresponding explicit method, and this is mainly due to the feedback nature in an implicit method [12]. Iserles suggested the following $R^{[1,2,H]}$ Runge-Kutta method which is the implicit midpoint rule. The value

\mathbf{y}_{n+1} is obtained by finding the midpoint between \mathbf{y}_{n+1} and the previous value \mathbf{y}_n , and implicitly depends on \mathbf{y}_{n+1} itself as defined in equation (2.7).

$$\mathbf{y}_{n+1} = \mathbf{y}_n + \mathbf{Tf}\left(\frac{1}{2}(\mathbf{y}_n + \mathbf{y}_{n+1})\right) \quad (2.7)$$

Another example of a multistep method which is $R^{[1,2,H]}$ is the trapezoidal integration rule which yields

$$\mathbf{y}_{n+1} = \mathbf{y}_n + \frac{T}{2}[\mathbf{f}(\mathbf{y}_n) + \mathbf{f}(\mathbf{y}_{n+1})] \quad (2.8)$$

or

$$\mathbf{y}_{n+1} - \frac{T}{2}\mathbf{f}(\mathbf{y}_{n+1}) = \mathbf{y}_n + \frac{T}{2}\mathbf{f}(\mathbf{y}_n) \quad (2.9)$$

This is recognised as the bilinear mapping method which is widely used for the digital filter design for linear systems. We see that to compute \mathbf{y}_{n+1} from \mathbf{y}_n involves the functional evaluation $\mathbf{f}(\mathbf{y}_{n+1})$ and in general the solution of a non-linear algebraic equation. So one may use Newton's method to solve equation (2.9). In a special case (linear systems), equation (2.9) can be solved as follows. Equation (2.1) can be described in a matrix form if the function $\mathbf{f}(\mathbf{y})$ is linear, so

$$\dot{\mathbf{y}} = \mathbf{A}\mathbf{y} \quad (2.10)$$

where \mathbf{A} is an $n \times n$ matrix. So equation (2.9) can be written as

$$\mathbf{y}_{n+1} - \frac{T}{2}\mathbf{A}\mathbf{y}_{n+1} = \mathbf{y}_n + \frac{T}{2}\mathbf{A}\mathbf{y}_n \quad (2.11)$$

and

$$\left[\mathbf{I} - \frac{T}{2}\mathbf{A}\right]\mathbf{y}_{n+1} = \left[\mathbf{I} + \frac{T}{2}\mathbf{A}\right]\mathbf{y}_n \quad (2.12)$$

where \mathbf{I} is the identity matrix. By rearranging (2.12) it becomes

$$\mathbf{y}_{n+1} = \left[\mathbf{I} - \frac{T}{2}\mathbf{A}\right]^{-1} \left[\mathbf{I} + \frac{T}{2}\mathbf{A}\right]\mathbf{y}_n \quad (2.13a)$$

Note that this equation is also valid for non-linear systems by allowing the matrix \mathbf{A} to vary with time. In this case, equation (2.13a) becomes

$$\mathbf{y}_{n+1} = \left[\mathbf{I} - \frac{T}{2} \mathbf{A}_{n+1} \right]^{-1} \left[\mathbf{I} + \frac{T}{2} \mathbf{A}_n \right] \mathbf{y}_n \quad (2.13b)$$

where \mathbf{A}_n is the n-th iterated time varying matrix. Equation (2.13b), in general, may not be directly solved since \mathbf{A}_{n+1} usually depends on \mathbf{y}_{n+1} . Thus, this requires numerical methods for solving (2.13b) such as Newton's method.

Details of the bilinear method will be given by examples in the next section. Also, in this report, we shall investigate the use of this bilinear mapping method for non-linear systems. We begin by briefly reviewing linear systems.

3. LINEAR SYSTEMS

3.1 Homogeneous System

In this section, the advantages of the bilinear method for mapping a continuous-time system to a discrete-time system are demonstrated on a simple second order linear system, and compared to the Euler method and the 4-th order Runge-Kutta method. We begin with an unforced (homogeneous) system. Consider an ordinary differential system.

$$\ddot{x} + \omega_0^2 x = 0. \quad (3.1)$$

The solution is

$$x = A \cos(\omega_0 t + \phi). \quad (3.2)$$

Equation (3.1) can be described in the form (2.1) as

$$\begin{aligned} \dot{x} &= y \\ \dot{y} &= -\omega_0^2 x \end{aligned} \quad (3.3)$$

Equation (3.3) will be mapped using three different methods - the Euler method, the 4-th order Runge-Kutta method and the bilinear method. Both the Euler method and Runge-Kutta method are explicit in the expression (2.6) and so are easy to solve. The bilinear method is implicit, i.e., a value y_{n+1} is a solution of (2.9). Generally, for a multi-dimensional set of equations one may use Newton's method to find a root at every iteration and in general is very complex. However the bilinear mapping of equation (3.3) in particular is easily converted to a simple set of difference equations because the system is linear. So the solution can be obtained from equation (2.13a), i.e., equation (3.3) can be written as

$$\begin{bmatrix} x_{n+1} \\ y_{n+1} \end{bmatrix} = \frac{1}{1 + \left(\frac{T}{2} \omega_0\right)^2} \begin{bmatrix} 1 - \left(\frac{T}{2} \omega_0\right)^2 & T \\ 1 & 1 - \left(\frac{T}{2} \omega_0\right)^2 \end{bmatrix} \begin{bmatrix} x_n \\ y_n \end{bmatrix} \quad (3.4)$$

Alternatively, bilinear mapping of equation (3.3) can be obtained from the following [6]. Assume that we know y_n in equation (2.9). So, the right-hand side of (2.9) can be easily obtained. Thus, equation (3.3) can be written as

$$x_{n+1} - \frac{T}{2}y_{n+1} = x_n + \frac{T}{2}y_n = c_1$$

$$\therefore x_{n+1} = c_1 + \frac{T}{2}y_{n+1} \quad (3.5a)$$

$$y_{n+1} + \frac{T}{2}\omega_0^2 x_{n+1} = y_n - \frac{T}{2}\omega_0^2 x_n = c_2$$

$$\therefore y_{n+1} = c_2 - \frac{T}{2}\omega_0^2 x_{n+1} \quad (3.5b)$$

By rearranging (3.5a) and (3.5b), we get the following equation

$$\therefore y_{n+1} = \frac{c_2 - \frac{T}{2}\omega_0^2 c_1}{1 + \left(\frac{T}{2}\omega_0\right)^2} \quad (3.6)$$

One can compute equation (3.6) and substitute the result in (3.5a) and then repeat this procedure to get the trajectory. From equation (3.2) the exact solution for this system is known. We know also their digital and analogue frequencies (ω_d and ω_a) are related by

$$\omega_d = \frac{2}{T} \tan^{-1} \left(\frac{\omega_a T}{2} \right) \quad (3.7)$$

where ω_d is the digital angular frequency by the bilinear method and

ω_a is the analogue angular frequency (original differential system).

By setting 'T = 0.1', 'initial conditions $x_0 = 1, y_0 = 0$ ', and varying the value ω_0 , where $\omega_0 = 2\pi f_0$, the effect of discretisation is investigated. Details of error estimation (accuracy) and stability analysis of a numerical method for the first order differential equation ($\dot{y} = \lambda y$) can be found in references [11 - 13]. The results of the first order systems can be used for higher order systems in the form of equation (2.10), since equation (2.10) can be represented in the form of $\dot{\mathbf{z}} = \mathbf{A}\mathbf{z}$ (or $\dot{z}_i = \lambda_i z_i$) if the system is linear and the matrix 'A' is diagonalisable, where $\mathbf{\Lambda}$ is the diagonal eigenvalue matrix of the matrix \mathbf{A} in equation (3.10) [11]. In this report, we demonstrate the effect of step size by example.

When the Euler method is applied to this system, it turns out that the method is very unstable and results in amplitude modulated signal even when f_0 is very small, i.e., the step size is relatively small compared to the period of oscillation of the system. This amplitude modulated signal gives a set of frequency components and not a single spike as shown in Figure 1.

