Stability of linear and nonlinear delay-differential systems

Cardwell, E. A.

Award date:
1995

Awarding institution:
University of Bath

Stability of Linear and Nonlinear Delay–Differential Systems

submitted by

E. A. Cardwell

for the degree of Ph.D

of the

University of Bath

1995

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Summary

This research is concerned with the stability of linear and nonlinear control systems containing time-delay elements. All systems will be nonlinear to some extent and equally contain some element of delay. Often these are neglected and linear system models given for simplicity; here the aim is to extend theory for linear time-delay systems and nonlinear delay-free systems to the delayed nonlinear case.

Existing stability techniques for linear and nonlinear systems are reviewed and developed providing a foundation for later investigations.

The stability problem for linear time-delay systems is concerned with the position of roots of the characteristic equation in the complex plane; a temporal locus method is given for the paths of these roots as the delay is increased.

Particular attention is devoted to systems of Van der Pol form containing delay, and the effect of increasing delay on the oscillatory behaviour. Linear delay methods find application together with nonlinear delay-free techniques.

Problems of two delays, either commensurate or non-commensurate, are discussed. A predictor control scheme containing mismatch in plant or delay receives consideration and the temporal locus method is extended to this two delay problem.

Contents

List of Figures

Acknowledgements

> Be strong and courageous and do the work. Do not be afraid or discouraged, for the Lord God, my God, is with you. He will not fail you or forsake you until all the work for the service of the temple of the Lord is finished.
>
> 1 Chronicles, chapter 28 verse 20

I should like to express my gratitude to my supervisor, Dr. John Marshall, for all his help and insights concerning this research, for his unending enthusiasm and his willingness always to give of his time.

Thanks are also due to Dr. Keith Walton for his valuable contributions, to Dr. Fran Burstall for sorting out my LaTeX problems, and to members of the Control and Mathematics Groups and computer support team for their help.

I should like to acknowledge the financial assistance provided by the Engineering and Physical Sciences Research Council enabling this research to take place.

Thanks to everyone in 1W3.6 and 1W3.5 for their friendship and a great atmosphere in which to work, and also to the Bailey and Dimond families for providing me with two wonderful homes.

I am extremely grateful to my family for all their support from such a great distance. Thanks to my husband-to-be Keith, for his constant support, encouragement and help in so many ways, and for giving me such an excellent incentive of the wedding to get written up!

> To the only God our Saviour be glory, majesty, power and authority, through Jesus Christ our Lord, before all ages, now and for evermore! AMEN!
>
> Jude, verse 25

Chapter 1

Introduction

Control systems may be modelled by means of differential equations. The stability and dynamic behaviour of linear delay-differential equations, particularly dependence on delay values, are complicated, but much progress has been made in their understanding. Equally, extensive study has taken place into the stability of nonlinear delay-free systems. The aim of this research is to develop theory in these two areas and to analyse nonlinear delay-differential systems.

These nonlinear time-delay systems are assumed to be the concatenation of three types of system. A threefold approach will be taken. Initially, delay, linear dynamics and instantaneous nonlinear functions are considered separately (to be termed *unary systems*), before being taken in pairs (*binary systems*) and finally all three strands are combined (*tertiary systems*). This approach where the delay, nonlinearity and linear dynamics are in a sense separable will enable a control theory perspective to be taken and certain classical techniques to be exploited. Systems containing two (or more) delays will also be of interest.

Introductory material is included in each section where appropriate. The following describes the structure of this thesis.

Unary systems are the subject of chapter 2, which contains a review of classical stability methods for linear delay-free systems, and discussion of instantaneous nonlinear systems and those with delay, gain and no linear dynamics.

The second element of the threefold model is introduced in chapter 3. Such binary systems fall into three categories. For linear time-delay systems, existing stability methods are described, the direct method of Walton and Marshall [71] for stability with delay being particularly relevant. Approximation of delay elements by Padé approximant models [48] is discussed. A vector approach for critical stability finds application in nonlinear system stability. Consideration of the effect of delay on phase plane be-

haviour for second order systems is made. A method is described for the construction of temporal loci marking the paths of roots of the characteristic equation of a linear time-delay system as the delay is increased, which gives valuable insights into stability properties.

Nonlinear delay-free systems fall into the next category. Again a review is given of existing stability methods, both exact (conservative) and approximate, reducing the problem to a linear one in some way. Attention is focussed on two forms of nonlinearity: piecewise linear monotonic functions, in particular saturation, and analytic functions such as the squaring function. The latter form is present in the equation of the Van der Pol oscillator [68] which is the subject of much of the subsequent nonlinear analysis. Averaging methods are discussed for quasilinear differential equations (of which the Van der Pol equation is a special case) in which the nonlinear element is small; series expansions for the solution of such equations for small and large nonlinear terms are given. This section also includes the vector approach for nonlinear systems.

The third category of binary system is the nonlinear dynamics-free delay system which reduces to a standard dynamical systems problem.

The three elements of nonlinearity, dynamics and delay are brought together in chapter 4. Attention has been devoted to equations of Van der Pol form containing delay in state or velocity. Use is made of linear and nonlinear results obtained in chapter 3: the direct method and temporal locus inform stability of the corresponding linearised system which in turn has implications for stability of the nonlinear system, and averaging methods (including the vector approach) and series expansions are extended to incorporate small delays. Trajectories of the solution in the state-velocity plane, and in particular limit cycle behaviour, are considered.

The subject of chapter 5 is systems with two or more delays, which may be commensurate (related by an integer) or non-commensurate. These are considered within the framework of predictor schemes following the method due to O. J. M. Smith [60], [61] with mismatch in the plant and delay models. Linear single-delay techniques are extended to the two delay case, including the temporal locus for one delay fixed and the other varying. Finally a brief discussion is given of the addition of a saturating nonlinear element to such two delay problems and of a second delay to the Van der Pol systems considered in chapter 4.

Concluding remarks are made in chapter 6.

Chapter 2

Unary Systems

2.1 Introduction

The stability analysis of these unary systems is, in part, elementary or treated extensively elsewhere. The material contained in this chapter is included for completeness and as a review of standard methods. Many of the techniques described here are extended in later chapters, when another element of the threefold model is introduced. In particular the control of linear time-delay systems has a firm foundation in the classical theory of delay-free systems as shall be seen in section 3.1.

For systems with a linear model, the Laplace transform of the differential equation may be taken and the system components modelled by their transfer function. Stability is then determined by the location in the complex s-plane of the roots of the characteristic equation of the closed loop system. Stability holds for all roots in the left-half plane; roots on the imaginary axis correspond to oscillation and for any roots in the right-half plane there is instability. In what ensues, $s := \sigma + i\omega$ in terms of the real variables σ, ω.

This research has concentrated on physical systems, that is those for which the plant is of retarded type (having transfer function strictly proper, with denominator of greater degree than numerator) as opposed to advanced (numerator of greater degree than denominator). Neutral systems (with degrees equal) being atypical of physical systems are omitted.

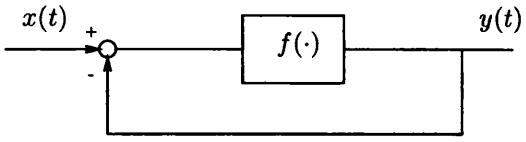


Figure 2-1: Instantaneous Nonlinear System

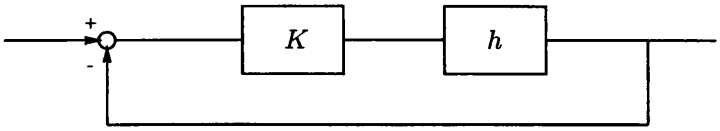


Figure 2-2: Delay System

2.2 Instantaneous Nonlinear Systems

The system in figure 2-1 shows an instantaneous nonlinearity (no linear dynamics) with unity negative feedback. Here

$$y(t) = f(x(t) - y(t))$$

and letting

$$g(x, y) = f(x - y) - y = 0$$

the system reduces to the standard numerical analysis steady state problem which may be solved by Newton or quasi-Newton methods. This is treated extensively elsewhere (see for example [6] and [47]).

2.3 Delay Systems

The system with gain K in series with a time-delay h and unity negative feedback (figure 2-2) has characteristic equation

$$1 + K\exp(-sh) = 0$$

This is a simple neutral system, the degree of the numerator and denominator of the transfer function of the 'plant' K being equal. As mentioned earlier, the emphasis of this research is on physical systems so in general will not be concerned with equations of neutral type. Chapter 2 of [35] gives detailed analysis of the stability of neutral linear delayed equations, including the behaviour as the delay is increased from zero and stability of the infinitely many new roots coming in from infinity arising from the delay. It may be shown that the above system is stable independent of delay for

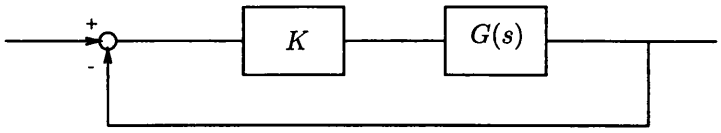


Figure 2-3: Linear System

$0 \leq K < 1$ but unstable for all values of delay if $K \geq 1$.

2.4 Linear Delay-Free Systems

Figure 2-3 shows the linear delay-free system with gain K, plant of transfer function $G(s)$ and unity negative feedback.

$$G(s) = \frac{C(s)}{A(s)}$$

where $C(s)$, $A(s)$ are polynomials in s, and $G(s)$ is proper. The system has the following characteristic equation:

$$F(s) := A(s) + KC(s) = 0 \tag{2.1}$$

The problem has been studied in great detail in classical control. Three stability methods are given here, one algebraic and two graphical. There are many excellent references, including [14], [53], [57] (and [11], [26], [44], [51], [73] amongst others).

2.4.1 The Routh-Hurwitz Stability Criterion

A century ago Routh [54] and Hurwitz [24] independently derived equivalent criteria for stability using an algebraic procedure. The methods give necessary and sufficient conditions for all roots of a polynomial to lie in the left half plane without determining the roots (hence relative stability/instability is not given).

Suppose a system has characteristic equation

$$s^n + b_{n-1}s^{n-1} + b_{n-2}s^{n-2} + \cdots + b_1 s + b_0 = 0 \tag{2.2}$$

First note that for all roots in the left half plane it is necessary that

$$b_i > 0 \qquad i = 0, \cdots n-1$$

since

$$b_i = (-1)^{n-i} \sum \textit{product of roots taken } (n-i) \textit{ at a time.}$$

Here is stated the criterion due to Hurwitz. Form the n by n array

$$\begin{pmatrix} b_{n-1} & 1=b_n & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ b_{n-3} & b_{n-2} & b_{n-1} & b_n & 0 & \cdots & \cdots & \cdots & 0 \\ b_{n-5} & b_{n-4} & b_{n-3} & b_{n-2} & b_{n-1} & b_n & 0 & \cdots & 0 \\ & & & & & & \ddots & & \\ & & & & & & & \ddots & \\ & & & & & & & & b_0 \end{pmatrix}$$

It is required that all principal minors be greater than zero:

$$b_{n-1} > 0; \quad b_{n-1}b_{n-2} - b_n b_{n-3} > 0 \quad \text{etc.}$$

Routh's method involves forming an array, the number of sign changes in its left hand column corresponding to the number of roots in the right half plane. It can be shown that this column can be expressed in terms of the principal minors of the Hurwitz criterion, and hence that the methods are equivalent (see for example [50]).

2.4.2 The Root Locus Method of Evans

The Evans root locus method [12], [13] is a graphical technique for the stability of the system in figure 2-3 as the gain K increases from zero to infinity. A plot is made in the complex s-plane of the roots of the characteristic equation of the closed-loop transfer function as K increases, using the open-loop poles and zeros which are readily available. The range of the gain, if any, for which the system is stable (all poles in the left half plane) may then be determined, and also any modifications to the controller necessary for stability for a desired range (for example phase advance or introduction of complex zeros).

Let

$$G(s) = \frac{C(s)}{A(s)} = \frac{\prod_{i=1}^{m}(s - z_i)}{\prod_{i=1}^{n}(s - p_i)}$$

where z_i, p_i are the open loop zeros and poles respectively. The characteristic equation is given by 2.1, so

$$-K = \frac{A(s)}{C(s)} \tag{2.3}$$

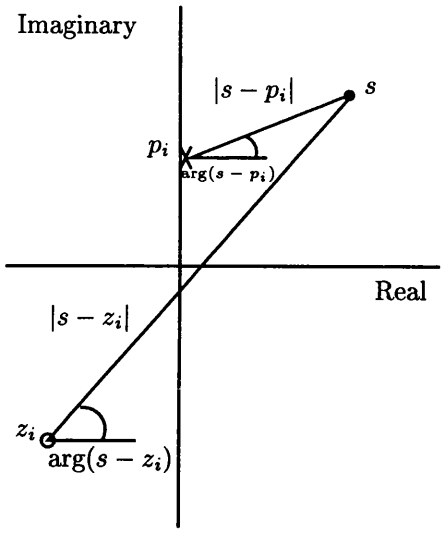


Figure 2-4: Argument and Modulus Rule

Taking arguments and moduli yields

The Angle Balance Rule

$$\sum_{i=1}^{n} \arg(s - p_i) - \sum_{i=1}^{m} \arg(s - z_i) = (2j + 1)\pi, \qquad j \epsilon \mathbb{N}$$

For a root s on the locus, the angles from the poles and zeros of the open-loop transfer function to s must satisfy the above equation.

The Modulus Rule

$$K = \frac{\prod_{i=1}^{n} |s - p_i|}{\prod_{i=1}^{m} |s - z_i|}$$

The gain at a point s on the locus is the ratio of *product of vector lengths from open loop poles to s* to *product of vector lengths from open loop zeros to s*.

Figure 2-4 (of the complex s-plane) illustrates this. There are many construction rules in addition to these.

1. There are n branches, corresponding to n roots.

2. The closed-loop transfer function is a real polynomial so complex roots occur in complex conjugate pairs and the plot is symmetrical with respect to the real axis.

3. At $K = 0$ the loci start at the open loop poles. m of the roots tend to the open loop zeros as $K \to \infty$ with the remainder 'going to infinity'.

4. Segments on the real axis form part of the locus if they lie to the left of an odd number of open loop poles and zeros, to satisfy the angle balance rule. (Real poles and zeros to the right of the point contribute an angle π; all others contribute nothing.)

5. Asymptotes to the $n - m$ branches which tend to infinity are equispaced in angle and intersect at the real point

$$\frac{\sum_{i=1}^{n} p_i - \sum_{i=1}^{m} z_i}{n - m}$$

   This follows from an expansion of 2.3 for K and s large in magnitude, and retention of only the dominant terms.

6. Branches depart from or arrive at multiple roots on the real axis at equispaced angles.

7. Intersection points of the locus with the imaginary axis (critical stability) may be determined by applying the Routh-Hurwitz stability criterion.

8. The angle of departure from a complex pole, or arrival at a complex zero, may be found using the angle balance rule.

In section 3.1 these rules will be extended to the linear time-delay case.

2.4.3 The Nyquist Stability Criterion

The Nyquist stability criterion [46] is a graphical method for determining the degree of stability or instability of a system with characteristic equation

$$1 + KG(s) = 0$$

and how the steady state or transient response may be improved.

The method is based on Cauchy's principle of the argument. Let $F(s)$ be a rational function of s and $\mathcal{C}$ be a simple closed contour, described counterclockwise and not passing through any poles or zeros of $F(s)$. Then the change in the argument of $F(s)$ (in a counterclockwise sense) is $2\pi(Z - P)$, Z being the number of zeros and P the number of poles enclosed by $\mathcal{C}$.

Now let $\mathcal{C}$ be the 'D' contour shown in figure 2-5 as $R \to \infty$ (the contour may be indented to the right of any imaginary poles or zeros) and use a reverse description ($\mathcal{C}$

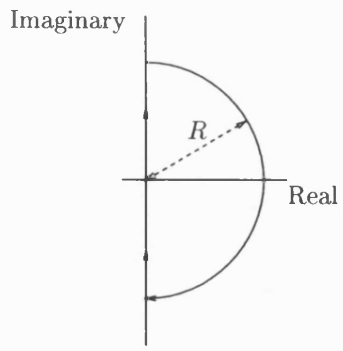


Figure 2-5: 'D' Contour

described clockwise and change in argument of $F(s)$ measured clockwise). The number of encirclements of the origin by the map of $\frac{1}{K} + G(s)$, or equivalently of the $-\frac{1}{K}$ point by $G(s)$, is required. Assuming knowledge of P, the number of open loop unstable poles of $G(s)$, gives Z, the number of closed loop unstable zeros. Note that the circumference of the 'D' maps to the origin since for physical systems, transfer functions are proper. In addition, $G(s)$ is real so that $G(-i\omega) = \overline{G(i\omega)}$. The Nyquist locus is a polar plot of the frequency response of the open-loop transfer function so it is not necessary to know $G(s)$ in polynomial form. The gain and phase as the frequency is increased from zero to infinity may be measured from the physical system provided that $G(s)$ is stable. The locus enables values of K ensuring stability to be determined; a series compensator may be included to pull the plot away from the $-\frac{1}{K}$ point. The inverse Nyquist plot is the corresponding Nyquist locus for $1/G(s)$ and is particularly useful for systems with delay (see sections 3.1.3 and 3.1.4).

2.4.4 The Mikhailov Stability Criterion

Mikhailov [41] developed another graphical approach, related to the Nyquist method, for the stability of a system with characteristic equation 2.1, where, as before, $A(s)$ and $C(s)$ are polynomials of order n, $m(< n)$ respectively. It is readily extendible to the corresponding delay case as shall be seen (section 3.1.4). If the equation is written in form 2.2 then as remarked previously all the coefficients must be strictly positive. For $s = i\omega$ writing

$$F(i\omega) = f_1(\omega) + if_2(\omega)$$

f_1 will be a polynomial in powers of ω^2 and f_2 a polynomial in odd powers of ω with common factor ω. The method involves a plot of $F(i\omega)$ as ω varies from zero to infinity. Note that if $C(s)$ is a constant c_0, the inverse Nyquist plot for $1 + \frac{Kc_0}{A(s)}$ results. There are two alternative formulations of the stability criterion.

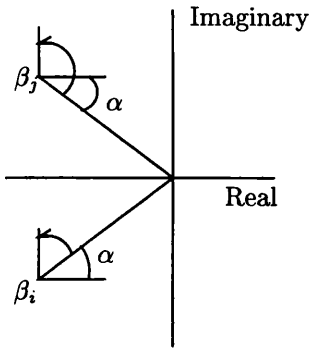


Figure 2-6: Contribution of Complex Roots

First Formulation A plot of $F(s)$ is made as s increases along the imaginary axis from 0 to $i\infty$. Then if F has no zeros on the imaginary axis, and $A(s)$ and $C(s)$ have real coefficients, then for a stable system the change in $\arg F(s) = n\pi/2$ where n is the degree of $A(s)$.

Proof For small values of $\omega > 0$, f_1 and f_2 are decreasing and increasing functions of ω respectively. So the $F(i\omega)$ curve moves anticlockwise from the point $(b_0, 0)$. Let the equation 2.2 have roots $\beta_1, \ldots, \beta_n$; for $s = i\omega$,

$$\prod_{i=1}^{n}(i\omega - \beta_i) = 0$$

Consider separately the following cases:

1. β_i real and negative. Then the angle of rotation of the vector from β_i to $i\omega$ as ω increases to infinity will be $\frac{\pi}{2}$ (counterclockwise).

2. β_i real and positive. Similarly the change in argument is $-\frac{\pi}{2}$.

3. $\beta_i = \overline{\beta_j}$ a pair of complex conjugate roots with negative real part. The contribution to the argument change from these two roots (from figure 2-6) is clearly $(\frac{\pi}{2} - \alpha)$ for β_i and $(\frac{\pi}{2} + \alpha)$ for β_j giving a total of π radians overall.

4. For a pair of complex conjugate roots with positive real part the contribution is $-\pi$.

So if $F(s)$ has q unstable roots then the change in argument equals

$$(n-q)\frac{\pi}{2} - q\frac{\pi}{2} = (n-2q)\frac{\pi}{2}$$ ■

Note that for a zero root, $b_0 = 0$ and the locus starts at the origin. For imaginary roots the locus passes through the origin at the critical frequency.

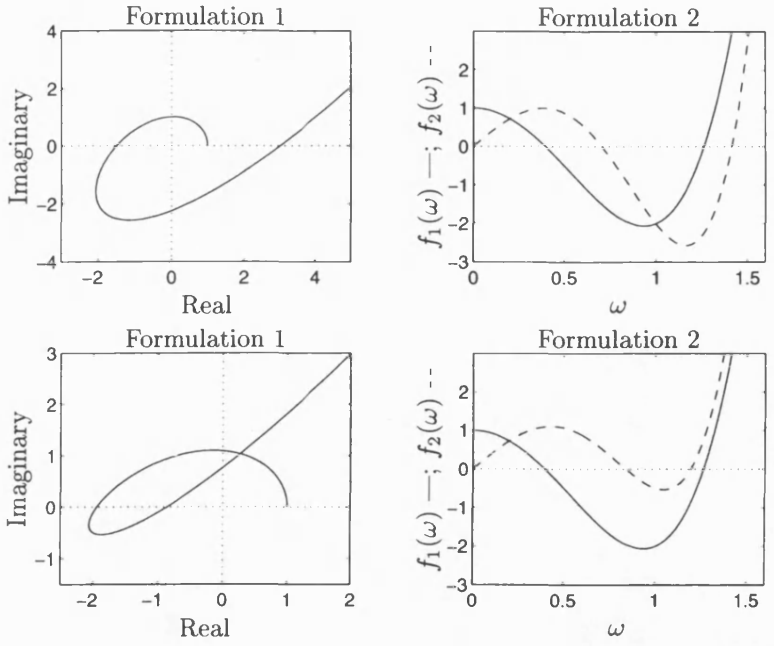


Figure 2-7: Mikhailov Plots for $4s^5 + 4s^4 + 10s^3 + 7s^2 + 4s + 1 = 0$ (above) and $4s^5 + 4s^4 + 8.5s^3 + 7s^2 + 4s + 1 = 0$ (below)

Second Formulation The system is stable if and only if the zeros of $f_1(\omega)$ and $f_2(\omega)$ are all real (simple) and alternate. (This may be determined from a plot of $f_1(\omega)$ and $f_2(\omega)$ against ω.) This follows directly from the first formulation, as for stability the locus intersects the real and imaginary axes alternately.

Figure 2-7 shows typical Mikhailov plots for the two formulations for a stable and unstable system. It can be seen that for the first example

$$4s^5 + 4s^4 + 10s^3 + 7s^2 + 4s + 1 = 0$$

the first plot encircles the origin anticlockwise by $5\pi/2$ and the roots of $f_1(\omega)$ and $f_2(\omega)$ interleave one another so the system is stable, whereas for

$$4s^5 + 4s^4 + 8.5s^3 + 7s^2 + 4s + 1 = 0$$

the plot fails to encircle the origin and the roots of $f_1(\omega)$ and $f_2(\omega)$ do not alternate, hence the system is unstable. The graphs for either formulation are difficult to plot and different values of the gain or any coefficient of A or C will modify the curves in possibly a quite complicated way. See [15], [33] and [52] for further details.

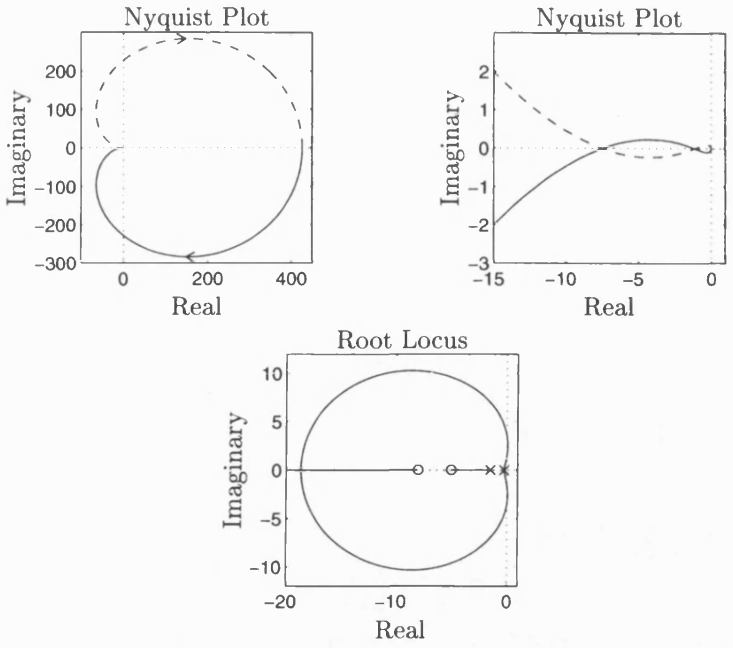


Figure 2-8: Conditional Stability: $1 + K\frac{(s+5)(s+8)}{(s+\frac{3}{2})(s+\frac{1}{4})^2} = 0$

2.4.5 Conditional Stability

A conditionally stable system is one for which increasing the gain gives rise to alternation between stability and instability. This is illustrated by the following example of a characteristic equation.

$$1 + K\frac{(s+5)(s+8)}{(s+\frac{3}{2})(s+\frac{1}{4})^2} = 0$$

Figure 2-8 shows the Nyquist and Evans locus plots (the second one being the critical region of the Nyquist). From the Routh-Hurwitz criterion it may be shown that the system is stable for $0 \leq K < 0.1337$ and $K > 0.8807$. For further examples see [67].

This behaviour is analogous with that observed for delayed linear and nonlinear systems when increasing the delay gives rise to 'stability windows' or stability switching. This will be discussed in section 3.1 and chapter 4.

Chapter 3

Binary Systems

3.1 Linear Time-Delay Systems

3.1.1 Introduction

In this section will be considered the stability of a plant $G(s)$ in series with gain K and delay h to which unity negative feedback is applied. A pure time-delay of h seconds with input $x(t)$ has output $x(t-h)$, a time-shifted version of the input, and so has transfer function $\exp(-sh)$ by the delay theorem of the Laplace transform. The system is shown in figure 3-1. Here, as before, $G(s) = C(s)/A(s)$ where $C(s)$ and $A(s)$ are real polynomials in s and $G(s)$ is strictly proper, attention again being restricted to retarded systems. The characteristic equation is as follows:

$$1 + KG(s)\exp(-sh) = 0 \tag{3.1}$$

or

$$F(s,h) := A(s) + KC(s)\exp(-sh) = 0 \tag{3.2}$$

In some cases, K will be taken to be unity without loss of generality (equivalent to multiplication of the coefficients of C by a factor K).

The main new feature of time-delay systems is that there are infinitely many roots of the characteristic equation as the complex exponential is periodic. Stability methods

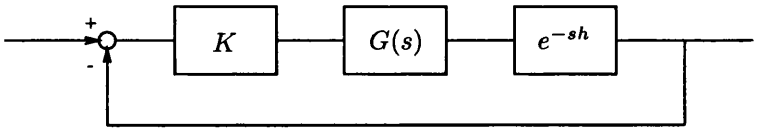


Figure 3-1: Linear Time-Delay System

such as the Routh-Hurwitz criterion are no longer valid because the characteristic equation is no longer algebraic. However the Nyquist and Mikhailov criteria and root locus method of Evans may be extended to the corresponding delayed case.

This section contains a review of stability methods, a vector formulation for stability, a technique for the temporal loci of a time-delay system, in which the delay is the free parameter (as opposed to the gain in the Evans method), a discussion of approximate methods, applications to cost functionals and phase plane representation of time-delay systems.

3.1.2 Nyquist Plots for Time-Delay Systems

The Nyquist criterion remains valid in the delayed case since the complex exponential function is analytic. The Nyquist diagram for the delay term alone is the circle of unity radius, centred on the origin, and starting at the point $(1, 0)$:

$$|\exp(-i\omega h)| = 1, \qquad \arg(\exp(-i\omega h)) = -\omega h$$

Note that there are infinitely many clockwise encirclements as ω increases from zero to infinity. The simplicity of the locus for pure time-delay is lost as soon as linear dynamics are introduced. For example, consider an integrator $1/s$ in series with delay and unity negative feedback. The number of encirclements of the $-1/K$ point by $\exp(-i\omega h)/s$ (which has modulus $1/\omega$ and argument $-\omega h - \pi/2$) is required. Figure 3-2 shows the plot for $h = 1$ (solid line for $\omega > 0$, dashed for $\omega < 0$). It may be seen that increasing the value of K or h makes the system increasingly unstable with more roots of the characteristic equation entering the right-half plane. The locus traverses the negative real axis (for the smallest value of gain) when $\omega h = \pi/2$ with gain $2h/\pi$. So $K < \pi/2h$ is required for stability. Here a very complicated locus results despite the delay and delay-free transfer functions each having simple loci. The figure also gives a further example,

$$1 + K\frac{b}{s^2 + s + b^2}\exp(-sh) = 0$$

where $b > 0$ is a parameter to be varied ($b = 1.0692$ in this case), and the delay is again unity.

3.1.3 Satché's Method

The problem of the complexity of the Nyquist locus for time-delay systems may be resolved to some extent by introduction of a dual Nyquist method, and utilising the inverse Nyquist locus. Satché [56] first introduced this technique for systems with delay.

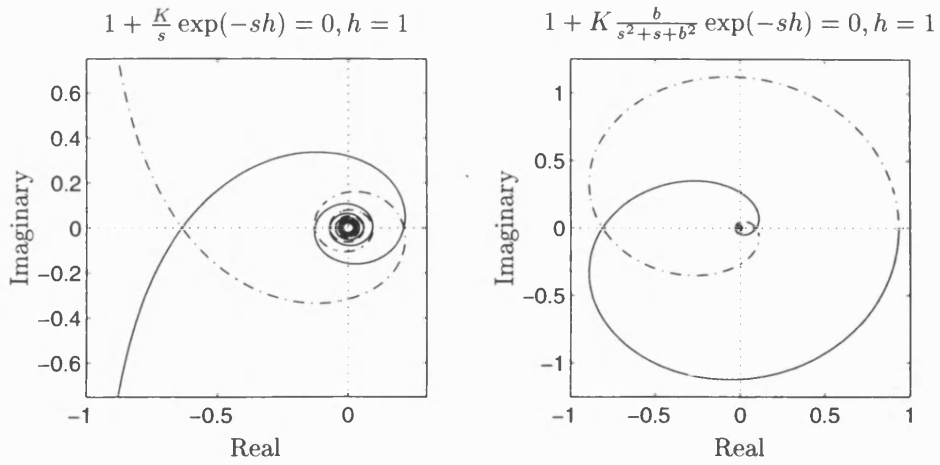


Figure 3-2: Nyquist Delay Loci

For stability of a system described by equation 3.1 the locus of $KG(s)\exp(-sh)$ must not enclose the $(-1, 0)$ point. The characteristic equation is rewritten in the form

$$-K\exp(-sh) = \frac{1}{G(s)}$$

and the two Nyquist plots for $G(s)^{-1}$ and $-K\exp(-sh)$ are made, giving two mappings of the 'D' contour of figure 2-5. That for $-K\exp(-sh)$ is simply the circle centred on the origin, radius K, starting at $(-K, 0)$ and described clockwise, with the interior of the 'D' mapping to the interior of the circle. The inverse Nyquist plot for G is often plotted easily, particularly if $C(s)$ is constant. If the curves do not intersect, there will be no value of ω for which there are purely imaginary roots of the characteristic equation, and either

- One plot completely encloses the other, in which case there will be instability for all values of delay. There are values of s in the right-half plane mapping to the interior of the circle which are also right-half plane roots of $G(s)^{-1}$.
  or
- The plots do not overlap; this implies no right hand poles and hence stability independent of delay.

Intersection of the curves requires more careful consideration, as the system may be stable or unstable. Two approaches may be taken.

- Critical values of the delay may be found by equating the values of the frequency for the two curves at each of the intersection points and distinguishing between

stabilising and destabilising values. Intersection at the same value of frequency corresponds to a closed-loop pole on the imaginary axis.

- For fixed delay, suppose that the locus of $G(i\omega)^{-1}$ lies within the circle radius K for $\omega_1 \leq \omega \leq \omega_2$. Stability results if the part of the locus for $-K\exp(-i\omega h)$ lies completely outside the area bounded by the locus for $G(i\omega)^{-1}$, *for the same range of frequency.* This may also be viewed in terms of the vector from the locus of $G(i\omega)^{-1}$ to the delay locus at the same value of ω. The behaviour of the vector as the loci pass through the intersection point is observed. The vanishing of the vector to zero corresponds to a root on the imaginary axis; if the nett rotation of the vector is 2π there would be encirclement of the $(-\frac{1}{K}, 0)$ point on the original locus and hence instability; nett rotation of zero implies stability. Further details may be found in [15] and [33].

The procedure will be made clearer by example. The same systems as for the delayed Nyquist locus are treated for comparison. In the first example, the 'D' contour maps to the same 'D'. The loci intersect on the imaginary axis at $(0, K)$, with $\omega = \frac{\pi}{2h}$ on the delay locus and $\omega = K$ on the inverse Nyquist plot for s. Hence critical stability occurs at $h = \frac{\pi}{2K}$. If $K < \frac{\pi}{2h}$ the 'D' locus will have passed out of the circle before the delay locus reaches the intersection point and the system is stable. For the second order system, with $K = b$, the inverse Nyquist plot begins at $\omega = 0$ at $(b^2, 0)$ and cuts the imaginary axis, and the circle radius b, at frequency $\omega = b$. (The semicircle of the 'D' maps to a contour at infinity to the right of the finite portion shown.) For critical stability, $|G(i\omega)| = b$ and the following polynomial results:

$$(\omega^2 - b^2)(\omega^2 - (b^2 - 1)) = 0$$

(compare with the W-polynomial discussed in section 3.1.10), giving solutions $\omega = b$ or $\sqrt{b^2 - 1}$. The corresponding values of frequency on the delay locus are $\omega = \frac{\pi}{h}(\frac{1}{2} + 2n)$ and $\frac{1}{h}(-\cos^{-1}(\frac{1}{b}) + (2n + 1)\pi)$. For the inverse Nyquist plot to escape from the circle before the delay locus reaches the first intersection point, $h < \frac{\pi}{2b}$ is required. Note that instability may follow stability; this may be checked by examining successive revolutions of the delay locus. See figure 3-3.

Advantages of Satché's method are that stability independent of delay may be seen clearly and that a complicated Nyquist plot is reduced to two simple ones. The method does not extend easily to systems with more than one delay, even commensurate ones (related by integers), since it relies on separating the delay term from everything else. However some progress on this problem has been made (this is described in section 5.1.4). Further references include [7], [15] and [33].

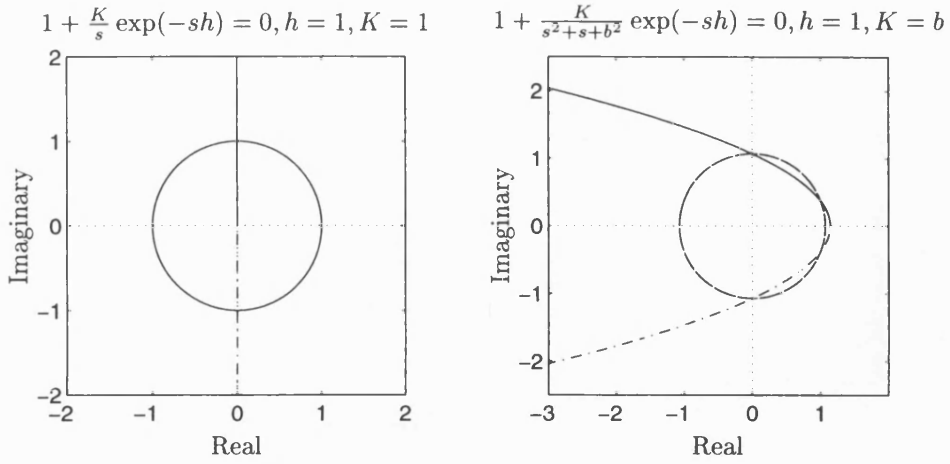


Figure 3-3: Satché Loci

3.1.4 Mikhailov's Method in the Delayed Case

The method outlined in section 2.4.4 extends to systems of delayed type with characteristic equation 3.2. If $C(s)$ is constant, the delay term may be thought of as addition of a vector modulus K to the locus for $A(i\omega)$. So if the curve for $A(i\omega)$ describes an anticlockwise path with change in argument of $\frac{n\pi}{2}$ as before, and in addition lies outside the circle radius K centre the origin, the system is stable. In this case the plot reduces to the Satché locus. For $C(s)$ a polynomial the equation may again be written in real and imaginary form

$$F(i\omega, h) = f_1(\omega, h) + if_2(\omega, h)$$

with

$$\exp(-i\omega h) = \cos(\omega h) - i\sin(\omega h)$$

and the plots for either formulation made. Detailed proofs of the validity of the criterion in the delayed case are given in [15]; see also [33] and [52]. The loci are complicated and for all but the simplest examples require computer plotting. Again, perturbation of the gain, delay or other parameters may affect the curves in a complicated manner.

3.1.5 The Phase-Angle Locus Method of Chu

The phase-angle locus method due to Chu [8], [9] analyses the stability with gain K for fixed delay of the system 3.2. References [15] and [33] also contain a discussion of this method. This is a root locus technique based on the Evans method. The characteristic equation is rewritten

$$\frac{C(s)}{A(s)}\exp(-sh) = -\frac{1}{K}$$

Equating arguments and moduli yields

$$\begin{aligned} \arg\left(\frac{C(s)}{A(s)}\right) - \omega h &= (2q+1)\pi \quad q\epsilon\mathbb{N} \\ \left|\frac{C(s)}{A(s)}\right| \exp(-\sigma h) &= \frac{1}{K} \end{aligned} \tag{3.3}$$

The phase-angle locus is a curve joining points of equal phase. Horizontal lines corresponding to $-\omega h$ for equispaced ω (h is fixed) and the angle locus of $C(s)/A(s)$ are both plotted in the complex plane. When the nett argument at the intersection of the curves is an odd multiple of π then the corresponding point lies on the root locus of the delay system. The loop gain K corresponding to each point on the locus may then be calculated from the magnitude relation. The loci are symmetrical with respect to the real axis so only the portion in the upper half plane need be drawn. Different branches of the root loci are found for different values of q in equation 3.3; there will be an infinite number of these.

As with other graphical techniques, the curves are difficult to construct and here two sets of plots must be made. In this case the complication arises from the phase-angle locus of $C(s)/A(s)$, the degree of difficulty depending on the complexity of the transfer function. One solution is to use superposition of the phase-angle loci corresponding to each pole and zero of $C(s)/A(s)$. The method is illustrated by the following example:

$$1 + \frac{K}{s(s+1)} \exp(-sh) = 0$$

Taking arguments and moduli

$$\begin{aligned} \arg\left(\frac{1}{s(s+1)}\right) - \omega h &= (2q+1)\pi \\ K &= |s(s+1)| \exp(\sigma h) \end{aligned}$$

The phase-angle loci for the delay are simply a set of straight lines parallel to the real axis. Superposition of the pole at the origin and that at $(-1,0)$ yields the angle locus for $\frac{1}{s(s+1)}$ since

$$\begin{aligned} \arg\left(\frac{1}{s(s+1)}\right) &= -\arg(s) - \arg(s+1) \\ &= -\tan^{-1}\left(\frac{\omega}{\sigma}\right) - \tan^{-1}\left(\frac{\omega}{\sigma+1}\right) \end{aligned}$$

Thus the loci for the pole at the origin are straight lines radiating from the origin, and those for the pole at $(-1,0)$ are lines radiating from the $(-1,0)$ point. Once the family

of phase-angle loci for $\frac{1}{s(s+1)}$ has been constructed, the root locus may be determined by finding the intersection points of these with the delay loci for which the nett argument is π (figure 3-4). Here the dash-dot lines correspond to the contribution of the pole at the origin and the dashed lines that of the pole at $(-1,0)$. The solid curves show the combined loci for these two poles, curves of equal phase. The loci for the delay (taken here to be unity) are dotted horizontal lines. In each case the curves are plotted in increments of $\pi/8$. Figure 3-5 shows the resultant root locus for $0 \leq K \leq 45$. (This was plotted using a high order Padé approximant; these will be discussed in section 3.1.7, and details of the program generating this plot are to be found in appendix A.2.) (The plot is of similar form to that of $\sigma = \omega \cot(\omega h)$, the Quadratrix of Hippias of Elis, 430 B.C. [30].)

It is of interest to determine points of intersection with the imaginary axis, when the system is critically stable. The pole at the origin contributes an angle $-\frac{\pi}{2}$, the pole at $(-1,0)$, an angle ω, and the delay, $-\omega h$. Thus critical stability occurs when

$$\begin{aligned} -\frac{\pi}{2} - \omega - \omega h &= (2q+1)\pi \\ \omega &= \frac{1}{1+h}\left(2q+\frac{1}{2}\right)\pi \end{aligned}$$

and the gain is given by

$$K = |s(s+1)|\exp(-sh)|_{s=i\omega} = \omega\sqrt{\omega^2+1}$$

It is clear that for $G(s)$ more complicated than considered here, the loci are very difficult to construct. This may be overcome by computer plotting. Further references include [14], [15], [33] and [66], the latter giving various construction rules for the phase-angle loci of $G(s)$ and several examples of these for low-order systems. The method may of course be implemented in the delay-free case with

$$\arg\left(\frac{C(s)}{A(s)}\right) = (2q+1)\pi$$

and

$$\left|\frac{C(s)}{A(s)}\right| = \frac{1}{K}$$

3.1.6 The Direct Method

The direct, or W-method of Walton and Marshall was introduced in [71] and is also discussed in [35]. It is an algebraic method for the stability with h of the system with characteristic equation 3.2, and finds extensions in the multiple delay case as shall be

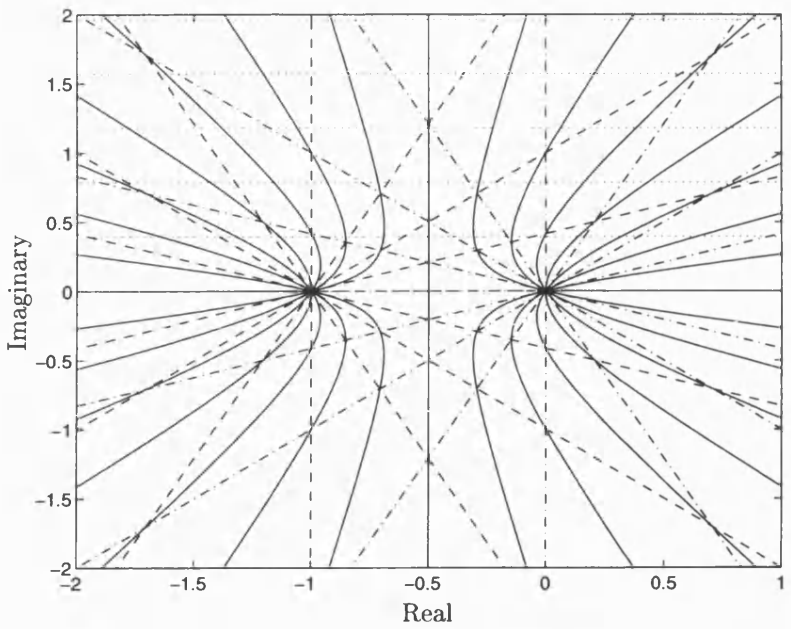


Figure 3-4: Phase-Angle Loci for $1 + \frac{K}{s(s+1)} \exp(-sh) = 0, h = 1$

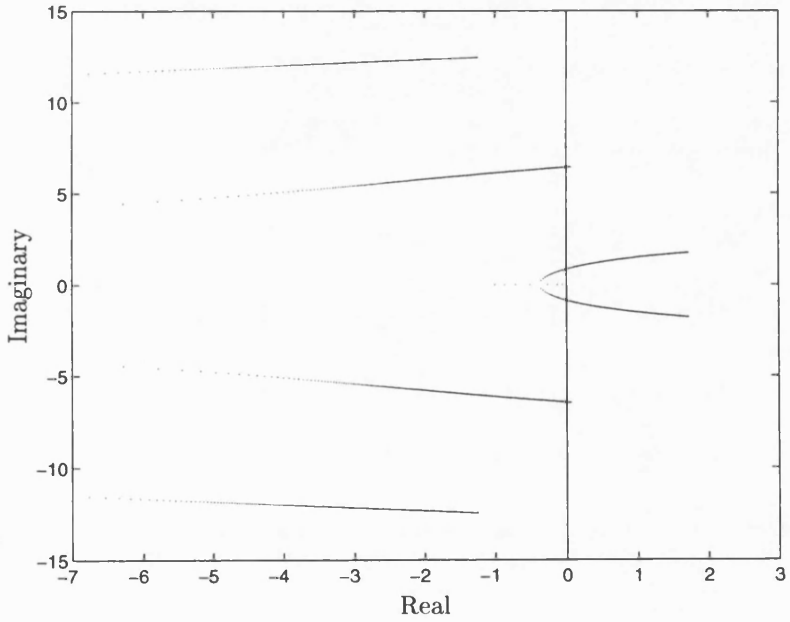


Figure 3-5: Chu Locus for $1 + \frac{K}{s(s+1)} \exp(-sh) = 0, h = 1$

seen in sections 5.1.2 and 5.1.3. Firstly the location of the roots is determined for $h = 0$. The second step is to consider h infinitesimally small and positive and the position of the infinitely many new roots. Points at which the loci cross or touch the imaginary axis are found together with the corresponding values of delay. Stabilising and destabilising values are then distinguished and ranges of delay giving rise to stability or instability are concluded. The method is now described in detail.

Step 0 $h = 0$. The equation 3.2 reduces to

$$A(s) + C(s) = 0$$

This is a delay-free stability problem and the standard techniques described in section 2.4 may be used to locate the roots. Some of these may lie in the right-half plane but this does not necessarily imply instability for all values of delay, as the delay may have a stabilising influence, as shall be seen.

Step 1 h infinitesimally small and positive. Infinitely many new roots appear due to the periodic nature of the complex exponential function. These must come in from infinity or $\exp(-sh)$ would be approximately unity and they would not be new roots. *Attention is restricted to retarded systems,* with $A(s)$ of greater degree than $C(s)$. Consequently the characteristic equation may only be satisfied for these roots if $\exp(-sh)$ is large, so s has negative real part and all the new roots lie in the left-half plane. (For advanced systems, $C(s)$ has strictly greater degree than $A(s)$, all new roots lie in the right-half plane and the system is unstable for all positive h. Neutral systems, for which the polynomials have equal degree, require more careful consideration. Even if the new roots lie in the left-half plane, it is possible for the system to be unstable. This is discussed in detail in chapter 2 of [35].)

Step 2 Crossing points on the imaginary axis (if any). Note that if $s = i\omega$ is a root of F then $s = -i\omega$ is also. Hence at a crossing point,

$$A(i\omega) + C(i\omega)\exp(-i\omega h) = 0 \tag{3.4}$$

and

$$A(-i\omega) + C(-i\omega)\exp(i\omega h) = 0$$

The exponential term is eliminated from these equations to give a polynomial in ω^2 of

degree equal to that of $A(s)$.

$$W(\omega^2) := A(i\omega)A(-i\omega) - C(i\omega)C(-i\omega) = 0 \tag{3.5}$$

Solutions for ω^2 may be found by standard means. Only non-negative solutions for ω^2 are required as ω is real; without loss of generality only solutions for non-negative ω are considered. If there are no solutions of the W-polynomial, there are no crossing points and so no change in the stability of the system as the delay is increased; the system remains stable/unstable according to whether it is stable/unstable respectively at zero delay. Also to be noted is that $W(\omega^2)$ is independent of h, which implies a *finite* number of crossing points on the imaginary axis. For the special case of $\omega = 0$ being a solution of 3.5, either $A(0) + C(0) = 0$ and $s = 0$ is a solution for all values of delay (unstable system) or $A(0) - C(0) = 0$ and $\exp(-i\omega h) = -1$ which implies infinite h. The latter will have the origin as a limit point as $h \to \infty$.

Step 3 Corresponding values of h. Any real solutions for h may be calculated from the relation

$$\exp(i\omega h) = -\frac{C(i\omega)}{A(i\omega)}$$

(Neither A nor C vanish at $s = i\omega$, as this would imply from equation 3.4 that they both vanish and it is assumed that all their common factors have been removed.) At each of the crossing points there will be an infinite number of critical delay values (multiples of $2\pi/\omega$ may be added to the smallest critical value).

Step 4 If there are critical values of h then it is necessary to distinguish between destabilising values, when the root passes from the left- to right-half plane, stabilising values, the roots passing from right to left, and a root merely touching the imaginary axis from the right or the left. At each critical value, two complex conjugate roots cross or touch and so the number of unstable roots in the right-half plane may be determined. The path of roots in the neighbourhood of the imaginary axis is given by the sign of $\mathrm{Re}(\partial s/\partial h)$:

$$\mathrm{Re}\left(\frac{\partial s}{\partial h}\right)\Big|_{s=i\omega} \begin{cases} > 0 & \text{at destabilising values (root crossing from left to right)} \\ < 0 & \text{at stabilising values (root crossing from right to left)} \\ = 0 & \text{higher derivatives must be examined.} \end{cases}$$

The following theorem shows that it is unnecessary to calculate $\mathrm{Re}(\partial s/\partial h)$ but that its sign may be inferred from the W-polynomial.

Sign Theorem (Walton and Marshall [71])

$$\operatorname{sgn}\left(\operatorname{Re}\left(\frac{\partial s}{\partial h}\right)\right) = \operatorname{sgn}\left(\frac{\partial W}{\partial \omega^2}\right)$$

where $s = i\omega$ for ω^2 a positive root of equation 3.5, and 3.2 holds.

Proof

$$\frac{\partial s}{\partial h} = -\frac{\partial F}{\partial h}\Big/\frac{\partial F}{\partial s} = \frac{sC(s)\exp(-sh)}{A'(s) + \exp(-sh)(C'(s) - hC(s))} \tag{3.6}$$

on solutions of 3.2. The prime denotes differentiation with respect to s. Simplifying,

$$\frac{\partial s}{\partial h} = \frac{-s}{h + \left(\frac{A'}{A} - \frac{C'}{C}\right)} \tag{3.7}$$

At $s = i\omega$ a critical value,

$$\begin{aligned}\frac{\partial s}{\partial h} &= \frac{-i\omega}{h + \left(\frac{A'}{A} - \frac{C'}{C}\right)\Big|_{s=i\omega}} \\ &= \frac{-\omega}{(\text{squared modulus})}\left(\operatorname{Im}\left(\frac{A'}{A} - \frac{C'}{C}\right) + i\left(h + \operatorname{Re}\left(\frac{A'}{A} - \frac{C'}{C}\right)\right)\right)\end{aligned} \tag{3.8}$$

Therefore, taking $\omega > 0$,

$$\operatorname{sgn}\left(\operatorname{Re}\left(\frac{\partial s}{\partial h}\right)\right) = \operatorname{sgn}\left(-\operatorname{Im}\left(\frac{A'}{A} - \frac{C'}{C}\right)\right)$$

Denoting $A(-s)$ by $\overline{A}$ etc.,

$$\begin{aligned}\frac{\partial W}{\partial \omega^2} &= \frac{\partial W}{\partial s}\frac{\partial s}{\partial \omega^2} \\ &= -\frac{1}{2s}\frac{d}{ds}(A\overline{A} - C\overline{C}) \\ &= -\frac{1}{2i\omega}(A'\overline{A} - A\overline{A'} - C'\overline{C} + C\overline{C'}) \\ &= \frac{i}{2\omega}A\overline{A}\left(\frac{A'\overline{A}}{A\overline{A}} - \frac{C'\overline{C}}{C\overline{C}} - \frac{A\overline{A'}}{A\overline{A}} + \frac{C\overline{C'}}{C\overline{C}}\right)\end{aligned}$$

using $A\overline{A} = C\overline{C}$. The term in brackets is a difference of conjugates, so

$$\frac{\partial W}{\partial \omega^2} = -\frac{A\overline{A}}{\omega}\operatorname{Im}\left(\frac{A'}{A} - \frac{C'}{C}\right)$$

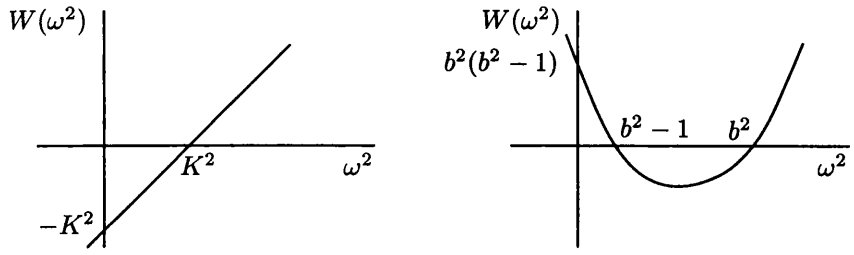


Figure 3-6: W-polynomials for $s + K\exp(-sh) = 0$ and $s^2 + s + b^2 + b\exp(-sh) = 0$

and for $\omega > 0$

$$\operatorname{sgn}\left(\frac{\partial W}{\partial \omega^2}\right) = \operatorname{sgn}\left(-\operatorname{Im}\left(\frac{A'}{A} - \frac{C'}{C}\right)\right)$$

is obtained as before. ■

Note that the sign is independent of h, so at a particular value of ω the crossing will always be in the same direction for successive critical values of h. For low order systems, $W(\omega^2)$ may be sketched easily and the sign of the gradient at its roots determined. Positive gradient corresponds to positive $\partial s/\partial h$ and a destabilising value of ω, negative gradient to a stabilising value, and if W touches the ω axis (from above), the root in the s-plane touches the imaginary axis (from the left). For large ω, $W(\omega^2) > 0$ so the highest root is destabilising followed by stabilising, etc. as ω is decreased.

Examples

1.

$$F(s,h) = s + K\exp(-sh) = 0$$

At $h = 0$, there is a single root at $s = -K$ and hence the system is stable for all positive gain. The equation is of retarded type so for small h all new roots appear in the left-half plane. $W(\omega^2) = \omega^2 - K^2$ and as can be seen from figure 3-6 the root $\omega = \pm K$ is destabilising. Corresponding values of h are given by

$$\begin{aligned} \exp(i\omega h) &= i\frac{K}{\omega} \\ h &= \frac{\pi}{2K} + \frac{2q\pi}{K} \qquad q\epsilon\mathbb{N} \end{aligned} \tag{3.9}$$

The system is stable for $0 \leq h < \frac{\pi}{2K}$ when two poles enter the right-half plane. Successive pairs enter the right-half plane at values of h given in equation 3.9.

2.

$$F(s,h) = s^2 + s + b^2 + b\exp(-sh) = 0$$

where $b > 1$ is a parameter to be varied. At $h = 0$, $s^2+s+b^2+b = 0$, a quadratic with positive coefficients and the system is stable.

$$\begin{aligned} W(\omega^2) &= (-\omega^2 + b^2 + i\omega)(-\omega^2 + b^2 - i\omega) - b^2 \\ &= (\omega^2 - (b^2-1))(\omega^2 - b^2) = 0 \end{aligned}$$

with roots $\omega_1 = \pm\sqrt{(b^2-1)}$, $\omega_2 = \pm b$. Figure 3-6 shows that ω_1 is stabilising and ω_2 destabilising. The corresponding values of h are given by

$$\exp(i\omega h) = \frac{-b}{-\omega^2 + b^2 + i\omega}$$

resulting in

$\omega_1 = \sqrt{b^2-1}$	$h_{1,q} = \frac{1}{\sqrt{b^2-1}}(-\cos^{-1}(\frac{1}{b}) + (2q+1)\pi), q\epsilon\mathbb{N}$	(stabilising)
$\omega_2 = b$	$h_{2,q} = \frac{1}{b}\left(\frac{\pi}{2} + 2q\pi\right)$	(destabilising)

The system is stable at $h = 0$ and becomes unstable at $h_{2,0}$, stable again at $h_{1,0}$ when the two right-half plane poles pass back into the left-half plane. If the values of $h_{2,q}$ and $h_{1,q}$ continue to alternate there will be repeated switching of stability. The intervals for which the system is stable are termed ***stability windows.*** However, failure to alternate and hence permanent instability will ensue since the spacing in h between destabilising values, equal to $\frac{2\pi}{b}$, is smaller than that between the stabilising values, equal to $\frac{2\pi}{\sqrt{b^2-1}}$. This will be true in general as for retarded systems the greatest critical value of ω is destabilising and the spacing is $\frac{2\pi}{\omega}$ in h. It will be shown that increasing the parameter b increases the number of stability windows.

The condition for n windows is that the 'width' of window n is negative (for the smallest value of n), i.e.

$$h_{2,n} - h_{1,n-1} < 0$$

So it is necessary to find n, the nearest higher integer, to satisfy

$$\frac{\pi}{\omega_2}\left(2n + \frac{1}{2}\right) = \frac{1}{\omega_1}\left(-\cos^{-1}\left(\frac{1}{b}\right) + (2n+1)\pi\right)$$

The number of windows is given by

$$n = \left[-\frac{1}{4} + b(b + \sqrt{b^2 - 1}) \left(\frac{3}{4} + \frac{1}{2\pi} \cos^{-1} \left(\frac{1}{b} \right) \right) \right] \tag{3.10}$$

where $[\cdot]$ denote rounding up to the nearest integer. For $b = 1$ there will be one window, for $b = \sqrt{2}$, three windows, for $b = 2$, seven windows and for $b = 20$, 793 windows.

3. Interchange of the stabilising influence of damping and the destabilising influencence of delay.

$$s^2 + a_1 s + \exp(-sh) = 0$$

This system has delay-free poles at $s = -\frac{a_1}{2} \pm \sqrt{(\frac{a_1}{2})^2 - 1}$, so is unstable for $a_1 < 0$, oscillatory for $a_1 = 0$ and stable for $a_1 > 0$. Destabilising values are given by

$$\omega = \sqrt{\tfrac{a_1}{2} + \sqrt{(\tfrac{a_1}{2})^2 + 1}}, \qquad \sin(\omega h) = a_1 \omega$$

This example was considered in [32] with particular reference to pairs of values h and a_1 corresponding to critical stability and oscillation. As the damping increases so does the range of delay for which the system remains stable. Balance between stabilising and destabilising influences is a theme developed in chapter 4.

Further examples are given in section 3.1.10, where the direct method is applied to finding the temporal loci of time-delay systems.

Advantages of the direct method are clear. The method is non-graphical and reduces an infinite problem to a finite one, that of finding the roots of the (generally simple) W-polynomial. The effect of changing delay or other parameters is evident, and stability independent of delay may be concluded after step 2. The same examples have been considered here as for the Satché and Mikhailov methods for comparison. Extensions to systems with commensurate and non-commensurate delays are possible (see sections 5.1.2 and 5.1.3).

3.1.7 The Padé Approximant

Approximation of the nonrational exponential function $\exp(-sh)$ by a rational one would enable delay-free stability techniques to be used. A truncated Taylor series for $\exp(-sh)$ gives much false information in reducing the number of roots from being infinite to finite and may introduce spurious right-half plane roots not present in the original equation (see [7] and [33]).

The Padé approximant [48] for the exponential function is a continued fraction expansion. Further references include [3] and [19]. It also involves truncating an expansion for $\exp(-sh)$, and hence reduces the problem to a finite one, yet the transfer function approximating the exponential has the same modulus. Typically, low frequencies are of interest and so an approximation that is accurate for s in the region of the origin is required. The Padé approximant for $\exp(-sh)$ is based on matching its power series expansion with a ratio of polynomials of degree m (numerator) and n (denominator) to the highest number of terms possible. The result is an (m, n) order Padé approximant denoted $P_{(m,n)}(sh)$. There will be $(m+n+1)$ free coefficients to be chosen so the Padé approximant will match the power series for the exponential function to at least $(m + n + 1)$ terms, i.e.

$$\exp(-sh) = P_{(m,n)}(sh) + \mathcal{O}((sh)^{m+n+1})$$

A table of these polynomials for various m, n is a Padé table. For reasons that shall become apparent later, attention is restricted to diagonal entries, that is approximants for which $m = n$. For example, to find $P_{(1,1)}(sh)$ let

$$\exp(-sh) = \frac{1 + \alpha(sh)}{\beta + \gamma(sh)} + \mathcal{O}((sh)^3)$$

So

$$(\beta + \gamma(sh)) \left(1 - sh + \frac{(sh)^2}{2!} - \frac{(sh)^3}{3!} + \cdots\right) = 1 + \alpha(sh) + \mathcal{O}((sh)^3)$$

and $\alpha = -\frac{1}{2}$, $\beta = 1$, $\gamma = \frac{1}{2}$.

$$P_{(1,1)}(sh) = \frac{1 - \frac{sh}{2}}{1 + \frac{sh}{2}}$$

Similarly

$$\begin{aligned} P_{(2,2)}(sh) &= \frac{1 - \frac{sh}{2} + \frac{(sh)^2}{12}}{1 + \frac{sh}{2} + \frac{(sh)^2}{12}} \\ P_{(3,3)}(sh) &= \frac{1 - \frac{sh}{2} + \frac{(sh)^2}{10} - \frac{(sh)^3}{120}}{1 + \frac{sh}{2} + \frac{(sh)^2}{10} + \frac{(sh)^3}{120}} \end{aligned}$$

It may be shown that in general

$$\exp(-sh) = P_{(n,n)}(sh) + \mathcal{O}((sh)^{2n+1})$$

where

$$P_{(n,n)}(sh) = \frac{P_n(-sh)}{P_n(sh)}; \qquad P_n(sh) = \sum_{j=0}^{n} \frac{n!(2n-j)!}{(2n)!(n-j)!}(sh)^j$$

so that

$$P_{(n,n)}(-sh) = \frac{1}{P_{(n,n)}(sh)} \tag{3.11}$$

(see [3]). In addition it may be shown that the denominator polynomials of the diagonal entries for the exponential have all roots lying in the left-half plane and are hence stable. An important property shaped by the exponential function and shared by the diagonal Padé approximants leads from equation 3.11.

$$(P_{(n,n)}(i\omega h)P_{(n,n)}(-i\omega h))^{\frac{1}{2}} = |P_{(n,n)}(i\omega h)| = 1 = |\exp(-i\omega h)|$$

and replacing the exponential by the Padé approximant in the characteristic equation 3.2,

$$A(s) + C(s)P_{(n,n)}(sh) = 0$$

Also

$$A(-s) + C(-s)P_{(n,n)}(-sh) = 0$$

and the same W-polynomial, hence the same critical values of ω, result when the exponential is approximated by the diagonal Padé approximant. The corresponding values of h will however be different. Any function $Q(s)$ such that $Q(s)Q(-s) = 1$ will have this property and such functions are termed *all-pass*. Clearly this includes $\exp(-sh)$, 1 and all systems with transfer function of form

$$Q(s) = \frac{E(s) - O(s)}{E(s) + O(s)}$$

where E, O are even, odd polynomials in s respectively. Stability results and cost functional analysis for all-pass systems are discussed further in [25] and [35]. These also contain comparisons between a delayed system and the corresponding system with a Padé approximant delay model.

Padé approximants as approximations to the exponential function have the advantage of stable transfer functions and the diagonal approximants the all-pass property. There will usually be agreement in stability properties for small delay; this may be improved by increasing the order of the Padé approximant. A comparison of stability for the systems $s + \exp(-sh) = 0$ and $s + P_{(n,n)}(sh) = 0$ for various n is given in the examples of section 3.1.10.

Minmax Error - Modification of the Padé Approximant

The diagonal Padé approximant, $P_{n,n}(sh)$, is a good approximation to the exponential function for low values of ωh, after which it diverges substantially away. An attempt is made to modify the Padé polynomials in such a way that the error is 'evened out' over the range of ω to be considered, using the minmax error criterion: the error e_2 in the argument between $\exp(-sh)$ and taking an nth order (all-pass) approximation

$$\exp(-sh) \approx \frac{1 - a_1 sh + a_2(sh)^2 - \cdots + (-1)^n a_n (sh)^n}{1 + a_1 sh + a_2(sh)^2 + \cdots + a_n(sh)^n} =: Q(sh) \tag{3.12}$$

has the moduli of the maximum and minimum values equal. (Compare with Chebyshev polynomials.) This may be achieved by suitable choice of the polynomial coefficients $a_i, i = 1, \ldots n$. So, taking arguments of both sides of (3.12) at $s = i\omega$,

$$e_2(\omega h) = \omega h - \arg Q(sh)\,|_{s=i\omega} \tag{3.13}$$

and the a_i are found by requiring that

$$\left| \max_{0 \le \omega \le \omega_m} e_2(\omega) \right| = \left| \min_{0 \le \omega \le \omega_m} e_2(\omega) \right|$$

where $[0, \omega_m]$ is the range of ω under consideration. Initially the delay h was taken to be unity and the case $n = 2$ was considered, with (3.13) reducing to

$$e_2(\omega) = \omega - 2\tan^{-1}\left(\frac{a_1 \omega}{1 - a_2 \omega^2}\right)$$

It may be seen under comparison with the corresponding Padé approximant case that the error has been smoothed out over the range of ω rather than starting small and increasing (as with the Padé approximant).

3.1.8 A Vector Approach

Consider the linear delay-differential equation

$$\sum_{k=0}^{n} a_k x^{(k)}(t) + \sum_{k=0}^{m} c_k x^{(k)}(t - h) = g(t) \tag{3.14}$$

Critical stability occurs when $x(t) = a\sin(\omega t)$ or equivalently when $s = \pm i\omega$ are roots of the characteristic equation, found from the Laplace transform of equation 3.14:

$$\sum_{k=0}^{n} a_k s^k + \exp(-sh) \sum_{k=0}^{m} c_k s^k = 0$$

Equation 3.14 may be viewed in terms of mutually perpendicular $\sin(\omega t)$ and $\cos(\omega t)$ components, and vectors corresponding to each term drawn in $(\sin(\omega t), \cos(\omega t))$ space. The resultant vector being zero corresponds to critical stability; $x(t) = a\sin(\omega t)$ and the equation 3.14 is satisfied. $\dot{x}(t)$ corresponds to a vector in the $\cos(\omega t)$ direction etc. and the delay terms a rotation in argument of $-\omega h$ radians (anticlockwise).

That resolving in the $\sin(\omega t)$, or $\cos(\omega t)$ direction is equivalent to equating the real, or imaginary part of the characteristic equation to zero at $s = i\omega$ is clear.

$$\begin{aligned}
x^{4k}(t) &= a\omega^{4k}\sin(\omega t) & (i\omega)^{4k} &= \omega^{4k} \\
x^{4k+1}(t) &= a\omega^{4k+1}\cos(\omega t) & (i\omega)^{4k+1} &= i\omega^{4k+1} \\
x^{4k+2}(t) &= -a\omega^{4k+2}\sin(\omega t) & (i\omega)^{4k+2} &= -\omega^{4k+2} \\
x^{4k+3}(t) &= -a\omega^{4k+3}\cos(\omega t) & (i\omega)^{4k+3} &= -i\omega^{4k+3}
\end{aligned}$$

Similarly for the delayed terms. This is related to Mikhailov's method in the critically stable case, when the plot of $F(i\omega)$ passes through the origin. Eliminating h from the two equations will give rise to the W-polynomial of the direct method.

This method quickly provides critical values of ω and h and no vector balance possible implies stability independent of delay. However destabilising and stabilising values are not easily distinguished nor the number of stability windows. The method does find extension to averaging methods in nonlinear systems as shall be seen.

Examples

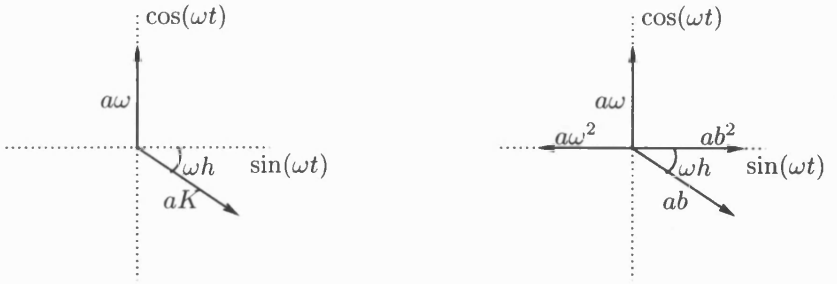
1. Integrator and delay.

$$\dot{x}(t) + Kx(t-h) = 0$$

In the transform domain this becomes

$$X(s) = \frac{x(0)}{s + K\exp(-sh)}$$

The characteristic equation of this system was analysed by the direct method in section 3.1.6 resulting in $W(\omega^2) = \omega^2 - K^2$ and instability at $\omega = K$, $h = \frac{\pi}{2K} + \frac{2q\pi}{K}$. Figure 3-7 shows the vectors in $(\sin(\omega t), \cos(\omega t))$ space. Resolving, the system will be critically stable at $\omega h = \frac{\pi}{2} + 2q\pi$, $\omega = K$ as before.



$\dot{x}(t) + Kx(t-h) = 0$ $\qquad$ $\ddot{x}(t) + \dot{x}(t) + b^2x(t) + bx(t-h) = 0$

Figure 3-7: Vector Diagrams

2.

$$\ddot{x}(t) + \dot{x}(t) + b^2x(t) + bx(t-h) = 0$$

In the transform domain

$$X(s) = \frac{\dot{x}(0) + (s+1)x(0)}{s^2 + s + b^2 + b\exp(-sh)}$$

Resolving the vectors in figure 3-7,

$$\omega^2 = b^2 + b\cos(\omega h) \qquad \omega = b\sin(\omega h)$$

Eliminating h results in the same W-polynomial as before (the second example of section 3.1.6) and hence the same critical values of ω and h.

3.1.9 The Method of Steps

The method of steps is a method in the time-domain for the solution of a delay-differential equation (see for example [35]). The error at time t is expressed in terms of that during the previous intervals of duration h, where h is the delay in the system. An analogous method considers steps in the Laplace transform domain. The following example illustrates the procedure.

Consider an integrator with zero initial condition in series with delay h, to which unity negative feedback and input $H(t)$ are applied (figure 3-8). Denoting the error by $e(t)$,

$$\dot{y}(t) = H(t) - y(t-h), \qquad e(t) = H(t) - \int_0^t e(t)\,dt\Big|_{t-h}$$

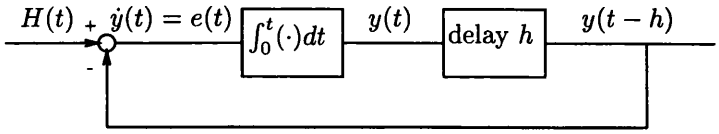


Figure 3-8: Integrator and Delay

The output is zero for the first h seconds, so for

$$\begin{aligned}
t < h \qquad & \dot{y}(t) = H(t) \\
& y(t) = tH(t) \\
h \leq t < 2h \qquad & \dot{y}(t) = H(t) - (t-h)H(t-h) \\
& y(t) = tH(t) - \tfrac{(t-h)^2}{2!}H(t-h) \\
2h \leq t < 3h \qquad & \dot{y}(t) = H(t) - (t-h)H(t-h) + \tfrac{(t-2h)^2}{2!}H(t-2h) \\
& y(t) = tH(t) - \tfrac{(t-h)^2}{2!}H(t-h) + \tfrac{(t-2h)^3}{3!}H(t-2h)
\end{aligned}$$

This process is repeated to yield

$$e(t) = \sum_{n=0}^{\infty}(-1)^n \frac{(t-nh)^n}{n!}H(t-nh)$$

In general such simple time-domain solutions are not available.

3.1.10 The Temporal Locus

Distribution of the roots in the linear delay case is a classical problem. In this section the root locus method of Evans (for stability with gain in the delay-free system) and Chu (stability with gain with a single fixed delay) are extended to the case of the 'free' parameter being the delay itself. Again, the system in figure 3-1 is considered, with characteristic equation 3.2, and the gain taken to be unity. Firstly a method for constructing the root loci for the above system, based on a numerical solution by Newton's method, is given. Several rules for finding the prominent features of the loci directly from the characteristic equation are outlined, these features being illustrated by examples. Finally approximate solution methods using the Padé approximant are discussed. An application of temporal loci to finite horizon cost functionals is given in appendix B.

Method for Construction

In this section is described a method for obtaining the root locus for the time delay system of figure 3-1. First stability for $h = 0$ is examined using standard methods

(for example Routh-Hurwitz) and the closed loop delay-free poles are found. Any part of the locus emanating from the delay-free poles will be termed the primary locus. Secondly the direct method is employed to find stabilising and destabilising values of ω, corresponding values of h and stability windows. This information will subsequently be used to provide starting points for iteration using Newton's method for systems. It is possible to use the delay as a parameter, incrementing h slightly at each step and solving by Newton's method, using the previous root as a starting 'guess'. This has the disadvantage of having step size between calculated locus points fixed by h. The sensitivity of the locus with respect to h, $\partial s/\partial h$, will vary greatly (see rule 7 in the following section) so that the step size will need to be small when the sensitivity is great, yet larger when the sensitivity is small. One way of overcoming this is to use an Euler-Newton Continuation algorithm in which an extra parameter is introduced, that of the arc length covered along the locus in (σ, ω, h) space. This ensures that points are equispaced and avoids problems of convergence when $\partial s/\partial h$ is large without making the step length inordinately small. In addition, the next point on the locus is predicted, rather than simply using the previous root as a starting point for the iteration. For the primary locus, iteration begins at each closed loop delay-free pole, increasing h (and hence arc length) from zero. For other loci, the starting point is their first destabilising value and iteration is performed backwards $h \to 0$ and forwards $h \to \infty$. Of course the locus need only be found for $\omega \geq 0$ then reflected in the σ axis. Details of the algorithm and a program listing in MATLAB [37] are to be found in appendix A.1.

The Padé approximant will be a good approximation for the delay term for low values of h and exhibits the all-pass property, so that the correct results will be obtained for stabilising and destabilising values of ω (although the corresponding values of h may be inaccurate). For a low order approximant it is particularly important that the parity of the order of the exact equation for $h > 0$ and that with the Padé model should be equal, so that the model mirrors the behaviour of the exact equation. If this is not the case, there will either be a spurious root from $-\infty$ in the model or a true negative real locus will be lacking. MATLAB has the capability of dealing with high order Padé approximants and so a program has been written to utilise this, together with the direct method, to plot the temporal loci for (3.2) with the delay term replaced by a high order Padé approximant. These plots will give the prominent features of the locus quickly and are reliable for low h. They are also useful in the case of stability independent of delay; here there are no crossing points of the imaginary axis and hence no starting points available for finding secondary and other loci. The Padé approximant version provides approximate starting points for subsequent accurate solution. A similar approach has been used to plot the Chu loci for the same characteristic equation as the gain is

increased from zero. Further details are to be found in appendix A.2 and a comparison of results obtained with Padé approximants of various order and the Euler-Newton method is given in the following examples.

The report [65] gives various other approaches for obtaining the root loci of time-delay systems.

Rules for Locus Construction

1. **Behaviour at $h = 0$.** Loci start at roots of $A(s) + C(s) = 0$ (delay-free poles) and stability may be determined by classical delay-free techniques, for example Routh-Hurwitz.

2. **Symmetry.** The locus is symmetrical with respect to the real axis; if s is a root of equation 3.2 then so is its conjugate.

3. **h infinitesimally small and positive.** Since the system is retarded, infinitely many new roots appear at infinity in the left half plane, as discussed earlier. Infinitely many loci result from the presence of the complex exponential function in the characteristic equation.

4. **Points at which the loci cross the imaginary axis.** Stabilising and destabilising values of ω and corresponding values of h are found using the direct method. Each successive locus to cross at a particular ω_c is spaced in h by $2\pi/\omega_c$.

5. **Argument and modulus rule.** Writing

$$\frac{C(s)}{A(s)} = -\exp(sh)$$

and taking arguments and moduli results in the angle balance rule and modulus rule for the delayed case.

$$\sum_{i=1}^{m} \arg(s - z_i) - \sum_{i=1}^{n} \arg(s - p_i) = \omega h + (2q + 1)\pi, \quad q \epsilon \mathbb{N} \tag{3.15}$$

and

$$\exp(\sigma h) = \frac{\prod_{i=1}^{m} |s - z_i|}{\prod_{i=1}^{n} |s - p_i|}$$

where z_i, p_i are the roots of C, A respectively.

Angle of departure from the delay-free roots of $A + C$ may be found from equation 3.15 with $h = 0$.

6. **Real axis part of the locus.** As in the Evans case, a *necessary* condition that a point on the real axis forms part of the locus is that it lies to the left of an odd number of roots of A and C. For when s is real, $\omega = 0$ and for finite h the delayed angle balance law 3.15 reduces to the delay-free version. Real roots of A and C to the right of the point contribute an angle π; those to the left contribute zero; and complex conjugate pairs also contribute nothing. However, since h disappears from equation 3.15 when $\omega = 0$, it must be verified that such real points result in *positive* h using the modulus rule.

   Note that if $C(s)$ is negative constant, equation 3.15 becomes

$$-\sum_{i=1}^{n} \arg(s - p_i) = \omega h + 2q\pi$$

   and points to the *right* of an odd number of roots of A may form part of the locus. A similar argument applies when the leading term of $C(s)$ is negative.

7. **Sensitivity with respect to the delay.** From equation 3.7 it may be seen that $\partial s/\partial h$ is infinite on the remote loci where s is infinite and h is infinitesimally small. This gradient is also infinite in the neighbourhood of a multiple root when the denominator of equation 3.6 vanishes (see rule 12). $\partial s/\partial h \to 0$ when $h \to \infty$ (see rule 10).

8. **Gradients of the loci.** The gradient of the locus $\partial\sigma/\partial\omega$ may be calculated to find turning points on the curves, points at which the gradient is infinite and the gradient of the locus at $s = i\omega$ for a destabilising or stabilising value of ω and the corresponding delays. One way of achieving this is by using equation 3.7 on solutions of the characteristic equation 3.2, and

$$\frac{\partial s}{\partial h} = \frac{\partial(\sigma + i\omega)}{\partial h}$$

   so

$$\frac{\partial \omega}{\partial \sigma} = \frac{\mathrm{Im}\,(\partial s/\partial h)}{\mathrm{Re}\,(\partial s/\partial h)}$$

9. **Gradient at $s = i\omega$.** The gradient $\partial\sigma/\partial\omega$ becomes increasingly steep negatively, or positively, at successive destabilising, or stabilising values of ω respectively (for $\omega > 0$; the reverse holds for $\omega < 0$).

   Proof

$$\frac{\partial \omega}{\partial \sigma} = \frac{\mathrm{Im}\,(\partial s/\partial h)}{\mathrm{Re}\,(\partial s/\partial h)}$$

and combining with equation (3.8),

$$\frac{\partial\omega}{\partial\sigma} = \frac{h + \mathrm{Re}\left(\frac{A'}{A} - \frac{C'}{C}\right)}{\mathrm{Im}\left(\frac{A'}{A} - \frac{C'}{C}\right)} \tag{3.16}$$

Now

$$\mathrm{Re}\left(\frac{\partial s}{\partial h}\right) = \frac{-\omega \mathrm{Im}\left(\frac{A'}{A} - \frac{C'}{C}\right)}{(\text{squared modulus})} \begin{cases} < 0 \text{ for } s = i\omega \text{ at stabilising values} \\ > 0 \text{ for } s = i\omega \text{ at destabilising values} \end{cases}$$

Therefore the denominator of equation (3.16) will be positive for stabilising, and negative for destabilising, values (for $\omega > 0$). Since $\mathrm{Re}\left(\frac{A'}{A} - \frac{C'}{C}\right)$ is a function of ω and therefore has a particular (fixed) value for a given stabilising or destabilising value of ω, it may be concluded that, for $\omega > 0$, $\partial\omega/\partial\sigma$ increases for stabilising and decreases for destabilising values of h for successive critical h values. ■

Note that the first destabilising gradient value may not necessarily be negative. See for example the locus of

$$s^4 + 3s^3 + 5s^2 + 2s + 4 + (s^3 + 3s^2 + 2s + 4)\exp(-sh) = 0$$

(given in the following examples).

10. **Asymptotic behaviour as $h \to \infty$.** There are three possibilities. First solutions are sought in the left-half plane, that is for $\mathrm{Re}(s) < 0$. Then $\exp(-sh)$ is large and positive in absolute value, so for the characteristic equation (3.2) to be satisfied, $C(s) = 0$. Therefore loci will tend to the left-half plane roots of C, spiralling in due to the trigonometric functions in the complex exponential. Similarly for solutions in the right-half plane, $\mathrm{Re}(s) > 0$ so $\exp(-sh) \to 0$ in absolute value as $h \to \infty$ and $A(s) = 0$, so loci spiral to the roots of A in the right-half plane. One locus will end at each of the left-half plane roots of C and right-half plane roots of A. The third possibility is that $\mathrm{Re}(s) \to 0$, in which case $s \to 0$ as other imaginary solutions are the critical crossing points for finite h.

$$\exp(sh) \to -\frac{C(0)}{A(0)} = -\frac{c_0}{a_0}$$

(for a_0, c_0 non-zero constant terms of A and C). Roots will approach from the left-half plane if the number of stabilising and destabilising values is equal *(i)*, but from the right-half plane if the number of destabilising values is one greater *(ii)*. (For retarded systems, the highest root of $W(\omega^2)$ is destabilising.)

Asymptotic solutions as $h \to \infty$ for $a_0, c_0 \neq 0$. The characteristic equation may be approximated in the limit by

$$a_0 + c_0 \exp(-sh) = 0$$

Hence

$$\begin{aligned} \exp(\sigma h)\sin(\omega h) &= 0 \\ \exp(\sigma h)\cos(\omega h) &= -\frac{c_0}{a_0} \\ \omega h = n\pi \quad &\text{and} \quad \cos(\omega h) = \pm 1 \end{aligned}$$

If a_0, c_0 are of like sign then

$$\cos(\omega h) = -1, \quad \omega = \frac{(2n+1)\pi}{h}, \quad \sigma = \frac{1}{h}\log\left(\frac{c_0}{a_0}\right)$$

and if of different sign,

$$\cos(\omega h) = 1, \quad \omega = \frac{2n\pi}{h}, \quad \sigma = \frac{1}{h}\log\left(-\frac{c_0}{a_0}\right)$$

For a given fixed (large) value of delay h, the roots on successive loci form chains, with equal real part and with imaginary part equally spaced.

In addition roots will approach from the left-half plane, right-half plane or very near the imaginary axis according as $|c_0/a_0|$ is less than, greater than, or equal to unity respectively (corresponding to *(i)* and *(ii)* above). Now $W(0) = a_0^2 - c_0^2$ so the case $|c_0/a_0| = 1$ corresponds to $\omega = 0$ being a solution of the W-polynomial and here the imaginary part of the roots dominates the real part in the limit $h \to \infty$. For the modulus strictly greater or less than unity it is the real part which dominates.

11. **Uniqueness of the locus.** Loci do not cross at different values of h except at critical values on the $i\omega$ axis.

Proof Let

$$A(s) + C(s)\exp(-sh_1) = A(s) + C(s)\exp(-sh_2) = 0$$

The case $C(s) = 0$ corresponds to infinite h and is treated in rule 10. For $C(s) \neq 0$, dividing,

$$\exp(s(h_2 - h_1)) = 1$$

Taking real and imaginary parts and writing $s = \sigma + i\omega$,

$$\exp(\sigma(h_2 - h_1))\cos(\omega(h_2 - h_1)) = 1 \tag{3.17}$$

$$\exp(\sigma(h_2 - h_1))\sin(\omega(h_2 - h_1)) = 0 \tag{3.18}$$

Solving these equations,

$$\begin{aligned} \sin(\omega(h_2 - h_1)) &= 0 \\ \exp(\sigma(h_2 - h_1)) &= 1 \\ \omega(h_2 - h_1) &= 2n\pi \\ \sigma(h_2 - h_1) &= 0 \\ h_2 = h_1 \quad &\text{or} \quad \sigma = 0 \end{aligned}$$

It is known that the crossing points of $\sigma = 0$ are stabilising or destabilising values or the origin (as $h \to \infty$; see rule 10) so that loci may only cross (for different values of h) at critical values of ω. ■

(Note this proof also gives $h_2 - h_1 = 2n\pi/\omega$ at a critical value agreeing with the direct method.) This does not discount the presence of multiple roots.

12. **Multiple roots** (loci meeting at same value of h). These may occur at any position in the complex plane. For n roots,

$$\frac{\partial^j F(s,h)}{\partial s^j} = 0, \qquad j = 1, \ldots, n-1$$

and these equations may be solved simultaneously with the characteristic equation. Many examples have been studied with double roots on the positive and negative real axis and also double complex roots at $h = 0$ (these examples follow).

13. **Stability independent of delay.** If a system is stable for $h = 0$ and there are no real positive roots ω^2 of the polynomial $W(\omega^2)$, then there can be no crossing by roots of the $i\omega$ axis and 'stability independent of delay' holds. However since successive remote loci pack into the region bounded by the primary locus and the imaginary axis, in reality a small perturbation of system parameters may result in the system becoming unstable for a given value of h. Instead of immediately concluding stability independent of delay if a system is stable for $h = 0$ and has no destabilising values of the W-polynomial, it would be sensible to find values of h such that the loci are confined to the left-half plane and may not enter sectors such as those shown in figure 3-9, bounded by an angle α to the imaginary axis.

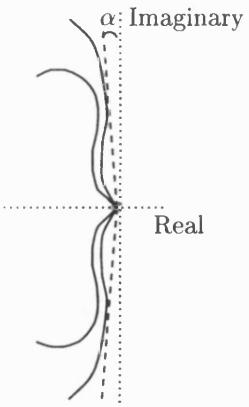


Figure 3-9: Stability Independent of Delay?

Examples

The technique described above has been employed for a variety of examples which illustrate various interesting features of the temporal loci. Roots were found using a MATLAB program and also plotted with MATLAB. Details of the programs are given in appendix A. In each of the plots using Euler-Newton continuation, the primary locus is shown by a solid line, the secondary, dashed and the tertiary, dash-dotted. The Padé approximant loci are plotted as dots equispaced in h.

First Order Examples

1. Integrator in series with delay.

$$s + K\exp(-sh) = 0 \tag{3.19}$$

This example was shown (section 3.1.6 example 1) to be stable at $h = 0$ and to have a destabilising value of ω at

$$\omega = K, \quad h_q = \tfrac{\pi}{2K} + \tfrac{2q\pi}{K}$$

Some of the properties of the loci will be explored here.

There is a root of A at the origin, hence the negative real axis lies to the left of an odd number of roots of A and C. The modulus rule for real s reduces to

$$\exp(\sigma h) = \frac{K}{|\sigma|}$$

resulting in positive h, and real solutions for s, for $\sigma \leq -K$.

To evaluate the gradient at a point (σ, ω), use equation 3.7:

$$\frac{\partial s}{\partial h} = \frac{-s}{h + 1/s} = -\frac{\sigma^2 - \omega^2 + 2i\sigma\omega}{1 + h\sigma + ih\omega}$$

which gives

$$\frac{\partial \omega}{\partial \sigma} = \frac{\omega(h\omega^2 + h\sigma^2 + 2\sigma)}{h\sigma^3 + \sigma^2 + h\sigma\omega^2 - \omega^2}$$

Even this low order system has a complicated gradient expression, however this reduces on the imaginary axis to

$$\frac{\partial \omega}{\partial \sigma} = \frac{h\omega^3}{-\omega^2} = -hK = -\left(\frac{\pi}{2} + 2q\pi\right)$$

Successive loci approach destabilising values of ω increasingly steeply.

Multiple roots will be found when

$$s + K\exp(-sh) = 0 \quad \text{and} \quad 1 - hK\exp(-sh) = 0$$

This has (real) solution $s = -1/h$, $s = -1/(Ke)$ and $h = Ke$. It is a double root as $\partial^2 F(s,h)/\partial s^2|_{s=-1/(Ke)} \neq 0$. A locus coming in from $-\infty$ will meet the primary locus from $s = -K$ at the double root and depart at an angle $\pi/2$. The pair of real roots for $0 < h < Ke$ may be found as follows. Let

$$\sigma_1 + K\exp(-\sigma_1 h) = 0 = \sigma_2 + K\exp(-\sigma_2 h)$$

and $r = \sigma_1/\sigma_2$. It may be shown that σ_1 and σ_2 are given by

$$\sigma_1 = -Kr^{\frac{r}{r-1}}, \quad \sigma_2 = -Kr^{\frac{1}{r-1}}$$

letting r range over the interval $(0, 1)$. The delay is given by

$$h = -\frac{1}{\sigma_1}\ln\left(-\frac{\sigma_1}{K}\right) = -\frac{1}{\sigma_2}\ln\left(-\frac{\sigma_2}{K}\right)$$

Successive remote loci cross the imaginary axis at $\omega = 1$. There are no left-hand plane roots of C or right-hand plane roots of A so as $h \to \infty$ all loci tend to the origin from the right-half plane. The locus for $K = 1$ is shown in figure 3-10, with a three-dimensional plot in (σ, ω, h) space given in figure 3-11. (The critical values are marked.)

2. Double real root in both left- and right-half planes.

$$s - 1 + 2\exp(-sh) = 0 \tag{3.20}$$

At $h = 0$, the system is stable with roots $s = -1$. The direct method yields one destabilising value of ω at $\omega = \sqrt{3}$, $h = \frac{\pi}{\sqrt{3}}\left(\frac{1}{3} + 2q\right)$. The gradients of the loci at this value are given by

$$\frac{\partial\omega}{\partial\sigma} = \frac{h + \mathrm{Re}\left(\frac{1}{i\omega}\right)}{\mathrm{Im}\left(\frac{1}{i\omega}\right)} = -\omega h = -\left(\frac{\pi}{3} + 2q\pi\right)$$

The modulus and argument rules together give $\sigma \leq -1$ and $0 \leq \sigma \leq 1$ as solutions. The following pair of double real roots may be found by differentiating the characteristic equation with respect to s: $\sigma = -3.3111$, $h = 0.2320$; $\sigma = 0.6266$, $h = 2.6783$. The primary locus from $s = -1$ joins a locus from $-\infty$ at the first double root, bifurcates at an angle $\frac{\pi}{2}$ to the real axis, meets the imaginary axis at $\omega = \sqrt{3}$ and forms the second double root. One branch tends to 1 and the other, together with all remote loci, tends to the origin as $h \to \infty$.

3. Example arising from a population model.

$$s - \alpha + (1 + \alpha)\exp(-K^2T)\exp(-sT) + K^2 = 0 \quad \alpha, K \epsilon \mathbb{R} \tag{3.21}$$

In the paper [5] an integro-differential reaction-diffusion population model is described, and pages 1682-1684 discuss instability in the model induced by delay. Writing $s_1 = s + K^2$, the equation 3.21 may be transformed to the usual form $A(s) + C(s)\exp(-sh) = 0$:

$$s_1 - \alpha + (1 + \alpha)\exp(-s_1T) = 0 \tag{3.22}$$

The locus for the roots of 3.21 is found by simply translating the plot for 3.22 to the left by K^2 units (hence increasing the magnitude of K has a stabilising effect). It may be seen that 3.22 is a generalisation of the first and second examples 3.19 and 3.20. Firstly the equation is stable at $T = 0$ with single real root $s_1 = -1$. $W(\omega^2) = \omega^2 - 1 - 2\alpha$ giving one destabilising value of ω for $\alpha > -\frac{1}{2}$: $\omega = \sqrt{1 + 2\alpha}$, $T = \frac{1}{\sqrt{1+2\alpha}}\left(\tan^{-1}\frac{1+2\alpha}{\alpha} + 2q\pi\right)$; for $\alpha \leq -\frac{1}{2}$ there are no critical values and stability independent of delay may be concluded. As $h \to \infty$, roots all tend to the origin unless $\alpha > 0$ in which case one root terminates at $s = \alpha$. In summary the following α-dependent behaviour is then expected.