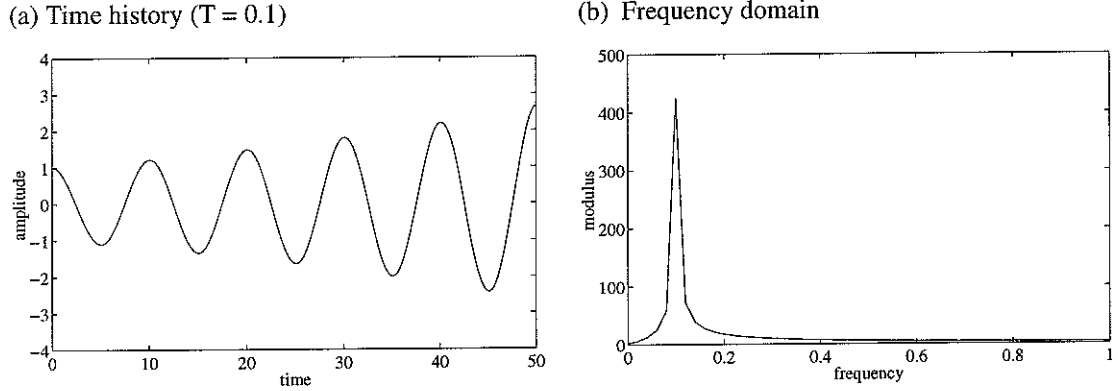


Figure 1. Solutions of the system (3.1) using the Euler method ($f_0 = 0.1$)

Next, Equation (3.3) is discretised by using the 4-th order Runge-Kutta method and the bilinear method. When f_0 is small ($f_0 = 0.5$) the Runge-Kutta method gives an almost exact solution as shown in Figure 2 while the bilinear method gives a slightly lower frequency as expected, i.e., from equation (3.7) the analogue frequency ($f_0 = 0.5$) results in digital frequency ($f_d = 0.496$), and this is shown in Figure 3. This lowering frequency is analogous to the period elongation described in [10].

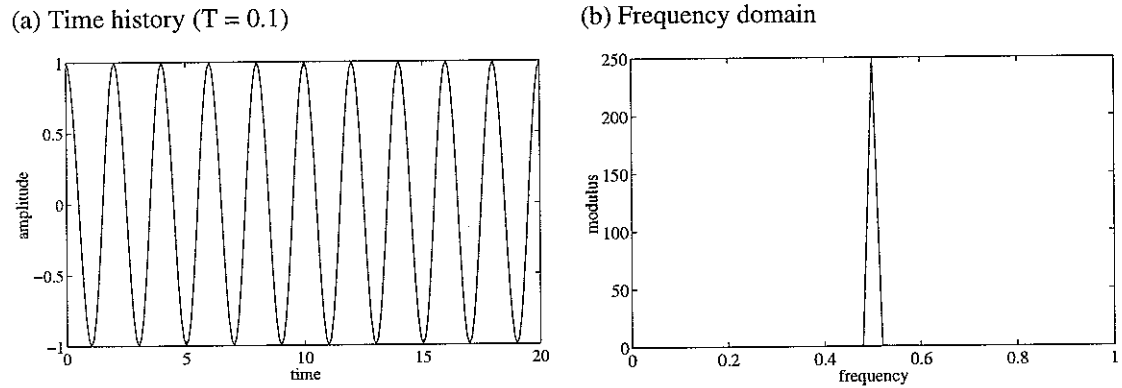


Figure 2. Solutions of the system (3.1) using the Runge-Kutta method ($f_0 = 0.5$)

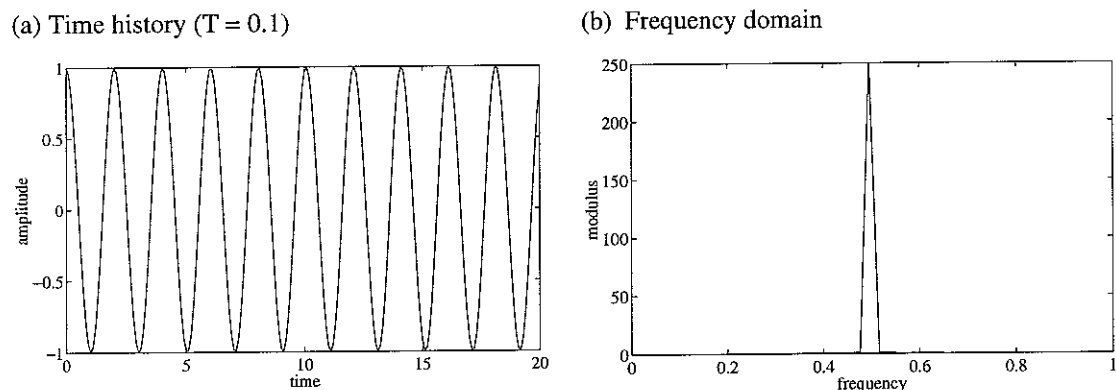


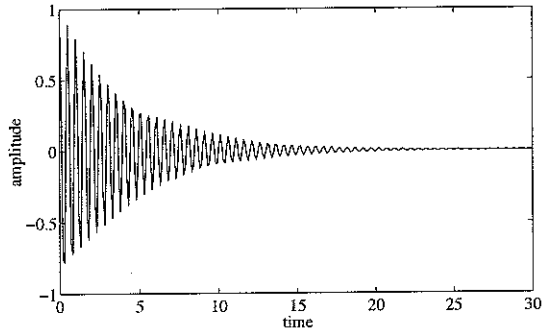
Figure 3. Solutions of the system (3.1) using the bilinear method ($f_0 = 0.5$, $f_d = 0.496$) where, f_d is the digital frequency mapped by bilinear method.

When f_0 is increased the Runge-Kutta method gives a very erroneous result which finally goes to a fixed point (note that the system being simulated has *no* damping). Since the original solution is oscillating and the numerical method is multi-step, it is very difficult to prove the reason of the erroneous result. On the other hand the bilinear method gives a solution which one can expect, i.e., from equation (3.7) the results are always anticipated. The results of the case ' $f_0 = 2$ ' are shown in Figure 4 (Runge-Kutta) and Figure 5 (bilinear).

If the parameter f_0 is increased further, when ' $f_0 > 4.5$ ', the Runge-Kutta method becomes unstable and also gives a spurious frequency component, and this is shown in Figure 6 in the case of ' $f_0 = 4.51$ '. In this case, we can imagine that the Runge-Kutta method with the above step size is completely outside the stability region (numerical errors are *unbounded* (Figure 6a)). Also, it gives erroneous frequency components (Figure 6b).

However, the bilinear method gives the solution as expected from equation (3.7) even when f_0 is beyond the folding frequency (since $T = 0.1$, the folding frequency is 5), and the case of ' $f_0 = 6$ ' is shown in Figure 7.