(a) $\alpha > 0$. The real axis forms the locus for $0 < \sigma < \alpha$ and $\sigma < -1$. The primary locus from -1 meets a locus from $-\infty$ at a negative real double root. There is one destabilising value of ω. A positive double real root is formed in the interval $(0, \alpha)$ and one branch tends to the origin with the other to $s = \alpha$. This is similar to the behaviour of equation 3.20.

(b) $-1/2 < \alpha \leq 0$. This follows a similar pattern to equation 3.19, with a double negative real root and one destabilising value of ω.

(c) $\alpha = -1/2$ forms the boundary between the existence and non-existence of a positive real root of $W(\omega^2) = 0$. There is a double negative real root but no destabilising value of ω.

(d) $-1 < \alpha < -1/2$. Behaviour is as for $\alpha = -1/2$.

(e) $\alpha = -1$. The problem becomes a trivial delay-free one with single root $s = -1$.

(f) $\alpha < -1$. Here $C(s) = -1$ and points to the right of the root of A at -2 may form part of the locus. The primary locus traverses the interval $(-1, 0)$ to $(0, 0)$ and other loci move in from infinity towards the origin.

Cases (3a), (3b), (3d) and (3f) are illustrated in figure 3-12.

Second Order Examples

1. Multiple stability windows.

$$s^2 + s + b^2 + b\exp(-sh) = 0$$

The critical values of ω and corresponding values of h for this system were found in section 3.1.6 to be as follows:

$$\begin{array}{lll} \omega_1 = \sqrt{b^2-1} & h_{1,q} = \frac{1}{\sqrt{b^2-1}}(\tan^{-1}(-\sqrt{b^2-1}) + 2q\pi), q\epsilon\mathbb{N} & \text{(stabilising)} \\ \omega_2 = b & h_{2,q} = \frac{1}{b}\left(\frac{\pi}{2} + 2q\pi\right) & \text{(destabilising)} \end{array}$$

and the number of stability windows is given by equation 3.10.

A and C have no real roots therefore there is no real value of s satisfying the characteristic equation. All loci tend to the origin from the left-half plane as $h \to \infty$, there being no left-half plane roots of C or right-half plane roots of A. The gradient at $s = i\omega$ is found to be

$$\frac{\partial\omega}{\partial\sigma} = \begin{cases} \frac{2b^2-1}{\sqrt{b^2-1}} + \frac{b^2}{b^2-1}\left(\tan^{-1}\left(-\sqrt{b^2-1}\right) + 2q\pi\right) & \text{for } \omega_1 \quad \text{(increasing)} \\ -2b - \left(\frac{\pi}{2} + 2q\pi\right) & \text{for } \omega_2 \quad \text{(decreasing)} \end{cases}$$

It was of interest to investigate the limiting case between n and $n+1$ stability windows by choice of b. For n windows plus a point window,

$$h_{2,n} - h_{1,n-1} = 0$$

This was solved for various n and figures 3-13 and 3-15 show the case of one window plus point window ($n = 1, b = 1.0692$), with greater detail of the critical region in figure 3-14.

2. Stability independent of delay.

$$s^2 + s + 2 + \exp(-sh) = 0$$

In this case the polynomial $W(\omega^2) = \omega^4 - 3\omega^2 + 3$ has no positive real roots, and since the system is stable for $h = 0$ the system is stable independent of delay. All loci tend to the origin from the left-half plane as $h \to \infty$ and there are no real solutions. It may be noted that the minimum of $W(\omega^2)$ occurs at $\omega = \sqrt{\frac{3}{2}}$, which corresponds to the turning point of each successive locus, when $\frac{\partial \omega}{\partial \sigma}$ is infinite. The primary locus may be found as usual by starting at the delay-free poles of $s = -\frac{1}{2} \pm i\frac{\sqrt{11}}{2}$. Remote loci may be found by either fixing h and performing a search on an appropriate grid or by using the Padé approximant to locate all roots for a particular h. See figures 3-16 and 3-17.

3. Two negative real double roots.

$$s^2 + s + 5 + 3s\exp(-sh) = 0$$

It may be shown that: for $h = 0$ two complex conjugate loci start at $s = -2 \pm i$; there are two negative double real roots on the primary locus; the entire negative real axis forms part of the locus; there is one stabilising and one destabilising value of ω with two stability windows; all roots tend to the origin as $h \to \infty$. Figure 3-18 shows the loci plotted for $0 \leq h \leq 1.8$ using an 11th order Padé approximant.

Higher Order Examples

1. Spirals in right-half plane.

$$s^4 + 2s^3 + 3s^2 + 4s + 5 + (s^3 + 2s^2 + 3s + 4)\exp(-sh) = 0$$

This system possesses delay-free unstable poles with roots at $h = 0$ at $-1.6673 \pm 0.9330i$ and $0.1673 \pm 1.5613i$.The polynomial $W(\omega^2)$ has no real roots so the system remains unstable for all values of delay. However this example is of interest as it exhibits the property of spirals from the delay-free unstable poles to the right-half plane roots of A at $0.2878 \pm 1.4161i$. There is a real root of C at -1.6506 and a locus from $-\infty$ will terminate here. In addition C has left-half plane roots at $-0.1747 \pm 1.5469i$ and a remote locus spirals to each of these roots; this is discussed further in the next example. Figure 3-19 shows the primary locus of this system and figures 3-20 and 3-21 closer detail of the spiral.

2. Spirals from remote loci.

$$s^4 + 3s^3 + a_2 s^2 + 2s + 4 + (s^3 + 3s^2 + 2s + 4)\exp(-sh) = 0$$

for various a_2, for example:

a_2	Roots of A	Roots of C	Roots of $A + C$	$\lvert\omega_c\rvert$
5.0	2 LHP, 2 RHP	3 LHP (1 real)	2 LHP, 2 RHP	> 1
6.0	2 LHP, 2 RHP	3 LHP (1 real)	2 LHP, 2 Imaginary	$= 1$
6.5	2 LHP, 2 RHP	3 LHP (1 real)	4 LHP	< 1
7.0	4 LHP	3 LHP (1 real)	4 LHP	< 1

where ω_c is in each case the critical value. Although this system is unstable at $h = 0$ it illustrates the property that is is possible for higher order loci rather than the primary locus to be those approaching the left-half plane roots of C and the right-half plane roots of A as $h \to \infty$.

Plots are shown in figures 3-22 and 3-23 for the case $a_2 = 5.0$, with figure 3-24 showing the detail if the left-half plane spiral. The primary locus is a solid curve and successive q-loci are dashed, dash-dotted, dotted and the sequence repeats. In figure 3-25 each locus is marked with points corresponding to successive critical values of h.

$h_0 = 5.5643$	is marked by	$+$
$h_1 = 11.6488$		$*$
$h_2 = 17.7334$		o
$h_3 = 23.8180$		$\times$
$h_4 = 29.9025$		$+$

Similar features are demonstrated for a_2 taking values such as those above.

3. Triple real root.

$$s^5 + 4s^4 + 6s^3 + 6s^2 + 3s + 2 + (s^3 + 3s^2 + 3s + 1)\exp(-sh) = 0$$

There is a triple real root as $h \to \infty$ at $s = -1$. Loci approach at equispaced angles of $2\pi/3$.

4. Double complex (delay-free) roots.

$$s^4 + 4s^3 + 8s^2 + 4s + 2 + (4s + 2)\exp(-sh) = 0$$

Also two pairs of double real roots are present.

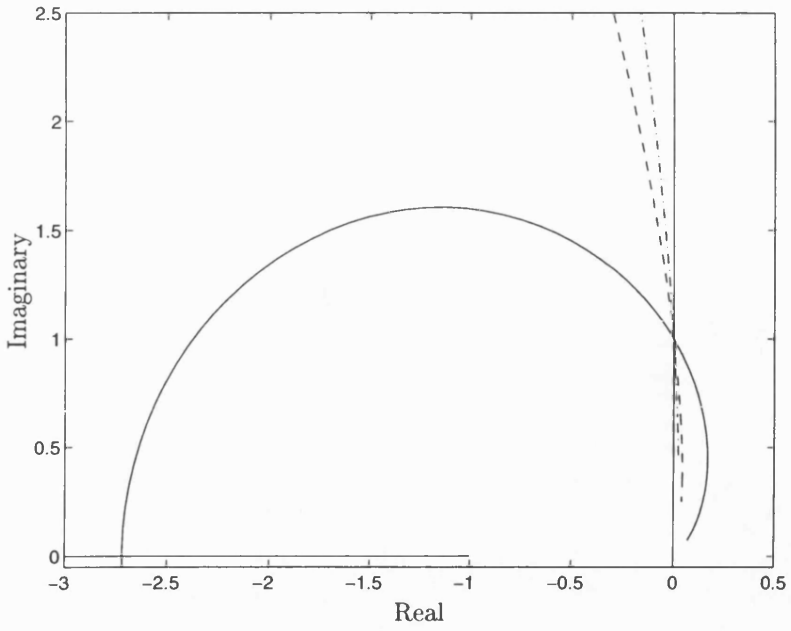


Figure 3-10: $s + \exp(-sh) = 0$

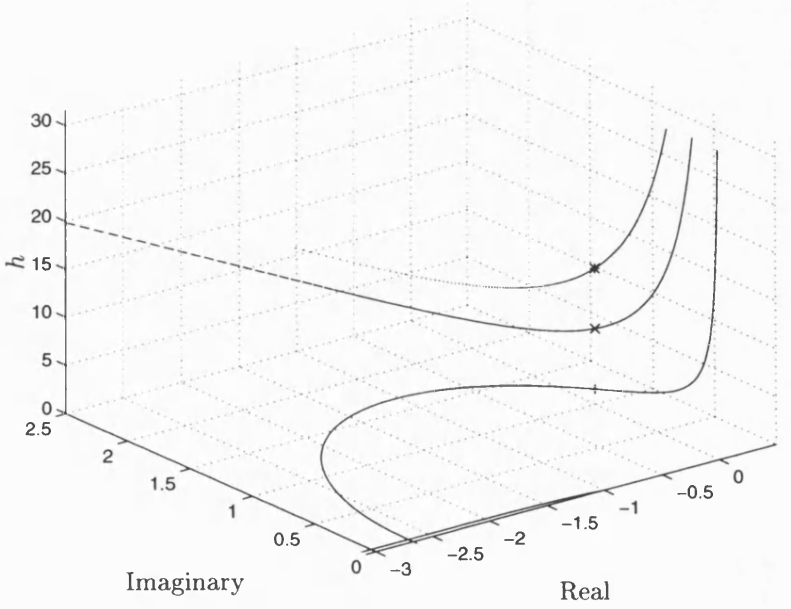


Figure 3-11: $s + \exp(-sh) = 0$

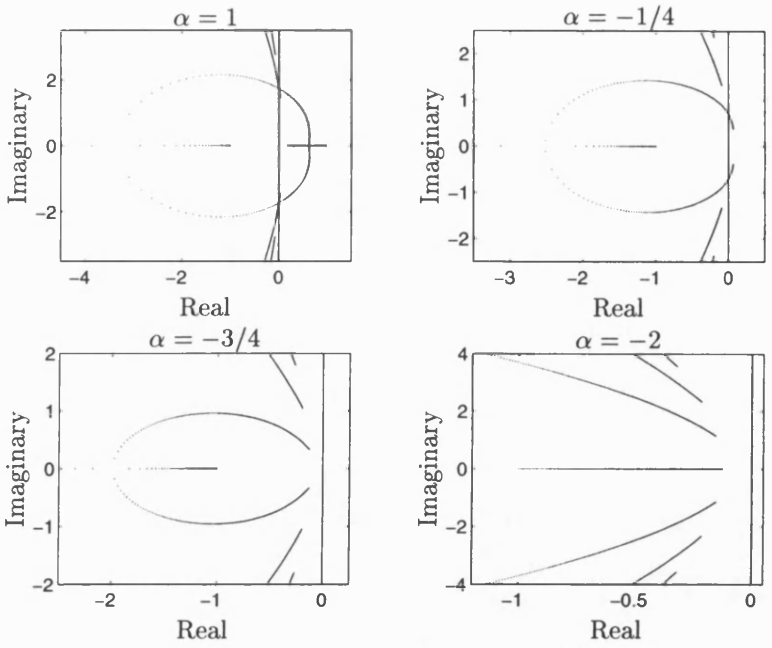


Figure 3-12: $s_1 - \alpha + (1 + \alpha)\exp(-s_1 T) = 0$

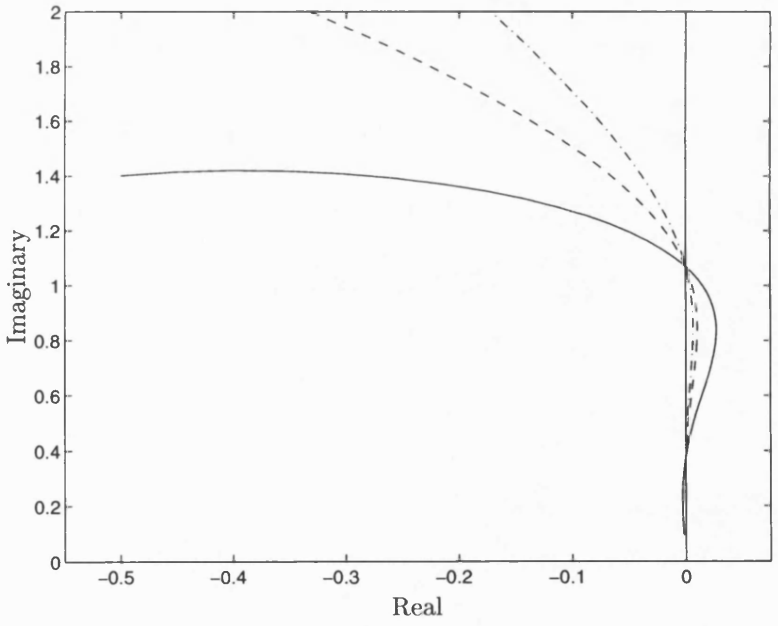


Figure 3-13: $s^2 + s + b^2 + b\exp(-sh) = 0$, $b = 1.0692$

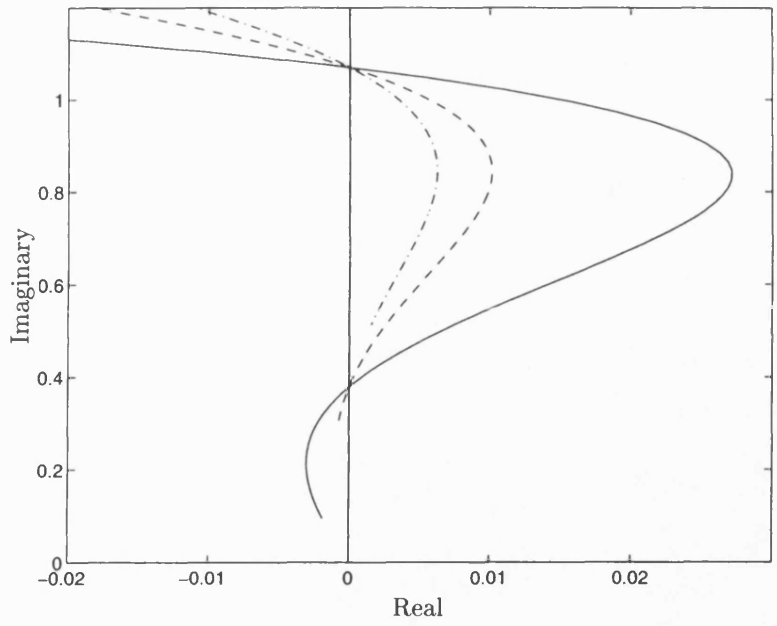


Figure 3-14: $s^2 + s + b^2 + b\exp(-sh) = 0$, $b = 1.0692$

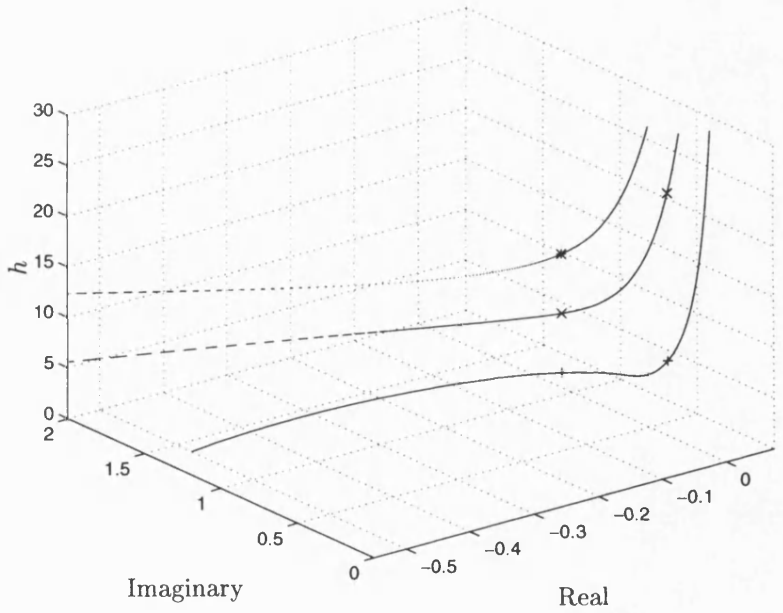


Figure 3-15: $s^2 + s + b^2 + b\exp(-sh) = 0$, $b = 1.0692$

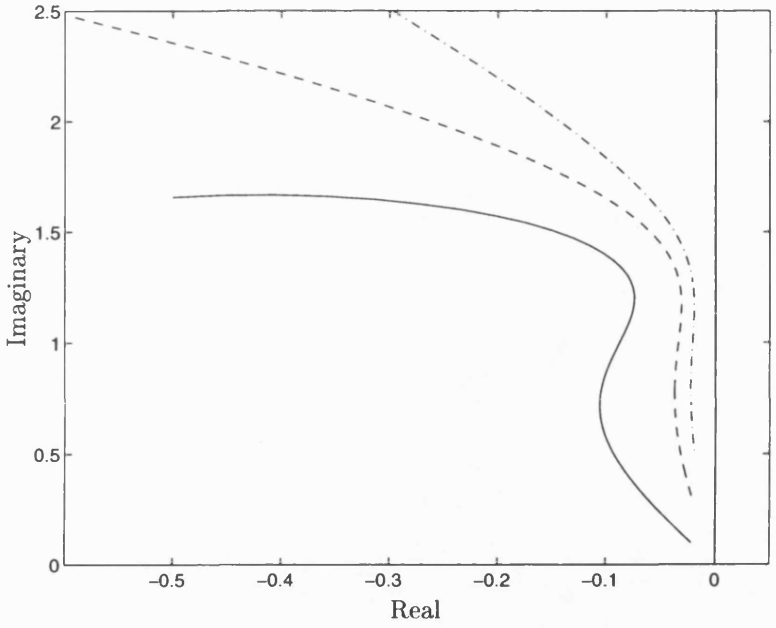


Figure 3-16: $s^2 + s + 2 + \exp(-sh) = 0$

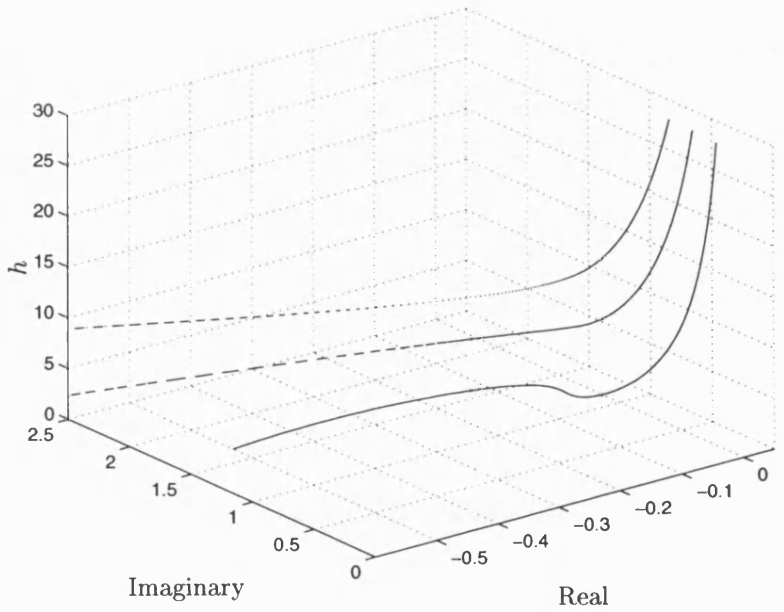


Figure 3-17: $s^2 + s + 2 + \exp(-sh) = 0$

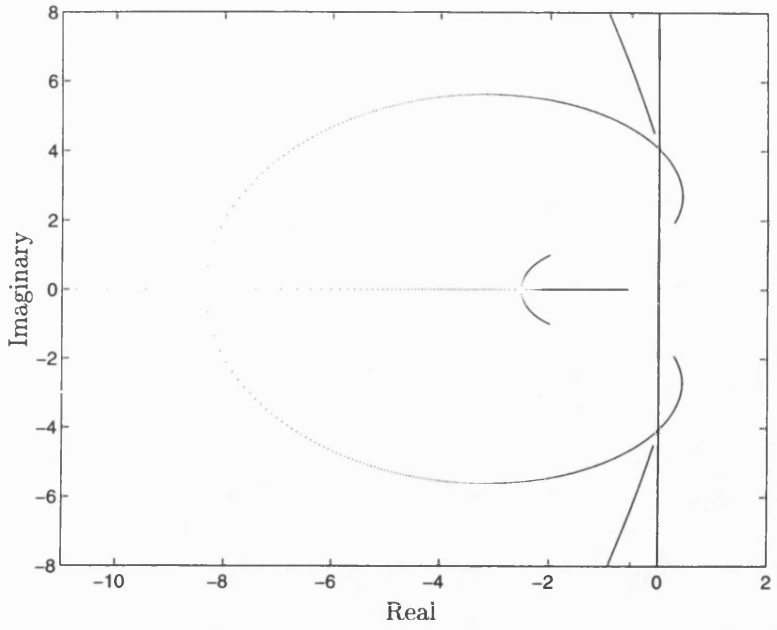


Figure 3-18: $s^2 + s + 5 + 3s\exp(-sh) = 0$

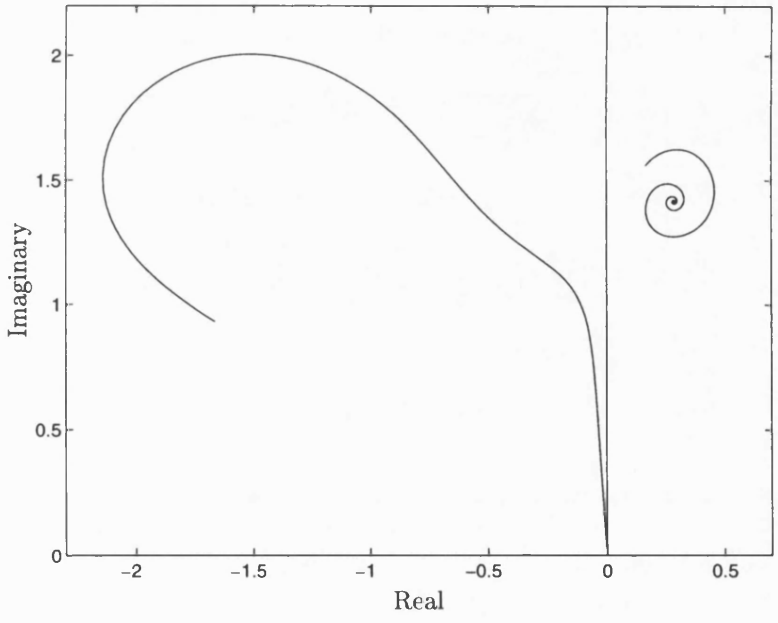


Figure 3-19: $s^4 + 2s^3 + 3s^2 + 4s + 5 + (s^3 + 2s^2 + 3s + 4)\exp(-sh) = 0$

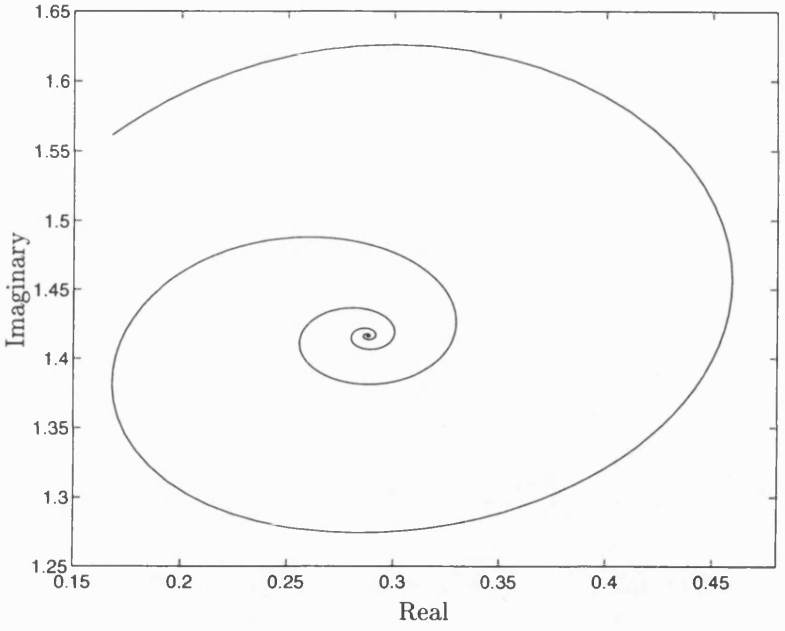


Figure 3-20: $s^4 + 2s^3 + 3s^2 + 4s + 5 + (s^3 + 2s^2 + 3s + 4)\exp(-sh) = 0$

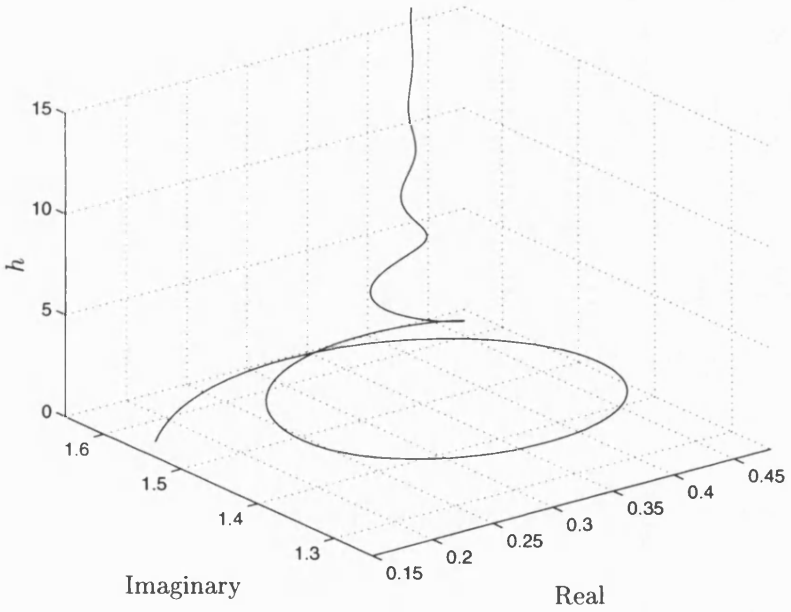


Figure 3-21: $s^4 + 2s^3 + 3s^2 + 4s + 5 + (s^3 + 2s^2 + 3s + 4)\exp(-sh) = 0$

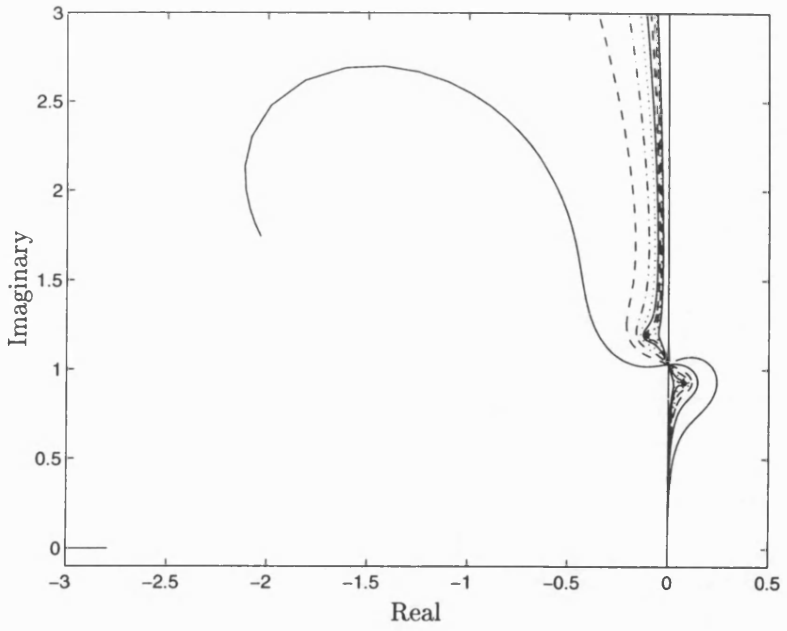


Figure 3-22: $s^4 + 3s^3 + 5s^2 + 2s + 4 + (s^3 + 3s^2 + 2s + 4)\exp(-sh) = 0$

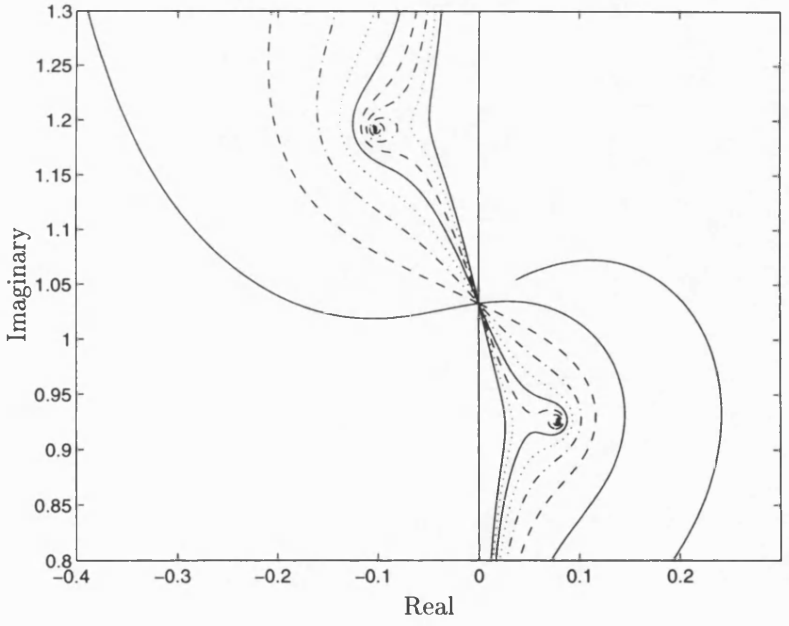


Figure 3-23: $s^4 + 3s^3 + 5s^2 + 2s + 4 + (s^3 + 3s^2 + 2s + 4)\exp(-sh) = 0$

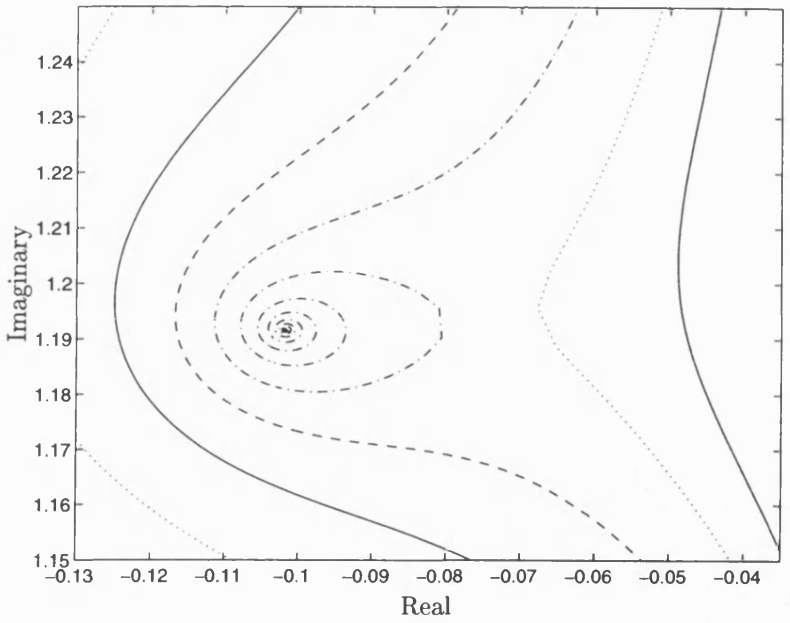


Figure 3-24: $s^4 + 3s^3 + 5s^2 + 2s + 4 + (s^3 + 3s^2 + 2s + 4)\exp(-sh) = 0$

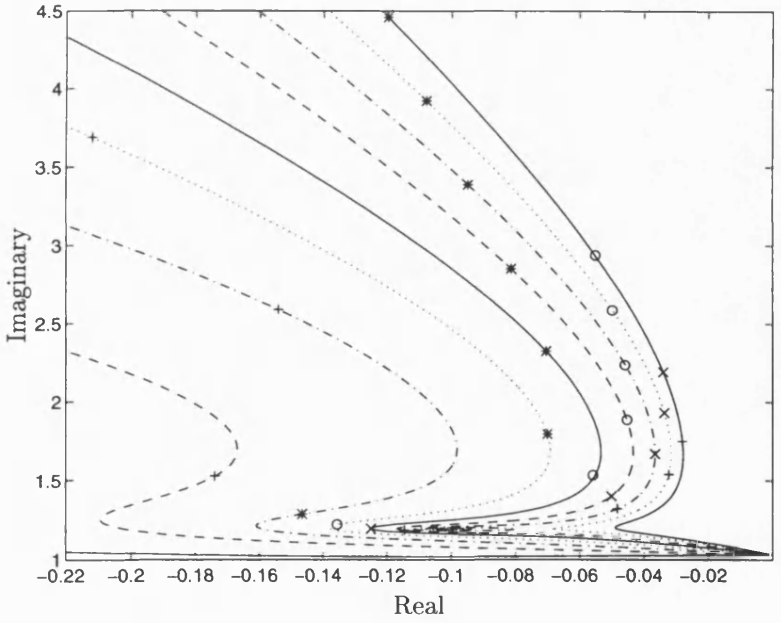


Figure 3-25: $s^4 + 3s^3 + 5s^2 + 2s + 4 + (s^3 + 3s^2 + 2s + 4)\exp(-sh) = 0$

Accuracy of the Padé Approximant Delay Model

The equation

$$s + \exp(-sh) = 0$$

has an even number of roots for $h > 0$; a locus from $-\infty$ meets the locus from $s = -1$ and other remote loci are conjugate. The Padé approximant model

$$s + P_{(n,n)}(sh) = 0 \tag{3.23}$$

for various $n \epsilon \mathbb{N}$ will be even for n odd (the order being equal to the product of the order of the polynomial A with that of the approximant, $n+1$ in this case). Figure 3-26 shows the plots of the temporal loci for $0 \leq h \leq 15$ using Padé approximants as follows. (The scale is the same as for figure 3-10 for comparison.)

- For $n = 1$ the locus meets the imaginary axis at $\omega = 1$, $h = 2$. There is a spurious positive double real root and $s = 0$ and $s = 1$ form limits as $h \to \infty$.
- For $n = 2$ the model equation possesses three roots and the spurious locus from $-\infty$ may be seen.
- $n = 5$ gives a more accurate picture. The tertiary locus should however approach the imaginary axis in the area bounded by the secondary locus and the imaginary axis.
- $n = 11$ accurately reflects the true behaviour for the low-order q-loci.

Table 3.1 charts the position of roots of the exact equation and that with Padé approximants of order $1 \leq n \leq 11$ at values of delay corresponding to critical stability. The improvement for high order approximants is clear. (The order of equation 3.23 means that there exist only a primary locus for $n = 1$ and $n = 2$, and primary and secondary loci for $n = 3$ and $n = 4$. Hence the locus points corresponding to these are not available.)

The Padé approximant model is unfavourable in that it overestimates the range in delay for which the system is stable. Stability windows results using a Padé model should be treated with caution.

	Primary Locus Roots $h = \frac{\pi}{2}$	Secondary Locus Roots $h = \frac{5\pi}{2}$	Tertiary Locus Roots $h = \frac{9\pi}{2}$
$\exp(-sh)$	$\pm\, i$	$\pm\, i$	$\pm\, i$
$P_{(1,1)}$	$-0.13661977236758 \pm 1.12007793591937i$	———	———
$P_{(2,2)}$	$-0.00344765948275 \pm 1.00523646093229i$	———	———
$P_{(3,3)}$	$-0.00006132010823 \pm 1.00009623553949i$	$-0.41487722735688 \pm 1.57662005304059i$	———
$P_{(4,4)}$	$-0.00000061476106 \pm 1.00000096565309i$	$-0.02565316718252 \pm 1.11546972816770i$	———
$P_{(5,5)}$	$-0.00000000388581 \pm 1.00000000610381i$	$-0.00288621787284 \pm 1.01963211202718i$	$-0.72296920485057 \pm 1.94447908630890i$
$P_{(6,6)}$	$-0.00000000001692 \pm 1.00000000002658i$	$-0.00034061467015 \pm 1.00260699074453i$	$-0.06480613521589 \pm 1.29023432289649i$
$P_{(7,7)}$	$-0.00000000000005 \pm 1.00000000000008i$	$-0.00003181563492 \pm 1.00024910857459i$	$-0.00977770634277 \pm 1.08205452524960i$
$P_{(8,8)}$	$-0.00000000000000 \pm 1.00000000000000i$	$-0.00000221614654 \pm 1.00001740110084i$	$-0.00185290818847 \pm 1.02152232910174i$
$P_{(9,9)}$	$0.00000000000000 \pm 1.00000000000000i$	$-0.00000011801459 \pm 1.00000092686883i$	$-0.00035171407997 \pm 1.00469856268503i$
$P_{(10,10)}$	$-0.00000000000000 \pm 1.00000000000000i$	$-0.00000000496995 \pm 1.00000003903524i$	$-0.00005781437952 \pm 1.00080754521758i$
$P_{(11,11)}$	$-0.00000000000000 \pm 1.00000000000000i$	$-0.00000000017025 \pm 1.00000000133613i$	$-0.00000769525543 \pm 1.00010858048639i$

Table 3.1: Roots of $s + \exp(-sh) = 0$ at successive critical delay values compared with Padé approximant models

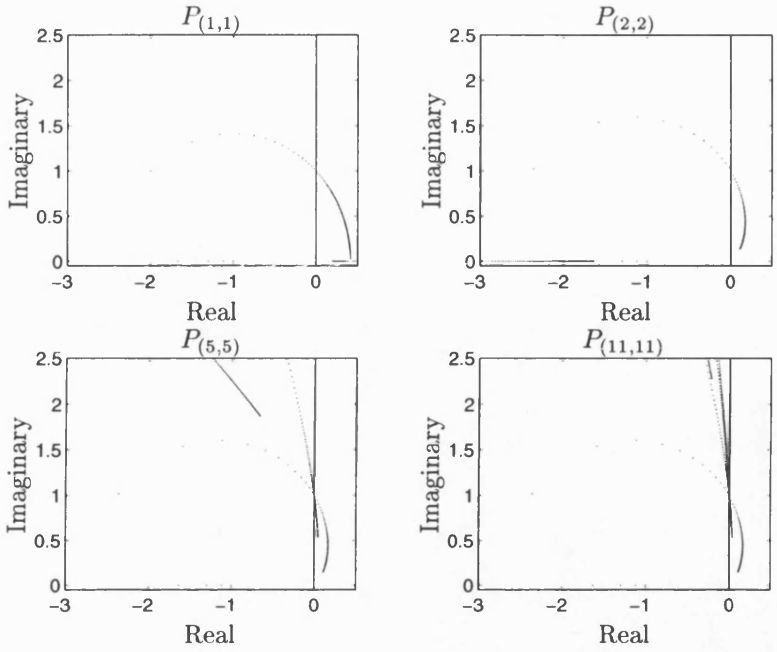


Figure 3-26: Comparison of Padé Approximants for $s + P_{(n,n)}(sh) = 0$

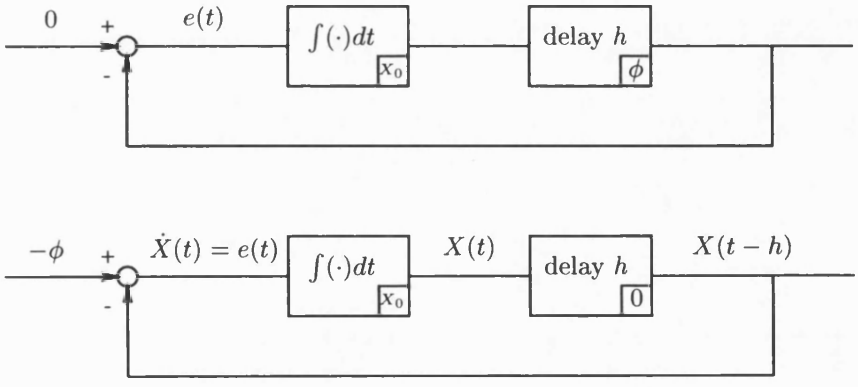


Figure 3-27: Phase Plane Representation of Time-Delay System

3.1.11 Phase Plane Representation of Time-Delay Systems

The phase plane is discussed in greater detail in section 3.2.6. Here its application to time-delay systems is discussed.

The system

$$\dot{e}(t) = -e(t-h)$$

is considered for various values of the delay h, which may be represented as in the upper diagram of figure 3-27, or equivalently as in the lower one, where ϕ is the stored function in the delay for $-h < t < 0$ which provides the input for $0 < t < h$, and X_0 is the initial condition. Due to there being a stored function in the delay, rather than just an initial condition, it may be thought not possible to represent the error $e(t)$ and $\dot{e}(t)$ of a time-delay system in phase space. However, by appropriate choice of ϕ this may be done. The aim is to choose ϕ such that $e(t)$ is an eigenfunction, i.e. related to a single real mode or complex pair of modes in phase space, so that the result is a single curve in the phase space. For a given value of h, poles of the transfer function are found, either on the primary locus or whichever is of interest. From these, $e(t)$ and $\dot{e}(t)$, and hence the phase plane trajectory in $(e, \dot{e})$ space, may be obtained. Now

$$\bar{\dot{X}} = s\bar{X} - X_0 = -\bar{\phi} - \exp(-sh)\bar{X}$$

so

$$E(s) = \frac{-s\bar{\phi} - \exp(-sh)X_0}{s + \exp(-sh)}$$

and

$$\bar{\phi}(s) = \frac{1}{s}\left(-(s + \exp(-sh))E(s) - \exp(-sh)X_0\right)$$

Also

$$\phi(t) = -e(t)(H(t) - H(t-h))$$

and X_0 is the output from the delay when $t = h$:

$$X_0 = -e(t)\,|_{t=h}$$

So the problem becomes find $E(s)$ (or $e(t)$), and hence $\phi(t)$, such that

$$\begin{aligned} \mathcal{L}\{-\phi(t)\} &= \mathcal{L}\{e(t)(H(t) - H(t-h))\} \\ &= \frac{\exp(-sh)}{s}X_0 + \left(1 + \frac{\exp(-sh)}{s}\right)E(s) \end{aligned}$$

As might be expected, for $s = \sigma + i\omega$ a root of the characteristic equation,

$$E(s) = \frac{s - \sigma}{(s - \sigma)^2 + \omega^2}$$

so

$$\begin{aligned} e(t) &= \exp(\sigma t)\cos(\omega t) \\ \dot{e}(t) &= \sigma \exp(\sigma t)\cos(\omega t) - \omega \exp(\sigma t)\sin(\omega t) \end{aligned}$$

Some examples are given for points on the primary locus of $s + \exp(-sh) = 0$:

- $h = 0$ Single real root $s = -1$ and the phase plane is a single node.
- $0 < h < \frac{1}{e}$ Pair of real roots. Stable node.
- $h = \frac{1}{e}$ Double root at $s = -e$. Here

  $$e(t) = t\exp(\sigma t)\cos(\omega t)$$

  Stable node.
- $\frac{1}{e} < h < \frac{\pi}{2}$ Pair of complex conjugate roots with negative real part. Stable spiral.
- $h = \frac{\pi}{2}$ Pair of imaginary roots $s = \pm i$. Centre.
- $h > \frac{\pi}{2}$ Pair of complex conjugate roots with positive real part. Unstable spiral.

Similar portraits apply to points on secondary and other loci.

In general, introduction of a delay h to a second order linear system yields an equation of form

$$a_2\ddot{x}(t) + a_1\dot{x}(t) + a_0x(t) + c_1\dot{x}(t-h) + c_0x(t-h) = g(t)$$

which is now infinite dimensional. The trajectory of a solution in $(x, \dot{x})$ space not only depends on its current position and velocity (i.e. on $x(t)$ and $\dot{x}(t)$) but also on that of h seconds earlier ($x(t-h)$ and $\dot{x}(t-h)$). *This means that trajectories may cross in the phase plane.* (If there is no crossing of trajectories, representation of the system by a second order model may be possible.) This behaviour occurs routinely and is exemplified in later chapters.

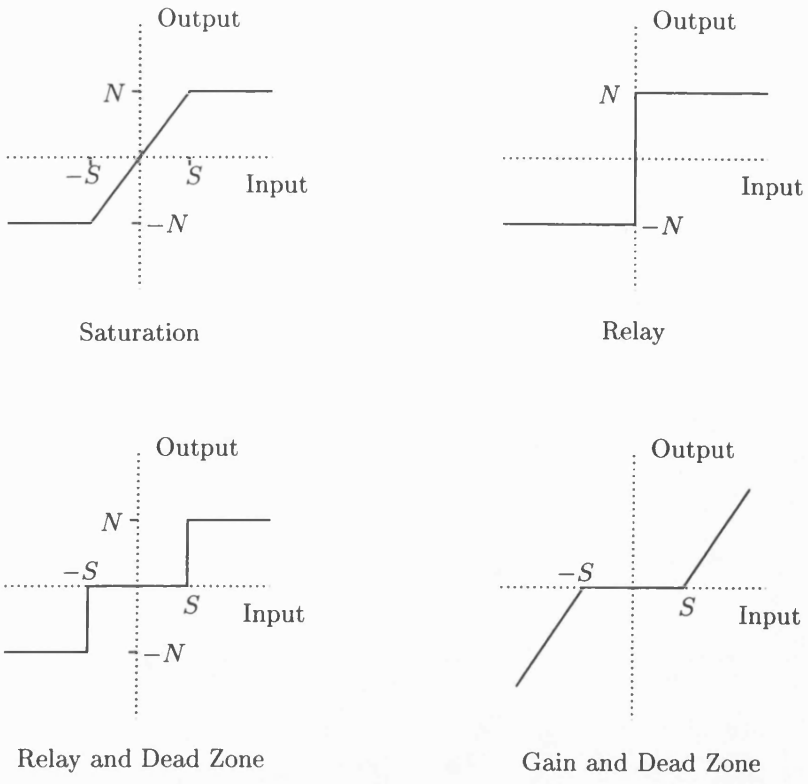


Figure 3-28: Piecewise Linear Nonlinearities

3.2 Nonlinear Delay-Free Systems

3.2.1 Introduction

In practice, all physical systems are nonlinear to some extent, whether the nonlinearity is intrinsic or has been deliberately introduced. So-called linear systems are approximations to the true behaviour, with systems only linear over a certain range of operation. There are many forms of nonlinearity. These include functions such as products of variables, powers and trigonometric functions. Such functions are analytic, possess convergent Taylor series and hence may be linearised, enabling linear control techniques to be applied. Piecewise-linear functions are not analytic at all points but behave linearly over certain intervals. The equations governing behaviour may be solved separately in each interval and joined at the boundaries. This type of nonlinearity includes saturation, dead-zone, relay etc., which are illustrated in figure 3-28. The nonlinearities described so far have been 'memoryless' in that the output depends only on the input and not on the history of the system. Multivalued 'functions' are nonlinearities which model hysteresis and backlash for example. Attention here will

concentrate on the first two of these types. In particular, for functions such as products of variables, those with frequency response of zero integral over a period will be considered. This property will be important in the Fourier analysis of sections 3.2.3 and 3.2.11. Piecewise linear functions will be monotone increasing and of odd symmetry (as in figure 3-28). 'Hump' nonlinearities are expected to give rise to chaos (see for example [39]).

A new feature of nonlinear systems is that the stability behaviour depends on the initial functions and the input; for linear systems, the behaviour may be described independently of these by the system parameters (using a transfer function for example). Superposition properties for combining the response to various inputs no longer hold. The frequency response of a linear system (the steady state on applying a sinusoidal input) will be a sinusoid of the same frequency but with different amplitude and phase. The output of nonlinear systems with sinusoidal input will usually contain components of other frequencies, or harmonics, not necessarily including the input frequency. The order of series components will be important in nonlinear systems; in the linear case they could be interchanged leaving system behaviour unaffected. Linear systems are either stable or unstable (globally), whereas a nonlinear system may be stable only for certain inputs or initial conditions, yet for others be unstable, or periodic.

In general, a nonlinear system will be described by an equation of form

$$A_n(t)y^{(n)}(t) + A_{n-1}(t)y^{(n-1)}(t) + \cdots + A_1(t)\dot{y}(t) + A_0(t)y(t) = x(t) \qquad (3.24)$$

where $x(t)$ represents the input and $y(t)$ the output, $A_0, \ldots, A_n$ are functions of time (for a linear system, they are constant) and the superscript $^{(i)}$ denotes the ith time derivative.

The following sections give an outline of various stability techniques applicable to nonlinear systems. The approach taken will depend on the nature of the nonlinearity and on other system components. There are many good references on this subject, including [1], [2], [4], [10], [57], [59] and [72].

3.2.2 Linearisation

For nonlinearities represented by analytic functions linearisation is possible. Clearly the closeness to linearity about a certain operating point will determine the accuracy of the approximation. Suppose that the coefficients of equation 3.24 are functions of the output then the following equation results

$$A_n y^{(n)}(t) + A_{n-1}y^{(n-1)}(t) + \cdots + A_1\dot{y}(t) + A_0 y(t) + \mu g(y(t), \dot{y}(t), \cdots, y^{(n-1)}(t)) = x(t)$$

Here the constant μ measures the degree of nonlinearity of the nonlinear function g. The solution $y(t)$ to this equation may be written as a power series in μ with coefficients functions of time. The leading term, independent of μ, is the solution to the corresponding linear problem, which for μ small will be a good approximation to the solution of the nonlinear problem for the purposes of controller design.

Linearisation is a technique also applicable when considering small perturbations about a reference state such as an equilibrium point. Using a state-space representation, the equation 3.24 may be written as n first degree nonlinear equations. Writing

$$x_1(t) = x(t), \quad x_2(t) = \dot{x}_1(t), \quad \text{etc.}$$

then

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{f}(\boldsymbol{x}(t), t), \quad \boldsymbol{x}(t_0) = \boldsymbol{x}_0, \quad \boldsymbol{x} \epsilon \mathbb{R}^n$$

for $\boldsymbol{f} : \mathbb{R}^{n+1} \to \mathbb{R}^n$ an analytic nonlinear function. By a simple change of coordinates if necessary, it may be assumed that the origin is an isolated equilibrium point of this equation, i.e. $\boldsymbol{f}(\mathbf{0}, 0) = \mathbf{0}$. Restricting attention to autonomous systems

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}) \tag{3.25}$$

and expanding $\boldsymbol{f}$ in a Taylor series about the origin,

$$\begin{aligned} \dot{\boldsymbol{x}} &= \boldsymbol{f}(\mathbf{0}) + J_0 \boldsymbol{x} + \mathcal{O}(\|\boldsymbol{x}\|^2) \\ &= J_0 \boldsymbol{x} \end{aligned}$$

to first order in $\boldsymbol{x}$. The problem has now been reduced to a linear one in the neighbourhood of the origin, enabling the implementation of linear techniques.

3.2.3 The Describing Function

As was remarked earlier, the frequency-response of a nonlinear system contains components at various harmonics. However, the frequency-response remains of interest as the output is usually dominated by a small number of frequencies. The describing function method is an approximation technique in which the fundamental frequency only is considered. The input-output relation, dependent on the amplitude of the input, is given by the describing function, this being the analogue of the transfer function for linear systems. Again the problem is reduced to a linear one.

In a linear system, the response to a sinusoidal input

$$x(t) = a\cos(\omega t) \tag{3.26}$$

is a sinusoid of the same frequency ω but with different amplitude and phase:

$$y(t) = \alpha\cos(\omega t) + \beta\sin(\omega t)$$

and likewise for an input $ia\sin(\omega t) = ia\cos(\omega t - \pi/2)$ the output equals

$$y(t) = i\alpha\cos(\omega t - \pi/2) + i\beta\sin(\omega t - \pi/2) = i\alpha\sin(\omega t) - i\beta\cos(\omega t)$$

By linearity, input $a\exp(-i\omega t)$ results in output equal to the difference of the two expressions above.

$$y(t) = \left(\frac{\alpha + i\beta}{a}\right) a\exp(-i\omega t)$$

Here $(\alpha+i\beta)/a$ is the complex gain factor. The describing function method uses Fourier series to give an analogue of this for nonlinear systems.

Considering nonlinearities of form

$$y(t) = f(x(t)) \tag{3.27}$$

the output on application of a sinusoidal input 3.26 is

$$y(t) = f(a\cos(\omega t)) = \frac{\alpha_0}{2} + \sum_{j=1}^{\infty}(\alpha_j\cos(j\omega t) + \beta_j\sin(j\omega t))$$

expanding as a Fourier series, where $f(a\cos(\omega t))$ is periodic with period $2\pi/\omega$, and

$$\alpha_j = \frac{\omega}{\pi}\int_{-\frac{\pi}{\omega}}^{\frac{\pi}{\omega}} f(a\cos(\omega t))\cos(j\omega t)\,dt$$

$$\beta_j = \frac{\omega}{\pi}\int_{-\frac{\pi}{\omega}}^{\frac{\pi}{\omega}} f(a\cos(\omega t))\sin(j\omega t)\,dt$$

The terms $\alpha_1\cos(\omega t)$ and $\beta_1\sin(\omega t)$ form the first harmonic with complex gain factor $(\alpha_1 + i\beta_1)/a$, with the other terms the additional harmonics introduced by the nonlinearity. (It is assumed that $\frac{\omega}{\pi}\int_{-\frac{\pi}{\omega}}^{\frac{\pi}{\omega}} f(a\cos(\omega t))\,dt = 0$ implying $\alpha_0 = 0$.) α_1 and β_1 are functions of the amplitude a of the input.

When a linear gain factor is required to replace the nonlinearity, it may be shown that the mean squared error between approximate and true output is minimised by the first harmonic approximation above. The minimum over α, β is required of the

function

$$I = \frac{\omega}{2\pi} \int_{-\frac{\pi}{\omega}}^{\frac{\pi}{\omega}} (y(t) - \tilde{y}(t))^2 \, dt$$

where

$y(t) = f(x(t))$ is the true output and

$\tilde{y}(t) = \alpha \cos(\omega t) + \beta \sin(\omega t)$ is the approximate output.

Now for the minimum

$$0 = \frac{\partial I}{\partial \alpha} = -\frac{\omega}{\pi} \int_{-\frac{\pi}{\omega}}^{\frac{\pi}{\omega}} (f(a\cos(\omega t)) - \alpha \cos(\omega t) - \beta \sin(\omega t)) \cos(\omega t) \, dt$$

which implies

$$\int_{-\frac{\pi}{\omega}}^{\frac{\pi}{\omega}} f(a\cos(\omega t)) \cos(\omega t) \, dt = \alpha \int_{-\frac{\pi}{\omega}}^{\frac{\pi}{\omega}} \cos^2(\omega t) \, dt = \frac{\pi \alpha}{\omega}$$

So

$$\alpha = \frac{\omega}{\pi} \int_{-\frac{\pi}{\omega}}^{\frac{\pi}{\omega}} f(a\cos(\omega t)) \cos(\omega t) \, dt = \alpha_1$$

Similarly $\partial I / \partial \beta = 0$ implies

$$\beta = \frac{\omega}{\pi} \int_{-\frac{\pi}{\omega}}^{\frac{\pi}{\omega}} f(a\cos(\omega t)) \sin(\omega t) \, dt = \beta_1$$

For a 'memoryless' nonlinearity, $\beta = 0$ and $\tilde{y}(t) = \alpha \cos(\omega t)$, an output in phase with the input, and the describing function, or equivalent gain, is equal to

$$\frac{\alpha}{a} = \frac{\omega}{\pi a} \int_{-\frac{\pi}{\omega}}^{\frac{\pi}{\omega}} f(a\cos(\omega t)) \cos(\omega t) \, dt \tag{3.28}$$

If a nonlinear function can be represented as the sum of nonlinear functions whose describing functions are known, the required describing function is the sum of the individual describing functions.

Example: Saturation

The nonlinearity shown in the first diagram of figure 3-28 is considered. Here

$$y(t) = \begin{cases} -N & \text{for} \quad x(t) \leq -S \\ \frac{N}{S} x(t) & \text{for} \quad -S \leq x(t) \leq S \\ N & \text{for} \quad x(t) \geq S \end{cases} \tag{3.29}$$

Figure 3-29 shows the input and output waveforms for $x(t) = a\cos(\omega t)$ and $a > S$.

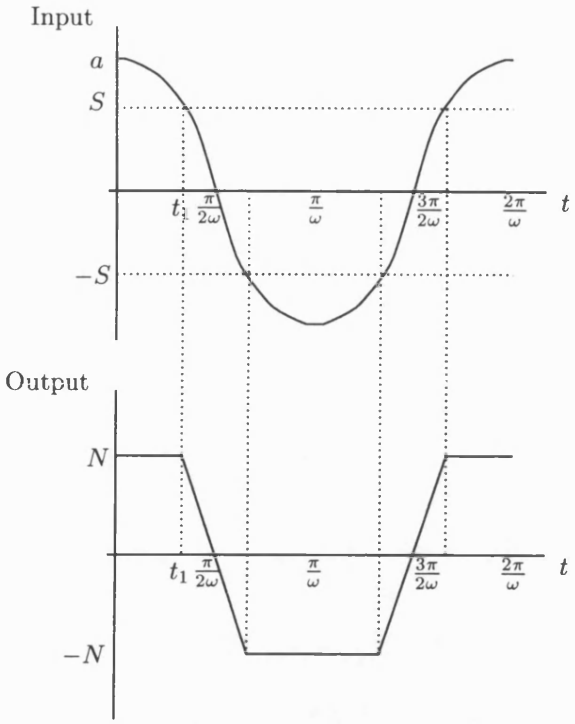


Figure 3-29: Input and Output Waveforms for Saturation

The output is an even, $2\pi/\omega$-periodic function of time. The describing function is given by equation 3.28, but from the symmetry of the system,

$$\begin{aligned}\frac{\alpha}{a} &= \frac{4\omega}{\pi a}\int_0^{\frac{\pi}{2\omega}} f(a\cos(\omega t))\cos(\omega t)\,dt \\ &= \frac{4\omega N}{\pi a}\left(\int_0^{t_1}\cos(\omega t)\,dt + \frac{a}{S}\int_{t_1}^{\frac{\pi}{2\omega}}\cos^2(\omega t)\,dt\right)\end{aligned}$$

where

$$t_1 = \frac{1}{\omega}\cos^{-1}\left(\frac{S}{a}\right)$$

$$\begin{aligned}\frac{\alpha}{a} &= \frac{4\omega N}{\pi a}\left(\frac{1}{\omega}\sin(\omega t_1) + \frac{a}{2S}\left(\frac{\pi}{2\omega} - t_1 - \frac{1}{2\omega}\sin(2\omega t_1)\right)\right) \\ &= \frac{N}{S}\left(1 + \frac{2S}{\pi a}\sqrt{1-\frac{S^2}{a^2}} - \frac{2}{\pi}\cos^{-1}\left(\frac{S}{a}\right)\right) \qquad (3.30)\end{aligned}$$

This is only a function of the input amplitude a and not the frequency ω. When $a = S$, the critical amplitude for the onset of saturation, the describing function 3.30 reduces

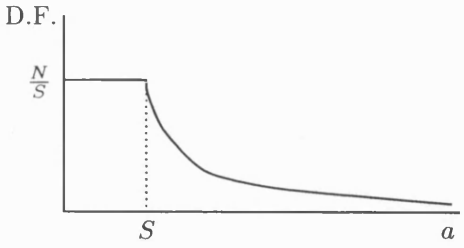
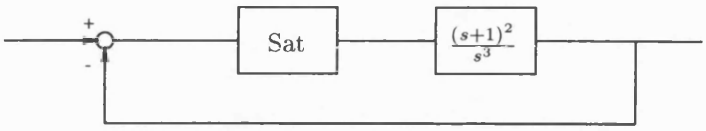


Figure 3-30: Describing Function for Saturation



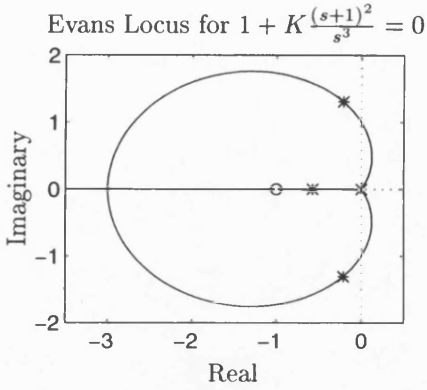


Figure 3-31: Root Locus with Saturation

to N/S. For a much larger than S, the output approximates a square wave, with describing function

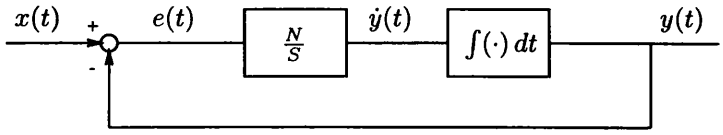
$$N\left(\frac{1}{S}+\frac{2}{\pi a}-\frac{1}{S}+\frac{2}{\pi a}\right)=\frac{4N}{\pi a}$$

to first order in S/a. ($\cos^{-1}(S/a) = \pi/2 - S/a + \cdots$.) Figure 3-30 shows a plot of the describing function against input amplitude.

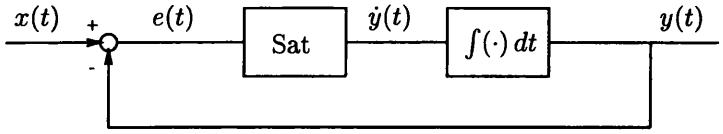
Application of Describing Function and Root Locus Methods to Nonlinear Systems Analysis

An equivalent gain perspective and Evans root locus techniques may be combined to analyse the stability of systems with linear dynamics and nonlinearities such as those shown in figure 3-28 (see [14]). This approach will be clarified by example.

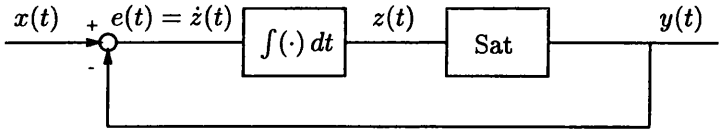
The system in figure 3-31 of saturation in series with linear dynamics and unity negative feedback will be considered. The saturation function is as shown in figure 3-



System 1



System 2



System 3

Figure 3-32: Integrator with Saturation

28 and described by equation 3.29. Here N and S are taken to be 1 for simplicity, so that for inputs of amplitude less than or equal to unity the saturation gain is unity. Considering the saturation in terms of equivalent gain, the root locus for the system is plotted varying a gain (K say) at the location of the saturation (see figure 3-31). At a gain $K = 1$ (for input to the saturation of amplitude less than 1), the roots are located at $s = -0.5698$ and $s = -0.2151 \pm 1.3071i$ as marked by asterisks on the locus plot. As the amplitude of the input to the saturation is increased, the equivalent gain decreases (as discussed in the describing function analysis above; see in particular figure 3-30). This corresponds to decreasing the gain on the root locus and moving along the curve from the $K = 1$ points towards the origin. The system becomes more lightly damped until at equivalent gain of $1/2$ the complex loci cut the imaginary axis with unity frequency and oscillation ensues. As the gain is decreased further the system becomes unstable. A similar approach might be applied to the Chu locus for systems with delay.

3.2.4 Piecewise Linear Functions

As was mentioned in the introduction 3.2.1, a nonlinearity which behaves linearly over certain intervals permits linear analysis on each interval and hence *exact* solutions are found. (If the nonlinearity approximates a piecewise linear function, this provides a method of approximate solution.) In this section step input $AH(t)$ is applied to

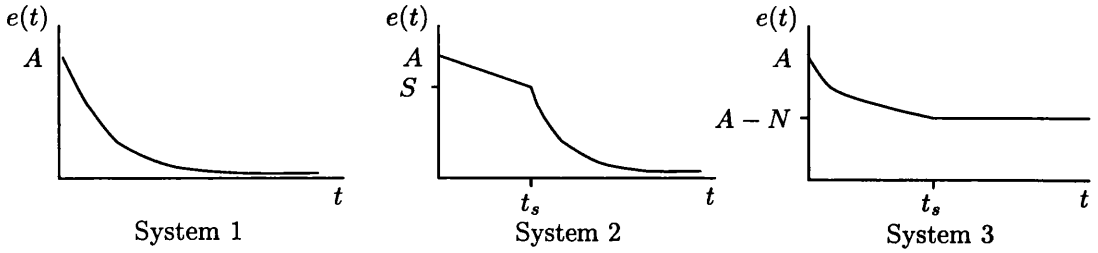


Figure 3-33: Error Comparison for Saturation exhibiting Wind-up

saturation in series with an integrator and unity negative feedback, and with reversal of the series components, as in figure 3-32, systems 2 and 3. Comparison is made with the corresponding linear case in which the saturating element is replaced by a gain of N/S (figure 3-32, system 1).

System 1

Zero initial conditions on the integrator are assumed. The equation governing the linear system is

$$\dot{y}(t) + \frac{N}{S}y(t) = \frac{NA}{S}$$

with solution

$$y(t) = A(1 - \exp(-Nt/S))$$

The error $e(t) = A\exp(-Nt/S)$ decays to zero as $t \to \infty$ and is shown in figure 3-33.

System 2

The output from the saturating element is given by

$$\dot{y}(t) = \begin{cases} -N & \text{for} \quad e(t) \leq -S \\ \frac{N}{S}e(t) & \text{for} \quad -S \leq e(t) \leq S \\ N & \text{for} \quad e(t) \geq S \end{cases}$$

where

$$e(t) = x(t) - y(t) = AH(t) - y(t)$$

Suppose that $A > S$ then initially the amplifier will be saturated; for this region of operation, $\dot{y}(t) = N$ so $y(t) = Nt$. Let t_s be the time at which the amplifier becomes unsaturated. Then $e(t_s) = S = A - Nt_s$ and $t_s = (A - S)/N$. During the unsaturated region,

$$\dot{y}(t) = \frac{N}{S}(AH(t) - y(t))$$

which has solution

$$y(t) = A - S\exp\left(-\frac{N}{S}(t - t_s)\right)$$

(The amplifier will remain unsaturated for $t \geq t_s$ as $e(t) = S\exp\left(-\frac{N}{S}(t - t_s)\right) \leq S$.) Finally the output $y(t)$ is given by

$$y(t) = \begin{cases} Nt & \text{for } 0 \leq t \leq t_s = (A - S)/N \\ A - S\exp\left(-\frac{N}{S}(t - t_s)\right) & \text{for } t \geq t_s \end{cases}$$

Also $y(t) \to A$ as $t \to \infty$. The error

$$e(t) = \begin{cases} A - Nt & \text{for } 0 \leq t \leq t_s \\ S\exp\left(-\frac{N}{S}(t - t_s)\right) & \text{for } t \geq t_s \end{cases}$$

decays to zero as $t \to \infty$ and is illustrated in figure 3-33. If $S < A$ the amplifier remains unsaturated throughout and the system reduces to the linear system 1.

System 3

The output is given by

$$y(t) = \begin{cases} -N & \text{for } z(t) \leq -S \\ \frac{N}{S}z(t) & \text{for } -S \leq z(t) \leq S \\ N & \text{for } z(t) \geq S \end{cases}$$

and the error $e(t) = AH(t) - y(t)$. Initially the amplifier is unsaturated so the equations governing the system are

$$\dot{z}(t) = AH(t) - \frac{N}{S}z(t)$$

with solution

$$z(t) = \frac{SA}{N}(1 - \exp(-Nt/S))$$

Saturation takes place when $z = S$ at a time

$$t_s = \frac{S}{N}\ln\left(\frac{A}{A - N}\right),$$

being real and positive if $A > N$. In the saturated region, $\dot{z}(t) = A - N$ so that

$$z(t) = (A - N)t + S\left(1 - \left(\frac{A - N}{N}\right)\ln\left(\frac{A}{A - N}\right)\right)$$

which grows without bound. This phenomenon associated with saturating amplifiers is known as *wind-up*. Here the error

$$e(t) = \begin{cases} A\exp(-Nt/S) & \text{for} \quad 0 \leq t \leq t_s = \frac{S}{N}\ln\left(\frac{A}{A-N}\right) \\ A - N & \text{for} \quad t \geq t_s \end{cases}$$

takes a constant value for $t \geq t_s$ (see figure 3-33). If $A < N$ saturation is avoided, the system again reducing to the linear one.

These examples also emphasize the effect of interchanging series elements in non-linear systems; for linear systems equivalence would be preserved.

3.2.5 Frequency Domain Criteria

Other stability methods deserving mention, for completeness, but which will not be developed further here include frequency domain criteria. Details may be found in many texts such as [4], [10], [45], [57] and [59].

Popov's Method

V. M. Popov developed an extension of the Nyquist criterion. The method gives sufficient, but not necessary, conditions for asymptotic stability of time-invariant control systems consisting of a nonlinear element in series with linear dynamics with unity negative feedback. The input-output relation of the nonlinearity must be restricted to a sector such that the output is bounded by a constant times the input, say

$$0 \leq \frac{f(x)}{x} < \alpha < \infty$$

for all $x \neq 0$ and $f(0) = 0$, where x is the input to the nonlinear function f. The restriction on the linear transfer function $G(s)$ is that the output response to an initial condition of $G(s)\exp(nt)$ must converge to zero (for some n positive or negative). For the error of the system to converge to zero, Popov's criterion requires

$$\mathrm{Re}(G(i\omega)) > -\frac{1}{\alpha} + \omega m\ \mathrm{Im}(G(i\omega))$$

for all $\omega \geq 0$ and some $m \epsilon \mathbb{R}$. So for all $\omega \geq 0$, the Nyquist plot of $G(i\omega)$ must lie to the right of a straight line, known as the Popov line, given by

$$\mathrm{Im}(G(i\omega)) = \frac{1}{\omega m}\left(\mathrm{Re}(G(i\omega)) - \frac{1}{\alpha}\right)$$

This is illustrated in figure 3-34.

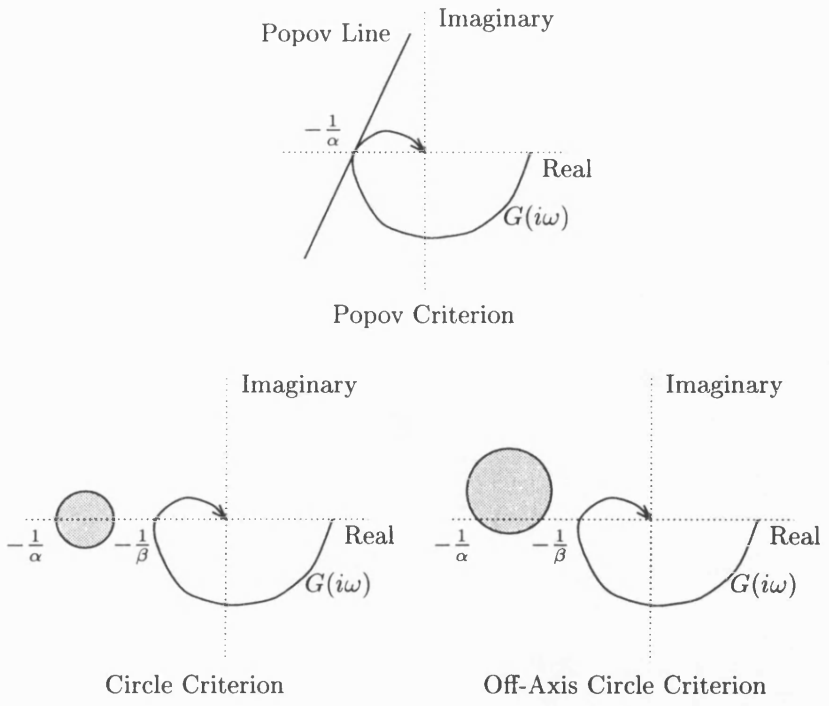


Figure 3-34: Frequency Domain Stability Criteria

Circle Criteria

Extensions of the Popov criterion are the circle criterion and off-axis circle criterion. The circle criterion applies to linear dynamics in series with a sector nonlinearity $f(\cdot)$ such that $f(0) = 0$ and

$$\alpha < \frac{f(x)}{x} < \beta < \infty$$

for all $x \neq 0$. Stability holds if the Nyquist criterion is satisfied for the associated linear system (which must itself be stable) with the rôle of the critical point being taken by the disc with diameter on the real axis between $-1/\alpha$ and $-1/\beta$, shown in figure 3-34. (The Nyquist diagram for $G(i\omega)$, the transfer function of the linear part, must not encroach upon this disc, and its encirclements relative to it must be as the associated linear system would require for the critical point.) This condition is sufficient but not necessary and may be improved by further restrictions on the nonlinearity to give the off-axis circle criterion. It is required that the nonlinearity should satisfy a slope restriction in addition to the above sector bound:

$$\alpha < \frac{\partial f(x)}{\partial x} < \beta$$

Here the Nyquist plot of $G(i\omega)$ for $\omega \geq 0$ is required to avoid an arbitrary circle passing through $-1/\alpha$ and $-1/\beta$ (see again figure 3-34).

3.2.6 The Phase Plane

The phase plane is a graphical representation of the transient behaviour of a nonlinear equation of second order which may or may not admit direct solution. The position $x(t)$ and 'velocity' $\dot{x}(t)$ are plotted in the plane at a given time t; a curve is traced as t varies. Suppose the system is described by

$$\ddot{x}(t) = f(x(t), \dot{x}(t)) = f(x_1(t), x_2(t))$$

or

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} x_2 \\ f(x_1, x_2) \end{pmatrix}$$

The trajectories in the phase plane are given by

$$\frac{dx_2}{dx_1} = \frac{f(x_1, x_2)}{x_2} \tag{3.31}$$

(Of course if the equation is soluble the trajectories are trivially found from the solution.) A set of trajectories plotted for various inputs or initial conditions is known as a phase portrait. Through each point in the phase plane passes a unique trajectory, with the exception of equilibrium points $x_2 = 0 = f(x_1, 0)$. This is in contrast to systems with delay, for which the loci may cross in the phase plane as remarked earlier. A direct consequence of equation 3.31 is that the vector $(x_2, f(x_1, x_2))$ drawn at the point (x_1, x_2) is tangent to the solution trajectory passing through (x_1, x_2).

The method of isoclines is a technique for obtaining phase plane trajectories graphically. An isocline is a line on which trajectories have an equal constant slope, i.e. for which $f(x_1, x_2)/x_2 = c$, constant. Points satisfying this equation are found for various c and the trajectories constructed by moving from one isocline to the next.

Special features of the phase portrait are to be expected in the neighbourhood of any equilibrium points. A linearisation of the differential equation in this region gives useful insights into the local behaviour. Considering the linear time-invariant system

$$\ddot{x}(t) + a_1\dot{x}(t) + a_2x(t) = 0$$

a standard solution is $x(t) = C\exp(st)$, $\dot{x}(t) = Cs\exp(st)$, giving a line $x_2 = sx_1$ in the phase plane, s being a root of the characteristic equation $s^2 + a_1s + a_0 = 0$. The phase portrait depends on the nature of the characteristic roots. The following table

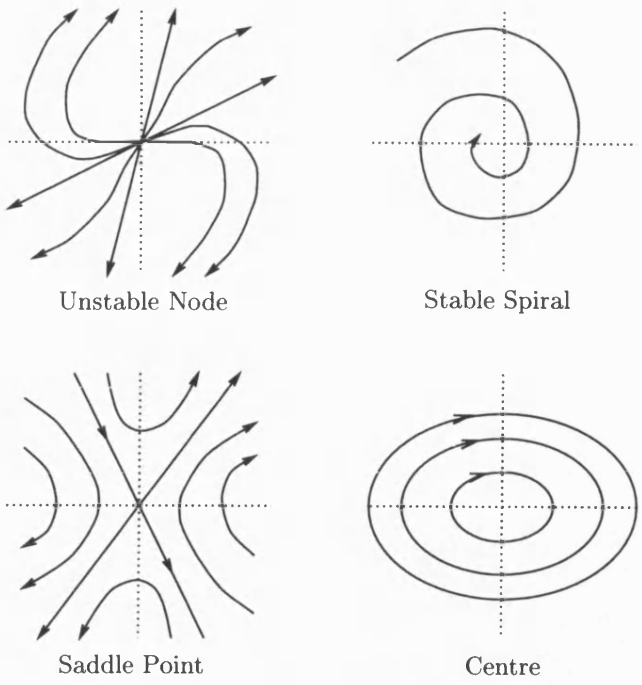


Figure 3-35: Phase Plane Characteristics in Neighbourhood of Equilibria

gives the qualitative behaviour, which is illustrated in figure 3-35.

Roots	Origin	
Real and negative	Stable node	(sink)
Real and positive	Unstable node	(source)
Real and opposite sign	Saddle point	
Complex conjugate with negative real part	Stable spiral	(sink)
Complex conjugate with positive real part	Unstable spiral	(source)
Imaginary	Centre	

A feature unique to nonlinear systems is that of limit cycles, representing sustained oscillations. These are periodic solutions of form

$$\begin{pmatrix} x_1(t+T) \\ x_2(t+T) \end{pmatrix} = \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix}$$

where T is the period, and are distinct from the centre. A limit cycle may either be stable, so that nearby trajectories converge, or unstable so that they diverge. Further details may be found in any standard nonlinear dynamical systems or control text. The following theorem deserves special mention however, giving sufficient conditions for the

existence of a limit cycle for nonlinear equations of *second* order.

Poincaré Bendixson Theorem

> Take a closed bounded set in the phase plane, not containing any equilibria. A trajectory starting within the set for $t = 0$ and remaining therein for all $t \geq 0$ is either a closed orbit or tends to a closed orbit (limit cycle) as $t \to \infty$.

The time elapsed between two points on a particular trajectory in the phase plane may be determined as follows. $x_2(t) = \dot{x}_1(t)$ implies

$$\int_{t_1}^{t_2} dt = t_2 - t_1 = \int_{x_1(t_1)}^{x_1(t_2)} \frac{1}{x_2(t)} dx_1(t)$$

If expressions for the positions on the trajectories are available, this integral may be evaluated directly. Alternatively, the phase portrait of $1/x_2(t)$ may be plotted against $x_1(t)$ and the area under the curve calculated numerically.

3.2.7 Lyapunov Stability Theory

A. M. Lyapunov developed two methods for stability of nonlinear dynamical systems taking an energy perspective. [49] gives an overview of Lyapunov's work and developments during the past century, and applications of the second method to control system analysis and design are made in [28].

The First Method of Lyapunov

The first method addresses the stability of a system for which the solution of the differential equation representing it is available, or for which the solution may be expanded as a power series. Considering for example a nonlinear function of form 3.27, this may be expanded about a particular point as a Taylor series, with the first two terms representing the linearisation of the nonlinearity about the operating point. Lyapunov proved that if the roots of the characteristic equation of this linearisation are not purely imaginary, they reflect the stability behaviour of the original nonlinear differential equation in the neighbourhood of the operating point. This method is particularly important when considering stability of an equilibrium point.

The Second Method of Lyapunov

This method does not require solution of the differential equation and gives sufficient conditions for stability for the n-dimensional time-invariant nonlinear system repre-

sented in state-space form 3.25. As remarked earlier, the origin may be assumed to be an equilibrium point.

Different forms of stability are distinguished for nonlinear systems as mentioned before. An equilibrium point is classified as *stable* if solutions starting within a δ-neighbourhood of the point remains within an ε-neighbourhood for all positive time. The point is *asymptotically stable* if in addition δ may be chosen such that the solution tends to the origin as time $t \to \infty$. These are local stability definitions. An equilibrium point is *globally asymptotically stable* if it is stable and solutions tend to the origin as $t \to \infty$ for any starting point.

The following preliminary definitions are also required. A scalar-valued function V defined on a neighbourhood B of the origin is *positive definite* if

- $\partial V/\partial x_i$, $i = 1, \ldots, n$ are continuous on B
- $V(\mathbf{0}) = 0$
- $V(\boldsymbol{x}) > 0$ for all $\boldsymbol{x} \epsilon B \backslash \{\mathbf{0}\}$

A positive definite function V defined on $\mathbb{R}^n$ is *radially unbounded* if given any $M > 0$ there exists an $r > 0$ such that $\|\boldsymbol{x}\| > r$ implies $V(\boldsymbol{x}) > M$.

It may be shown that surfaces of constant V, where V is positive-definite, form concentric shells around the origin. If $\dot{V} \leq 0$ along all trajectories in a neighbourhood of the origin, motion cannot take place away from the origin. Lyapunov's stability results are now given.

Lyapunov Stability Theorems

> Suppose that $\boldsymbol{x} = \mathbf{0}$ is an equilibrium of the system $\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x})$. Then if there exists a positive definite function V defined on some neighbourhood B of the origin and in addition
>
> **Local Stability** $\dot{V} \leq 0$ along trajectories in B, then the origin is stable.
>
> **Local Asymptotic Stability** $\dot{V} < 0$ along trajectories in $B \backslash \{\mathbf{0}\}$, then the origin is asymptotically stable.
>
> **Global Asymptotic Stability** V is defined on $\mathbb{R}^n$ and is radially unbounded, and $\dot{V} < 0$ along trajectories in $B \backslash \{\mathbf{0}\}$, then the origin is globally asymptotically stable.
>
> **Instability** If V satisfies the first two conditions of the definition for positive definite functions above, but takes negative values arbitrarily near the origin (somewhere), and $\dot{V} < 0$ along trajectories in $B \backslash \{\mathbf{0}\}$, then the origin is unstable.

Example: The Van der Pol Equation

Van der Pol's differential equation, representing a simple harmonic oscillator with non-linear damping (see [68]), has form

$$\ddot{x}(t) - \mu(1 - x^2(t))\dot{x}(t) + x(t) = 0 \qquad \mu > 0 \tag{3.32}$$

and will be discussed in detail in later sections (the remainder of this chapter and chapter 4). Negative damping occurs in the region $|x| < 1$ and positive damping in the regions $|x| > 1$. Connected to the Van der Pol equation is Rayleigh's equation

$$\ddot{x}(t) - \mu\left(1 - \frac{1}{3}\dot{x}^2(t)\right)\dot{x}(t) + x(t) = 0$$

arising from models of oscillations in a violin string. Van der Pol's equation is derived from this by differentiating and putting $y = \dot{x}$.

Returning to equation 3.32, this may be written as

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} x_2 \\ \mu(1 - x_1^2)x_2 - x_1 \end{pmatrix}$$

The origin may be shown to be unstable by taking

$$V(x_1, x_2) = -\left(\frac{1}{\mu} + \frac{\mu}{2}\right)x_1^2 + x_1x_2 - \frac{1}{\mu}x_2^2$$

which takes negative values near the origin (for example $x_2 = 0$), so that

$$\dot{V} = \left(-2\left(\frac{1}{\mu} + \frac{\mu}{2}\right)x_1 + x_2\right)\dot{x}_1 + \left(x_1 - \frac{2}{\mu}x_2\right)\dot{x}_2 = -x_1^2(1 + \mu x_1x_2) - x_2^2(1 - 2x_1^2)$$

which may be shown to be negative definite inside the circle of radius $\min\{1/\sqrt{2}, \sqrt{2/\mu}\}$ centre the origin. Thus the origin is unstable. The Poincaré Bendixson theorem may be applied to show that the Van der Pol equation has a periodic solution.

3.2.8 The Methods of Poincaré and Lindstedt

These two approximation methods (see for example [23]) apply to equations in which the nonlinear element is small; a series expansion of the solution is made in powers of the (small) coefficient of the nonlinear term. The methods are most clearly explained in terms of an example, in this case the Van der Pol equation 3.32, with initial conditions $x(0) = A$ and $\dot{x}(0) = B$.

Poincaré's Method

The solution $x(t)$ is written as a series expansion in powers of μ.

$$x = \sum_{j=0}^{\infty} \mu^j x_j$$

so

$$\dot{x} = \sum_{j=0}^{\infty} \mu^j \dot{x}_j\,, \qquad \ddot{x} = \sum_{j=0}^{\infty} \mu^j \ddot{x}_j$$

and the initial conditions become, without loss of generality,

$$x_0(0) = A\,, \quad \dot{x}_0(0) = B\,, \quad x_j(0) = 0 = \dot{x}_j(0) \quad j \geq 1$$

Then substituting back into the original differential equation,

$$\sum_{j=0}^{\infty} \mu^j \ddot{x}_j - \mu \left(1 - \left(\sum_{j=0}^{\infty} \mu^j x_j\right)^2\right) \sum_{j=0}^{\infty} \mu^j \dot{x}_j + \sum_{j=0}^{\infty} \mu^j x_j = 0$$

Terms in equal powers of μ are equated resulting in a series of linear second order equations. Equating terms in μ^0 gives $\ddot{x}_0 + x_0 = 0$ with solution $x_0 = A\cos(t) + B\sin(t)$. Equating terms in μ^1 and substituting the result for x_0, $\dot{x}_0$ gives a linear second order differential equation for x_1. The particular integral gives rise to so-called *secular* terms, being products of powers of t with trigonometric terms. These give inaccuracies for high t (or large μ) as the approximation contains these terms not present in the true solution. Explicitly, x_1 will contain terms in $t\sin(t)$ and $t\cos(t)$, so that when t is of order μ^{-1}, the series expansion will be a poor approximation for x as μx_1 will no longer be smaller than x_0.

Poincaré's method attempts to find a periodic solution to a nonlinear equation by insisting that it should have the same period as the corresponding linear equation with $\mu = 0$.

Lindstedt's Method

This is a similar approach to Poincaré's method but seeks to avoid the presence of secular terms by allowing both x and the frequency ω to vary with μ. Each x_j is just required to be periodic.

The transformation $\tau = \omega t$ is made. Then $\dot{x} = wx'$, $\ddot{x} = w^2 x''$, the prime denoting differentiation with respect to τ.

Returning to the example of the Van der Pol equation 3.32 and applying this trans-

formation,

$$\omega^2 x''(\tau) - \mu\omega(1 - x^2(\tau))x'(\tau) + x(\tau) = 0 \tag{3.33}$$

with initial conditions $x(0) = A$, $x'(0) = B/\omega$, but since time does not enter into equation 3.32 explicitly, the time origin may be fixed so that $x'(0) = 0$. It is assumed that

$$x = \sum_{j=0}^{\infty} \mu^j x_j\,, \qquad \omega = 1 + \sum_{j=1}^{\infty} \mu^j \omega_j$$

Here $\omega_0 = 1$ since the solution for $\mu = 0$ is simple harmonic motion of unity frequency. The initial conditions are

$$x_0(0) = A\,, \quad x_0'(0) = 0\,, \quad x_j(0) = 0 = x_j'(0) \quad j \geq 1$$

Substituting the series into the Van der Pol equation,

$$\left(1 + \sum_{j=1}^{\infty} \mu^j \omega_j\right)^2 \sum_{j=0}^{\infty} \mu^j x_j'' - \mu \left(1 + \sum_{j=1}^{\infty} \mu^j \omega_j\right) \left(1 - \left(\sum_{j=0}^{\infty} \mu^j x_j\right)^2\right) \sum_{j=0}^{\infty} \mu^j x_j' + \sum_{j=0}^{\infty} \mu^j x_j = 0$$

The same procedure of equating terms of equal power in μ is applied. Again the term in μ^0 gives $x_0'' + x_0 = 0$ with solution $x_0 = A\cos(\tau) = A\cos(\omega t)$. The equation to be solved for terms in μ^1 is

$$x_1'' + x_1 = 2\omega_1 A\cos(\tau) + A\left(\frac{A^2}{4} - 1\right)\sin(\tau) + \frac{A^3}{4}\sin(3\tau)$$

The first two terms on the right-hand side will give rise to secular terms. To avoid this, they are set to zero, so that $\omega_1 = 0$ and $A = 2$ ($A = 0$ being the trivial steady state solution $x = 0$). Now solving the equation for x_1 and applying the initial conditions gives

$$x_1 = \frac{3}{4}\sin(\tau) - \frac{1}{4}\sin(3\tau) + C\cos(\tau)$$

(C an arbitrary constant to be determined from subsequent equations). Proceeding to the terms in μ^2 and eliminating secular terms as before gives $C = 0$ and $w_2 = -1/16$. Finally the series expansions for $x(t)$ and ω to order μ^2 are found to be

$$\begin{aligned} x(t) &= 2\cos(\omega t) + \frac{\mu}{4}(3\sin(\omega t) - \sin(3\omega t)) \\ &\quad -\frac{\mu^2}{96}(13\cos(\omega t) - 18\cos(3\omega t) + 5\cos(5\omega t)) + \cdots \qquad (3.34) \\ \omega &= 1 - \frac{\mu^2}{16} + \cdots \qquad (3.35) \end{aligned}$$

(Note that considering the dominant terms only the motion approximates a sinusoid of unity frequency and amplitude 2; this is consistent with the result obtained via the method of Kryloff and Bogoliuboff in section 3.2.10.)

Lindstedt's method gives accurate solutions for small μ by deriving many terms in the asymptotic expansion but only finds periodic solutions. No information is given on the stability of such a periodic solution nor how quickly solutions may tend to it.

3.2.9 The Series Expansions of Shohat

Shohat [58] implements Lindstedt's method for Van der Pol's equation for expansions of the solution and frequency in powers of the *small* parameter μ. The final part of the paper develops new series valid for all values of μ, large and small.

Again equation 3.32 is transformed to equation 3.33 by the substitution $\tau = \omega t$, and the equation

$$\omega^2 x''(\tau) - \mu\omega\rho(1 - x^2(\tau))x'(\tau) + x(\tau) = 0 \tag{3.36}$$

is considered. ρ is a new parameter in powers of which x and ω are to be expanded; subsequently it is replaced by unity. Initial conditions are $x(0) = A$, $x'(0) = 0$ and it is assumed that

$$\begin{aligned} x &= \sum_{j=0}^{\infty} \rho^j x_j \\ \mu &= \frac{1}{1+\gamma-\rho} \\ \omega &= (1+\gamma-\rho)\sum_{j=0}^{\infty} \rho^j \omega_j \end{aligned}$$

γ constant. So for $\rho = 1$, $\mu\ (= 1/\gamma)$ may take any value depending on the value of γ.