(a) Time history ($T = 0.1$)



(b) Frequency domain

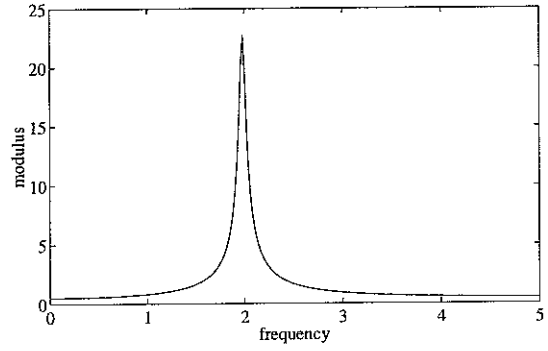
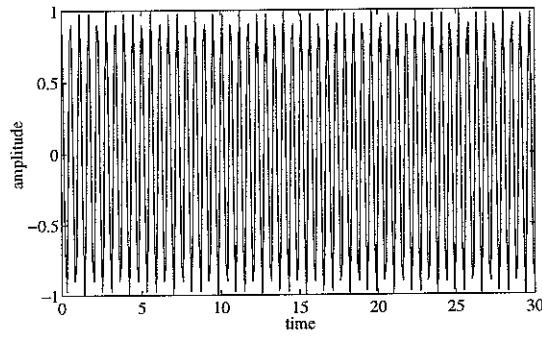


Figure 4. Solutions of the system (3.1) using the Runge-Kutta method ($f_0 = 2$)

(a) Time history ($T = 0.1$)



(b) Frequency domain

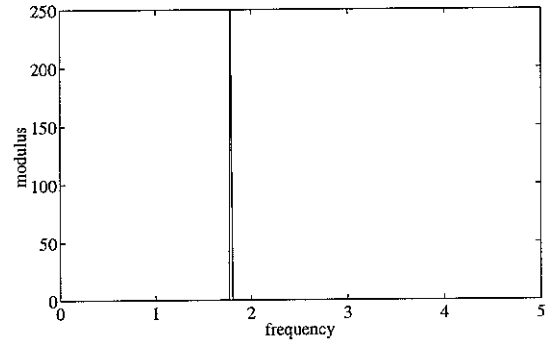
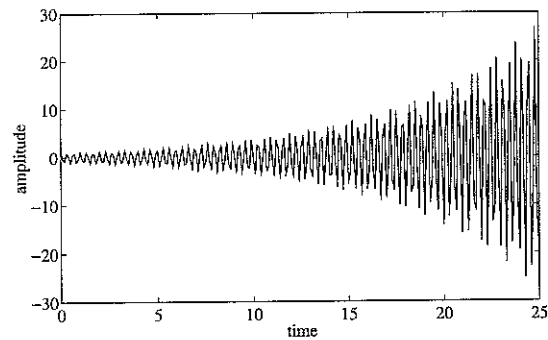


Figure 5. Solutions of the system (3.1) using the bilinear method ($f_0 = 2$, $f_d = 1.79$)

(a) Time history ($T = 0.1$)



(b) Frequency domain

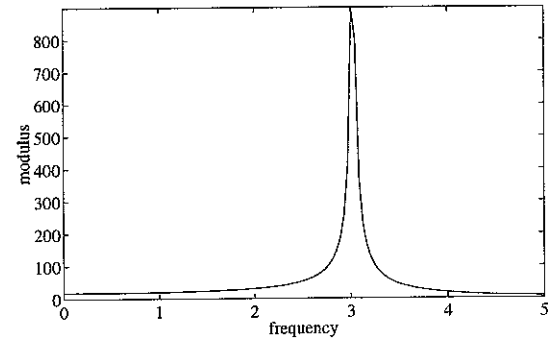


Figure 6. Solutions of the system (3.1) using the Runge-Kutta method ($f_0 = 4.51$)

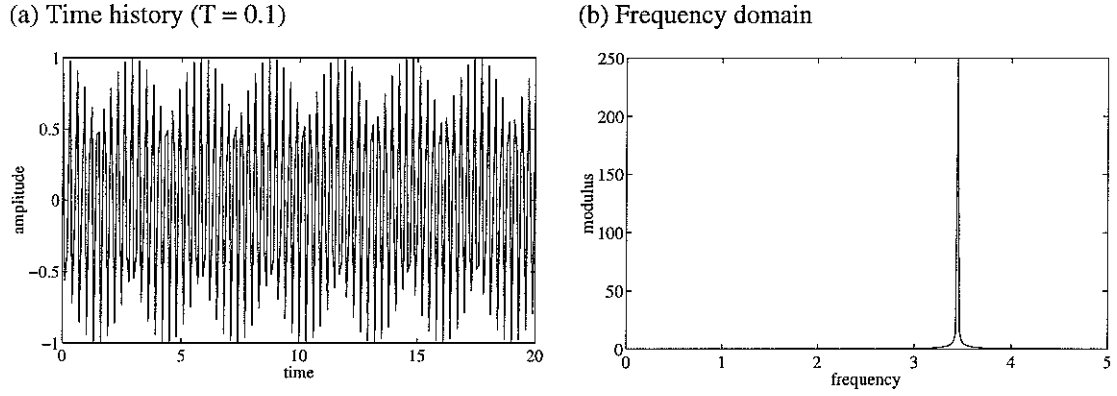


Figure 7. Solutions of the system (3.1) using the bilinear method ($f_0 = 6$, $f_d = 3.45$)

3.2 The Effect of Forcing

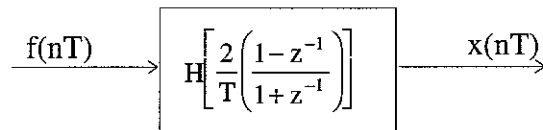
In the previous section, we only considered a homogeneous system. In this section, we include a forcing term, and demonstrate how the forcing term affects the bilinear mapping method. We now include a harmonic forcing term in equation (3.1), e.g.,

$$\ddot{x}(t) + \omega_0^2 x(t) = f(t), \text{ where } f(t) = A \cos \omega t \quad (3.8)$$

The right hand side of equation (3.8) is an input to the system. This input-output relationship can be described by the Laplace transformation which becomes ' $X(s) = H(s)F(s)$ ' for zero initial conditions. $H(s)$ is the transfer function of the system. The continuous system function $H(s)$ can be transformed to the discrete-time system function $H(z)$ by using the bilinear transform, i.e.,

$$H(z) = H(s), \text{ where } s = \frac{2}{T} \left(\frac{1 - z^{-1}}{1 + z^{-1}} \right) \quad (3.9)$$

The difference equation corresponding to $H(z)$ can be obtained by applying the trapezoidal integration rule (2.7) to the differential equation corresponding to $H(s)$, and if the input-output relationship of the system is described pictorially, then



Notice that the system is bilinearly mapped, but the input is not.

From the above illustration (input-output relationship), it is obvious that if the input is aliased by discretising coarsely, then the output is also aliased. Thus, in the case of equation (3.8), the sampling frequency ($1/T$) should be selected more than twice the highest input frequency in order to avoid aliasing.

One should also be aware that the frequency relationship equation (3.7) is only for the imaginary axis in the 's-Plane'. In other words, when the system possesses a large damping and the step size is very large, one may not be able to find the exact analogue frequency where the maximum amplitude occurs from the corresponding (maximum amplitude) digital frequency by using equation (3.7). Because poles in the s-Plane which are located other than on the imaginary axis do not follow equation (3.7) when they are mapped onto the z-plane, there may be an amplitude distortion when the step size and damping are very large, as shown in the Appendix. However in general, for small damping we may still use equation (3.7) since the frequency difference between a damped system and an undamped system is very small and a very large step size is unrealistic in practice. More details on this matter and the mapping from the 's-plane' to the 'z-plane' by the bilinear method is shown in the Appendix.