The usual pattern is followed of substitution into the differential equation 3.36, equating terms of equal power in ρ, solving the linear second order differential equations (eliminating secular terms) and finally setting $\rho = 1$. This leads to

$$\begin{aligned} x(t) &= 2\cos(\omega t) + \frac{\mu}{4(1+\mu)}(3\sin(\omega t) - \sin(3\omega t)) \\ &\quad + \frac{\mu^2}{4(1+\mu)^2}\left(-\frac{1}{2}\cos(\omega t) + \frac{3}{4}\cos(3\omega t) + \frac{5}{24}\cos(5\omega t) + 3\sin(\omega t) - \sin(3\omega t)\right) + \cdots \\ \omega &= \frac{1}{1+\mu} + \frac{\mu}{(1+\mu)^2} + \frac{15}{16}\frac{\mu^2}{(1+\mu)^3} + \frac{13}{16}\frac{\mu^3}{(1+\mu)^4} + \cdots \end{aligned} \tag{3.37}$$

Note that, for μ small, expanding these series for x and ω in powers of μ yields the same series expansions as obtained via Lindstedt's method 3.34, 3.35. Differentiating

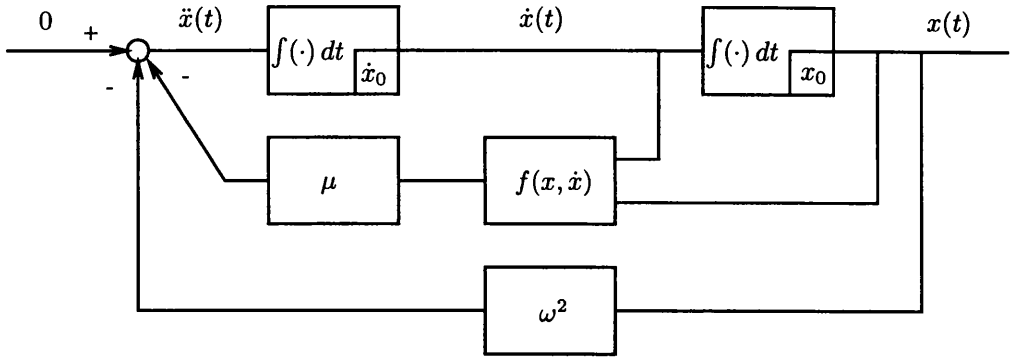


Figure 3-36: Quasi-Linear System

the series for $x(t)$ gives a series for $\dot{x}(t)$, the latter being less accurate than the former due to the extra factor of ω (approximate). As μ increases from zero to infinity, the frequency ω decreases from 1 to zero implying an increase in period from 2π to infinity.

3.2.10 The Method of Kryloff and Bogoliuboff

Kryloff and Bogoliuboff [29] developed an approximate method of solution for quasi-linear differential equations, having form

$$\ddot{x}(t) + \mu f(x(t), \dot{x}(t)) + \omega^2 x(t) = 0 \tag{3.38}$$

for f a nonlinear function and μ a constant measuring the degree of nonlinearity present. This is shown in system form in figure 3-36. Their method of the first approximation is one of averaging, which, like Lindstedt's method, seeks to avoid the presence of secular terms of form $tx(t)$. It is based on the fact that for $\mu = 0$ the equation reduces to simple harmonic motion of frequency ω:

$$x(t) = a\sin(\omega t + \phi) \tag{3.39}$$

$$\dot{x}(t) = a\omega\cos(\omega t + \phi) \tag{3.40}$$

for arbitrary constants a and ϕ (amplitude and phase respectively). Kryloff and Bogoliuboff attempt to find a solution of the differential equation 3.38 of form 3.39 with derivative 3.40 but in which a and ϕ are functions of time. Solutions are sought such that

$$x(t) = a(t)\sin(\omega t + \phi(t))$$

Differentiating this expression,

$$\dot{x}(t) = \dot{a}(t)\sin(\omega t + \phi(t)) + a(t)(\omega + \dot{\phi}(t))\cos(\omega t + \phi(t))$$

and to preserve 3.40,

$$\dot{a}(t)\sin(\omega t + \phi(t)) + a(t)\dot{\phi}(t)\cos(\omega t + \phi(t)) = 0 \tag{3.41}$$

$$\dot{x}(t) = a(t)\omega\cos(\omega t + \phi(t))$$

Differentiating again,

$$\ddot{x}(t) = \dot{a}(t)\omega\cos(\omega t + \phi(t)) - a(t)\omega(\omega + \dot{\phi}(t))\sin(\omega t + \phi(t)) \tag{3.42}$$

Substituting into the original differential equation

$$\dot{a}(t)\omega\cos(\omega t+\phi(t)) - a(t)\dot{\phi}(t)\omega\sin(\omega t+\phi(t)) = -\mu f(a\sin(\omega t+\phi(t)), a\omega\cos(\omega t+\phi(t)))$$

This equation and 3.41 when solved simultaneously give equations describing the variations of amplitude and phase with time.

$$\dot{a}(t) = -\frac{\mu}{\omega}f(a\sin(\omega t + \phi), a\omega\cos(\omega t + \phi))\cos(\omega t + \phi) \tag{3.43}$$

$$\dot{\phi}(t) = \frac{\mu}{a\omega}f(a\sin(\omega t + \phi), a\omega\cos(\omega t + \phi))\sin(\omega t + \phi) \tag{3.44}$$

The single second order differential equation 3.38 in $x(t)$ has been reduced to two first order differential equations in $a(t)$ and $\phi(t)$. So far no approximations have been made and these are *exact* equations.

The derivatives $\dot{a}(t)$ and $\dot{\phi}(t)$ are small, being proportional to the small quantity μ, and are therefore slowly varying functions of time. Considering a period $T = 2\pi/\omega$, the quantity $\omega t + \phi(t)$ will increase by approximately 2π whereas a and ϕ will not have changed appreciably.

Expanding the right-hand sides of equations 3.43 and 3.44 as a Fourier series,

$$\dot{a}(t) = -\frac{\mu}{\omega}\left(\frac{A_0(a)}{2} + \sum_{n=1}^{\infty}(A_n(a)\cos(n(\omega t + \phi)) + B_n(a)\sin(n(\omega t + \phi)))\right) \tag{3.45}$$

$$\dot{\phi}(t) = \frac{\mu}{a\omega}\left(\frac{C_0(a)}{2} + \sum_{n=1}^{\infty}(C_n(a)\cos(n(\omega t + \phi)) + D_n(a)\sin(n(\omega t + \phi)))\right) \tag{3.46}$$

For the purposes of integration, a and ϕ are taken to be constant over the period

$[t, t+T]$ and equal to the values $a(t)$ and $\phi(t)$. Then

$$A_n(a) = \frac{\omega}{\pi}\int_t^{\frac{2\pi}{\omega}+t} f(a\sin(\omega t+\phi), a\omega\cos(\omega t+\phi))\cos(\omega t+\phi)\cos(n(\omega t+\phi))\,dt$$

and similarly for the remaining coefficients. As shall be seen, only the leading terms $A_0(a)$ and $C_0(a)$ will be required. These are given by

$$\begin{aligned} A_0(a) &= \frac{\omega}{\pi}\int_t^{\frac{2\pi}{\omega}+t} f(a\sin(\omega t+\phi), a\omega\cos(\omega t+\phi))\cos(\omega t+\phi)\,dt \\ C_0(a) &= \frac{\omega}{\pi}\int_t^{\frac{2\pi}{\omega}+t} f(a\sin(\omega t+\phi), a\omega\cos(\omega t+\phi))\sin(\omega t+\phi)\,dt \end{aligned}$$

Now integrating the expressions 3.45 and 3.46 on the interval $[t, t+T]$ and again taking a and ϕ as constant, all terms bar the leading one disappear so that

$$\frac{a(t+T)-a(t)}{T} = -\frac{\mu}{2\omega}A_0(a), \qquad \frac{\phi(t+T)-\phi(t)}{T} = \frac{\mu}{2a\omega}C_0(a)$$

The left-hand sides of these expressions are replaced by $\dot{a}(t)$ and $\dot{\phi}(t)$ since the differences $a(t+T)-a(t)$ and $\phi(t+T)-\phi(t)$ are assumed small. Finally the equations of the first approximation result:

$$\dot{a}(t) = -\frac{\mu}{2\pi}\int_0^{\frac{2\pi}{\omega}} f(a\sin(\omega t+\phi), a\omega\cos(\omega t+\phi))\cos(\omega t+\phi)\,dt \qquad (3.47)$$

$$\dot{\phi}(t) = \frac{\mu}{2\pi a}\int_0^{\frac{2\pi}{\omega}} f(a\sin(\omega t+\phi), a\omega\cos(\omega t+\phi))\sin(\omega t+\phi)\,dt \qquad (3.48)$$

$\dot{a}(t)$ and $\dot{\phi}(t)$ are therefore approximated by their *average* values over a period T, with the amplitude a and phase ϕ taken to be constant during this averaging process.

The equations 3.47 and 3.48 show how properties of the function f influence the amplitude and phase of the approximation. For example, if f only contains powers of x, then $\dot{a}(t) = 0$ and the amplitude remains unaffected although there is a change in phase. Conversely if f only contains powers of $\dot{x}$ then the phase is unaffected but the amplitude affected. The analytical solution of 3.47 and 3.48 may be possible in some cases, giving an approximate expression for x; alternatively, steady-state oscillations may be investigated by setting $\dot{a} = 0 = \dot{\phi}$.

Kryloff and Bogoliuboff in their book [29] give a physical justification for considering this first approximation as well as methods for higher order approximations. The latter have not been treated here as the motivation for study is the extension to the delayed case; the higher approximations, being more complicated, do not readily admit such an extension. Further references on averaging methods include [27], [42], [55] and [69].

Example: The Van der Pol Oscillator

Consider again the Van der Pol equation

$$\ddot{x}(t) - \mu(1 - x^2(t))\dot{x}(t) + \omega^2 x(t) = 0 \tag{3.49}$$

Here $f(x, \dot{x}) = (x^2 - 1)\dot{x}$. For the first approximation,

$$\begin{aligned}
\dot{a}(t) &= -\frac{\mu a}{2\pi}\int_0^{\frac{2\pi}{\omega}} (a^2 \sin^2(\omega t + \phi) - 1)\cos^2(\omega t + \phi)\, dt \\
&= \frac{\mu a}{2\pi}\int_0^{\frac{2\pi}{\omega}} \frac{1}{2}(\cos(2(\omega t + \phi)) + 1) - \frac{a^2}{8}(1 - \cos(4(\omega t + \phi)))\, dt \\
&= \frac{\mu a}{2\omega}\left(1 - \frac{a^2}{4}\right)
\end{aligned} \tag{3.50}$$

and

$$\dot{\phi}(t) = \frac{\mu}{2\pi}\int_0^{\frac{2\pi}{\omega}} (a^2 \sin^2(\omega t + \phi) - 1)\sin(\omega t + \phi)\cos(\omega t + \phi)\, dt = 0$$

From equation 3.50, the amplitude a at the steady state ($\dot{a}(t) = 0$) is 2 (discounting the trivial solution $x = 0$). Solving the first order separable differential equation gives the result

$$a(t) = \frac{a_0 \exp(\mu t/2\omega)}{\sqrt{1 + \frac{1}{4}a_0^2(\exp(\mu t/\omega) - 1)}}$$

where $a(0) = a_0$. So if $a_0 = 0$ no oscillation will develop, but for $a_0 > 0$, $a(t) \to 2$ as $t \to \infty$, the amplitude increasing for $0 < a_0 < 2$ and decaying for $a_0 > 2$. Also $\phi(t) = \phi(0) = \phi_0$ so the first approximation solution, for μ small, is

$$x(t) = \frac{a_0 \exp(\mu t/2\omega)\sin(\omega t + \phi_0)}{\sqrt{1 + \frac{1}{4}a_0^2(\exp(\mu t/\omega) - 1)}}$$

A program has been written to solve equations of form 3.49 numerically for given starting position x_0 and velocity $\dot{x}_0$ so that comparison may be made with the first approximation of Kryloff and Bogoliuboff, particularly in the case where the first order differential equations 3.47 and 3.48 cannot be solved analytically. Also of interest is a comparison between the estimate of the amplitude a of any limit cycle as $t \to \infty$ and the true steady state behaviour. This will be discussed in greater detail in chapter 4, since the program has been written to incorporate solutions of Van der Pol type equations with added delay, of which the problem in question here is a special case. Figure 3-37 shows trajectories in the $(x, \dot{x})$ phase plane described for initial amplitudes less than and greater than 2 together with the limit cycle as $t \to \infty$, for the case $\mu = 0.1$. The

predicted limit cycle from the method of Kryloff and Bogoliuboff is marked, being the (dotted) circle centre the origin of radius 2. The corresponding results for $\mu = 1/4$ are given in figure 3-38; the greater accuracy of the approximation for the smaller value of μ is evident. The larger value of μ increasing the damping element present is also clear.

3.2.11 The Vector Approach for Nonlinear Systems

The vector approach as described in section 3.1.8 is extendible to systems with nonlinear elements described by functions $f(\cdot)$ by an approximation method. As before, periodic solutions are sought of form $x(t) = a\sin(\omega t)$, $\dot{x}(t) = a\omega\cos(\omega t)$. The nonlinear function f is evaluated for x and its derivatives of this form and expanded as a Fourier series. All but the first harmonic is neglected resulting in an expression linear in $\sin(\omega t)$ and $\cos(\omega t)$, as with the describing function. Vectors corresponding to these nonlinear components together with any linear elements present in the equation may be plotted on the vector diagram as before, and the vectors resolved in the $\sin(\omega t)$ and $\cos(\omega t)$ directions.

For quasi-linear equations 3.38, it may be shown that the vector approach is equivalent to the Kryloff and Bogoliuboff method. The terms $\ddot{x}(t)$, $\omega^2 x(t)$ represent a vector of magnitude $-\omega^2$, ω^2 respectively in the $\sin(\omega t)$ direction, the resultant of which is zero. The term $\mu f(x(t), \dot{x}(t))$ when expanded as a first harmonic Fourier series has components

$$\frac{\mu\omega}{\pi}\int_t^{\frac{2\pi}{\omega}+t} f(a\sin(\omega t), a\omega\cos(\omega t))\sin(\omega t)\,dt \qquad \text{in the } \sin(\omega t) \text{ direction}$$

$$\frac{\mu\omega}{\pi}\int_t^{\frac{2\pi}{\omega}+t} f(a\sin(\omega t), a\omega\cos(\omega t))\cos(\omega t)\,dt \qquad \text{in the } \cos(\omega t) \text{ direction}$$

(Here the time origin has been fixed so that $x(0) = 0$.) The constant term in the Fourier expansion must be zero to satisfy equation 3.38. This will impose restrictions on the nonlinearity for the existence of periodic solutions. The $\cos(\omega t)$ and $\sin(\omega t)$ components must both be zero; this corresponds to the steady state solution of the Kryloff and Bogoliuboff method, setting $\dot{a}(t)$ of equation 3.47 (equivalently the $\cos(\omega t)$ vector component) to zero and $\dot{\phi}(t)$ of equation 3.48 (equivalently the $\sin(\omega t)$ vector component) to zero.

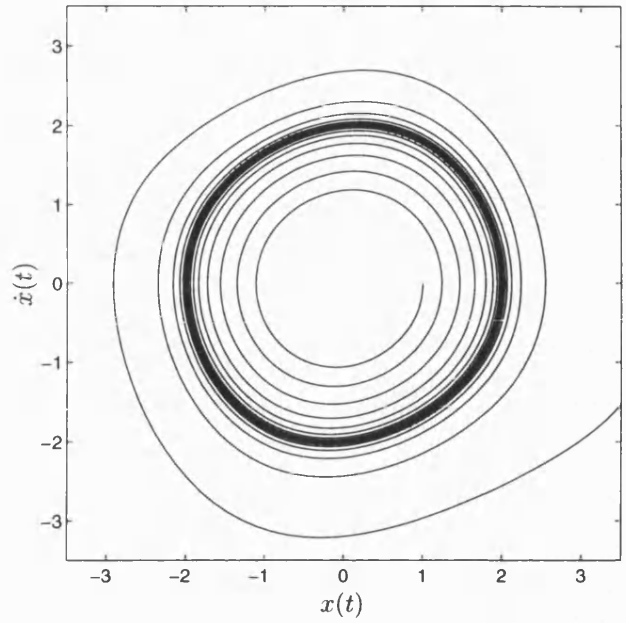


Figure 3-37: Phase Portrait for Van der Pol Equation: $\mu = 0.1$

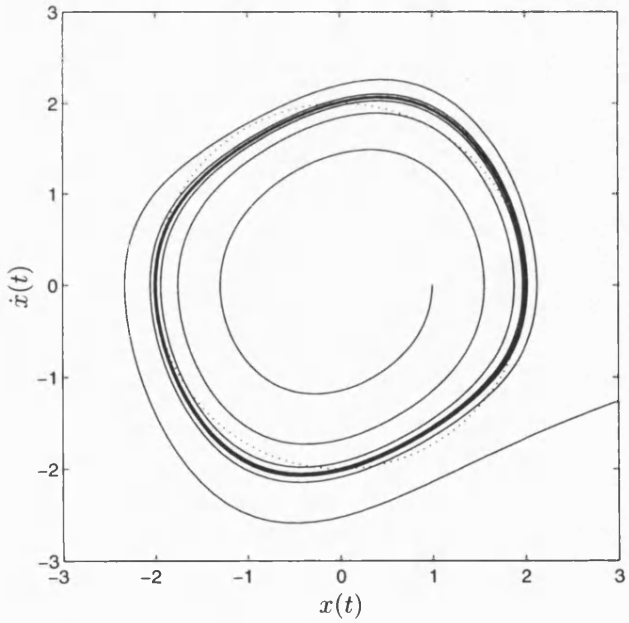


Figure 3-38: Phase Portrait for Van der Pol Equation: $\mu = 1/4$

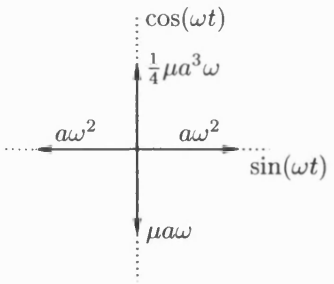


Figure 3-39: Vector Diagram for Van der Pol Equation

Example: The Van der Pol Equation

Returning to equation 3.49, the nonlinear term $\mu a^3\omega \sin^2(\omega t)\cos(\omega t)$ has Fourier series $\mu a^3\omega(\cos(\omega t) - \cos(3\omega t))/4$. So neglecting all but the first harmonic,

$$-a\omega^2 \sin(\omega t) - \mu a\omega \cos(\omega t) + \frac{1}{4}\mu a^3\omega \cos(\omega t) + a\omega^2 \sin(\omega t) = 0$$

The vector diagram is then as in figure 3-39. This gives the approximate amplitude of a limit cycle to be 2 as before. The method gives a faster alternative to the Kryloff and Bogoliuboff method for the existence and approximate amplitude of limit cycles yet the information contained in equations 3.47 and 3.48 for the transient amplitude and phase is lost.

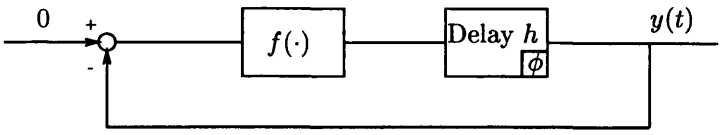


Figure 3-40: Nonlinear Delay System

3.3 Nonlinear Dynamics-Free Delay Systems

For completeness, a brief discussion is given of the form of system shown in figure 3-40. Unity negative feedback is applied to an instantaneous nonlinearity $f(\cdot)$ in series with a delay duration h in which is stored an initial function $\phi(t)$. The output is given by

$$y(t) = \begin{cases} \phi(t) & 0 \leq t < h \\ f(-y(t-h)) & t \geq h \end{cases}$$

This may be viewed in terms of a Poincaré map. Considering only the value of the initial function at $t = 0$,

$$y(0) = \phi(0)\,, \quad y(h) = f(-y(0)) = f(-\phi(0))\,, \quad y(2h) = f(-y(h)) = f(-f(-\phi(0)))$$

etc. so that the output at time values equal to multiples of h is found from the discrete map

$$y(nh) = f(-y((n-1)h))$$

Similarly the output at values of time between integer multiples n of h may be found from a map

$$y(t) = y(nh + \tau) = f(-y((n-1)h + \tau))\,, \quad 0 \leq \tau < h$$

The solution of this equation is a standard problem treated in any dynamical systems text (see for example [64] and [69]. [40] gives a general overview of chaos and [39] a treatise of dynamical systems in the framework of feedback systems.) Each point on the initial function $\phi(t)$ may be viewed as a different starting point for iteration by the map. It is possible that the map may exhibit chaotic properties, in which case it is certain that the output $y(t)$ will also, being at each point the solution of the map with different starting value. However a map not exhibiting chaos may lead to an unpredictable output.

Consider the example of the logistic equation

$$f(x) = \mu x(1-x) \tag{3.51}$$

($f(x)\epsilon[0,1]$ for all $\mu, x\epsilon\mathbb{R}$.) Adjacent points on the initial function do not necessarily remain together, even for low values of μ for which the map does not behave chaotically. For example, suppose μ is such that $(0,1)$ is in the basin of attraction of a period-4 solution of equation 3.51 ($\mu = 3.5$ say). Then adjacent points may tend to different points on this period-4 solution after each iteration.

In view of the extensive treatment elsewhere of the logistic equation, logistic systems and those without explicit dynamical element are not pursued further here. The book by Gopalsamy [16] considers delayed equations in population dynamics, in particular the delayed logistic equation. The emphasis here is on systems arising more naturally in physical and engineering control systems. Chaos may occur in control systems, as is discussed in section 5.2.

Chapter 4

Tertiary Systems

4.1 Introduction

In this chapter the three strands of chapter 3 (linear time-delay systems, delay-free nonlinear dynamic systems and instantaneous nonlinear systems with delay) are brought together for the study of nonlinear delay-differential systems. Delay together with saturation is discussed briefly in section 5.2.2. Attention has focussed on delayed Van der Pol type systems, demonstrating very rich behaviour as shall be seen.

4.2 Modified Van der Pol Systems Incorporating Delay

Research has concentrated on three examples:

1. Extended Van der Pol system with delay in state. This is described by the differential equation

$$\ddot{x}(t) - (\mu_1 - \mu_0 x^2(t-h))\dot{x}(t) + x(t-h) = 0 \tag{4.1}$$

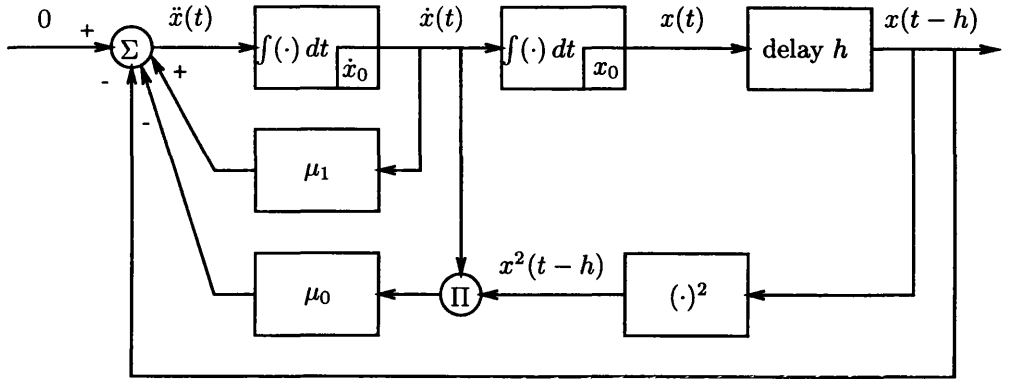
   The state has delay h and the additional parameter enables small perturbations from the corresponding linear case to be investigated.

2. Extended Van der Pol system with delay in linear state term only.

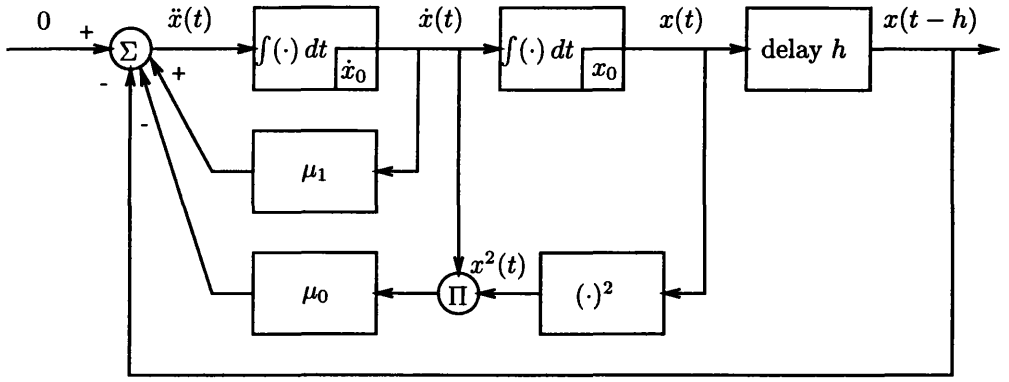
$$\ddot{x}(t) - (\mu_1 - \mu_0 x^2(t))\dot{x}(t) + x(t-h) = 0 \tag{4.2}$$

3. Extended Van der Pol system with delay in velocity term.

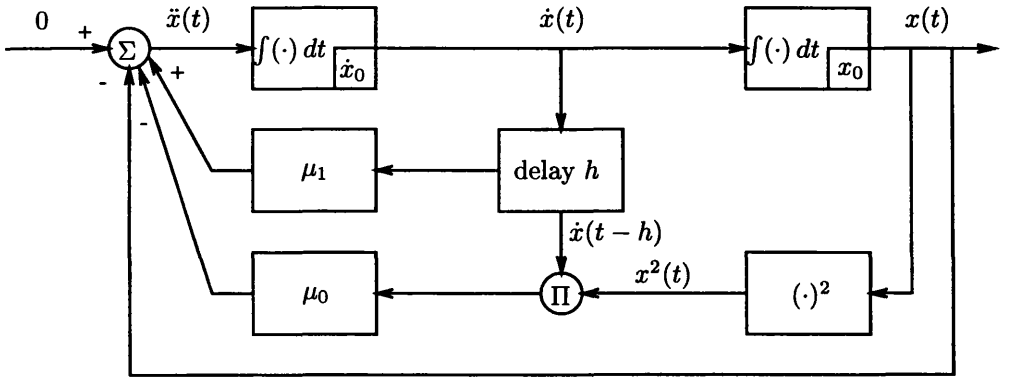
$$\ddot{x}(t) - (\mu_1 - \mu_0 x^2(t))\dot{x}(t-h) + x(t) = 0 \tag{4.3}$$



$$\ddot{x}(t) - (\mu_1 - \mu_0 x^2(t-h))\dot{x}(t) + x(t-h) = 0$$



$$\ddot{x}(t) - (\mu_1 - \mu_0 x^2(t))\dot{x}(t) + x(t-h) = 0$$



$$\ddot{x}(t) - (\mu_1 - \mu_0 x^2(t))\dot{x}(t-h) + x(t) = 0$$

Figure 4-1: Modified Van der Pol Systems with Delay

For the system representation of these three examples see figure 4-1. Initial conditions on the integrators are $\dot{x}(0) = \dot{x}_0$ and $x(0) = x_0$. Input to the systems is taken to be zero; the systems are driven by the initial conditions. The examples will now be examined in detail; extension of the features considered to similar problems will be clear.

A numerical solution of equations of this type has been formulated, employing Runge-Kutta and predictor-corrector methods, which enable comparison with the analytical results which follow. Plots of the state x and velocity $\dot{x}$ against time as well as trajectories in the 'phase' or $(x, \dot{x})$ plane are made; these illustrate behaviour described in the following sections. Details of the programs involved are to be found in appendix C.

4.2.1 Extended Van der Pol System with Delay in State

Extension of the Method of Kryloff and Bogoliuboff

Firstly the method of Kryloff and Bogoliuboff (section 3.2.10) is extended to two parameters μ_0 and μ_1 and small delays h. The analysis follows a similar pattern to before. For $\mu_0 = \mu_1 = h = 0$, the system reduces to simple harmonic motion. Therefore for small parameters, solutions are sought of form

$$\begin{aligned} x(t) &= a(t)\sin(\omega t + \phi(t)) \\ \dot{x}(t) &= a(t)\omega\cos(\omega t + \phi(t)) \end{aligned}$$

and again equation 3.41 must hold. Differentiating the $\dot{x}$ expression yields equation 3.42 for $\ddot{x}(t)$; this is substituted into the original differential equation and the analogue of equations 3.43 and 3.44 are

$$\begin{aligned} \dot{a}(t) &= \frac{1}{\omega}\cos(\omega t + \phi(t)) \\ &\quad \left(\begin{array}{l} a(t)\omega^2\sin(\omega t + \phi(t)) - a(t-h)\sin(\omega(t-h) + \phi(t-h)) + \\ (\mu_1 - \mu_0 a^2(t-h)\sin^2(\omega(t-h) + \phi(t-h)))a(t)\omega\cos(\omega t + \phi(t)) \end{array} \right) \\ \dot{\phi}(t) &= -\frac{1}{a(t)\omega}\sin(\omega t + \phi(t)) \\ &\quad \left(\begin{array}{l} a(t)\omega^2\sin(\omega t + \phi(t)) - a(t-h)\sin(\omega(t-h) + \phi(t-h)) + \\ (\mu_1 - \mu_0 a^2(t-h)\sin^2(\omega(t-h) + \phi(t-h)))a(t)\omega\cos(\omega t + \phi(t)) \end{array} \right) \end{aligned}$$

So far these are exact equations. Note that the frequency ω is approximately unity since if μ_0, μ_1 and h are all zero, the period of oscillation is 2π. The derivatives $\dot{a}(t)$

and $\dot{\phi}(t)$ are composed of terms proportional to μ_0, μ_1 (small) and the term

$$a(t)\omega^2 \sin(\omega t + \phi(t)) - a(t-h)\sin(\omega(t-h) + \phi(t-h))$$

which is also small for small delay. So $\dot{a}(t)$ and $\dot{\phi}(t)$ are small and slowly varying functions of time. As before, these derivatives are replaced by their average values over a period $T = 2\pi/\omega$, taking the amplitude a and phase ϕ to be constant during averaging. (The dummy variable q is introduced for this purpose.) The equations of the first approximation result:

$$\begin{aligned}
\dot{a}(t) &= \frac{1}{2\pi}\int_0^{\frac{2\pi}{\omega}} \cos(\omega q + \phi(t)) \\
&\quad \left(\begin{array}{l} a(t)\omega^2 \sin(\omega q + \phi(t)) - a(t-h)\sin(\omega(q-h) + \phi(t-h)) + \\ (\mu_1 - \mu_0 a^2(t-h)\sin^2(\omega(q-h) + \phi(t-h)))a(t)\omega\cos(\omega q + \phi(t)) \end{array} \right) dq \\
\dot{\phi}(t) &= -\frac{1}{2\pi a(t)}\int_0^{\frac{2\pi}{\omega}} \sin(\omega q + \phi(t)) \\
&\quad \left(\begin{array}{l} a(t)\omega^2 \sin(\omega q + \phi(t)) - a(t-h)\sin(\omega(q-h) + \phi(t-h)) + \\ (\mu_1 - \mu_0 a^2(t-h)\sin^2(\omega(q-h) + \phi(t-h)))a(t)\omega\cos(\omega q + \phi(t)) \end{array} \right) dq
\end{aligned}$$

Various integrals are required for solution. Details of the calculations are not given here.

- $\displaystyle\int_0^{\frac{2\pi}{\omega}} \cos^2(\omega q + \phi(t))\, dq = \int_0^{\frac{2\pi}{\omega}} \sin^2(\omega q + \phi(t))\, dq = \frac{\pi}{\omega}$
- $\displaystyle\int_0^{\frac{2\pi}{\omega}} \cos(\omega q + \phi(t))\sin(\omega(q-h) + \phi(t-h))\, dq = \frac{\pi}{\omega}\sin(\phi(t-h) - \phi(t) - \omega h)$
- $\displaystyle\int_0^{\frac{2\pi}{\omega}} \sin(\omega q + \phi(t))\sin(\omega(q-h) + \phi(t-h))\, dq = \frac{\pi}{\omega}\cos(\phi(t-h) - \phi(t) - \omega h)$
- $\displaystyle\int_0^{\frac{2\pi}{\omega}} \cos^2(\omega q + \phi(t))\sin^2(\omega(q-h) + \phi(t-h))\, dq =$
$$\frac{\pi}{4\omega}(2 - \cos(2(\phi(t-h) - \phi(t) - \omega h)))$$
- $\displaystyle\int_0^{\frac{2\pi}{\omega}} \sin(\omega q + \phi(t))\cos(\omega q + \phi(t))\sin^2(\omega(q-h) + \phi(t-h))\, dq =$
$$\frac{\pi}{4\omega}\sin(2(\phi(t-h) - \phi(t) - \omega h))$$

Using these integral results,

$$\dot{a}(t) = \frac{1}{2}\mu_1 a(t) - \frac{1}{2\omega}a(t-h)\sin(\phi(t-h) - \phi(t) - \omega h)$$

$$-\frac{1}{8}\mu_0 a(t)a^2(t-h)(2-\cos(2(\phi(t-h)-\phi(t)-\omega h))) \tag{4.4}$$

$$\begin{aligned}\dot{\phi}(t) &= -\frac{\omega}{2}+\frac{1}{2\omega}\frac{a(t-h)}{a(t)}\cos(\phi(t-h)-\phi(t)-\omega h)\\ &\quad +\frac{1}{8}a^2(t-h)\sin(2(\phi(t-h)-\phi(t)-\omega h))\end{aligned} \tag{4.5}$$

These equations do not admit analytical solution in the general case. In the steady state, $\dot{a}(t) = 0 = \dot{\phi}(t)$, so $a(t) = a = a(t-h)$ and $\phi(t) = \phi = \phi(t-h)$ and from the above equations,

$$\begin{aligned}a(t)\left(\mu_0 a^2(2-\cos(2\omega h))-4\left(\mu_1+\frac{1}{\omega}\sin(\omega h)\right)\right) &= 0\\ \mu_0 a^2\sin(2\omega h)+4\left(\omega-\frac{1}{\omega}\cos(\omega h)\right) &= 0\end{aligned}$$

Taking the first of these, $a = 0$ (trivial solution), or, provided $\mu_0 > 0$,

$$a^2 = \frac{4\left(\mu_1+\frac{1}{\omega}\sin(\omega h)\right)}{\mu_0(2-\cos(2\omega h))} \tag{4.6}$$

and from the second,

$$a^2 = \frac{4\left(\frac{1}{\omega}\cos(\omega h)-\omega\right)}{\mu_0\sin(2\omega h)} \tag{4.7}$$

(If $\mu_0 = 0$, the problem is a linear one and

$$\mu_1\omega+\sin(\omega h) = 0 \quad \text{and} \quad \omega^2-\cos(\omega h) = 0$$

so that $\omega^4+\mu_1^2\omega^2-1 = 0$, which is in fact the W-polynomial for this system. If $\sin(2\omega h) = 0$, either $\cos(\omega h) = 0$, $\omega = 0$ implying infinite delay, or $\sin(\omega h) = 0$ so that $\omega^2 = \cos(\omega h) = 1$, $\omega = 1$, $h = 2q\pi$ and $a^2 = 4\mu_1/\mu_0$.)

The two equations 4.6 and 4.7 give the amplitude a in terms of the known parameters μ_0, μ_1 and h and the unknown frequency ω. They may be solved for ω:

$$\frac{\mu_1+\frac{1}{\omega}\sin(\omega h)}{1+2\sin^2(\omega h)} = \frac{\frac{1}{\omega}\cos(\omega h)-\omega}{2\sin(\omega h)\cos(\omega h)}$$

$$\omega^2(1+2\sin^2(\omega h))+2\mu_1\omega\sin(\omega h)\cos(\omega h)-\cos(\omega h) = 0$$

an equation which may be solved numerically for ω, given μ_1 and h. Hence the steady state amplitude may be determined.

Sensitivity of the Limit Cycle with respect to μ_0, μ_1 and h

The parameter μ_0 in equation 4.1 represents positive damping. From equations 4.6 and 4.7 it may be seen that increasing the value of μ_0 corresponds to a decrease in the amplitude of the limit cycle and the parameter has a stabilising influence as expected. $-\mu_1\dot{x}(t)$ is a destabilising term and increasing μ_1 corresponds to an increase in limit cycle amplitude. Increasing μ_0 from zero gives a limit cycle moving in from infinity; increasing μ_1 from zero gives a limit cycle increasing in amplitude from the origin (when $h = 0$). From analysis of the linearised system later, increasing the delay from zero increases the instability of the system. It may be verified that $\partial a/\partial h$ is positive for small μ_1 and h, i.e. increasing the delay has the destabilising effect of increasing the amplitude of the limit cycle.

The Linearised System

In the neighbourhood of the origin in $(x, \dot{x})$ space the term $\mu_0 x^2(t-h)\dot{x}(t)$ will be small and motion and stability in this neighbourhood is governed by the linearised system

$$\ddot{x}(t) - \mu_1\dot{x}(t) + x(t-h) = 0$$

(Compare with the first method of Lyapunov in section 3.2.7.) Stability may be determined by the direct method for the characteristic equation

$$s^2 - \mu_1 s + \exp(-sh) = 0$$

This system was considered in section 3.1.6 and is unstable for $h = 0$ for all non-negative μ_1 with one destabilising value of ω. There is instability for all values of delay. This implies that the origin is an unstable fixed point of the delayed nonlinear differential equation and that all trajectories are repelled from the origin. This suggests that the limit cycle should it exist (for high values of μ_1 and h say) will be stable.

The Vector Approach

Again it is shown that the vector approach is (a much faster!) equivalent to the method of Kryloff and Bogoliuboff for systems of form 4.1, 4.2 and 4.3. The nonlinear terms in these three equations are such that their integral over a period is zero (for $x = a\sin(\omega t)$, $\dot{x} = a\omega\cos(\omega t)$) and therefore their Fourier series has zero constant term as shall be seen.

Returning to equation 4.1 and seeking periodic solutions of form $x(t) = a\sin(\omega t)$,

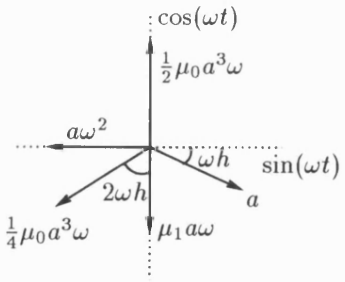


Figure 4-2: Vector Diagram for $\ddot{x}(t) - (\mu_1 - \mu_0 x^2(t-h))\dot{x}(t) + x(t-h) = 0$

$\dot{x}(t) = a\omega\cos(\omega t)$, the nonlinear term when expanded as a Fourier series becomes

$$\begin{aligned}
\mu_0 x^2(t-h)\dot{x}(t) &= \mu_0 a^3\omega \sin^2(\omega(t-h))\cos(\omega t) \\
&= \frac{1}{2}\mu_0 a^3\omega(1 - \cos(2\omega(t-h))\cos(\omega t) \\
&= \frac{1}{2}\mu_0 a^3\omega\left(\cos(\omega t) - \frac{1}{2}(\cos(3\omega t - 2\omega h) + \cos(\omega t - 2\omega h))\right)
\end{aligned}$$

and retaining only the first harmonic,

$$\mu_0 x^2(t-h)\dot{x}(t) \approx \frac{1}{2}\mu_0 a^3\omega\cos(\omega t) - \frac{1}{4}\mu_0 a^3\omega\cos(\omega t - 2\omega h)$$

The vector diagram is shown in figure 4-2. Resolving in the $\sin(\omega t)$ direction,

$$\omega^2 + \frac{1}{4}\mu_0 a^2\omega\sin(2\omega h) = \cos(\omega h) \tag{4.8}$$

Then for $\mu_0 \neq 0 \neq \sin(2\omega h)$,

$$a^2 = \frac{4\left(\frac{1}{\omega}\cos(\omega h) - \omega\right)}{\mu_0\sin(2\omega h)}$$

In the $\cos(\omega t)$ direction

$$\frac{1}{2}\mu_0 a^2\omega = \frac{1}{4}\mu_0 a^2\omega\cos(2\omega h) + \mu_1\omega + \sin(\omega h) \tag{4.9}$$

and for $\mu_0 \neq 0$,

$$a^2 = \frac{4\left(\mu_1 + \frac{1}{\omega}\sin(\omega h)\right)}{\mu_0(2 - \cos(2\omega h))}$$

These two expressions for the steady state amplitude are equivalent to equations 4.6 and 4.7 obtained by the extension of the Kryloff and Bogoliuboff method. In general,

for

$$\ddot{x}(t) + f(x(t), x(t-h), \dot{x}(t), \dot{x}(t-h), \mu_0, \mu_1) + x(t) = 0$$

or

$$\ddot{x}(t) + f(\mu_0, \mu_1, x(t), x(t-h), \dot{x}(t), \dot{x}(t-h)) + x(t-h) = 0$$

combining discussion of sections 3.1.8 and 3.2.11 it may be shown that resolving in the $\sin(\omega t)$, or $\cos(\omega t)$, direction is equivalent to equating the Kryloff and Bogoliuboff expression for $\dot{\phi}(t)$, or $\dot{a}(t)$, respectively to zero at the steady state. The disadvantage of the vector approach is the loss of information on the transient amplitude and phase contained in equations 4.4 and 4.5.

Special Cases and Examples

1. **Extended Delay-Free Van der Pol Equation**

$$\ddot{x}(t) - (\mu_1 - \mu_0 x^2(t))\dot{x}(t) + x(t) = 0 \tag{4.10}$$

Here $\mu_0 > 0$, $\mu_1 > 0$ and $h = 0$. Equations 4.4 and 4.5 become

$$\begin{aligned} \dot{a}(t) &= \frac{1}{2}\mu_1 a(t) - \frac{1}{8}\mu_0 a^3(t) \\ \dot{\phi}(t) &= \frac{1}{2}\left(\frac{1}{\omega} - \omega\right) \end{aligned}$$

with solution

$$\begin{aligned} a(t) &= \frac{a_0 \exp(\mu_1 t/2)}{\sqrt{1 + \frac{\mu_0}{4\mu_1} a_0^2 (\exp(\mu_1 t) - 1)}} \\ \phi(t) &= \frac{1}{2}\left(\frac{1}{\omega} - \omega\right) t + \phi_0 \end{aligned}$$

where $a_0 = a(0)$ and $\phi_0 = \phi(0)$, and steady state solution

$$a^2 = \frac{4\mu_1}{\mu_0}, \qquad \omega = 1$$

The amplitude grows with time to the steady state for $0 < a_0 < 2\sqrt{\mu_1/\mu_0}$ and decays with time for $a_0 > 2\sqrt{\mu_1/\mu_0}$. Finally

$$\tilde{x}(t) = a(t)\sin(\omega t) = \frac{a_0 \exp(\mu_1 t/2)\sin(\omega t)}{\sqrt{1 + \frac{\mu_0}{4\mu_1} a_0^2 (\exp(\mu_1 t) - 1)}} \tag{4.11}$$

$$\tilde{\dot{x}}(t) \;=\; a(t)\omega\cos(\omega t) = \frac{a_0\omega\exp(\mu_1 t/2)\cos(\omega t)}{\sqrt{1+\frac{\mu_0}{4\mu_1}a_0^2(\exp(\mu_1 t)-1)}} \tag{4.12}$$

$\tilde{x}(t)$ and $\tilde{\dot{x}}(t)$ denote the approximate expressions for state and velocity. Figures 4-3 and 4-4 illustrate the effects of the parameters μ_0 and μ_1 in the equation. Figure 4-5 compares the state and velocity with time of the numerical solution with the extension of Kryloff and Bogoliuboff's method in equations 4.11 and 4.12; it may be seen that there is very good agreement for $\mu_0 = \mu_1 = 0.1$.

The equation 4.10 may be transformed to standard Van der Pol form by the substitution

$$x(t) = \sqrt{\frac{\mu_1}{\mu_0}}\, y(t)$$

to give

$$\ddot{y}(t) - \mu_1(1-y^2(t))\dot{y}(t) + y(t) = 0$$

Then by the expansions 3.37, curtailing these series at the number of terms given,

$$x(t) = \sqrt{\frac{\mu_1}{\mu_0}}\left(2\cos(\omega t) + \frac{\mu_1}{4(1+\mu_1)}(3\sin(\omega t) - \sin(3\omega t)) + \frac{\mu_1^2}{4(1+\mu_1)^2}\left(-\frac{1}{2}\cos(\omega t) + \frac{3}{4}\cos(3\omega t) + \frac{5}{24}\cos(5\omega t) + 3\sin(\omega t) - \sin(3\omega t)\right)\right)$$

$$\omega = \frac{1}{1+\mu_1} + \frac{\mu_1}{(1+\mu_1)^2} + \frac{15}{16}\frac{\mu_1^2}{(1+\mu_1)^3} + \frac{13}{16}\frac{\mu_1^3}{(1+\mu_1)^4}$$

Note that ω is independent of μ_0. Thus results may be compared with those of equation 4.11 for small μ_0 and μ_1. Figure 4-6 gives a comparison of the numerically calculated state with that given by the series above for $\mu_0 = 0.1$, $\mu_1 = 1$. (One is a delayed version of the other as the series expression gives the approximate steady state solution from $t = 0$.)

The differential equation may be rewritten

$$\frac{d\dot{x}}{dx} = -\frac{x}{\dot{x}} + \mu_1 - \mu_0 x^2$$

The gradient $d\dot{x}/dx$ gives the direction of flow in the $(x, \dot{x})$ plane. Here this represents a circle centre the origin with perturbation in gradient of $\mu_1 - \mu_0 x^2$; μ_1 is destabilising and μ_0 stabilising as expected. The direction of flow across circles centre the origin of radius 1, 2 and 3 is shown in figure 4-7, along with trajectories, for $\mu_0 = \mu_1 = 1/4$ (being the same example as that of figure 3-38).