4. NONLINEAR SYSTEMS

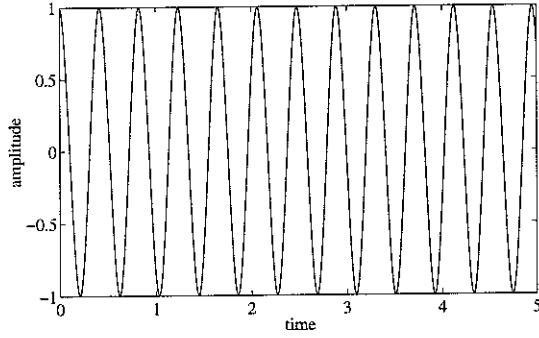
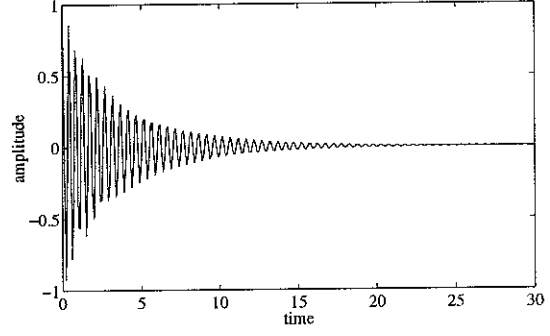
The previous section has summarised and given some numerical results for linear systems, and the usefulness of the bilinear mapping method has been emphasised. In this section, a non-linear term is included together with a study of whether the bilinear mapping method has similar properties for a non-linear system. So, the use of the bilinear method for a non-linear system is investigated by applying the trapezoidal rule to a simple known non-linear system and compared to the Runge-Kutta method. Consider a non-linear differential system without a forcing term

$$\ddot{x} + \omega_0^2 x + \epsilon x^3 = 0 \quad (4.1)$$

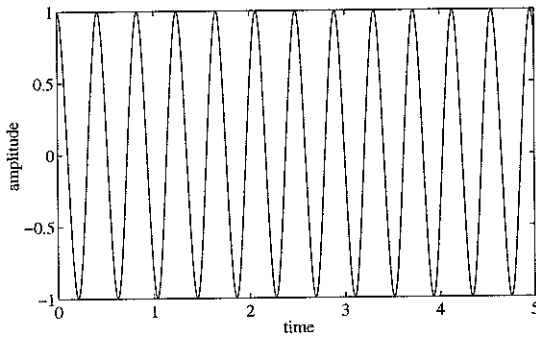
Equation (4.1) can be expressed in the form (2.1) as

$$\begin{aligned} \dot{x} &= y \\ \dot{y} &= -\omega_0^2 x - \epsilon x^3 \end{aligned} \quad (4.2)$$

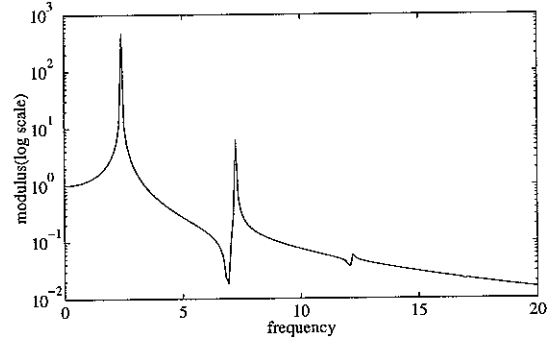
Equation (4.2) can be mapped by using the Runge-Kutta method and the bilinear method in the same way as described in the section 3. The results obtained by these two methods are shown by setting the initial conditions ($x_0 = 1$, $y_0 = 0$), $\epsilon = 100$ and $f_0 = 2$, where $\omega_0 = 2\pi f_0$. The Runge-Kutta method is very sensitive to the step size 'T'. Thus when the step size is relatively large compared to the inherent system's oscillation, this method gives an erroneous result as shown in Figure 8(b). Note that the original system does not have damping. Thus, the step size is set to be sufficiently small for the Runge-Kutta method (Figure 8(a)) to get close to the original system (i.e., it is assumed that the result of the Runge-Kutta method with step size 'T = 0.01' represents the true system (4.1), because it is observed that the solution obtained from the Runge-Kutta method with variable step size is almost the same as the Figure 8(a), although it is not shown here).

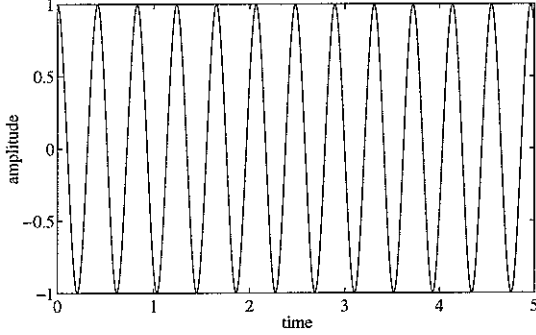
(a) Time history ($T = 0.01$)(b) Time history ($T = 0.1$)Figure 8. Solutions of the system (4.1) using the Runge-Kutta method ($f_0 = 2$)

The bilinear method is investigated by varying the step size to see the effect of discretisation. With step size ' $T = 0.01$ ' solutions from the Runge-Kutta method and the bilinear method are almost the same (see Figures 9 and 10). The frequency (Hz) components from the Runge-Kutta method are $f_{R1} \approx 2.4$, $f_{R3} \approx 7.2$ and $f_{R5} \approx 12$, where f_{R1} is the fundamental frequency and the others are third and fifth harmonics (only odd harmonics are present in this case, i.e., $f_{R3} = 3 \times f_{R1}$ and $f_{R5} = 5 \times f_{R1}$), and, the frequency components by the bilinear method are $f_{d1} \approx 2.4$, $f_{d3} \approx 7.2$ and $f_{d5} \approx 12$, where f_{d1} is the fundamental frequency and the others are third and fifth harmonics. These results are shown in Figure 9 and Figure 10 respectively.

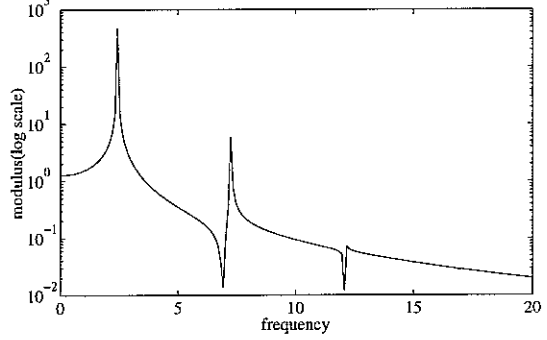
(a) Time history ($T = 0.01$)

(b) Frequency domain (log scale)

Figure 9. Solutions of the system (4.1) using the Runge-Kutta method ($f_{R1} \approx 2.4$, $f_{R3} \approx 7.2$ and $f_{R5} \approx 12$)

(a) Time history ($T = 0.01$)

(b) Frequency domain (log scale)

Figure 10. Solutions of the system (4.1) using the bilinear method ($f_{d1} \approx 2.4$, $f_{d3} \approx 7.2$, $f_{d5} \approx 12$ and $T=0.01$)

In the case of ' $T = 0.05$ ', the results of the bilinear method are shown in Figure 11, and the frequency components are $f_{d1} \approx 2.3$, $f_{d3} \approx 6.9$ and $f_{d5} \approx 8.5$ (aliased version of $f_{d5} \approx 11.5$). From the above two cases ($T=0.01$ and $T=0.05$), one can see that the fundamental frequency ($f_{R1} \approx 2.4$) is shifted to f_{d1} by equation (3.7), and, unlike the linear system, it is seen that aliasing does occur in the bilinear method for non-linear systems. This is caused by the fact that the *harmonic components do not follow equation (3.7)*. The harmonic components in the digital frequency domain (obtained using the bilinear method) are harmonics of the fundamental frequency in the digital frequency domain (which is computed from equation (3.7)), i.e., $f_{d3} = 3 \times f_{d1}$ and $f_{d5} = 5 \times f_{d1}$. Thus, the fifth harmonic f_{d5} is 11.5 Hz which is greater than the folding frequency 10 Hz and so appears at 8.5 Hz as an alias. This aliasing effect is more readily shown when the step size is increased further, and the results of the cases ' $T=0.1$ and $T=0.25$ ' are shown in Figures 12 and 13, and is summarised in Table 1 by comparing original frequencies (f_{Ri}) to the results from the bilinear method (f_{di}).