2. Linear Equation with Delay

$$\ddot{x}(t) - \mu_1 \dot{x}(t) + x(t-h) = 0$$

Here $\mu_0 = 0$, $\mu_1 > 0$ and $h > 0$. This equation has already been discussed; it is trivially unstable for $\mu_1 > 0$. The case $\mu_1 < 0$ is stable for zero delay and has one destabilising value of ω at $\sqrt{\frac{1}{2}(-\mu_1^2 + \sqrt{\mu_1^4 + 4})}$ with corresponding $h = -\frac{1}{\omega}\tan^{-1}(\frac{\mu_1}{\omega})$ from linear theory. Alternatively, the method of Kryloff and Bogoliuboff, which is exact in this linear case, yields the same results for critical ω and h (being equivalent to the vector approach which in turn is equivalent to equating the real and imaginary parts of the characteristic equation to zero). The above problem has been treated in detail in [33], in particular the interchange of the damping parameter μ_1 with the delay h to maintain critical stability, again the theme of balancing stabilising and destabilising influences. Note that critical stability in this case corresponds to a centre in the $(x, \dot{x})$ plane rather than a limit cycle.

3. The General Equation

Returning to equation 4.1, it remains to examine the flows in the $(x, \dot{x})$ plane. The delay is required to be small; flows on the approximate predicted limit cycle $x(t) = a\sin(\omega t)$, $\dot{x}(t) = a\omega\cos(\omega t)$ may be sought or the delay terms expanded as a Taylor series retaining terms to second order in parameters μ_0, μ_1 and h. On the approximate limit cycle,

$$x(t-h) = x(t)\cos(\omega h) - \frac{1}{\omega}\dot{x}(t)\sin(\omega h)$$

Then

$$\frac{d\dot{x}}{dx} = -\frac{x}{\dot{x}}\cos(\omega h) + \frac{1}{\omega}\sin(\omega h) + \mu_1 - \mu_0\left(x\cos(\omega h) - \frac{1}{\omega}\dot{x}\cos(\omega h)\right)^2$$

Alternatively, replacing $x(t-h)$ by

$$x(t) - h\dot{x}(t) + \frac{h^2}{2}\ddot{x}(t) + \mathcal{O}(h^3)$$

$$x^2(t-h) = x^2(t) - 2hx(t)\dot{x}(t) + h^2\dot{x}(t) + h^2x(t)\ddot{x}(t) + \mathcal{O}(h^3)$$

then to $\mathcal{O}(h^3)$,

$$\dot{x}\frac{d\dot{x}}{dx} = \left(\mu_1 - \mu_0\left(x^2 - 2hx\dot{x} + h^2\left(\dot{x}^2 + x\dot{x}\frac{d\dot{x}}{dx}\right)\right)\right)\dot{x} - x + h\dot{x} - \frac{h^2}{2}\dot{x}\frac{d\dot{x}}{dx}$$

and

$$\frac{d\dot{x}}{dx} = \frac{1}{1+\frac{h^2}{2}}\left(-\frac{x}{\dot{x}} + \mu_1 + h - \mu_0 x^2 + 2\mu_0 h x\dot{x}\right)$$

to second order in small parameters μ_0, μ_1, h. The second formulation has the advantage of describing the approximate flows anywhere in the $(x, \dot{x})$ plane.

Figure 4-8 shows the trajectories beginning inside and outside the stable limit cycle when $\mu_0 = \mu_1 = h = 0.1$. The extension of Kryloff and Bogoliuboff's method predicts a limit cycle of amplitude 2.801; this is shown by a dotted circle. The effect of varying the parameters μ_0 and μ_1 will be similar to that observed in the delay-free case. Distortion of the steady-state sinusoidal solution produced by large delays is illustrated in figure 4-9 (state and velocity against time) and figure 4-10 (the $(x, \dot{x})$ plane) again for $\mu_0 = \mu_1 = 0.1$. The limit cycle remains stable as expected.

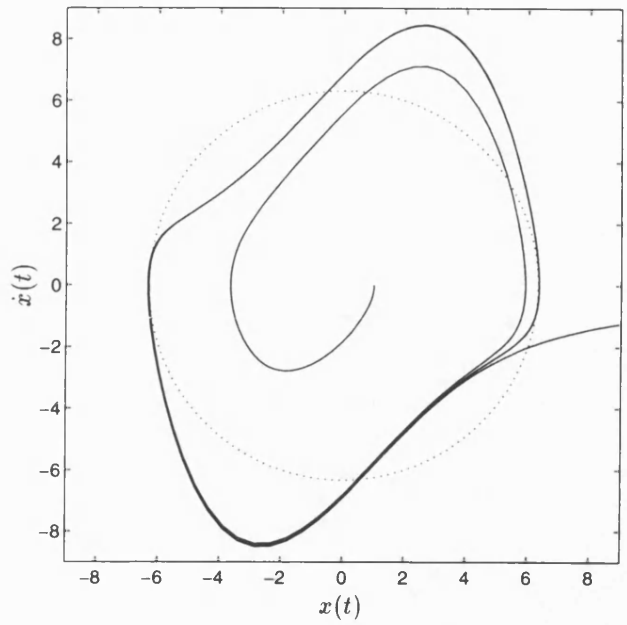


Figure 4-3: Trajectories for $\ddot{x}(t) - (\mu_1 - \mu_0 x^2(t))\dot{x}(t) + x(t) = 0, \quad \mu_0 = 0.1, \mu_1 = 1$

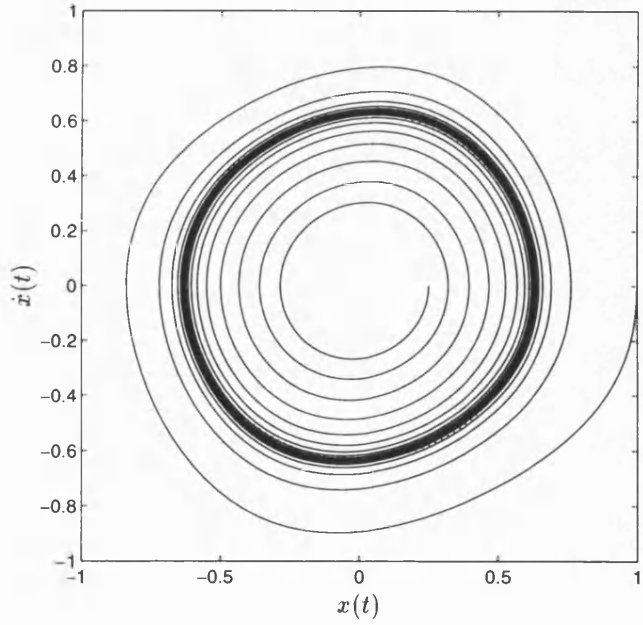


Figure 4-4: Trajectories for $\ddot{x}(t) - (\mu_1 - \mu_0 x^2(t))\dot{x}(t) + x(t) = 0, \quad \mu_0 = 1, \mu_1 = 0.1$

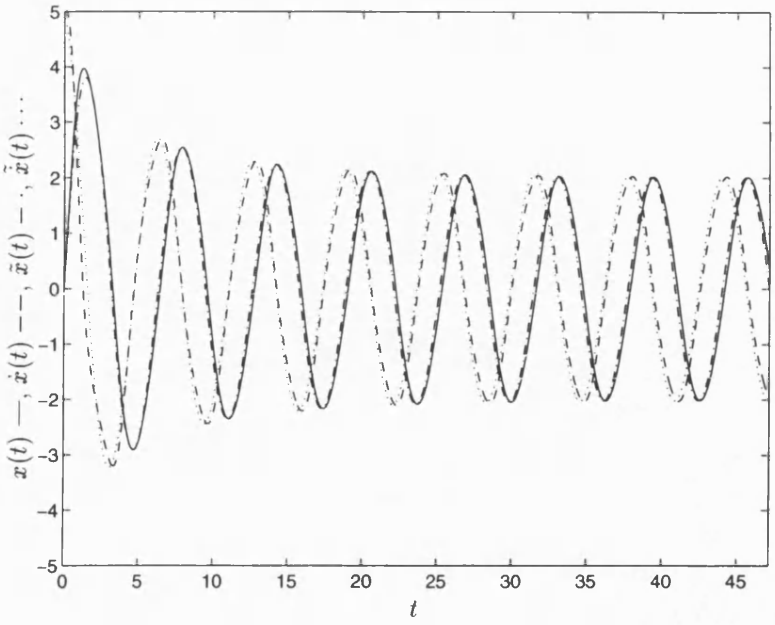


Figure 4-5: State and Velocity Comparisons for $\ddot{x}(t) - (\mu_1 - \mu_0 x^2(t))\dot{x}(t) + x(t) = 0$, $\mu_0 = 0.1$, $\mu_1 = 0.1$

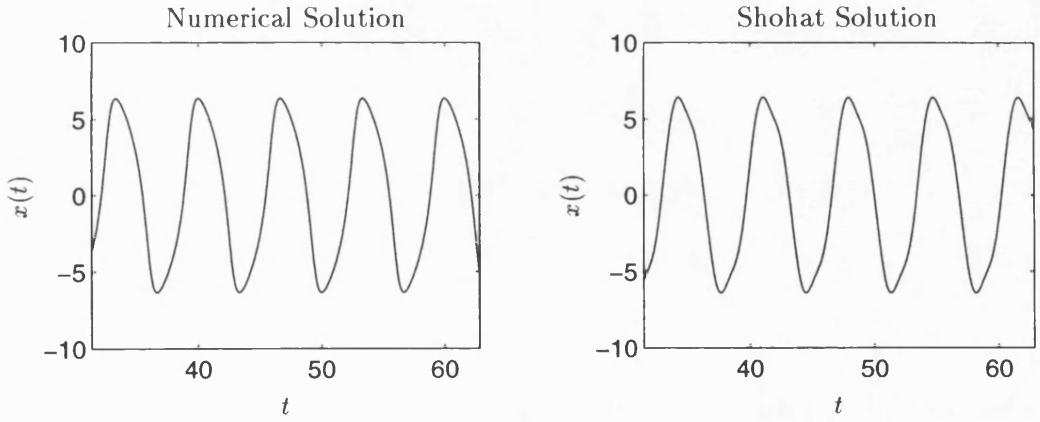


Figure 4-6: Comparison of Numerical and Shohat Results for $\ddot{x}(t) - (\mu_1 - \mu_0 x^2(t))\dot{x}(t) + x(t) = 0$, $\mu_0 = 0.1$, $\mu_1 = 1$

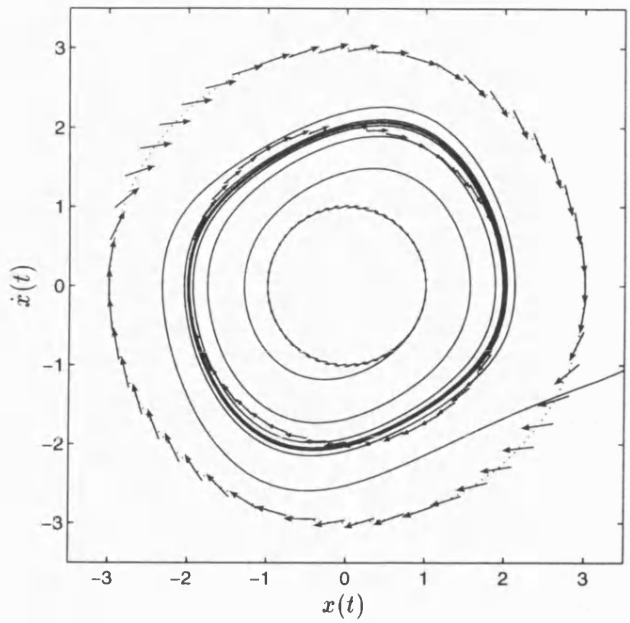


Figure 4-7: Trajectories and Flows for $\ddot{x}(t) - (\mu_1 - \mu_0 x^2(t))\dot{x}(t) + x(t) = 0$, $\mu_0 = 1/4$, $\mu_1 = 1/4$

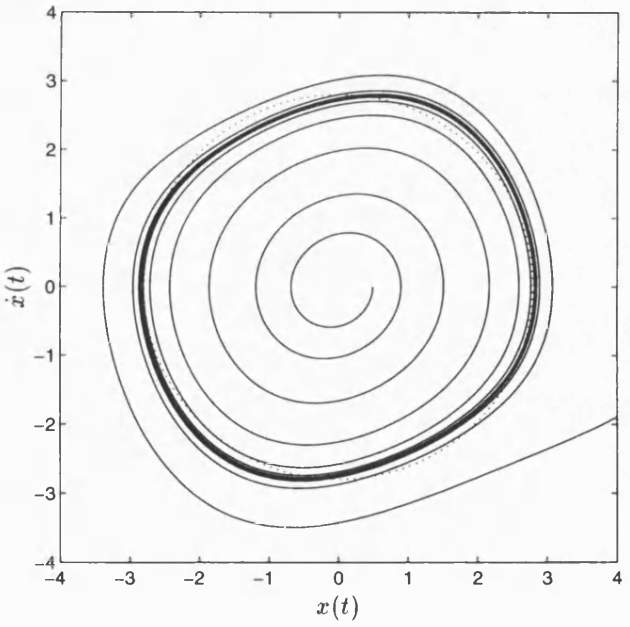


Figure 4-8: Trajectories for $\ddot{x}(t) - (\mu_1 - \mu_0 x^2(t-h))\dot{x}(t) + x(t-h) = 0$, $\mu_0 = \mu_1 = h = 0.1$

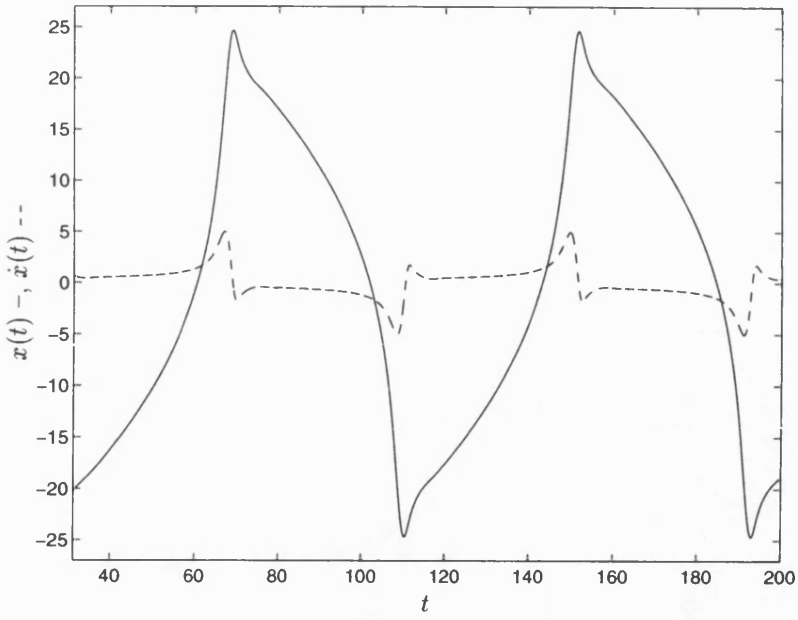


Figure 4-9: State and Velocity for $\ddot{x}(t) - (\mu_1 - \mu_0 x^2(t-h))\dot{x}(t) + x(t-h) = 0$, $\mu_0 = \mu_1 = 0.1, \quad h = 2\pi$

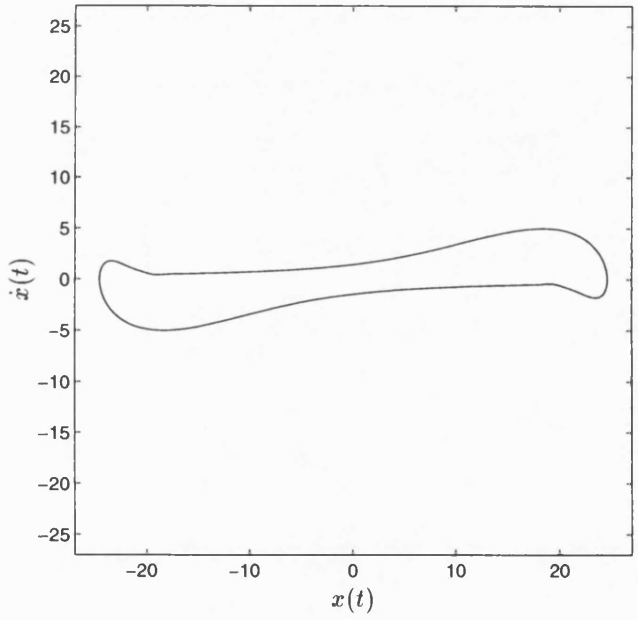


Figure 4-10: Periodic Solution for $\ddot{x}(t) - (\mu_1 - \mu_0 x^2(t-h))\dot{x}(t) + x(t-h) = 0$, $\mu_0 = \mu_1 = 0.1, \quad h = 2\pi$

4.2.2 Extended Van der Pol System with Delay in Linear State Term

This equation has been the subject of research by Grafton [17], [18], Hale [21] and later Hale and Lunel [22]. The paper [17] gives a proof of a theorem giving sufficient conditions for the existence of a nontrivial periodic solution, of period greater than $2h$, for equations of form

$$\ddot{x}(t) + f(x(t))\dot{x}(t) + g(x(t-h)) = 0$$

where the delay h is strictly positive. The equation in question here is clearly a special case of this. The conditions of the theorem are as follows.

1. $f(x) = -\varepsilon + f_1(x)$ is continuous for all x with $f_1(0) = 0$, $F(x) = \int_0^x f(x)\,dx \to \infty(-\infty)$ monotonically as $x \to \infty(-\infty)$.
2. $g(x) = x(1+g_1(x))$ is nondecreasing, $g(x)$ is continuous, $g_1(0) = 0$ and $xg(x) > 0$ for all nonzero x.
3. $F^{-1}(y)g(F^{-1}(y))/y \to 0$ as $|y| \to \infty$.
4. If $0 < \omega < \pi/2h$ solves $\omega^2 = \cos(\omega h)$ then $\varepsilon\omega + \sin(\omega h) > 0$.

Conditions 1 and 2 are clearly satisfied by equation 4.2 and are similar to those required in the delay-free case (see [31] for example). Condition 4 ensures that at least one pair of roots of the linearised system are located in the right-half complex plane and hence that the origin is unstable. A quicker and more elegant alternative to this would be to apply the direct method to the equation in question (compare condition 4 with the real and imaginary parts of the characteristic equation). Condition 3 is necessary in the proof in place of some boundedness condition on solutions (conditions 1 and 2 being sufficient in the delay-free case).

The Vector Approach

The approximate limit cycle amplitude and frequency may be determined by an extension of Kryloff and Bogoliuboff's method following the pattern and integral results of section 4.2.1; here the equivalent vector method is used. (The first order differential equations for $a(t)$ and $\phi(t)$ do not admit analytical solution due to the delay present so no information is lost with the alternative vector method.) In this case the nonlinear term when evaluated at $x(t) = a\sin(\omega t)$, $\dot{x}(t) = a\omega\cos(\omega t)$ and expanded as a Fourier series becomes

$$\frac{\mu_0 a^3 \omega}{4}(\cos(\omega t) - \cos(3\omega t))$$

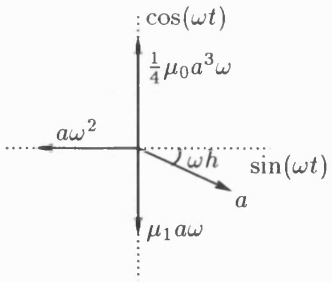


Figure 4-11: Vector Diagram for $\ddot{x}(t) - (\mu_1 - \mu_0 x^2(t))\dot{x}(t) + x(t-h) = 0$

so that, neglecting all but the first harmonic, the equation to be satisfied is

$$-a\omega^2 \sin(\omega t) + a\omega \left(\frac{\mu_0 a^2}{4} - \mu_1\right) \cos(\omega t) + a \sin(\omega(t-h)) = 0$$

The vector diagram appears in figure 4-11 and resolving in the $\sin(\omega t)$ and $\cos(\omega t)$ directions yields

$$\omega^2 = \cos(\omega h) \qquad \mu_1 \omega + \sin(\omega h) = \frac{\mu_0 a^2 \omega}{4}$$

and provided $\mu_0 \neq 0$,

$$a^2 = \frac{4}{\mu_0}\left(\mu_1 + \frac{1}{\omega}\sin(\omega h)\right)$$

The frequency is independent of the parameters μ_0 and μ_1; the amplitude grows with the destabilising μ_1 and shrinks with the stabilising μ_0 as expected. For very small h, $\frac{1}{\omega}\sin(\omega h) \approx h$ and again the amplitude grows with the introduction of the destabilising delay.

Figures 4-12 and 4-13 show state and velocity plots against time and the limit cycle in the $(x, \dot{x})$ plane for parameter values $\mu_0 = \mu_1 = h = 0.1$. The predicted amplitude of $\sqrt{\frac{4}{\mu_0}\left(\mu_1 + \frac{1}{\omega}\sin(\omega h)\right)} = 2.8725$ is shown by the dotted circle in figure 4-13. The corresponding plots for $h = 7$ are given in figures 4-14 and 4-15. As expected the limit cycle has grown in average amplitude and remains stable (see below).

The Linearised System

The linearised system is given by

$$\ddot{x}(t) - \mu_1 \dot{x}(t) + x(t-h) = 0$$

which is equivalent to the linearised system of the equation with delay in all state terms. Thus the origin of the $(x, \dot{x})$ plane is unstable for all values of delay. Again this suggests that the limit cycle is stable and attracts nearby trajectories.

The Expansions of Shohat

As there is no (approximate) analytical solution available for $x(t)$, the series expansions of Shohat are of particular interest. However the delay term again causes difficulties and it is necessary to restrict attention to small h (though large μ_0, μ_1 are still acceptable) in order to transform the retarded differential equation to standard Van der Pol form. The delay term is expanded as a Taylor series to $\mathcal{O}(h^3)$. Then

$$\ddot{x}\left(1+\frac{h^2}{2}\right) - (\mu_1 + h - \mu_0 x^2(t))\dot{x}(t) + x(t) = 0$$

to $\mathcal{O}(h^3)$. So

$$\ddot{x}\left(1+\frac{h^2}{2}\right) - (\mu_1 + h)\left(1 - \frac{\mu_0}{\mu_1 + h}x^2(t)\right)\dot{x}(t) + x(t) = 0$$

The transformation

$$x(t) = \sqrt{\frac{\mu_1 + h}{\mu_0}}\, y\left(\frac{1}{\sqrt{1+\frac{h^2}{2}}}\, t\right)$$

is made. Then

$$\sqrt{\frac{\mu_1 + h}{\mu_0}}\,\ddot{y} - (\mu_1 + h)\sqrt{\frac{\mu_1 + h}{\mu_0(1+\frac{h^2}{2})}}\,(1 - y^2)\dot{y} + \sqrt{\frac{\mu_1 + h}{\mu_0}}\, y = 0$$

and finally

$$\ddot{y} - \frac{\mu_1 + h}{\sqrt{1+\frac{h^2}{2}}}\,(1 - y^2)\dot{y} + y = 0$$

The series expansions 3.37 then become

$$x(t) = \sqrt{\frac{\mu_1 + h}{\mu_0}}\left(2\cos(\omega t) + \frac{\beta}{4(1+\beta)}(3\sin(\alpha t) - \sin(3\alpha t))\right.$$
$$\left. + \frac{\beta^2}{4(1+\beta)^2}\left(-\frac{1}{2}\cos(\alpha t) + \frac{3}{4}\cos(3\alpha t) + \frac{5}{24}\cos(5\alpha t) + 3\sin(\alpha t) - \sin(3\alpha t)\right) + \cdots\right)$$
$$\alpha = \frac{1}{\sqrt{1+\frac{h^2}{2}}}\left(\frac{1}{1+\beta} + \frac{\beta}{(1+\beta)^2} + \frac{15}{16}\frac{\beta^2}{(1+\beta)^3} + \frac{13}{16}\frac{\beta^3}{(1+\beta)^4} + \cdots\right)$$

where

$$\beta = \frac{\mu_1 + h}{\sqrt{1 + \frac{h^2}{2}}}$$

Figure 4-16 compares the numerical output with the value of state given by the series expansion for small parameters $\mu_0 = \mu_1 = h = 0.1$. (Again translation in time is necessary to account for the settling of the numerical solution to the steady state.) A phase plane comparison is not given due to the inaccuracy of the expression for $\dot{x}(t)$ found by differentiating the series expansion above (see earlier comment).

Such a Taylor method is further complicated if delay appears in one of the nonlinear terms when the Taylor expansion must be truncated to even lower order in h to maintain a second order differential equation.

Flows

For flows in the $(x, \dot{x})$ plane, again h is required to be small. Two approaches may be taken, either seeking flows solely on the approximate limit cycle predicted by the extension of Kryloff and Bogoliuboff's method, as for the previous example, or by expanding the delay term as a Taylor series to $\mathcal{O}(h^3)$. For the former method, $x(t-h)$ is approximated by $x(t)\cos(\omega h) - \frac{1}{\omega}\dot{x}(t)\sin(\omega h)$ and the differential equation becomes

$$\frac{d\dot{x}}{dx} = -\frac{x}{\dot{x}}\cos(\omega h) + \mu_1 - \mu_0 x^2 + \frac{1}{\omega}\sin(\omega h)$$

Using a Taylor expansion,

$$\frac{d\dot{x}}{dx} = \frac{1}{1 + \frac{h^2}{2}}\left(-\frac{x}{\dot{x}} + \mu_1 - \mu_0 x^2 + h\right)$$

Figure 4-17 shows the trajectories commencing outside and inside the limit cycle, again for the example $\mu_0 = \mu_1 = h = 0.1$, together with arrows representing the flow across circles inside, outside and at the predicted limit cycle amplitude of $a = 2.8725$, found using the Taylor method.

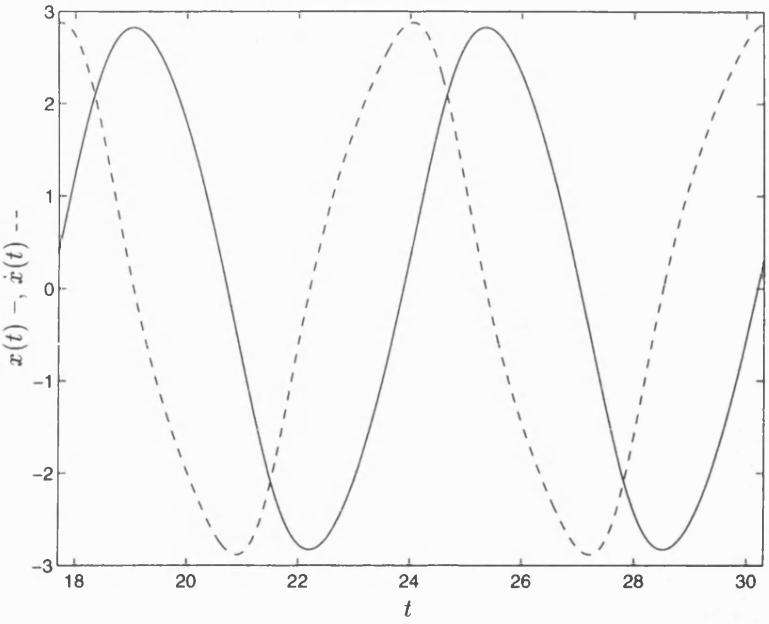


Figure 4-12: State and Velocity for $\ddot{x}(t) - (\mu_1 - \mu_0 x^2(t))\dot{x}(t) + x(t-h) = 0$, $\mu_0 = \mu_1 = h = 0.1$

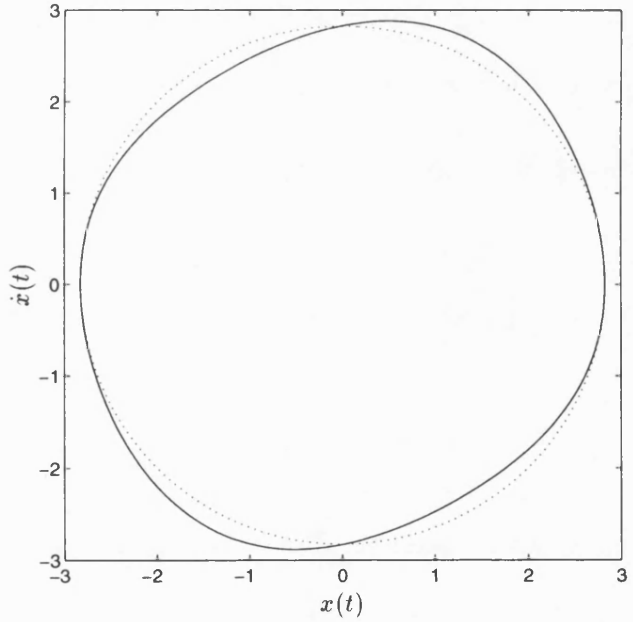


Figure 4-13: Limit Cycle for $\ddot{x}(t) - (\mu_1 - \mu_0 x^2(t))\dot{x}(t) + x(t-h) = 0$, $\mu_0 = \mu_1 = h = 0.1$

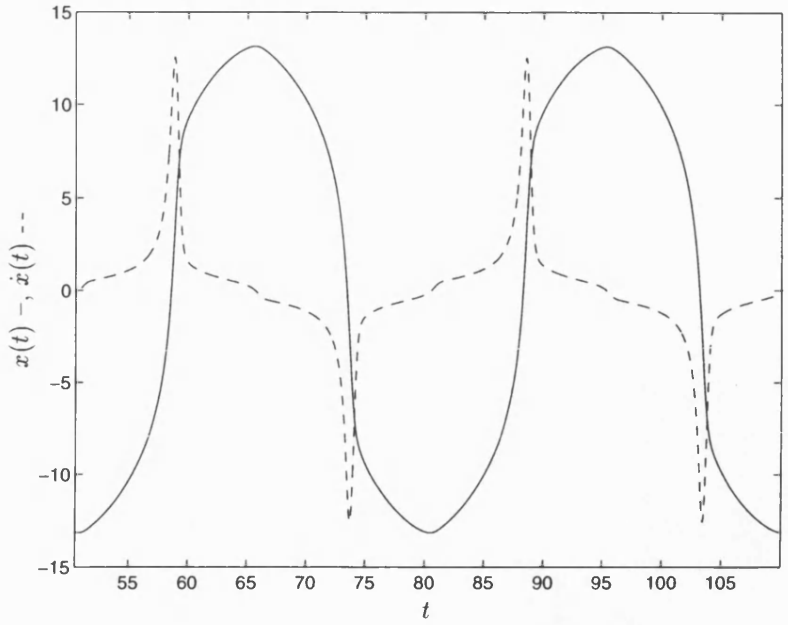


Figure 4-14: State and Velocity for $\ddot{x}(t) - (\mu_1 - \mu_0 x^2(t))\dot{x}(t) + x(t-h) = 0$, $\mu_0 = 0.1$, $\mu_1 = 0$, $h = 7$

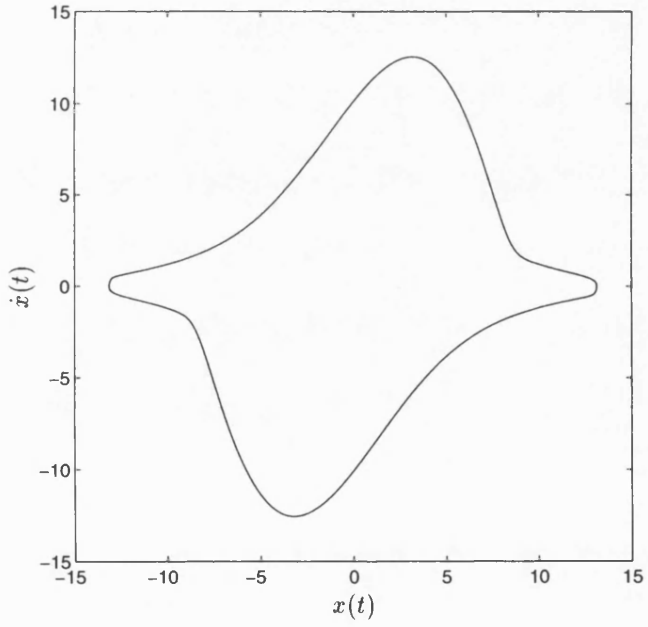


Figure 4-15: Limit Cycle for $\ddot{x}(t) - (\mu_1 - \mu_0 x^2(t))\dot{x}(t) + x(t-h) = 0$, $\mu_0 = 0.1$, $\mu_1 = 0$, $h = 7$

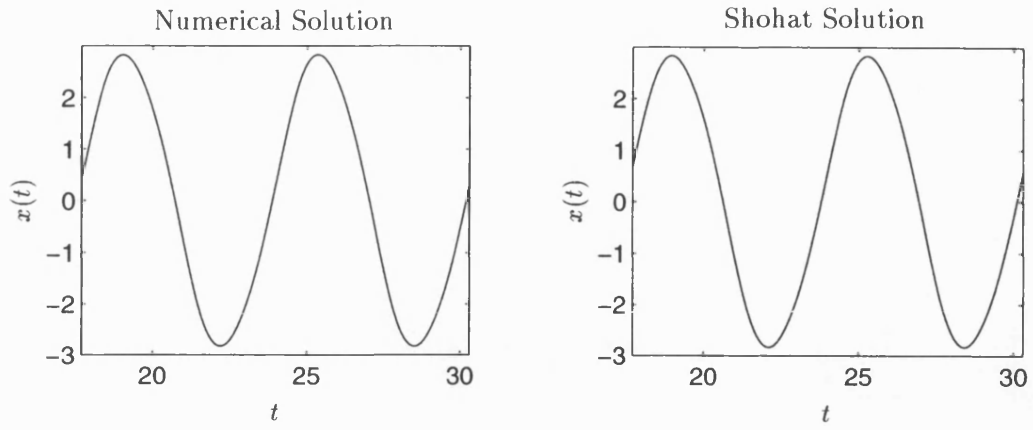


Figure 4-16: Comparison of Numerical and Shohat Results for
$\ddot{x}(t) - (\mu_1 - \mu_0 x^2(t))\dot{x}(t) + x(t-h) = 0, \quad \mu_0 = \mu_1 = h = 0.1$

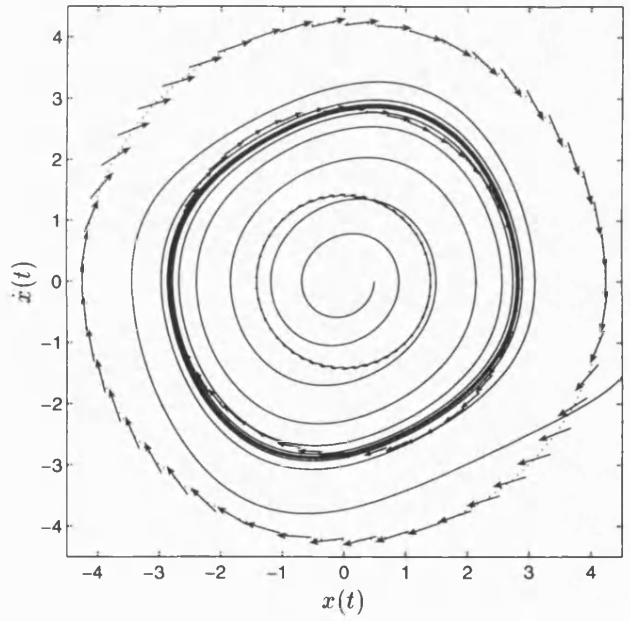


Figure 4-17: Trajectories and Flows for $\ddot{x}(t) - (\mu_1 - \mu_0 x^2(t))\dot{x}(t) + x(t-h) = 0$,
$\mu_0 = \mu_1 = h = 0.1$

4.2.3 Extended Van der Pol System with Delay in Velocity Term

The Vector Approach

Firstly the vector method gives approximate periodic solutions. The nonlinear term is evaluated at $x(t) = a\sin(\omega t)$, $\dot{x}(t) = a\omega\cos(\omega t)$ and expanded as a Fourier series.

$$\begin{aligned}
\mu_0 x^2(t)\dot{x}(t-h) &= \mu_0 a^3\omega\sin^2(\omega t)\cos(\omega(t-h)) \\
&= \mu_0 a^3\omega(\cos(\omega h)(\cos(\omega t) - \cos^3(\omega t)) + \sin(\omega h)\sin^3(\omega t)) \\
&= \frac{\mu_0 a^3\omega}{4}((\cos(\omega t) - \cos(3\omega t))\cos(\omega h) \\
&\qquad + (3\sin(\omega t) - \sin(3\omega t))\sin(\omega h)) \\
&= \mu_0 a^3\omega\left(\cos(\omega t)\cos(\omega h) - \frac{3}{4}\cos(\omega(t+h))\right)
\end{aligned}$$

neglecting all but the first harmonic terms. The differential equation becomes

$$-a\omega^2\sin(\omega t) - \mu_1 a\omega\cos(\omega(t-h)) + \mu_0 a^3\omega\cos(\omega t)\cos(\omega h) \\ -\frac{3}{4}\mu_0 a^3\omega\cos(\omega(t+h)) + a\sin(\omega t) = 0$$

The vector diagram is shown in figure 4-18. Resolving in the $\sin(\omega t)$ direction,

$$\omega^2 + \mu_1\omega\sin(\omega h) = 1 + \frac{3}{4}\mu_0 a^2\omega\sin(\omega h) \tag{4.13}$$

and in the $\cos(\omega t)$ direction,

$$\mu_0 a^2\omega\cos(\omega h) = \left(\mu_1 + \frac{3}{4}\mu_0 a^2\right)\omega\cos(\omega h)$$

so that for $\cos(\omega h) \neq 0$ and μ_0 and μ_1 strictly positive,

$$a^2 = \frac{4\mu_1}{\mu_0}$$

Substituting into equation 4.13,

$$\omega^2 - 2\mu_1\omega\sin(\omega h) - 1 = 0 \tag{4.14}$$

Note that the expression for a is independent of h so that (for h small) a limit cycle of constant amplitude $2\sqrt{\mu_1/\mu_0}$ is expected as h varies. μ_1 is again destabilising and μ_0

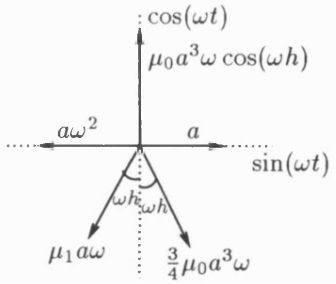


Figure 4-18: Vector Diagram for $\ddot{x}(t) - (\mu_1 - \mu_0 x^2(t))\dot{x}(t-h) + x(t) = 0$

stabilising. (If $\cos(\omega h) = 0$, $\sin(\omega h) = \pm 1$ and

$$\omega^2 \pm \left(\mu_1 - \frac{3}{4}\mu_0 a^2\right)\omega - 1 = 0$$

enabling determination of a in terms of h (h-dependent in this case). Preserving continuity, $a^2 = 4\mu_1/\mu_0$ is still a solution and equation 4.14 is again satisfied.)

The Linearised System

In the neighbourhood of the origin of the $(x, \dot{x})$ plane, $\mu_0 x^2(t)\dot{x}(t-h)$ will be small and behaviour and stability in this neighbourhood may be investigated by studying the corresponding linear system

$$\ddot{x}(t) - \mu_1 \dot{x}(t-h) + x(t) = 0$$

Taking Laplace transforms in order to apply the direct method, the characteristic equation is

$$s^2 + 1 - \mu_1 s \exp(-sh) = 0$$

For μ_1 small and positive, at $h = 0$ the system is unstable with a pair of right-half plane complex poles at

$$s = \frac{\mu_1}{2} \pm i\sqrt{1 - \left(\frac{\mu_1}{2}\right)^2}$$

(for $\mu_1 > 2$, a pair of positive real roots). The W-polynomial

$$W(\omega^2) = (1-\omega^2)^2 - \mu_1\omega^2 = \omega^4 - (2+\mu_1^2)\omega^2 + 1 = 0$$

has solutions

$$\omega = \pm\frac{\mu_1}{2} + \frac{1}{2}\sqrt{\mu_1^2 + 4}$$

using a difference of two squares factorisation and considering only positive ω. Also

$$\exp(i\omega h) = \frac{i\mu_1\omega}{1-\omega^2}$$

So the critical values of ω and h are given by

$$\omega_1 = -\tfrac{\mu_1}{2} + \tfrac{1}{2}\sqrt{\mu_1^2+4} \quad h_{1,q} = \tfrac{\pi}{\omega_1}\left(2q+\tfrac{1}{2}\right),\ q\epsilon\mathbb{N} \quad \text{(stabilising)}$$

$$\omega_2 = \tfrac{\mu_1}{2} + \tfrac{1}{2}\sqrt{\mu_1^2+4} \quad h_{2,q} = \tfrac{\pi}{\omega_2}\left(2q+\tfrac{3}{2}\right) \quad \text{(destabilising)}$$

(These critical values of ωh correspond to $\cos(\omega h) = 0$; at $\cos(\omega h) = 0$ information is lost in the vector approach on the amplitude of the limit cycle.) Note that

$$\omega_1\omega_2 = 1 \quad \text{and} \quad \omega_2 - \omega_1 = \mu_1 \tag{4.15}$$

where $\omega_1 < 1 < \omega_2$.

The system is initially unstable so that for n stability windows,

$$h_{1,n-1} < h_{2,n-1} \quad \text{and} \quad h_{1,n} > h_{2,n-1}$$

So it is necessary to find n, the nearest higher integer, to satisfy

$$\frac{\pi}{\omega_1}\left(2n+\frac{1}{2}\right) = \frac{\pi}{\omega_2}\left(2n+\frac{3}{2}\right)$$

The number of windows is given by

$$n = \left[\frac{\omega_2 - 3\omega_1}{4(\omega_1-\omega_2)}\right]$$

where $[\cdot]$ denote rounding up to the nearest integer. Using the properties 4.15 this may be expressed in terms of μ_1:

$$n = \left[\frac{1}{2}\left(\sqrt{\frac{1}{4}+\frac{1}{\mu_1^2}} - 1\right)\right]$$

n is real provided that $\mu_1 < 2/\sqrt{3} = 1.1547$; for $\mu_1 \geq 2/\sqrt{3}$ there will be no stability windows and instability for all values of delay. The following table illustrates the dependence on μ_1 of the number of stability windows.

μ_1	0.01	0.1	0.2	0.25	0.3	0.5	≥ 1.1547
n	50	5	3	2	2	1	0

In particular, for $\mu_1 = 0.25$, $\omega_1 = 0.8828$, $\omega_2 = 1.1328$ and the critical delay values are

$h_{1,q}$	$h=0$	1.7794		8.8989		
$h_{2,q}$			4.1600		9.7067	15.2534
No RHP roots	2	0	2	0	2	4

$h_{1,q}$	16.0143		23.1318		30.2493
$h_{2,q}$		20.8001		26.3468	
No RHP roots	2	4	2	4	2

Figures 4-19 and 4-20 show the temporal locus for this system.