The following non-rigorous argument helps to explain the frequency components that arise. Equation (4.1) is often solved approximately using perturbation methods when ε is small [7], i.e., the solution of (4.1) can be expressed as a power series

$$x(t) = x_0(t) + \varepsilon x_1(t) + \varepsilon^2 x_2(t) + \dots \quad (4.3)$$

and we perturb the frequency ω by letting

$$\omega_0^2 = \omega^2 - \varepsilon b_1 - \varepsilon^2 b_2 - \dots \quad (4.4)$$

where b_i are the corrective terms for the frequency and are dependent on the amplitude of the motion. This yields the hierarchy of equations

$$\ddot{x}_0 + \omega^2 x_0 = 0 \quad (4.5a)$$

$$\ddot{x}_1 + \omega^2 x_1 = b_1 x_0 - x_0^3 \quad (4.5b)$$

$$\ddot{x}_2 + \omega^2 x_2 = b_1 x_1 + b_2 x_0 - 3x_1 x_0^2 \quad (4.5c)$$

:

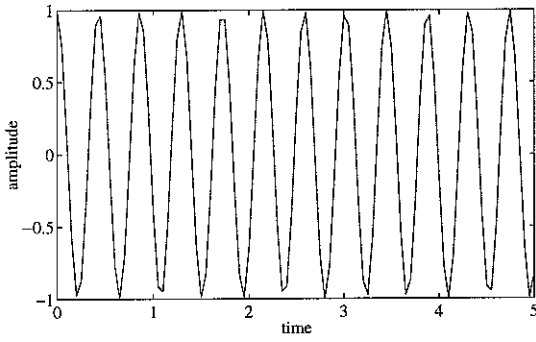
This shows that the fundamental frequency component shifts from ω_0 and the evaluations of the b_i (reference [7]) allow the modified fundamental frequency ω to be computed. Further the solution to (4.5a) acts as an input to (4.5b); the non-linear cubic term indicates the harmonic at 3ω . Assuming that applying the trapezoidal integration rule to (4.1) is equivalent to applying it to the set of equations (4.5), then (4.5a) will result in a digital oscillation occurring at $\omega_d = \frac{2}{T} \tan^{-1}\left(\frac{\omega T}{2}\right)$ (i.e., as equation (3.7)).

Equation (4.5b) is interpreted as a linear system driven by an input ' $b_1 x_0 - x_0^3$ ' which results in a digital frequency of ω_d *plus* a harmonic at $3\omega_d$. The right hand sides of (4.5) are like forcing terms as described in section 3 and are subject to aliasing leading to the results obtained in Figures 11 - 13. For example, ' $f_{d5} \approx 8.5\text{Hz}$ ' in Figure 11 is the aliased version of 11.5Hz , and similarly f_{d3} and f_{d5} in Figures 12 and 13 are the aliased versions as shown in Table 1. If the harmonic forcing term is included in equation (4.1), it can be written as

$$\ddot{x} + \omega_0^2 x = -\varepsilon x^3 + f(t), \text{ where } f(t) = A \cos \omega t \quad (4.6)$$

Thus, it can be interpreted that the whole right-hand side of equation (4.6) is an input whose effect is much the same as that described above (aliasing problem).

(a) Time history ($T = 0.05$)



(b) Frequency domain (log scale)

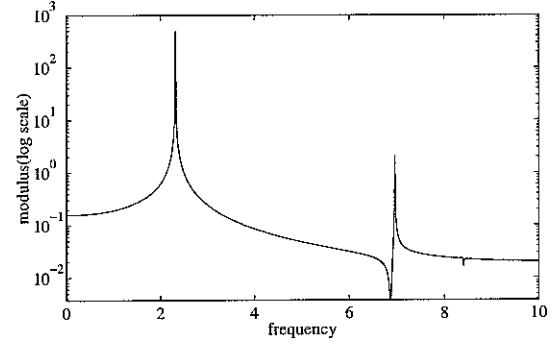
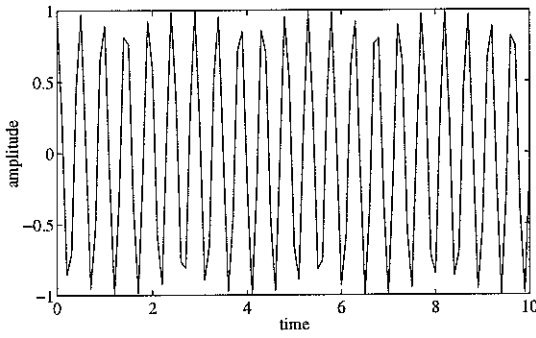


Figure 11. Solutions of the system (4.1) using the bilinear method ($f_{d1} \approx 2.3$, $f_{d3} \approx 6.9$, $f_{d5} \approx 8.5$ and $T=0.05$)

(a) Time history ($T = 0.1$)



(b) Frequency domain (log scale)

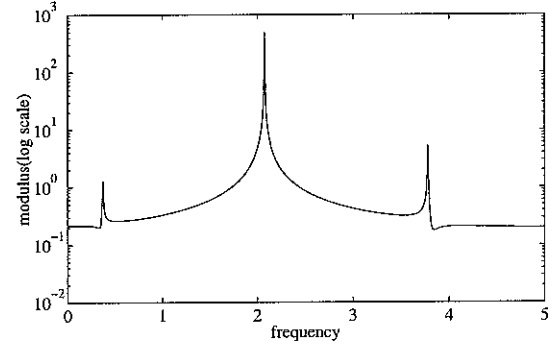
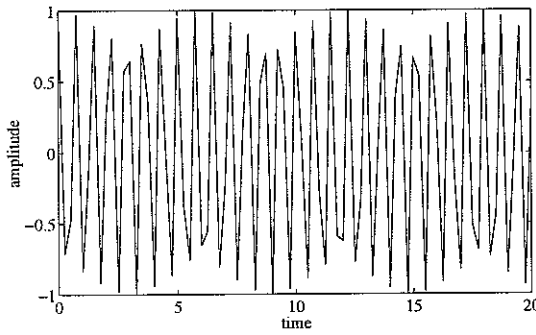


Figure 12. Solutions of the system (4.1) using the bilinear method ($f_{d1} \approx 2.06$, $f_{d3} \approx 3.82$, $f_{d5} \approx 0.3$ and $T=0.1$)

(a) Time history ($T = 0.25$)



(b) Frequency domain (log scale)

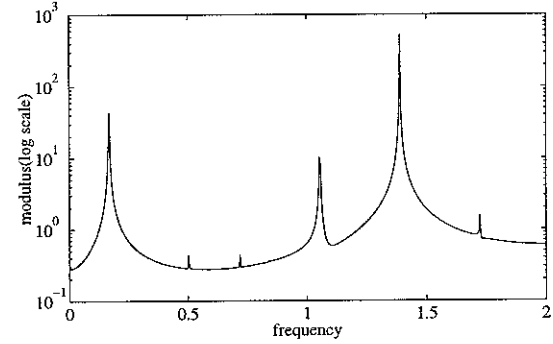


Figure 13. Solutions of the system (4.1) using the bilinear method ($f_{d1} \approx 1.38$, $f_{d3} \approx 0.14$, $f_{d5} \approx 1.1$ and $T=0.25$)

Original frequency	bilinear method (T=0.01)	bilinear method (T=0.05)	bilinear method (T=0.1)	bilinear method (T=0.25)
$f_{R1} \approx 2.4$	$f_{d1} \approx 2.4$	2.3	2.06	1.38
$f_{R3} \approx 7.2$	$f_{d3} \approx 7.2$	6.9	3.82* (6.18)	0.14* (4.14)
$f_{R5} \approx 12$	$f_{d5} \approx 12$	8.5* (11.5)	0.3* (10.3)	1.1* (6.9)

Table 1. Summary of the frequency components of the system (4.1).

* are the aliased version of ().

Consider the case that a small damping term is included in equation (4.1) so that

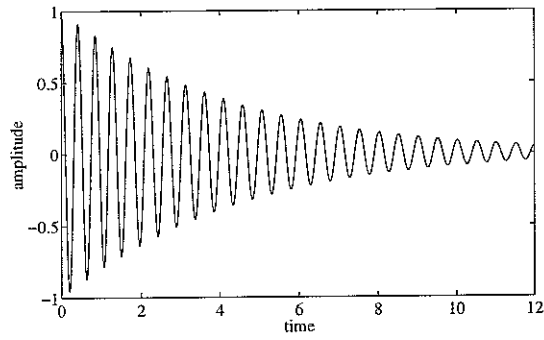
$$\ddot{x} + 2\zeta\omega_0\dot{x} + \omega_0^2 x + \epsilon x^3 = 0 \quad (4.7)$$

and this equation can be described in the form (2.1), i.e.

$$\begin{aligned} \dot{x} &= y \\ \dot{y} &= -2\zeta\omega_0 y - \omega_0^2 x - \epsilon x^3 \end{aligned} \quad (4.8)$$

where $\zeta = 0.02$ and the other parameters are the same as in the case of the undamped system. By changing step sizes, the results of the bilinear mapping method are shown in Figures 15 to 18, and the results are compared to the original system (Figure 14). It is assumed that the result of the Runge-Kutta method with step size (T=0.01) represents the true system (4.7). The results are very similar to the undamped case, i.e., the fundamental frequency follows equation (3.7) while the other harmonics do not and eventually alias when the step size is very large. The results are summarised for this system in Table 2 as for the undamped system by comparing original frequencies (f_{Ri}) to the results from the bilinear method (f_{di}).

(a) Time history ($T = 0.01$)



(b) Frequency domain (log scale)

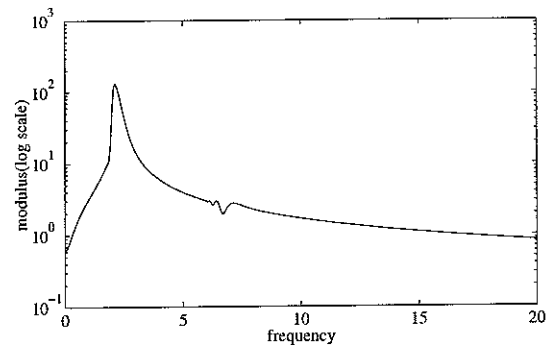
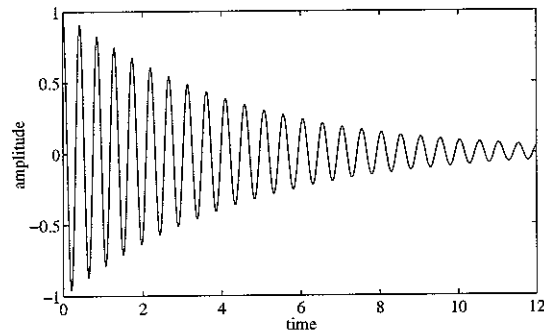


Figure 14. Solutions of the system (4.7) using the Runge-Kutta method ($f_{R1} \approx 2.2$, $f_{R3} \approx 6.6$ and $T=0.01$)

(a) Time history ($T = 0.01$)



(b) Frequency domain (log scale)

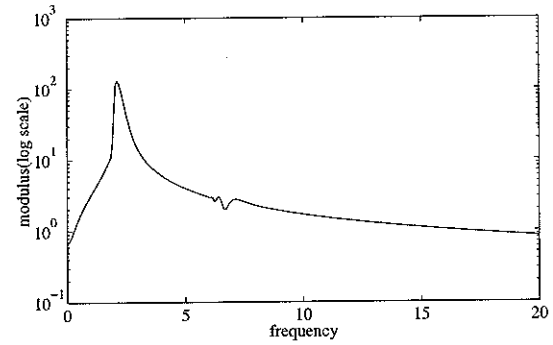
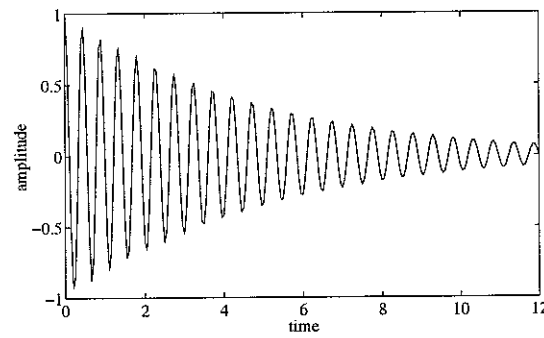


Figure 15. Solutions of the system (4.7) using the bilinear method ($f_{d1} \approx 2.2$, $f_{d3} \approx 6.6$ and $T=0.01$)

(a) Time history ($T = 0.05$)



(b) Frequency domain (log scale)

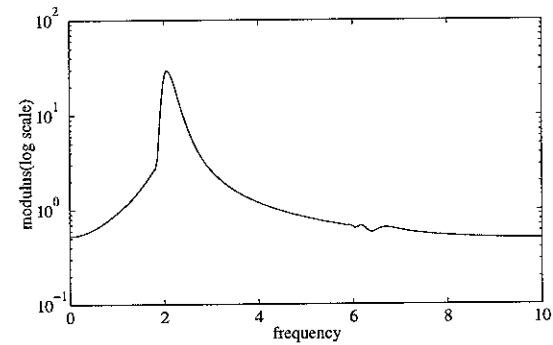
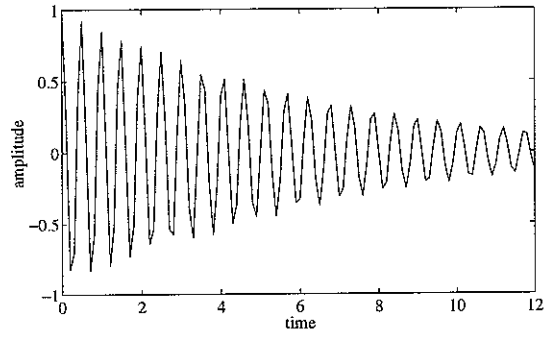
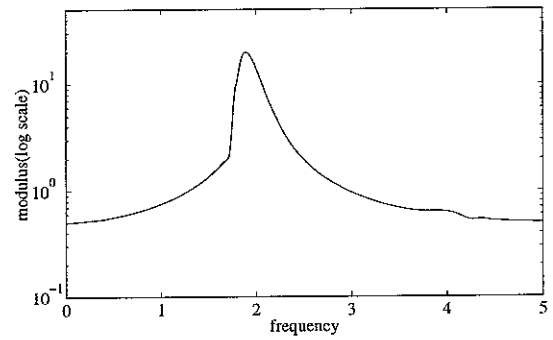


Figure 16. Solutions of the system (4.7) using the bilinear method ($f_{d1} \approx 2.1$, $f_{d3} \approx 6.3$ and $T=0.05$)

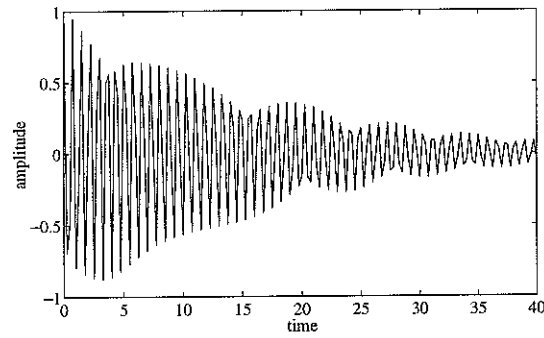
(a) Time history (T = 0.1)



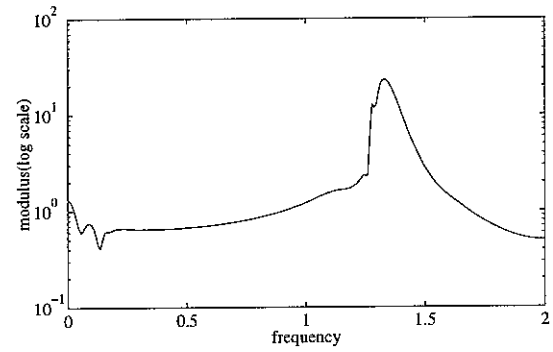
(b) Frequency domain (log scale)

Figure 17. Solutions of the system (4.7) using the bilinear method ($f_{d1} \approx 1.9$, $f_{d3} \approx 4.3$ and $T=0.1$)

(a) Time history (T = 0.25)



(b) Frequency domain (log scale)

Figure 18. Solutions of the system (4.7) using the bilinear method ($f_{d1} \approx 1.3$, $f_{d3} \approx 0.1$ and $T=0.25$)

Original	bilinear method	bilinear method	bilinear method	bilinear method
frequency	(T=0.01)	(T=0.05)	(T=0.1)	(T=0.25)
$f_{R1} \approx 2.2$	$f_{d1} \approx 2.2$	2.1	1.9	1.3
$f_{R3} \approx 6.6$	$f_{d3} \approx 6.6$	6.3	4.3* (5.7)	0.1* (3.9)

Table 2. Summary of the frequency components of the system (4.7).