Flows

Seeking flows in the $(x, \dot{x})$ plane, the differential equation may be rewritten

$$\frac{d\dot{x}(t)}{dx(t)} = -\frac{x(t)}{\dot{x}(t)} + (\mu_1 - \mu_0 x^2(t))\frac{\dot{x}(t-h)}{\dot{x}(t)} \tag{4.16}$$

For flows on the approximate predicted limit cycle the delay term may be written $\dot{x}(t-h) = \dot{x}(t)\cos(\omega h) + \omega x \sin(\omega h)$ so that

$$\frac{d\dot{x}}{dx} = -\frac{x}{\dot{x}} + (\mu_1 - \mu_0 x^2)\left(\cos(\omega h) + \frac{x}{\dot{x}}\omega \sin(\omega h)\right)$$

Using a Taylor series expansion for the delay term, and considering terms to second order in parameters μ_0, μ_1 and h to retain a second order differential equation, the approximate equation for the flows becomes

$$\frac{d\dot{x}}{dx} = \frac{1}{1 + h(\mu_1 - \mu_0 x^2)}\left(-\frac{x}{\dot{x}} + \mu_1 - \mu_0 x^2\right)$$

From the exact equation 4.16 it may be seen that for small μ_0, μ_1 and h the circular orbit dominates and that when $x = \pm\sqrt{\mu_1/\mu_0}$ the trajectories are tangential to the circles $d\dot{x}/dx = -x/\dot{x}$. Attention will be restricted to the case $\mu_0 = \mu_1 = \mu$ and trajectories in the $(x, \dot{x})$ plane considered for small h. When $x = \pm 1$, motion is in a clockwise sense and tangential to a circle centred on the origin; $d\dot{x}/dx = \mp 1/\dot{x}$. Nullclines are given by

$$\dot{x} = 0 \quad \text{on the } x \text{ axis}$$
$$\ddot{x} = 0 \quad \text{when } \dot{x} = \frac{x}{\mu(1-x^2)} \text{ and } x^2 \neq 1 + \frac{1}{\mu h}$$

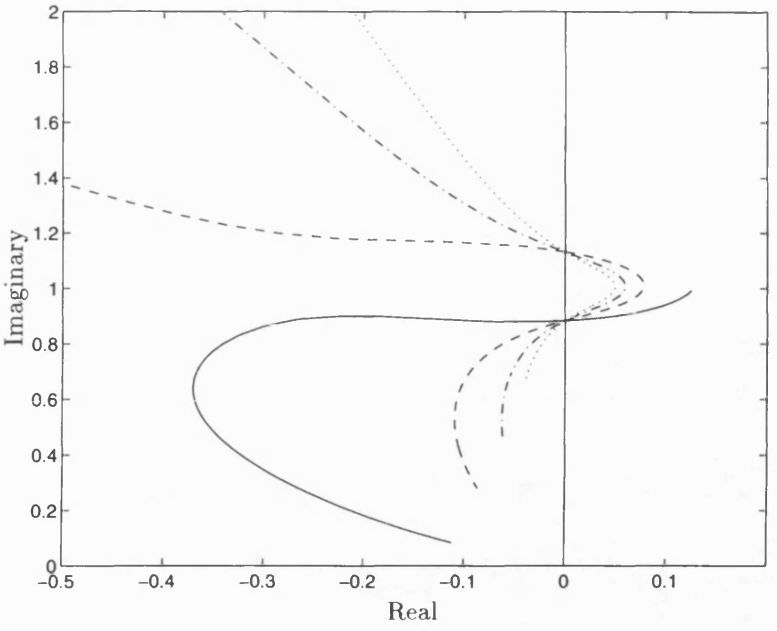


Figure 4-19: Temporal Locus of $s^2 + 1 - \mu_1 s \exp(-sh) = 0$, $\mu_1 = 1/4$

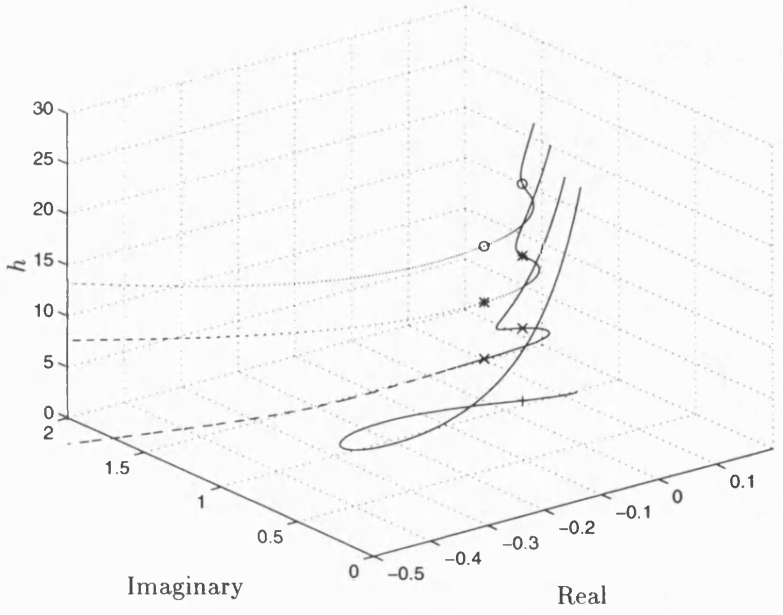


Figure 4-20: Temporal Locus of $s^2 + 1 - \mu_1 s \exp(-sh) = 0$, $\mu_1 = 1/4$

The gradient of the trajectories is infinite at $\dot{x} = 0$ and $x^2 = 1 + 1/(\mu h)$. The signs of $\dot{x}$ and $\ddot{x}$ determine the direction of flow ($\ddot{x}$ changing sign when $x = \pm\sqrt{1 + 1/(\mu h)}$ and across $\dot{x} = x/(\mu(1 - x^2))$). On $x = 0$, $d\dot{x}/dx = \mu/(1 + \mu h)$. Figure 4-21 shows the limit cycle together with the nullclines and direction of flow in each region, marked by arrows. The lines $x = \pm\sqrt{1 + 1/(\mu h)}$ act as 'barriers' which trajectories may not cross; even for small values of delay with stable limit cycle solutions, if the initial condition x_0 is chosen to be sufficiently large then it is expected that the solution will become unbounded in finite time. This stability boundary is reduced as h or μ are increased. Though this is an approximate method to second order in parameters, numerical calculations support this theory for h and μ small.

Delays of Multiple Period

Consider the system shown in figure 4-22. Initially the switch at point A is connected so that the system is equivalent to that of figure 4-1. Suppose that a steady state periodic orbit of period T arises from some set of initial conditions $(x_0, \dot{x}_0)$ (this may be stable or unstable). Then for large t,

$$\dot{x}(t - h - nT) = \dot{x}(t - h) \qquad n \epsilon \mathbb{N}$$

The output signal from the delay element h and that if the connection B is made instantaneously instead of A (where the delay h has been replaced by delay $h + nT$) are then identical. Therefore there will be the same periodic solution when h is incremented by any integer multiple of the period. This does not necessarily imply that the same stability properties of the limit cycle will be inherited by the new system. With the same initial conditions and initial delay of $h + nT$ the steady-state behaviour may be quite different. (There is then no contradiction when considering incrementing the delay by an integer number of periods so that it is brought into the region following 'failure to alternate'. The same periodic solution will exist but may be unstable.) Similar arguments apply to the examples of sections 4.2.1 and 4.2.2.

Behaviour with Increasing Delay

Since the effects of varying the two independent parameters μ_0 and μ_1 have been observed for the previous two examples, and these effects are to be expected in all similar delayed Van der Pol type equations, the specialisation $\mu_0 = \mu_1 = \mu$ will be made. This will give prominence to the additional features of this problem. The

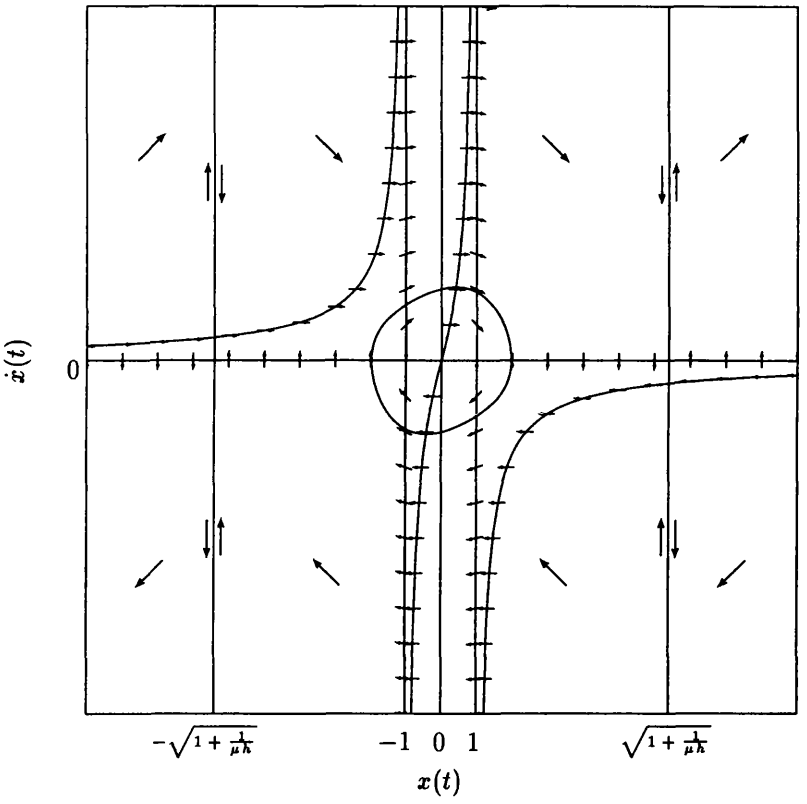


Figure 4-21: Flows for $\ddot{x}(t) - \mu(1 - x^2(t))\dot{x}(t-h) + x(t) = 0$

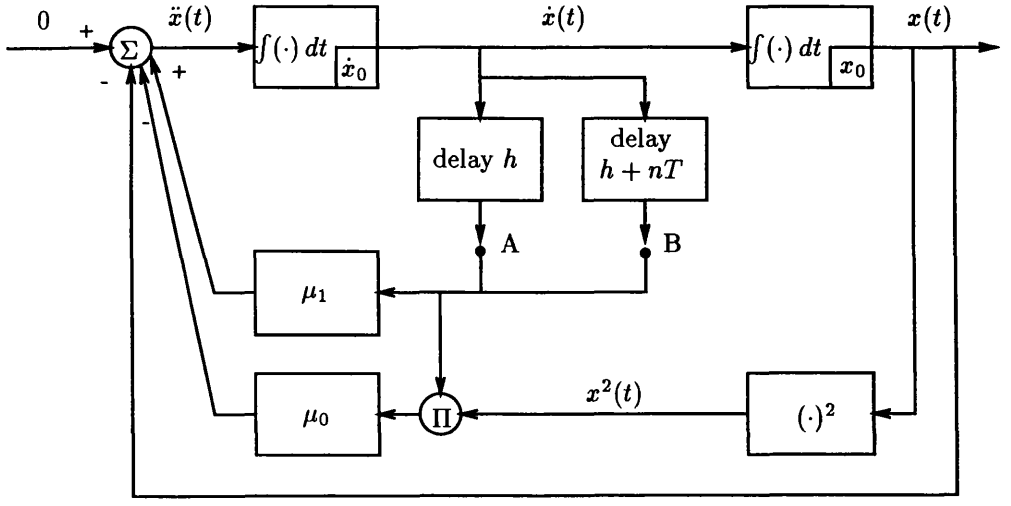


Figure 4-22: Modified Van der Pol System with Delay in State: Addition of Delay of Multiple Period

equation to be considered is then

$$\ddot{x}(t) - \mu(1 - x^2(t))\dot{x}(t - h) + x(t) = 0 \tag{4.17}$$

From the discussion of the stability properties of the linearised system, the origin will switch between being unstable, to stable, etc. until there is failure to alternate of the stabilising and destabilising h values and there are no more stability windows. It is expected that this will determine the stability of any limit cycle so that for the origin unstable, the limit cycle will be stable and vice versa. As eigenvalues traverse the imaginary axis from the left to right-half plane for the linearised system the limit cycle will turn from being weakly repelling to weakly attracting. For all eigenvalues distant from the imaginary axis the trajectories will approach or depart from the limit cycle more rapidly. The range of h following failure to alternate and the effect of four or more poles in the right-half plane of the linearised system will also be of interest. Many of the aforementioned results for the nonlinear system depend on small parameters μ and h; it is necessary to sacrifice a very small μ so that failure to alternate occurs for h not too large. Particular attention has been devoted to the example $\mu = 1/4$ for which there are two stability windows and instability for all values of delay greater than 9.7067 of the linearised system. The following section describes behaviour as the delay is increased from zero with reference to the figures illustrating the consistency of numerical with expected results.

Behaviour with h of $\ddot{x}(t) - \mu(1 - x^2(t))\dot{x}(t - h) + x(t) = 0$, $\mu = 1/4$

1. $h = 0$
   The problem reduces to the standard Van der Pol equation with limit cycle of approximate amplitude 2. This is illustrated in figure 3-38.

2. h small and positive
   The linearised system has two poles in the right-half plane implying instability in the neighbourhood of the origin. The method of Kryloff and Bogoliuboff predicts a limit cycle again of radius 2 which is expected to be stable since the origin is unstable. Plots of the state and velocity against time (approximately sinusoidal) are given in figure 4-23. Figure 4-24 shows trajectories in the $(x, \dot{x})$ plane starting near the origin and outside the limit cycle for $h = 0.1$, together with the circle of radius 2.

3. $0 < h < 1.7794$
   The linearised system remains unstable with two right-half plane poles; this is

reflected in instability in the neighbourhood of the origin and a stable limit cycle of approximate amplitude 2. As the delay is increased the distortion from the circular orbit becomes greater.

4. $h = 1.7794$
   The linearised system has a pair of imaginary axis poles. As h passes through this value the origin passes from being unstable to stable and the limit cycle from being stable to unstable. The rate at which trajectories approach or depart from the origin or limit cycle becomes very slow in the neighbourhood of this critical value.

5. $1.7794 < h < 4.1600$
   (h is now 'large' so that any results dependent on small h, for example the extension of Kryloff and Bogoliuboff's method, must be treated with caution.) This region forms the first stability window of the linearised system, with the origin being stable and the limit cycle unstable. Figure 4-25 illustrates the behaviour in the $(x, \dot{x})$ plane for $h = \pi$ with starting values $(x_0, \dot{x}_0) = (2, 0)$ and $(2.01, 0)$. The approximation of the circle radius 2 centre the origin to the limit cycle still holds.

6. $h = 4.1600$
   marks the transition between stability and instability of the linear system as two poles enter the right-half plane. The origin switches from being stable to unstable.

7. $4.1600 < h < 8.8989$
   The origin is unstable and an attracting limit cycle is present. See figure 4-26 for $h = 5$. The radius of convergence to the limit cycle for trajectories commencing outside has decreased as expected.

8. $h = 8.8989$
   The two right-half plane poles of the linearised system move back into the left-half plane and stability is resumed at the origin.

9. $8.8989 < h < 9.7067$
   The second stability window sees the nonlinear delayed system with a repellent limit cycle and stability to the origin. This behaviour is similar to point 5.

10. $9.7067 < h < 15.2534$
    Stability of the linearised system is lost at $h = 9.7067$; the limit cycle again becomes stable.

11. $h = 15.2534$

    This value of delay marks failure to alternate between stabilising and destabilising values of h in the linearised system, with an additional two poles entering the right-half plane giving four unstable roots.

12. $h \geq 15.2534$

    As h is increased from 15.2534, a periodic orbit which encircles the origin several times before repeating is formed. On examining the plots of x and $\dot{x}$ against time, *beats* are revealed as in the superposition of two sinusoids of different amplitude and frequency. A feature of these is that motion does not always take place in a clockwise sense about the origin in the $(x, \dot{x})$ plane but *loops* form, which become larger as h increases, intersect and pass through the origin. This is clarified by the plots (all for the steady state solution) which are given for $h = 15.2534$ (figures 4-27 and 4-28), $h = 15.5$ (figure 4-29), $h = 5\pi = 15.7080$ (figure 4-30), $h = 15.85$ (figures 4-31, 4-32 and 4-33, the latter showing the progression of the orbit with time over one period ($T \approx 63.5$)), $h = 16.0143$ (figure 4-34) and $h = 17$ (figure 4-35). The presence of the beats is evident from the corresponding plots of state and velocity against time. Loops correspond to the oscillation turning back before traversing the $x = 0$ or $\dot{x} = 0$ line. These properties persist even for $16.0143 < h < 20.8001$ for example, when a pair of poles of the linearised system has passed back into the left-half plane leaving a single unstable pair.

    Motion taking place in an anticlockwise sense about the origin does not contradict $\dot{x}$ being positive (negative) implying increasing (decreasing) x respectively. Trajectories must move from left to right in the upper half of the $(x, \dot{x})$ plane ($\dot{x}$ positive), from right to left in the lower-half plane and must cut the x axis with infinite gradient. This may be seen to be the case.

    It is conjectured that a form of synchronisation is taking place to induce the Lissajoux-type figures in the $(x, \dot{x})$ plane. [20] contains much information on synchronisation in (delay-free) systems in which one oscillator acts as a forcing function to another to induce periodic solutions. Since for this delayed Van der Pol example periodic orbits of this form are seen to occur for a continuous range of delay, it cannot be concluded that the frequencies of the two oscillations (being related to the position or frequency of the right-half plane linear roots perhaps) are related by an integer which is the requisite for Lissajoux motion. The hypothesis is made that for these large delay values the delay term is acting as a forcing function, considering an equivalent system where the stored function in the delay is replaced by zero stored function plus a forcing function (as in

section 3.1.11).

13. Large h

    The solution fills a bounded region centred on the origin and either fails to settle or forms a periodic orbit of greatly extended period. Figure 4-37 illustrates this (the 'misleading' information due to the stored function in the delay has been removed so that the solution shown is a long-time one); the orbit fills a region approximating a circle radius 2 centre the origin. Figure 4-36 shows the state and velocity against time with beats still occurring. Computational restrictions made numerical solution for very large h impractical using the predictor-corrector method of appendix C; however, calculations using SIMULINK suggest that this behaviour continues with a bounded orbit yet instability for large x_0. (h was considered up to 100.)

These results are summarised in table 4.1.

4.2.4 Summary

In this chapter, special cases of Van der Pol type systems with delay have been considered, namely delay in all state terms, delay in linear state term only and delay in velocity term. It was found that the application of the linear analysis of section 3.1 was appropriate to determine stability of the origin and behaviour in a neighbourhood thereof; this in turn had implications for the stability of any limit cycle. For small delays the method of Kryloff and Bogoliuboff was extended and limit cycles approximately circular, centred on the origin and of amplitude 2 were predicted, being attracting or repelling according as the linearised system was unstable or stable respectively. The effects of the parameters μ_0 and μ_1 and particularly h on the solution were investigated. For the first two examples the origin was unstable for all delay values and increasing the delay brought a distortion of the limit cycle from the circle. The linearisation of the third example was found to exhibit stability windows implying alternation between stability and instability of the origin and limit cycle of the nonlinear system. All the above behaviour has been verified by two independent numerical methods. As the delay is increased further beyond failure to alternate for the third system, the phenomenon of a periodic orbit encircling the origin was observed corresponding to beats in the state and velocity with time. For very large h the periodic orbit became unsettled.

It is expected that application of the techniques outlined in this chapter to other delayed Van der Pol systems will reveal similar behaviour. Clearly there is much scope for further investigation of such systems.

h	$h_{1,q}$	$h_{2,q}$	RHP Poles	Origin	Limit Cycle	Comment	Figures
0			2	Unstable	Stable	Standard Van der Pol with limit cycle	3-38
$0 < h < 1.7794$			2	Unstable	Stable	Limit cycle $a \approx 2$	4-23, 4-24
1.7794	1.7794		2 Imaginary				
$1.7794 < h < 4.1600$			0	Stable	Unstable	1st stability window, limit cycle repels	4-25
4.1600		4.1600	2 Imaginary				
$4.1600 < h < 8.8989$			2	Unstable	Stable	Attracting limit cycle	4-26
8.8989	8.8989		2 Imaginary				
$8.8989 < h < 9.7067$			0	Stable	Unstable	2nd stability window, limit cycle repels	
9.7067		9.7067	2 Imaginary				
$9.7067 < h < 15.2534$			2	Unstable	Stable	Attracting limit cycle	
15.2534		15.2534	2 RHP, 2 Imag.	Unstable	Stable	Failure to alternate, beats	4-27, 4-28
$15.2534 < h < 16.0143$			4	Unstable	Stable	Loops, beats	4-29–4-33
16.0143	16.0143		2 RHP, 2 Imag.	Unstable	Stable	Loops passed through origin, beats	4-34
$16.0143 < h \leq 17$			2	Unstable	Stable	Periodic orbit, beats	4-35
$h > \approx 18$			≥ 2	Unstable	Stable	Bounded unsettled orbit	4-36, 4-37

Table 4.1: Behaviour with h of $\ddot{x}(t) - \mu(1 - x^2(t))\dot{x}(t - h) + x(t) = 0$, $\mu = 1/4$

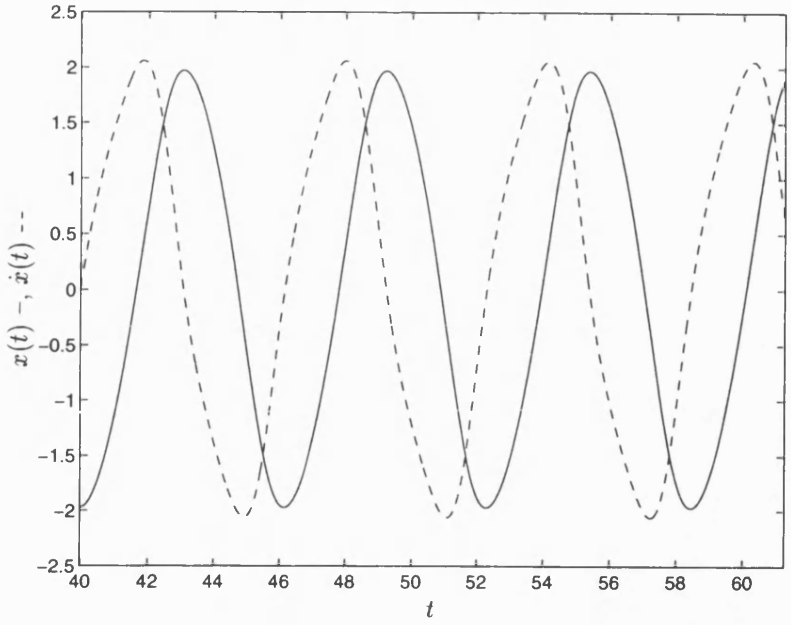


Figure 4-23: State and Velocity for $\ddot{x}(t) - \mu(1 - x^2(t))\dot{x}(t - h) + x(t) = 0$, $\mu = 1/4, \quad h = 0.1$

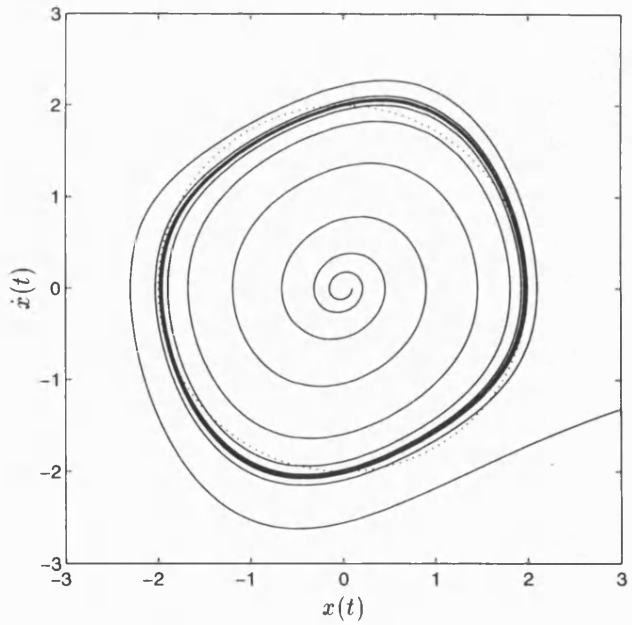


Figure 4-24: Trajectories for $\ddot{x}(t) - \mu(1 - x^2(t))\dot{x}(t - h) + x(t) = 0, \quad \mu = 1/4, \quad h = 0.1$

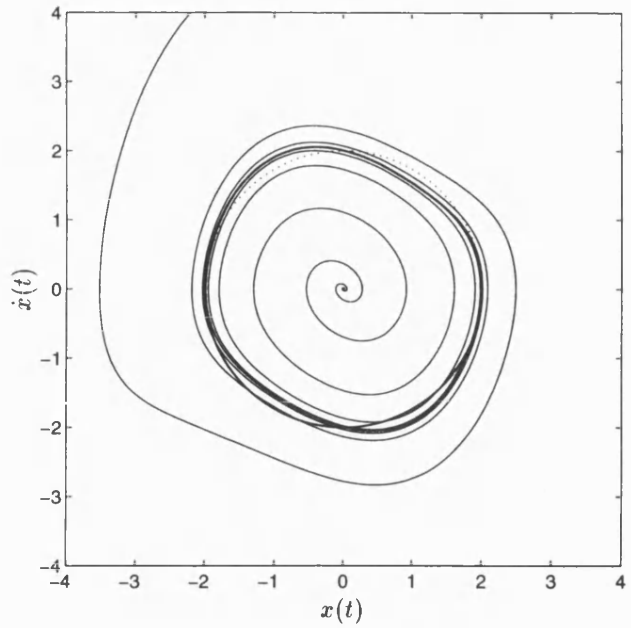


Figure 4-25: Trajectories for $\ddot{x}(t) - \mu(1 - x^2(t))\dot{x}(t - h) + x(t) = 0, \quad \mu = 1/4, \quad h = \pi$

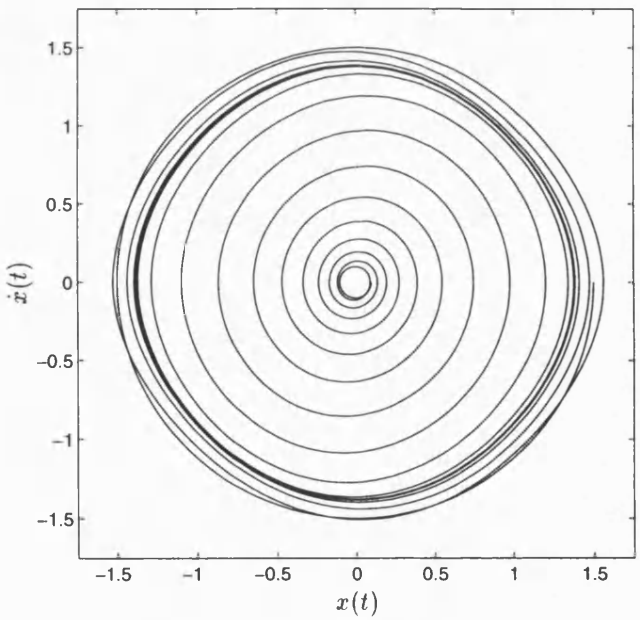


Figure 4-26: Trajectories for $\ddot{x}(t) - \mu(1 - x^2(t))\dot{x}(t - h) + x(t) = 0, \quad \mu = 1/4, \quad h = 5$

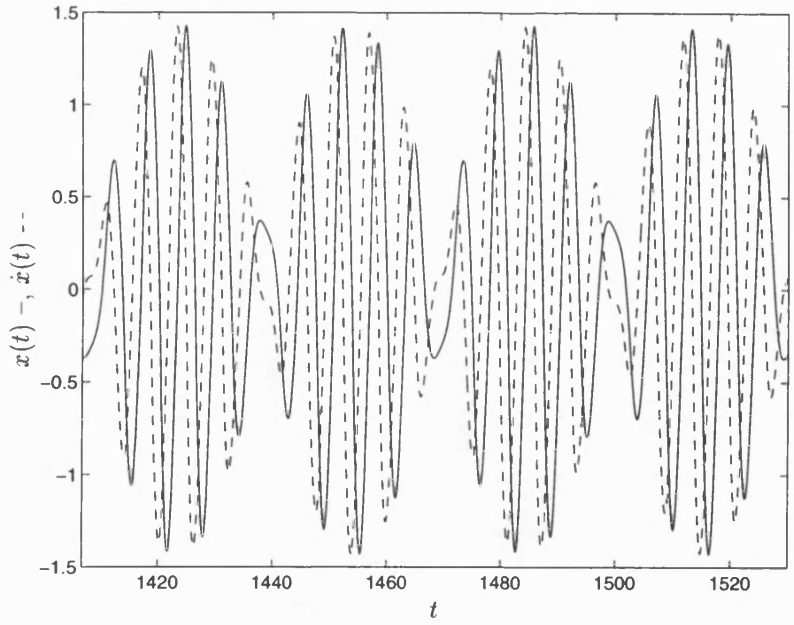


Figure 4-27: State and Velocity for $\ddot{x}(t) - \mu(1 - x^2(t))\dot{x}(t - h) + x(t) = 0$, $\mu = 1/4, \quad h = 15.2534$

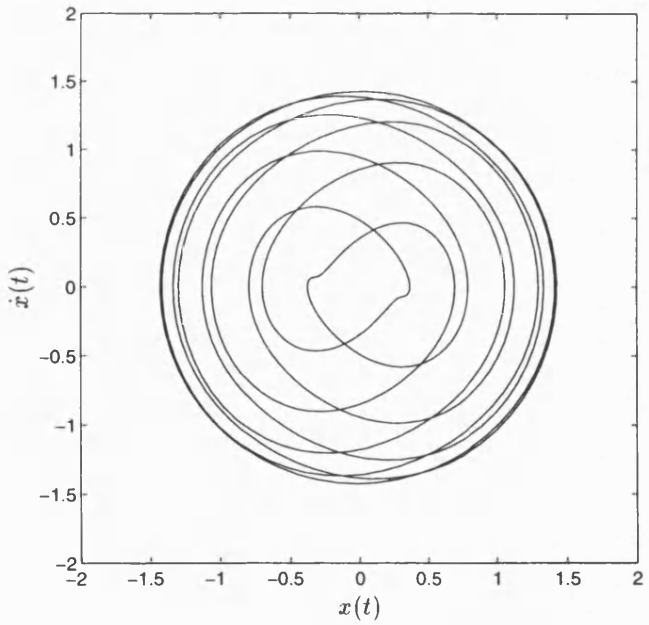


Figure 4-28: Periodic Orbit for $\ddot{x}(t) - \mu(1 - x^2(t))\dot{x}(t - h) + x(t) = 0$, $\mu = 1/4, \quad h = 15.2534$

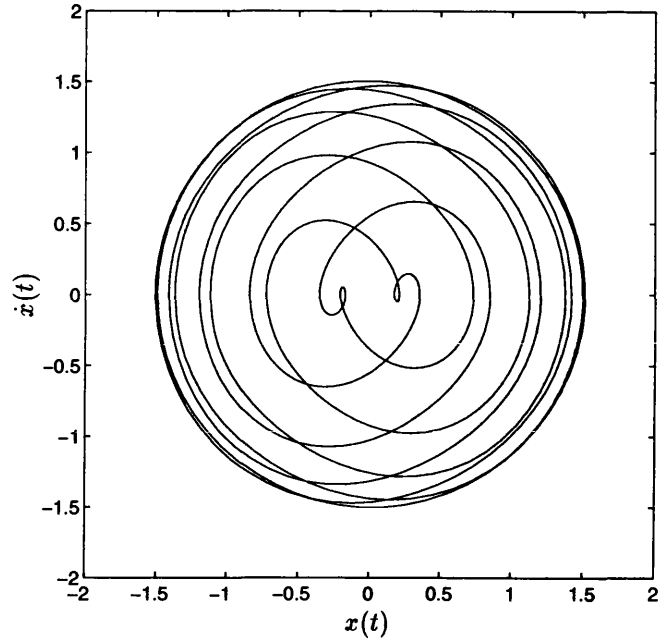


Figure 4-29: Periodic Orbit for $\ddot{x}(t) - \mu(1 - x^2(t))\dot{x}(t - h) + x(t) = 0$, $\mu = 1/4, \quad h = 15.5$

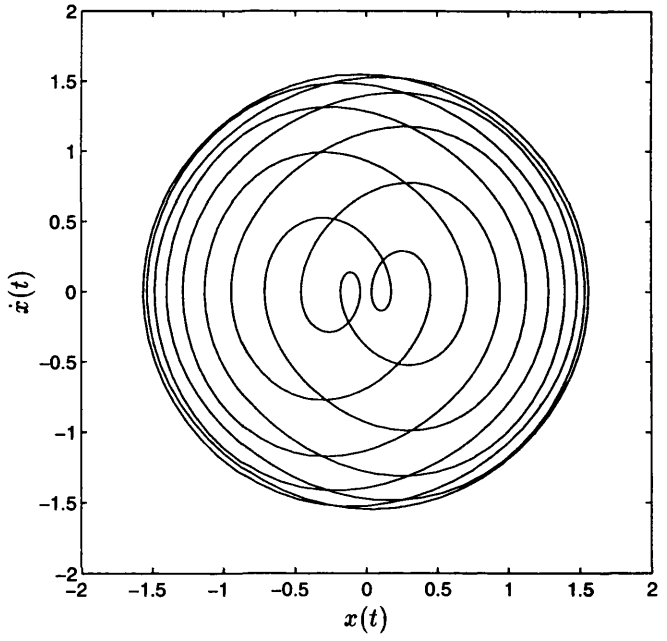


Figure 4-30: Periodic Orbit for $\ddot{x}(t) - \mu(1 - x^2(t))\dot{x}(t - h) + x(t) = 0$, $\mu = 1/4, \quad h = 5\pi = 15.7080$

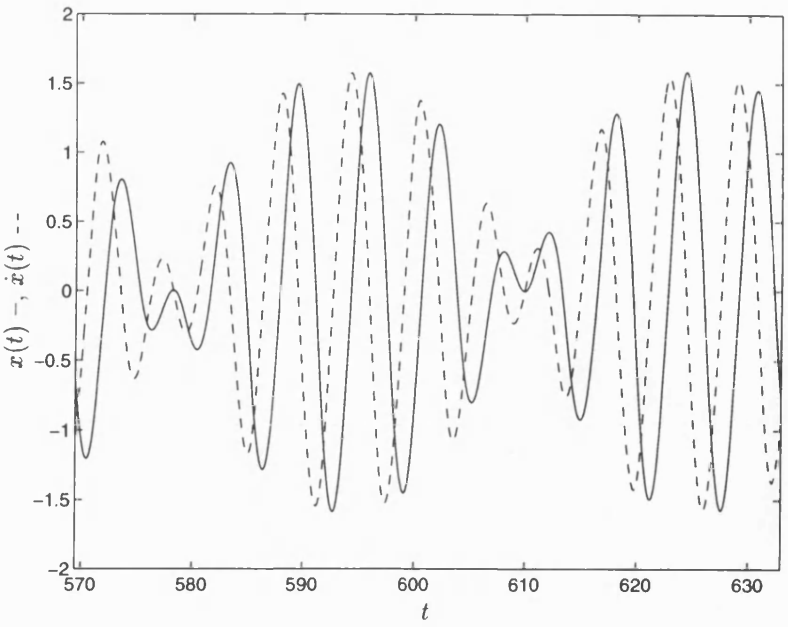


Figure 4-31: State and Velocity for $\ddot{x}(t) - \mu(1 - x^2(t))\dot{x}(t - h) + x(t) = 0$, $\mu = 1/4, \quad h = 15.85$

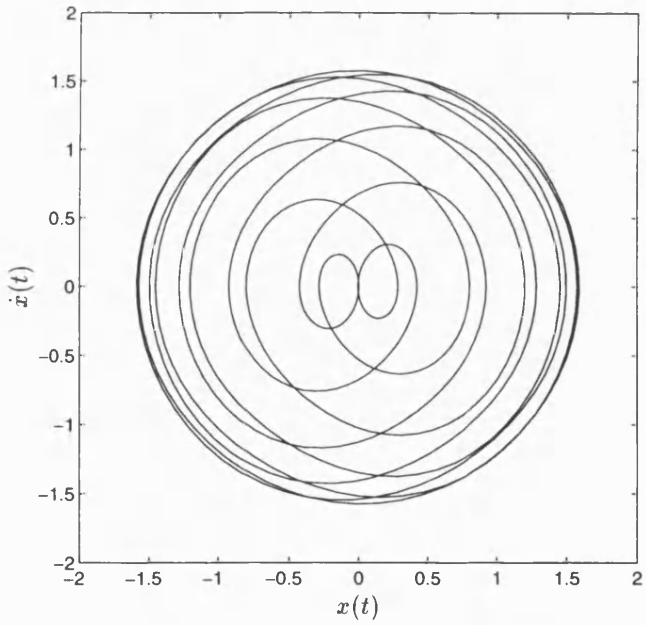


Figure 4-32: Periodic Orbit for $\ddot{x}(t) - \mu(1 - x^2(t))\dot{x}(t - h) + x(t) = 0$, $\mu = 1/4, \quad h = 15.85$

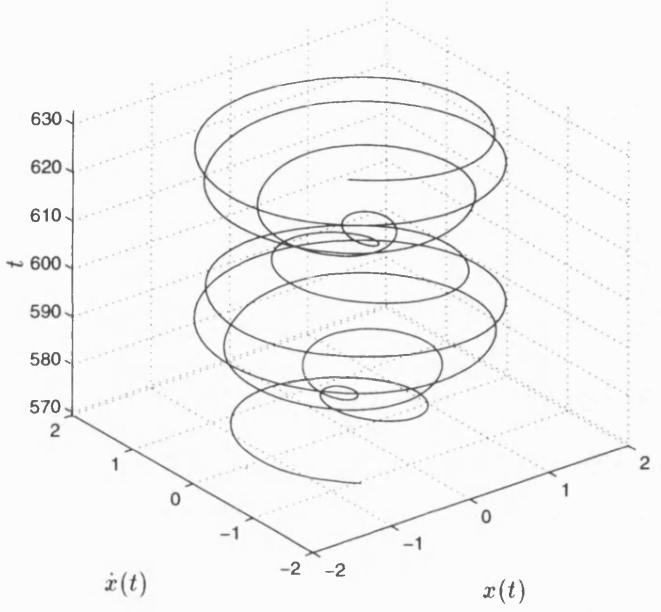


Figure 4-33: Periodic Orbit with Time for $\ddot{x}(t) - \mu(1 - x^2(t))\dot{x}(t - h) + x(t) = 0$, $\mu = 1/4, \quad h = 15.85$

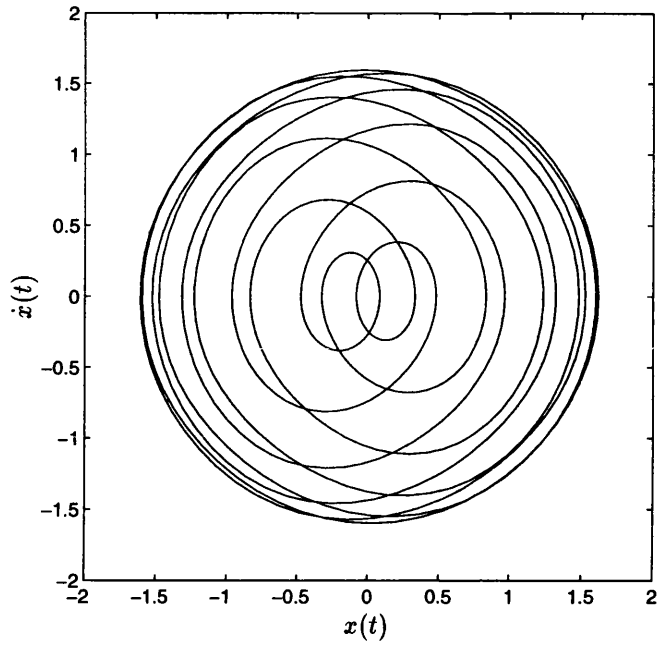


Figure 4-34: Periodic Orbit for $\ddot{x}(t) - \mu(1 - x^2(t))\dot{x}(t - h) + x(t) = 0$, $\mu = 1/4, \quad h = 16.0143$

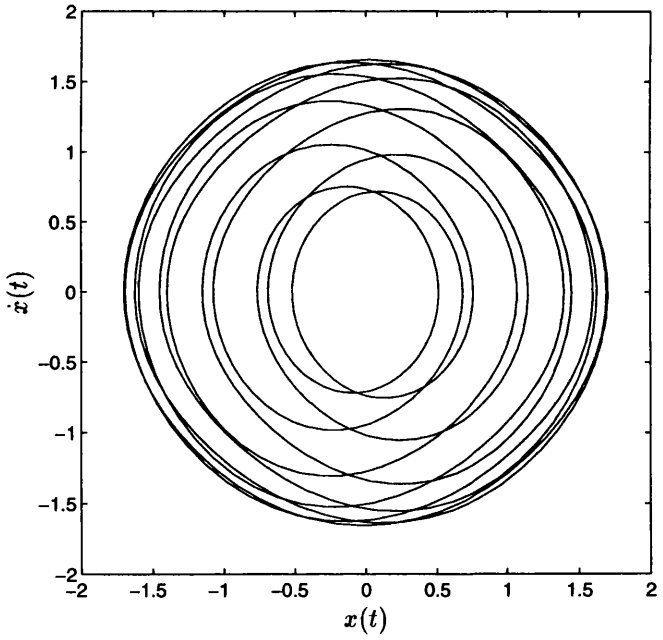


Figure 4-35: Periodic Orbit for $\ddot{x}(t) - \mu(1 - x^2(t))\dot{x}(t - h) + x(t) = 0$, $\mu = 1/4, \quad h = 17$

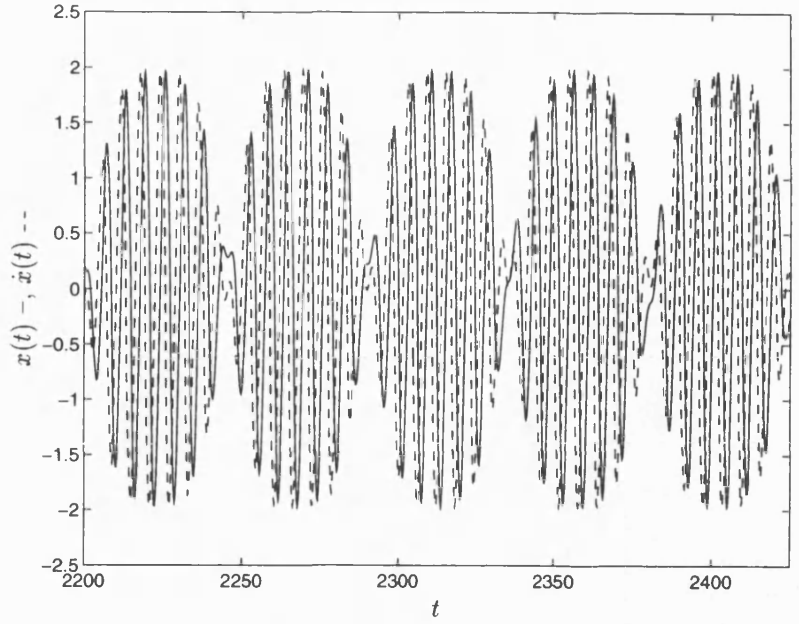


Figure 4-36: State and Velocity for $\ddot{x}(t) - \mu(1 - x^2(t))\dot{x}(t - h) + x(t) = 0$, $\mu = 1/4, \quad h = 21.4$

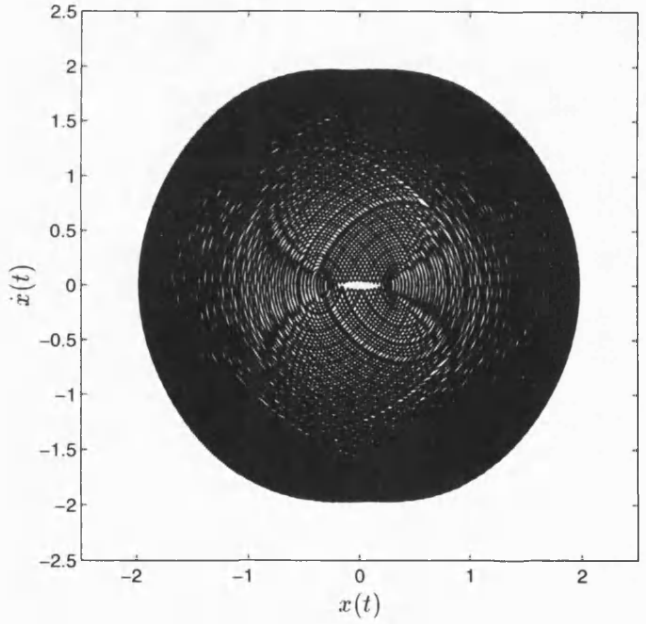


Figure 4-37: Unsettled Orbit for $\ddot{x}(t) - \mu(1 - x^2(t))\dot{x}(t - h) + x(t) = 0$, $\mu = 1/4, \quad h = 21.4$

Chapter 5

The Two–Delay Problem

5.1 Linear Two–Delay Problems

5.1.1 Smith's Method

O. J. M. Smith [60], [61] devised a method of control system design for systems with delay. The delay is, in principle, eliminated from the characteristic equation of the closed loop system, so that delay-free controller design techniques may be used. It is necessary to model the delay and the delay-free plant and clearly the accuracy of the model will determine the stability of the controlled system.

Attention here will be restricted to a system consisting of a stable delay-free plant $G(s)$ in series with a delay of duration h. Initial conditions are zero and there is no stored function in the delay. Knowledge of the corresponding delay-free controller $C(s)$ is assumed (for a particular class of inputs). A controller $C^*(s)$ is chosen to satisfy

Smith's Principle

> For a given input $X(s)$, $C^*(s)$ is chosen so that the output $Y(s)$ is the same as that of the corresponding delay-free system with inevitable delay h.

This is illustrated by figure 5-1, the upper diagram showing the Smith controller in series with plant and delay with unity negative feedback and the lower diagram the corresponding delay-free system with delay in measurement h. By equating the two transfer functions,

$$\frac{C^*(s)G(s)\exp(-sh)}{1+C^*(s)G(s)\exp(-sh)} = \frac{C(s)G(s)\exp(-sh)}{1+C(s)G(s)}$$

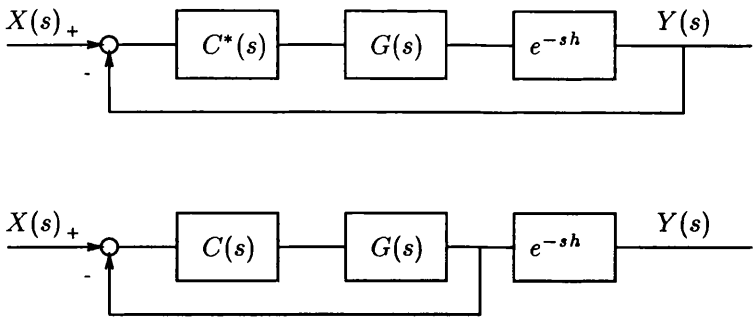


Figure 5-1: Smith's Principle

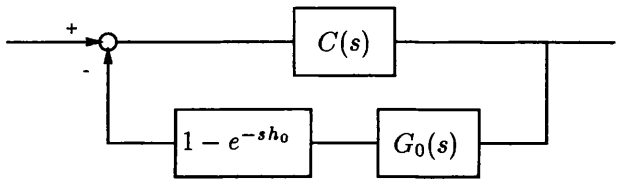


Figure 5-2: The Smith Controller $C^*(s)$

So

$$C^*(s) = \frac{C(s)}{1 + C(s)G(s)(1 - \exp(-sh))}$$

However, $G(s)$ and h are not known exactly so these are modelled by $G_0(s)$ and h_0 respectively. $G_0 \neq G$ is known as *parametric mismatch* and $h_0 \neq h$, *temporal mismatch.* Accounting for these modelling inaccuracies,

$$C^*(s) = \frac{C(s)}{1 + C(s)G_0(s)(1 - \exp(-sh_0))}$$

C^* is then equivalent to the system of figure 5-2. The upper diagram of figure 5-1 may then be shown to be equivalent to the system shown in figure 5-3, via simple block diagram manipulation.