* are the aliased version of ().

5. DISCUSSION

From the results of sections 3 and 4, it is shown that the use of the bilinear mapping method for non-linear systems is very useful. Since the fundamental frequency of the non-linear system obeys equation (3.7), one can understand the original differential system from the results of bilinear mapping method. Accordingly it may be possible to apply this to non-linear digital filter design with care, i.e., aliasing can occur in non-linear systems, and also the results of the bilinear method will give a good guide for the determination of the step sizes for other numerical methods, because the bilinear method is very stable even when the step size is quite large. For unknown systems one can first apply the bilinear method in order to get ' f_{d1} ' (fundamental frequency) and then convert this value to the analogue frequency by using equation (3.7). Then one can easily determine the step sizes for other sophisticated numerical methods since the inherent oscillation of the system may be understood from the results. One must be aware of the possibility that application of the bilinear method to a non-linear system does not ensure the accuracy of the true trajectories of the original system, i.e., the original frequency of the continuous system always becomes lower when the bilinear method is applied.

REFERENCES

1. A. Iserles 1994 Numerical Analysis Report, May DAMTP 1994/NA5. Dynamics of Numerics, University of Cambridge
2. A. T. Peplow 1989 MSc Thesis in Nonlinear Mathematics. Dynamics of Numerical Methods of Initial Value Problems, University of Bath
3. J. M. Sanz-Serna 1990 Applied Mathematics and Computation Report, Report 1990/3. Numerical ordinary differential equations vs. dynamical systems, Universidad de Valladolid, Spain
4. A. Iserles, A. T. Peplow and A. M. Stuart 1991 SIAM J. NUMER. ANAL. **28**(6), 1723-1751. A Unified Approach to Spurious Solutions Introduced by Time Discretisation. Part I: Basic Theory
5. D. F. Griffiths, P. K. Sweby and H. C. Yee 1992 IMA Journal of Numerical Analysis **12**, 319-338. On Spurious Asymptotic Numerical Solutions of Explicit Runge-Kutta Method
6. S. Z. Abdul-Hamid 1994 MSc Thesis in Sound & Vibration Studies. The Numerical Effect of Discretisation of Continuous Time Nonlinear Systems. Institute of Sound and Vibration Research, University of Southampton
7. A. V. Oppenheim and R. W. Schaffer 1989 Discrete-Time Signal Processing. Prentice Hall
8. F. S. Tse, I. E. Morse and R. T. Hinkle 1988 Mechanical Vibrations; Theory and Applications. Prentice Hall
9. E. Kreyszig 1993 Advanced Engineering Mathematics. 7-th ed, John Wiley & Sons
10. D. Stolle 1985 Letter to The Editor, Journal of Sound and Vibration. **180**(3), 513-518. A direct integration algorithm and the consequence of numerical stability
11. G. W. Gear 1971 Numerical Initial Value Problems in Ordinary Differential Equations. Prentice Hall
12. L. Lapidus and J. H. Seinfeld 1971 Numerical Solution of Ordinary Differential Equations. Academic Press

13. J. D. Lambert 1973 Computational Methods in Ordinary Differential Equations.
John Wiley & Sons

APPENDIX. BILINEAR MAPPING FROM S-PLANE TO Z-PLANE

When a complex plane (s-plane) is mapped to another complex plane (z-plane) using the bilinear method, the following relationship holds, i.e., three given distinct points s_1, s_2, s_3 can always be mapped onto three prescribed distinct points z_1, z_2, z_3 by bilinear transformation.

$$\frac{z - z_1}{z - z_3} \cdot \frac{z_2 - z_3}{z_2 - z_1} = \frac{s - s_1}{s - s_3} \cdot \frac{s_2 - s_3}{s_2 - s_1} \quad (\text{A } 1)$$

The transformation from the left half plane in the s-plane to the unit disk in the z-plane can be found by setting $s_1 = 0, s_2 = -2/T, s_3 = \infty, z_1 = 1, z_2 = 0$ and $z_3 = -1$, and this results in

$$z = \frac{1 + \frac{T}{2}s}{1 - \frac{T}{2}s} \quad (\text{A } 2)$$

Rearranging this expression, it is the same as equation (3.8), i.e.,

$$s = \frac{2}{T} \left(\frac{z - 1}{z + 1} \right) = \frac{2}{T} \left(\frac{1 - z^{-1}}{1 + z^{-1}} \right) \quad (\text{A } 3)$$

The Figure A-1 shows the details of the bilinear mapping from s-plane to z-plane. In a single degree of freedom (SDOF) linear vibration system, there is the following relationship,

$$\omega_d = \omega_n \sqrt{1 - \zeta^2} \quad \text{i.e.,} \quad (\text{A } 4)$$

$$\omega_d^2 + (\zeta \omega_n)^2 = \omega_n^2 \quad (\text{A } 5)$$

where ω_d is the damped angular natural frequency, ω_n is the undamped angular natural frequency and ζ is the damping ratio. Thus, equation (A 5) can be described in the s-plane as shown in Figure A-2. ω_n becomes the radius R in both Figures A-1 and Figure A-2.

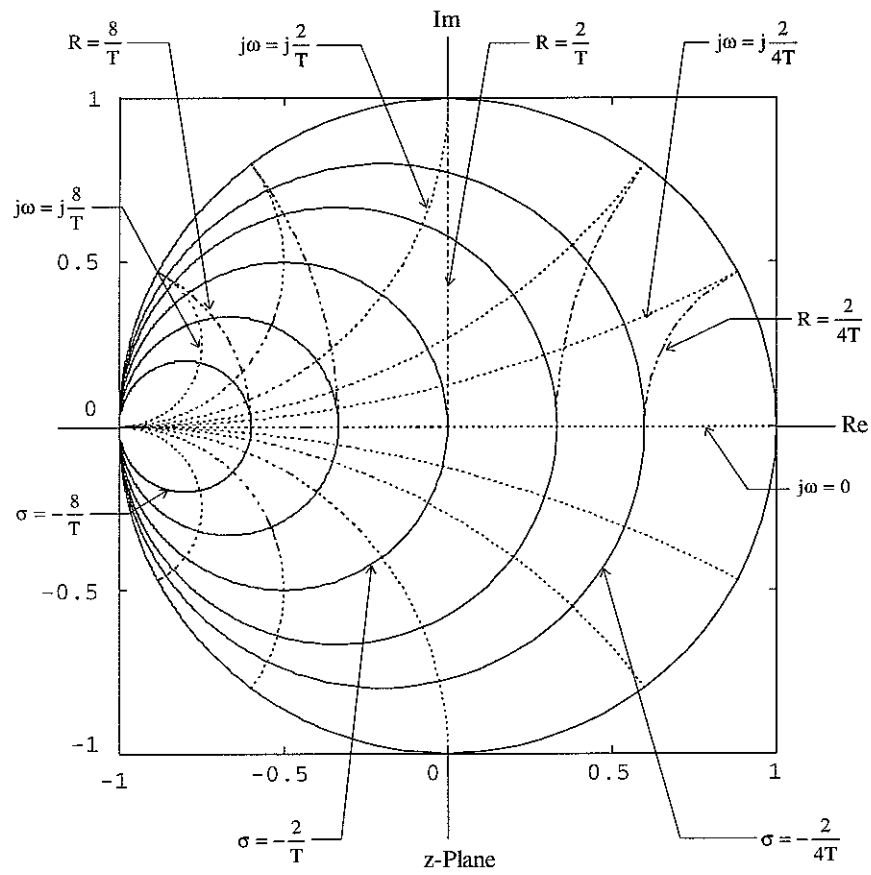
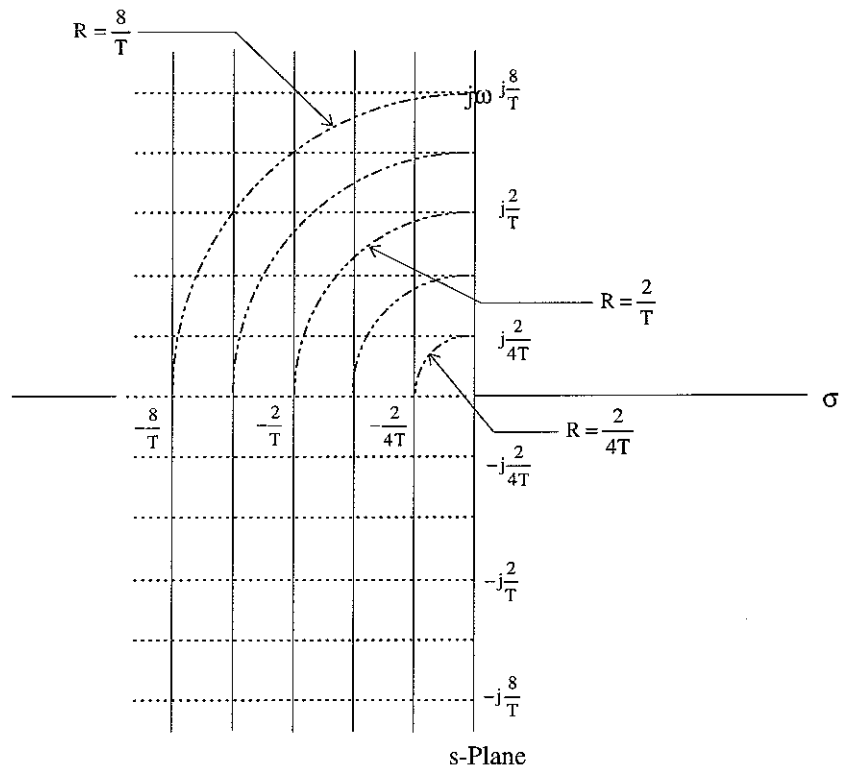


Figure A-1. Bilinear Mapping from the s-plane to the z-plane

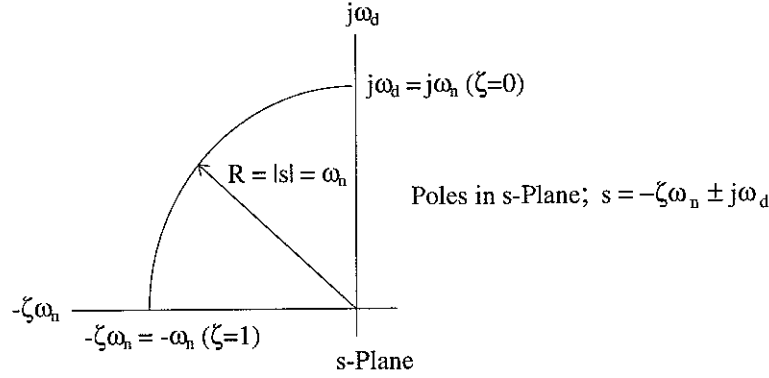


Figure A-2. s-plane representation of the relationship between ω_n and ω_d

From Figures A-1 and A-2, it is obvious that the damped natural frequency is the same as the undamped natural frequency in the z-plane when ω_n is equal to $2/T$. Since the radius ' $R=2/T$ ' in the s-plane is mapped to the imaginary axis vertically in the z-plane, the damped natural frequency evaluated in the z-plane is the same regardless of the damping ratio, where the arc of radius R is the location of pole in the s-plane. This is shown in Figure A-3 both in the time and frequency domains. Note that the resonant frequency (maximum amplitude) is slightly lower than the damped natural frequency, i.e., the resonant frequency (ω_{\max}) in the s-plane is ' $\omega_{\max} = \omega_n \sqrt{1 - 2\zeta^2}$ ', which is a little smaller than the ω_d especially when the damping ratio is very large as shown in Figure A-3(b) in the case of ' $\zeta = 0.4$ '. Finally, if ω_n is greater than $2/T$, the bilinear method gives a damped natural frequency which is larger than the undamped natural frequency. This may be interpreted as an amplitude distortion in the frequency domain, i.e., although the frequency components in the imaginary axis in the s-plane mapped onto the z-Plane by the relationship(3.7) the location of poles in z-Plane may alter the amplitude in the frequency domain when the step size and the damping are very large.

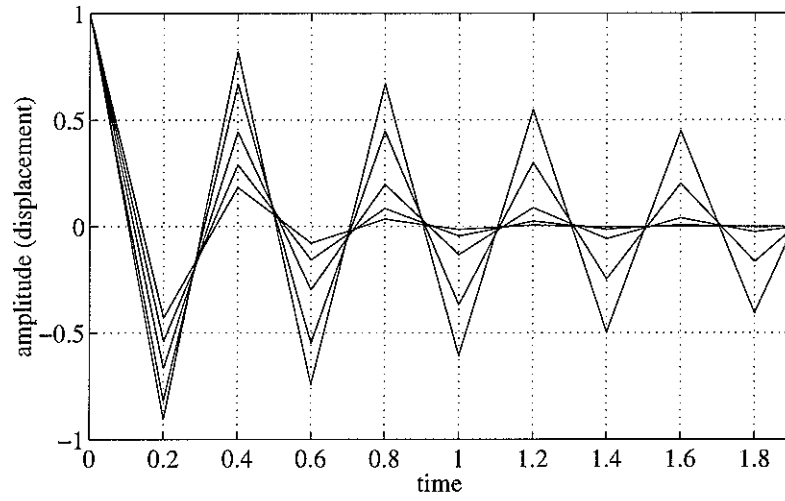


Figure A-3(a). Time history of a damped SDOF free vibration system using the bilinear method in the case of ' $\omega_n = 2/T$ ', where T is 0.1 and the damping ratios are ' $\zeta = 0.05, 0.1, 0.2, 0.3,$ and 0.4 '.

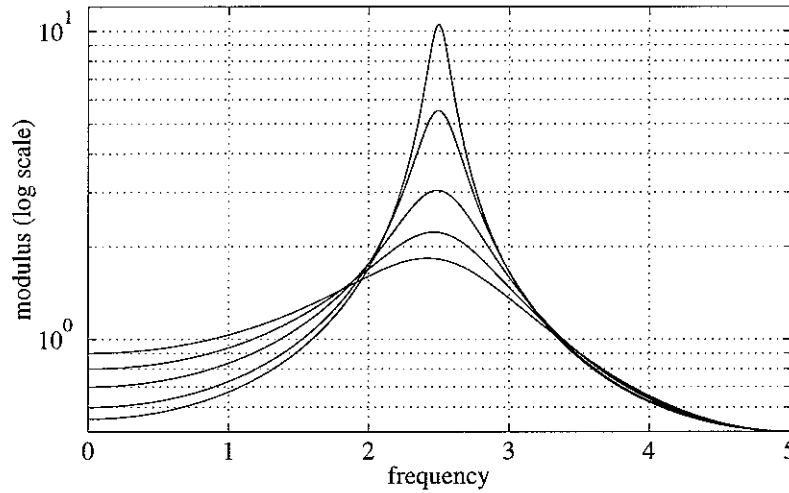


Figure A-3(b). Frequency characteristics of a damped SDOF free vibration system using the bilinear method in the case of ' $\omega_n = 2/T$ ', where T is 0.1 and the damping ratios are ' $\zeta = 0.05, 0.1, 0.2, 0.3,$ and 0.4 '.