To find the transfer function of this system (suppressing the explicit s-dependence),

$$E = X - (Y - CG_0 \exp(-sh_0)E) - CG_0E$$

so

$$\begin{aligned} Y &= CG\exp(-sh)E \\ &= \frac{CG\exp(-sh)(X - Y)}{1 - CG_0\exp(-sh_0) + CG_0} \end{aligned}$$

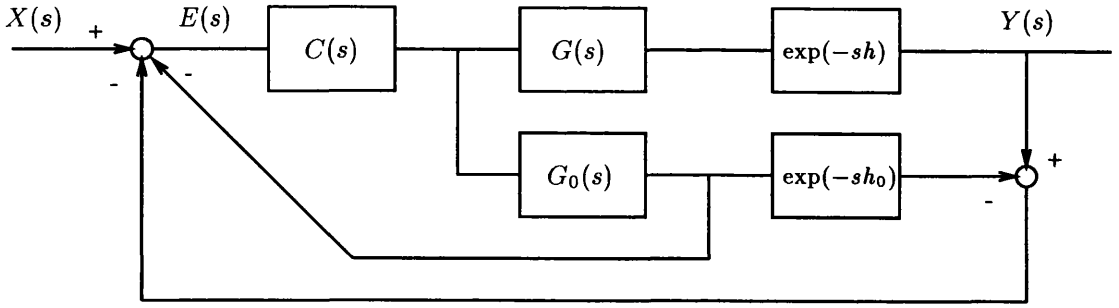


Figure 5-3: Smith's Method

and

$$\frac{Y}{X} = \frac{CG\exp(-sh)}{1 + CG_0 + C(G\exp(-sh) - G_0\exp(-sh_0))}$$

with characteristic equation

$$1 + CG_0 + C(G\exp(-sh) - G_0\exp(-sh_0)) = 0 \tag{5.1}$$

The system is known as a *Smith Predictor* as the output from the plant model is an approximate prediction for $y(t)$. In the matched case the signal at this point is $y(t+h)$, and the method makes the output from the plant of the upper diagram in figure 5-1 accessible. At matching, the signal around the lower feedback loop of figure 5-3 will be zero. In addition the transfer function reduces to

$$\frac{CG\exp(-sh)}{1 + CG}$$

as it should by Smith's principle, and the system is stable with $1 + CG = 0$ having all left-half plane roots.

Reference [33] gives a detailed treatment of the effects of mismatch, particularly system sensitivity and performance. It is somewhat surprising that performance may be *improved* by deliberate mismatch of the delay in certain cases. Stability analysis in both parametric and delay mismatch cases will be given in later sections. See also [15], [34] and [35].

5.1.2 Multiple Commensurate Delays

Commensurate delays are those related by an integer. Characteristic equations of form

$$F(s,h) := \sum_{j=0}^{n} A_j(s)\exp(-jsh) = 0 \tag{5.2}$$

will arise for example from multi-input, multi-output delayed systems. Walton and Marshall [71], [35] have extended the direct method for stability with single delay to the multiple commensurate delay case.

As before at $h = 0$ the equation becomes delay-free and classical methods may be employed to ascertain the number and location of unstable right-half plane roots, if any. Attention is restricted to retarded systems (for which the degree of $A_0(s)$ is greater than that of $A_j(s)$, $j = 1, \cdots n$) so that for h infinitesimally small and positive all new roots appear at infinity in the left-half plane.

To find crossing points of the imaginary axis, solutions are sought of the simultaneous equations 5.2 and

$$F(-s,h) = \sum_{j=0}^{n} A_j(-s)\exp(jsh) = 0$$

at $s = i\omega$. A recursive method is used to remove powers of $\exp(-sh)$ from these systematically, resulting in a polynomial in ω^2. This is solved for ω and corresponding critical delay values are found from the last equation (in single power of $\exp(-sh)$). It must then be verified that these solutions for ω and h of the equation in $\exp(-sh)$ are solutions of the original characteristic equation. An analogue of the sign theorem distinguishes stabilising and destabilising values. Details of the algebra, the generalised W-polynomial and sign theorem are to be found in [71] and [35]. Here is given an example for $n = 2$ to illustrate the basic procedure.

Example Smith predictor with parametric matching and model delay twice the plant delay.

$$F(s,h) = s + 1 + \exp(-sh) - \exp(-2sh) = 0$$

At $h = 0$, $s = -1$ and the system is stable. The characteristic equation is of retarded type so all new roots appear at infinity in the left-half plane as h increases from zero. For critical stability, common roots of $F(i\omega,h)$ and $F(-i\omega,h)$ are required:

$$\begin{aligned} i\omega + 1 + \exp(-i\omega h) - \exp(-2i\omega h) &= 0 \\ (-i\omega + 1)\exp(-2i\omega h) + \exp(-i\omega h) - 1 &= 0 \end{aligned}$$

Eliminating the term in $\exp(-2i\omega h)$ yields

$$(-i\omega + 1)(i\omega + 1 + \exp(-i\omega h)) + \exp(-i\omega h) - 1 = 0$$

$$\omega^2 + (-i\omega + 2)\exp(-i\omega h) = 0 \qquad (5.3)$$

and eliminating the term independent of h gives

$$\exp(-i\omega h) - \exp(-2i\omega h) + (i\omega + 1)((-i\omega + 1)\exp(-2i\omega h) + \exp(-i\omega h)) = 0$$

$$\omega^2 \exp(-i\omega h) + i\omega + 2 = 0 \tag{5.4}$$

Now combining equations 5.3 and 5.4 to remove the term in $\exp(-i\omega h)$ gives

$$W(\omega^2) = \omega^4 - \omega^2 - 4 = 0$$

with solutions $\omega = \pm\sqrt{\frac{1}{2}(1+\sqrt{17})}$. Substituting into equation 5.3

$$\begin{aligned} h &= -\tfrac{1}{\omega}\tan^{-1}\left(\tfrac{\omega}{2}\right) + \tfrac{1}{\omega}(2q+1)\pi \quad q\epsilon\mathbb{N} \\ &= 1.5412, 5.5513, \cdots \end{aligned}$$

It is as simple matter to check that these correspond to solutions of the original characteristic equation. As this is the only crossing point it may be assumed to be destabilising; alternatively the extended sign theorem may be used. (The critical h values may be seen to agree with the (h, h_0) diagram of section 5.1.3; the line $h_0 = 2h$ on the diagram enters the successive regions of two extra poles in the right-half plane at these values of h.)

5.1.3 Stability of Smith Predictor Systems via Direct Method

The direct method for single delays is extended to Smith predictor systems with temporal mismatch, and temporal and multiplicative parametric mismatch, where in each case the delays are non-commensurate. The commensurate delay case was discussed in the previous section and temporal matching reduces to the single-delay problem

$$1 + CG_0 + C(G - G_0)\exp(-sh) = 0$$

which may be analysed by single-delay techniques.

For temporal mismatch only, the characteristic equation becomes

$$1 + CG + CG(\exp(-sh) - \exp(-sh_0)) = 0$$

and the magnitude of CG will be important in stability studies since $1 + CG = 0$ is the mismatch-free characteristic equation. Similarly, for parametric mismatch only,

$$1 + CG + C(G - G_0)(\exp(-sh) - 1) = 0$$

and the magnitude of $\exp(-sh) - 1$, small for h small, will be important.

Temporal Mismatch

This problem is the subject of the paper [36] in which it is shown that *avenues* of stability about $h = h_0$ are to be found in diagrams of stability regions in the (h, h_0) plane. A summary of the techniques and examples of the paper is given here. [62] and [63] treat similar linear two-delay problems for stability regions from a Hopf bifurcation viewpoint.

The characteristic equation in question is 5.1.3 and it is convenient to write

$$H(s) = \frac{1 + C(s)G(s)}{C(s)G(s)}, \qquad \delta = h_0 - h$$

with $-h \leq \delta < \infty$ for non-negative delays. Note that $1/H(s)$ is strictly proper and stable, by Smith's principle. So

$$H(s) + \exp(-sh) - \exp(-sh_0) = 0$$

or

$$H(s) + \exp(-sh)(1 - \exp(-s\delta)) = 0$$

Stability at matching. Firstly, for $h_0 = h$, $H(s) = 0$ and for all h the characteristic equation has all roots in the left-half plane; the line $h_0 = h$ in the (h, h_0) plane will belong to a stable area.

Stability avenue about $\delta = 0$. To find a stability avenue $|\delta| < \delta_c$ for which the system remains stable about $h_0 = h$ for all h, a similar approach is taken as for the single-delay case. For roots on the imaginary axis,

$$\begin{aligned} H(i\omega) + \exp(-i\omega h)(1 - \exp(-i\omega\delta)) &= 0 \\ H(-i\omega) + \exp(i\omega h)(1 - \exp(i\omega\delta)) &= 0 \end{aligned}$$

must be satisfied simultaneously. Eliminating h,

$$\begin{aligned} \tilde{W}(\omega^2, \delta) &= H(i\omega)H(-i\omega) - (1 - \exp(-i\omega\delta))(1 - \exp(i\omega\delta)) \\ &= |H(i\omega)|^2 - 2(1 - \cos(\omega\delta)) = 0 \end{aligned} \tag{5.5}$$

If $\tilde{W}(\omega^2, \delta)$ is strictly positive for all $\omega^2 > 0$ for a given δ then there will be stability independent of h. $\delta = 0$ is one such value. The value of ω corresponding to the boundary of the stability region $0 \leq \delta < \delta_c$ will satisfy equation 5.5 and also $\partial\tilde{W}/\partial\omega^2 = 0$. Given

$H(s)$, these equations may be solved numerically for ω and δ_c.

An alternative approach is to consider

$$\frac{d\delta}{d\omega^2} = -\frac{\partial \tilde{W}}{\partial \omega^2} \Big/ \frac{\partial \tilde{W}}{\partial \delta}$$

when $\partial \tilde{W}/\partial \omega^2 = 0$ (since $\partial \tilde{W}/\partial \delta$ is negative in the region of this minimum). This implies $d\delta/d\omega = 0$. But from equation 5.5,

$$|H(i\omega)|^2 = 4\sin^2\left(\frac{\omega\delta}{2}\right)$$

so that

$$\delta_c = \min_{\omega}\left\{\frac{2}{\omega}\sin^{-1}\left(\frac{|H(i\omega)|}{2}\right)\right\}$$

Again given $H(s)$, δ_c may be found numerically. From equation 5.5 it follows that δ may be positive or negative and there will be stability for $|\delta| < \delta_c$. The corresponding values of h may be found from the original characteristic equation

$$\exp(i\omega_c h_c) = \frac{\exp(-i\omega_c\delta_c) - 1}{H(i\omega_c)}$$

of which there are infinitely many solutions equispaced in h by $2\pi/\omega_c$. (The successive h values will be different for δ positive and negative but will have the same spacing.) h_0 is then given by $h_c + \delta_c$.

Stability for $h = 0$ and $h_0 = 0$. It is of interest to determine stability behaviour for $h = 0$ and $h_0 = 0$ for stability on the axes of the (h, h_0) diagram. This gives two single-delay problems

$$H(s) + 1 - \exp(-sh_0) = 0 \quad \text{for} \quad h = 0 \tag{5.6}$$

$$H(s) - 1 + \exp(-sh) = 0 \quad \text{for} \quad h_0 = 0 \tag{5.7}$$

Stability avenue about $h = 0$ or $h_0 = 0$. If equation 5.6 is stable independent of h_0 a stability avenue about $h = 0$ may be sought. Eliminating h_0 from the simultaneous equations

$$\begin{aligned} H(i\omega) + \exp(-i\omega h) - \exp(-i\omega h_0) &= 0 \\ H(-i\omega) + \exp(i\omega h) - \exp(i\omega h_0) &= 0 \end{aligned} \tag{5.8}$$

gives

$$\tilde{W}(\omega^2, h) = (H(i\omega) + \exp(-i\omega h))(H(-i\omega) + \exp(i\omega h)) - 1 = 0 \tag{5.9}$$

which may be solved with $\partial\tilde{W}(\omega^2,h)/\partial\omega^2 = 0$ for the critical values of ω and h by one of the two approaches used for the avenue about $\delta = 0$. Similarly if equation 5.7 is stable independent of h, a stability avenue about $h_0 = 0$ is found by eliminating h, etc.

Stability regions in the (h, h_0) plane. Each critically stable point (h_c, h_{0c}) delimiting the stability avenue about $\delta = 0$ forms part of a curve in the plane corresponding to a solution of the characteristic equation on the imaginary axis. To find these curves, equation 5.8 may be split into real and imaginary parts and ω eliminated. Alternatively, ω may be taken as a parameter and for example equation 5.9 solved for h in terms of ω and equation 5.8 solved for h_0. Plots of the curves are given for the two examples which follow; crossing each curve from a stable region corresponds to two poles entering the right-half plane, and, as each envelope bounded by a curve is entered, a further two cross to the right-half plane.

Examples

1.

$$C(s) = 1, \qquad G(s) = G_0(s) = \frac{1}{s}$$

The characteristic equation becomes

$$s + 1 + \exp(-sh) - \exp(-sh_0) = 0 \tag{5.10}$$

with $H(s) = s + 1$ and stability at $h = h_0$. Here

$$\tilde{W}(\omega^2, \delta) = \omega^2 + 1 - 2(1 - \cos(\omega\delta))$$

so that

$$\delta_c = \min_{\omega} \left\{ \frac{2}{\omega} \sin^{-1} \left(\frac{\sqrt{\omega^2+1}}{2} \right) \right\} = 1.4775$$

So there will be a stability avenue for $|h - h_0| < \delta_c$.

For $h_0 = 0$ the characteristic equation becomes $s + \exp(-sh) = 0$ which was discussed in section 3.1.6; there is one destabilising value $\omega = 1$, with $h = \pi/2 + 2q\pi$. For $h = 0$, $s + 2 - \exp(-sh_0) = 0$. The system is stable at $h_0 = 0$ and $W(\omega^2) = \omega^2 + 3$ which has no real solution for ω, so stability independent of delay holds and the whole of the h_0 axis lies in a stable region. To find the stability avenue at $h = 0$,

$$\tilde{W}(\omega^2, h) = \omega^2 + 1 - 2\omega\sin(\omega h) + 2\cos(\omega h) = 0 \tag{5.11}$$

and

$$\frac{\partial \tilde{W}(\omega^2, h)}{\partial \omega^2} = 1 - \frac{h}{\omega}\sin(\omega h) - h\cos(\omega h) = 0$$

Figure 5-4 shows the (h, h_0) diagram with the stability avenue about $\delta = 0$ marked by a dotted line. (the dashed line $h_0 = 2h$ is also plotted for comparison with the commensurate delay example of section 5.1.2.) The stability avenue for h small is evident and the curves intersect the h axis at values $\pi/2 + 2q\pi$.

2.

$$C(s) = 1, \qquad G(s) = G_0(s) = \frac{1}{s+1}$$

This system is described by the characteristic equation

$$s^2 + s + 1 + \exp(-sh) - \exp(-sh_0) = 0 \tag{5.12}$$

with $H(s) = s^2 + s + 1$. Stability holds when $h = h_0$, with roots $s = \frac{1}{2}(1 \pm i\sqrt{3})$. Also

$$\tilde{W}(\omega^2, \delta) = \omega^4 - \omega^2 + 1 - 2(1 - \cos(\omega\delta))$$

and

$$\delta_c = \min_{\omega}\left\{\frac{2}{\omega}\sin^{-1}\left(\frac{\sqrt{\omega^4 - \omega^2 + 1}}{2}\right)\right\} = 1.0460$$

So there will be a stability avenue for $|h - h_0| < \delta_c$.

When $h_0 = 0$, $s^2 + s + \exp(-sh) = 0$, stable at $h = 0$ and destabilising at $\omega = \sqrt{\frac{1}{2}(\sqrt{5} - 1)}$, $h = (\tan^{-1}(1/\omega) + 2q\pi)/\omega = 1.1506, 9.1429, \cdots$. For $h = 0$, $s^2 + s + 2 - \exp(-sh_0) = 0$, stable independent of h_0, giving a stability avenue by $h = 0$.

The (h, h_0) diagram appears in figure 5-5, showing the stability avenue and behaviour for $h_0 = 0$ and h near zero.

Multiplicative Parametric Mismatch

Similar stability results are sought for the Smith predictor with plant $KG(s)$ and plant model $G(s)$, having characteristic equation

$$1 + CG + CG(K\exp(-sh) - \exp(-sh_0)) = 0$$

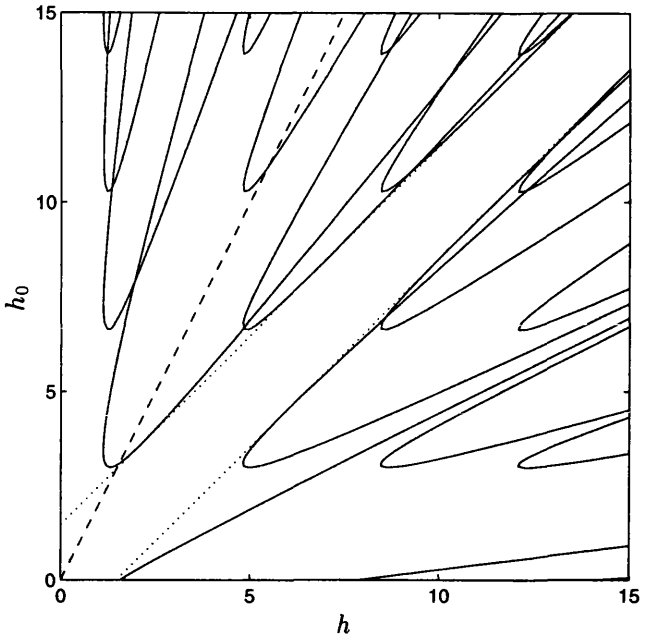


Figure 5-4: Stability Avenue for $s + 1 + \exp(-sh) - \exp(-sh_0) = 0$

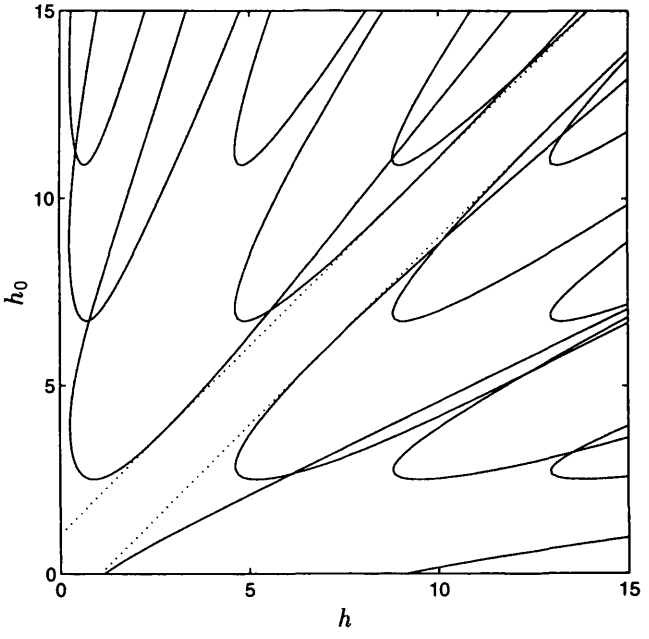


Figure 5-5: Stability Avenue for $s^2 + s + 1 + \exp(-sh) - \exp(-sh_0) = 0$

Again writing $H = (1+CG)/CG$ and $\delta = h_0 - h$,

$$H + \exp(-sh)(K - \exp(-s\delta)) = 0$$

For solutions on the imaginary axis,

$$\tilde{W}(\omega^2, \delta) = |H(i\omega)|^2 - (K-1)^2 - 4K\sin^2\left(\frac{\omega\delta}{2}\right) = 0$$

The question arises as to whether a stability avenue about $\delta = 0$ is still possible for $K \neq 1$. Firstly, at $\delta = 0$, $H + (K-1)\exp(-sh) = 0$, an equation which must be stable independent of h. If this condition holds, the width of the avenue may be found by solving $\tilde{W}(\omega^2, \delta) = 0$ and $\partial\tilde{W}(\omega^2, h)/\partial\omega^2 = 0$ for ω_c and δ_c or alternatively by using

$$\delta_c = \min_{\omega}\left\{\frac{2}{\omega}\sin^{-1}\left(\frac{1}{2}\sqrt{\frac{|H(i\omega)|^2 - (K-1)^2}{K}}\right)\right\}$$

5.1.4 Satché's Method for Two Delays

As was seen in section 3.1.3, Satché's method depends upon separating the delay term from everything else so that the simple Nyquist loci for the inverse plant and delay may be sketched separately and the intersection points found. Extension to multiple delay cases will only be possible if the characteristic equation may be rearranged in such a way that each delay term has a constant coefficient.

Consider the example of the characteristic equation of a Smith predictor system with parametric and temporal mismatch 5.1. Clearly for Satché's method to be applicable, there must either be temporal matching (a single-delay problem), or multiplicative parametric mismatch $G = KG_0$. In the latter case, defining

$$H_0(s) := \frac{1 + C(s)G_0(s)}{C(s)G_0(s)}$$

the characteristic equation may be rewritten

$$H_0(s) + K\exp(-sh) - \exp(-sh_0)$$

The Nyquist plot for $H_0(i\omega)$ is made, together with that for $\exp(-sh_0)$, being the unit circle starting at $(1,0)$ and described clockwise. Points on the locus for $H_0(i\omega)$ corresponding to a particular ω form the centre of a circle radius K. Intersection points of the two circles are then found. This is illustrated in figure 5-6.

ω is known and from the construction $\omega h_0 = \theta_0 + 2m\pi$ and $\omega h = \theta + 2n\pi$, $m, n \epsilon \mathbb{N}$,

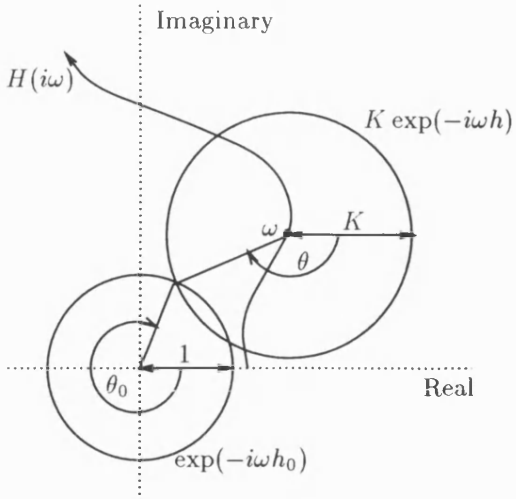


Figure 5-6: Satché's Method for Smith Predictor with Temporal and Multiplicative Parametric Mismatch

may be found for each critical value. The delay mismatch

$$\delta = h_0 - h = \frac{1}{\omega}(\theta_0 - \theta + 2\pi(m-n))$$

is minimised when $m = n$. Repeating the construction procedure for various ω gives δ_c, the minimum value of δ, and from this the width of the stability avenue may be calculated. Note that the circle of radius $K+1$ centred on the origin gives the limit of ω for critical stability and intersection. Similar 'catching-up' arguments apply as in the single delay case.

This process is laborious and requires accurate geometrical construction; the extension of the direct method in section 5.1.3 is a much faster analytic method.

Example Consider the equation 5.10. Here the Nyquist locus for $s+1$ is the line with real part equal to unity and closed to the right as a 'D'-contour, with the locus for $\exp(-sh_0)$ being the unit circle centre the origin. As ω increases from zero, there are circles centre $(1,\omega)$. For $\omega = 1$ for example, intersections occur at $\omega h = \pi/2$ or π, $\omega h_0 = 0$ or $3\pi/2$ respectively. (The former reduces to the single delay problem $s + \exp(-sh) = 0$ treated earlier.)

5.1.5 The Vector Approach for Multiple Delays

The vector approach has a simple extension to problems with more than one delay.

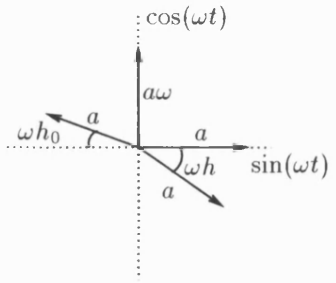


Figure 5-7: Vector Diagram for $\dot{x}(t) + x(t) + x(t-h) - x(t-h_0) = 0$

Example A Smith predictor with parametric matching

$$\dot{x}(t) + x(t) + x(t-h) - x(t-h_0) = 0$$

(corresponding to equation 5.10) has vector diagram as shown in figure 5-7. Resolving in the $\sin(\omega t)$ and $\cos(\omega t)$ directions,

$$\cos(\omega h_0) = \cos(\omega h) + 1 \quad \text{and} \quad \omega + \sin(\omega h_0) = \sin(\omega h)$$

Eliminating h_0,

$$\omega^2 - 2\omega \sin(\omega h) + 2\cos(\omega h) + 1 = 0$$

resulting in the same h_0-independent equation as before (equation 5.11). Given h, critical values of ω may be found from this equation and hence the corresponding values of h_0. As before, destabilising and stabilising values are not distinguished and the number of stability windows is not clear.

5.1.6 Temporal Locus for Two Delays

The method of section 3.1.10 has been extended to the two-delay problem, in particular for the roots of the characteristic equation 5.1 of a Smith predictor with both parametric and temporal mismatch. Other systems would involve a simple modification. The case of fixed h and h_0 increasing from zero has been considered. The analysis could equally well be done for h_0 fixed and h varying, or considering one of h, h_0 fixed and $\delta = h_0 - h$ varying. The latter does pose the problem of the treatment of negative delays for $h > h_0$.

Method and Rules for Locus Construction

The method of solution is again by Euler-Newton Continuation or Padé approximant, the latter requiring caution due to inaccuracies in the two delay models. Construction rules are given for the important characteristics, followed by simplifications for the special case $G = G_0$ (parametric matching). Examples include parametric matched and mismatched cases.

The characteristic equation 5.1 is rewritten in the following form

$$C_D G_D G_{0D} + C_N G_{0N} G_D + C_N G_N G_{0D} \exp(-sh) - C_N G_{0N} G_D \exp(-sh_0) = 0 \quad (5.13)$$

where

$$C = \frac{C_N}{C_D}, \quad G = \frac{G_N}{G_D}, \quad G_0 = \frac{G_{0N}}{G_{0D}}$$

all proper. To simplify notation, define

$$\left.\begin{aligned} C_1 &= C_N G_{0N} G_D \\ A_1 &= G_D(C_D G_{0D} + C_N G_{0N}) = C_D G_D G_{0D} + C_1 \\ A_2 &= C_N G_N G_{0D} \end{aligned}\right\} \quad (5.14)$$

Then

$$F(s, h, h_0) := A_1 + A_2 \exp(-sh) - C_1 \exp(-sh_0) = 0 \quad (5.15)$$

It is assumed that any common factors of A_1, A_2 and C_1 have been removed; these would give rise to spurious roots (an example being the parametrically matched case $G_0 = G$ where G_D is common). Since C, G and G_0 are all proper it follows that the new characteristic equation 5.15 is of retarded type; $C_D G_D G_{0D}$ is of higher order than $C_N G_{0N} G_D$ and $C_N G_N G_{0D}$.

The following rules give features of the loci and form the basis of computer calculation. Where a proof has not been given, it proceeds in the same manner as the single delay case in section 3.1.10.

1. $h = h_0 = 0$. The characteristic equation becomes $1 + CG = 0$ which is stable by Smith's principle. The left-half plane poles may be located by delay-free methods.

2. **Behaviour for $h_0 = 0$.**

$$C_D G_D + C_N G_N \exp(-sh) = 0$$

The problem reduces to a single delay one of retarded type and the direct method and techniques described in section 3.1.10 may be used to ascertain the stability

of the system and location of the (infinitely many) roots, using the previous rule for starting points and increasing h from zero. Hence there will be infinitely many primary loci.

3. **Symmetry.** The locus is again symmetrical with respect to the real axis.

4. h_0 **infinitesimally small and positive.** Arguments here are similar to the single-delay case on increasing the delay from zero (see also [71] and [35]). Infinitely many new roots from infinity are introduced in the left-half plane, with real part σ approximately equal and spaced by $2\pi/h_0$ in imaginary part ω. Therefore the number of unstable poles of the system remains the same as h_0 is increased from zero.

   Writing the characteristic equation

$$1 + CG_0 + CG\exp(-sh) - CG_0\exp(-\sigma h_0)(\cos(\omega h_0) - i\sin(\omega h_0)) = 0$$

   spirals are expected outwards from the single-delay poles due to the trigonometric terms. The effect will be more marked in the left-half plane where $\exp(-\sigma h_0)$ is greater than unity.

5. W^*, **the analogue of** $W(\omega^2)$ **in the single delay case.** For critical stability it is necessary to find a function $W^*(\omega, h)$ independent of h_0 (previously in section 5.1.3 $\tilde{W}$ was a function of ω^2 and δ). When complex conjugate roots cross the imaginary axis,

$$A_1(i\omega) + A_2(i\omega)\exp(-i\omega h) - C_1(i\omega)\exp(-i\omega h_0) = 0$$

   and

$$A_1(-i\omega) + A_2(-i\omega)\exp(i\omega h) - C_1(-i\omega)\exp(i\omega h_0) = 0$$

   will be satisfied simultaneously. Eliminating h_0,

$$\begin{aligned} W^*(\omega, h) \;=\; & A_1(i\omega)A_1(-i\omega) + A_2(i\omega)A_2(-i\omega) - C_1(i\omega)C_1(-i\omega) \\ & +2\mathrm{Re}(A_1(i\omega)A_2(-i\omega)\exp(i\omega h)) = 0 \end{aligned} \tag{5.16}$$

   (Note that: the first three terms on the right-hand side are functions of ω^2; $A_1(i\omega)A_2(-i\omega)$ has real part a function of ω^2, imaginary part ω times a function of ω^2; $\cos(\omega h)$ is a power series in ω^2 and $\sin(\omega h)$ a series in powers of ω^2 with common factor ω. So W^* may be thought of as a function of ω^2 and h.)

6. **Points at which the loci intersect the imaginary axis.** Given h, equation 5.16 may be solved numerically for ω. Successive destabilising values of h_0 are determined by substituting ω into the characteristic equation 5.15 and are separated at each crossing point by $2\pi/\omega$.

7. **Number of crossing points of the imaginary axis.** In the single delay case, $W(\omega^2)$ was a finite polynomial in ω^2 implying a finite number of critical values. In the two-delay case, W^* is given by equation 5.16 and A_1, A_2 and C_1 by equation 5.14. For large ω, W^* is dominated by $A_1(i\omega)A_1(-i\omega)$, a function of ω^2 so that $W^* \to \infty$ as $\omega \to \infty$ as in the single delay case. There will be a finite number of roots for ω as eventually $A_1(i\omega)A_1(-i\omega)$ will dominate the trigonometric terms. These critical values of ω will be finite also. Typical plots are given for the examples which follow.

8. **The 'Sign Theorem'.**

$$\operatorname{sgn}\left(\operatorname{Re}\left(\frac{\partial s}{\partial h_0}\right)\right) = \operatorname{sgn}\left(\frac{\partial W^*}{\partial \omega^2}\right)$$

where $s = i\omega$ for ω a root of equation 5.16, and 5.15 holds.
The proof is similar to that in section 3.1.6, taking h fixed. Destabilising and stabilising values of ω may be distinguished, corresponding to positive and negative slope of W^* respectively. A sketch of W^* against ω may prove helpful; the same result holds for the plot against ω as this represents a scaling of the ω^2 axis.

9. **Argument and modulus rule.** These have no direct analogue for the system with parametric mismatch as the delay terms cannot be separated from everything else.

10. **Real axis part of the locus.** Determination of segments of the real axis forming part of the locus is not as simple as in the single delay case, with no analogue of the argument rule. For real solutions $s = \sigma$ and positive h_0,

$$h_0 = -\frac{1}{\sigma}\ln\left(\frac{A_1(\sigma) + A_2(\sigma)\exp(-\sigma h)}{C_1(\sigma)}\right)$$

Positive h_0 yields the following condition

$$\frac{A_1(\sigma) + A_2(\sigma)\exp(-\sigma h)}{C_1(\sigma)} \begin{cases} \epsilon(0,1) & \text{for } \sigma \text{ real and positive} \\ = 1 & \text{for } \sigma = 0 \text{ or } h_0 = 0 \\ \epsilon(1,\infty) & \text{for } \sigma \text{ real and negative} \end{cases}$$

For a real locus from $-\infty$, $\exp(-\sigma h)$ will be large and positive, thus $A_2(\sigma)/C_1(\sigma)$ must be positive. Now

$$\frac{A_2(\sigma)}{C_1(\sigma)} = \frac{G_N(\sigma)G_{0D}(\sigma)}{G_{0N}(\sigma)G_D(\sigma)}$$

The sign of this expression is dependent on the sign of the coefficients of the leading terms in each of the four polynomials (two of which may be taken to be unity without loss of generality).

11. Sensitivity with respect to h_0.

$$\frac{\partial s}{\partial h_0} = -\frac{\partial F}{\partial h_0} \Big/ \frac{\partial F}{\partial s}$$

$$= \frac{-sC_1 \exp(-sh_0)}{A_1' + \exp(-sh)(A_2' - hA_2) - \exp(-sh_0)(C_1' - h_0 C_1)} \tag{5.17}$$

is infinite on the remote loci where s is infinite and h_0 is infinitesimally small. This gradient is also infinite in the neighbourhood of a multiple root when the denominator of equation 5.17 vanishes. $\partial s/\partial h_0 \to 0$ when $h \to \infty$ (see rule 15).

12. Gradients of the loci.

$$\frac{\partial \omega}{\partial \sigma} = \frac{\mathrm{Im}\,(\partial s/\partial h_0)}{\mathrm{Re}\,(\partial s/\partial h_0)}$$

At $s = i\omega$, the gradient becomes increasingly steep negatively, or positively, at successive destabilising, or stabilising values of ω respectively (for $\omega > 0$; the reverse holds for $\omega < 0$).

13. $h_0 = h > 0$.

$$1 + CG_0 + C(G - G_0)\exp(-sh) = 0$$

Single delay techniques may be applied. Infinitely many roots go to $-\infty$ and then come from $-\infty$ as h_0 passes through h.

14. Roots approaching/departing from $-\infty$. As already seen, this will be the case when h_0 increases from zero and when h_0 passes through h.

There will be a finite number of roots if there is no resultant delay in the characteristic equation for all possible solutions s at these delay values. This will occur in the matched case when $h_0 = h$.

15. Asymptotic Behaviour as $h_0 \to \infty$. There are three possibilities:

- left-half plane roots of $C_1(s)$

- right-half plane roots of the single delay equation

$$A_1(s) + A_2(s)\exp(-sh) = 0$$

- the origin.

Spirals are expected at the non-zero solutions.

16. **Uniqueness of the locus.** The loci do not cross at *different* values of h_0 except at critical values on the imaginary axis.

17. **Multiple roots.** Multiple roots are found by differentiating the characteristic equation partially with respect to s.

18. **Stability independent of h_0 for fixed h.** This follows if the system is stable at $h_0 = 0$ and there are no real solutions of $W^*(\omega, h) = 0$.

Both the Euler-Newton Continuation program and that using the Padé approximant are extensions of those for the single delay case described in appendix A.1. For solution by the continuation algorithm, the starting points on the loci at $h_0 = 0$ are found from temporal locus methods for one delay. Iteration begins from these and h_0 (or equivalently arc length along the curves in (σ, ω, h_0) space) is increased from zero. Given h, critical values of ω and corresponding delays h_0 may be found numerically from W^*, and these also provide starting points for iteration. For any portions of the loci not emanating from the single-delay roots or having no intersection with the imaginary axis, the Padé approximant will give approximate roots. The existence of infinitely many primary loci necessitates a high order approximant.

Parametric Matching and Multiplicative Mismatch

In this section is considered the special case $G(s) = KG_0(s)$, $K > 0$ (multiplicative parametric mismatch) and in particular $K = 1$ (parametric matching). The equations 5.1, 5.14 and 5.15 describing the system reduce to

$$1 + CG_0(1 + K\exp(-sh) - \exp(-sh_0)) = 0$$

$$\left.\begin{aligned} C_1 &= C_N G_{0N} \\ A_1 &= C_D G_{0D} + C_N G_{0N} = C_D G_{0D} + C_1 \\ A_2 &= KC_N G_{0N} = KC_1 \end{aligned}\right\}$$

$$F(s, h, h_0) = A_1 + C_1(K\exp(-sh) - \exp(-sh_0)) = 0$$

Properties carry through as before with the simplified polynomials above. Some of the rules admit simplification or special properties.

5. $W^*(\omega, h)$.

$$\begin{aligned} W^*(\omega,h) &= A_1(i\omega)A_1(-i\omega) + (K^2-1)C_1(i\omega)C_1(-i\omega) \\ &\quad +2K\mathrm{Re}(A_1(i\omega)C_1(-i\omega)\exp(i\omega h)) = 0 \end{aligned}$$

and with parametric matching,

$$W^*(\omega,h) = A_1(i\omega)A_1(-i\omega) + 2\mathrm{Re}(A_1(i\omega)C_1(-i\omega)\exp(i\omega h)) = 0$$

7. **Bound on critical values of ω.** Writing the characteristic equation

$$\frac{1+CG_0}{CG_0} + K\exp(-sh) - \exp(-sh_0) = 0,$$

defining $H_0 = (1+CG_0)/CG_0$, taking real and imaginary parts at $s = i\omega$ and squaring and adding yields

$$|H_0(i\omega)|^2 = K^2 + 1 - 2K\cos(\omega(h_0 - h))$$

The left-hand side of the above equation is a function of ω^2 whilst the right-hand side may be bounded in terms of K giving a useful bound on critical values of ω.

9. **Argument and modulus rule.** Writing

$$\frac{A_1(s)}{C_1(s)} = \exp(-sh_0) - K\exp(-sh)$$

and taking arguments and moduli results in the angle balance rule and modulus rule for the two delay case.

$$\sum_i \arg(s-P_i) - \sum_i \arg(s-Z_i) = \arg(\exp(-sh_0) - K\exp(-sh)) \qquad (5.18)$$

and

$$\frac{\prod_i |s-P_i|}{\prod_i |s-Z_i|} = \frac{}{\sqrt{\exp(-2\sigma h_0) + K^2\exp(-2\sigma h) - 2K\exp(-\sigma(h+h_0))\cos(\omega(h-h_0))}}$$

where P_i, Z_i are the roots of A_1, C_1 respectively.

Angle of departure from the $h_0 = 0$ roots may be found from equation 5.18 with $h = 0$, the right hand side becoming $-\omega h + (2q+1)\pi$, $q \epsilon \mathbb{N}$.

10. Real axis part of the locus. The above argument and modulus rule expressions are complicated. However for real axis solutions, $\omega = 0$ and

$$\sum_i \arg(s - P_i) - \sum_i \arg(s - Z_i) = \arg(\exp(-\sigma h_0) - K\exp(-\sigma h))$$

$$= \begin{cases} 2q\pi \\ (2q+1)\pi \end{cases} \text{ for } \exp(-\sigma h_0) - K\exp(-\sigma h) \begin{cases} > 0 \\ < 0 \end{cases}$$

where $q \epsilon \mathbb{N}$. From this, the real axis portion of the locus may be determined. The locus to the right, or left, of an odd number of roots of A_1 and C_1 according as $\exp(-\sigma h_0) - K\exp(-\sigma h)$ is positive, or negative, respectively. As in the single delay case this is only a necessary condition and it must be verified that the modulus rule also holds for positive h_0. In practice the condition for the general mismatched case may prove simpler, involving only one delay term.

From the analysis given for the mismatched case, there will always be a real locus from $-\infty$ since $A_2(\sigma)/C_1(\sigma) = K$ is positive.

13. $h_0 = h > 0$. In the matched case the characteristic equation becomes

$$1 + CG = 0$$

a finite polynomial which is stable by Smith's principle. There is no resultant delay in the equation and hence a finite number of roots. As h_0 passes through h the infinitely many roots of the characteristic equation go to and subsequently depart from infinity (apart from those passing through the roots of $1 + CG = 0$). All roots remain in the left-half plane as the system is retarded and the number of right-half plane roots is always finite. $h_0 = h$ corresponds to $\delta = 0$ and there will be an interval of stability about this value. The temporal loci may be studied in conjunction with the appropriate (h, h_0) diagrams, to compare regions of stability, the number of poles in the right-half plane, etc.

Examples

First the two stability avenue examples of section 5.1.3 are considered, so that comparison with the (h, h_0) diagrams is possible. These are both examples of parametric matching. The third example is one of multiplicative parametric together with temporal mismatch. The locus method is however applicable whatever the mismatch may

be.

In each of the locus plots, loci commence at $h_0 = 0$ at points marked '+', the position of the loci at $h_0 = h$ is marked '×' and any limit points as $h_0 \to \infty$ other than the origin are marked '•'.

1. First order with parametric matching.

$$s + 1 + \exp(-sh) - \exp(-sh_0) = 0$$

with

$$C(s) = 1, \qquad G(s) = G_0(s) = \frac{1}{s}$$

and

$$A_1 = s + 1, \quad A_2 = 1, \quad C_1 = 1$$

In the first example of section 5.1.3 it was seen that for both delays zero there is a single root $s = -1$. Also keeping h_0 zero and taking positive fixed h, $s + \exp(-sh) = 0$, giving the positions of the infinitely many roots. This system is stable for $h = 0$ with destabilising value $\omega = 1$, $h = (2q + 1/2)\pi$. The particular case $h = 3$ has been studied, a value for which there will be two right-half plane poles at $h_0 = 0$.

For critical stability,

$$W^*(\omega, h) = \omega^2 + 1 - 2\omega\sin(\omega h) + 2\cos(\omega h) = 0$$

and critical values are bounded by $\omega^2 = 1 - 2\cos(\omega(h_0 - h)) < 3$. The upper part of figure 5-8 shows a plot of W^* against ω for $h = 3$ on two scales. Numerical calculation yields $\omega_{(1)} = 0.5573$ (stabilising) and $\omega_{(2)} = 1.0354$ (destabilising) with corresponding delay values given in the following table:

$h_{0(1)}$	$h_0 = 0$	0.8128			12.0877		
$h_{0(2)}$			4.5519	10.6203		16.6887	22.7571
No RHP roots	2	0	2	4	2	4	6

There is one stability window $0.8128 < h_0 < 4.5519$.

The condition for real solutions σ reduces to

$$\sigma + \exp(-\sigma h) \begin{cases} \epsilon(-1, 0) & \text{for } \sigma \text{ real and positive} \\ \epsilon(0, \infty) & \text{for } \sigma \text{ real and negative} \end{cases}$$

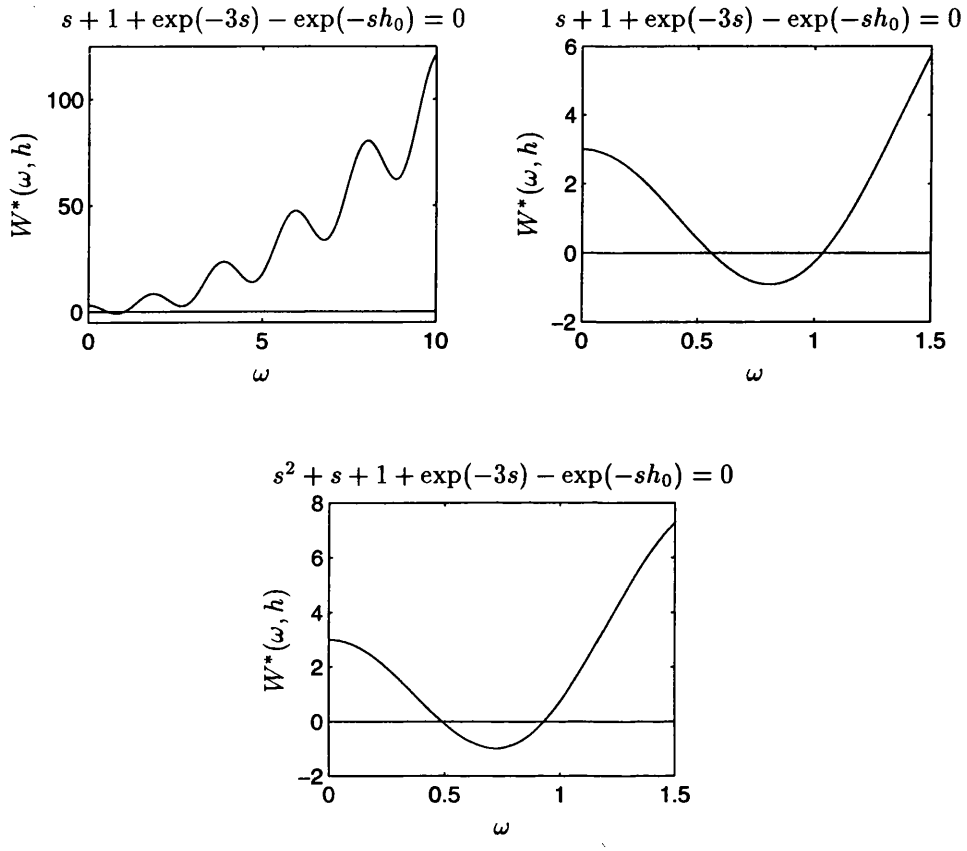


Figure 5-8: $W^*(\omega, h)$ for $s + 1 + \exp(-3s) - \exp(-sh_0) = 0$ and $s^2 + s + 1 + \exp(-3s) - \exp(-sh_0) = 0$

There is no solution for positive σ. A locus from $-\infty$ will exist as the system is matched parametrically. The range of σ will be dependent on the value of h; in particular for $h = 3$ the entire negative real axis forms part of the locus.

When $h_0 = h$, there is no resultant delay in the characteristic equation and $s = -1$. One root passes through this point with the remainder approaching and departing from $-\infty$. The (h, h_0) diagram is given in figure 5-10, with the line $h = 3$ marked for comparison with critical stability results obtained above.

As $h_0 \to \infty$, all roots tend to the origin from the left-half plane, there being no right-half plane roots of $s+1+\exp(-sh) = 0$ for any h, and $C_1(s)$ being constant.

A double negative real root occurs at $s = -1.2639$, $h_0 = 2.9953$. Stability independent of h_0 will hold for small h, as seen previously.

Figures 5-11 and 5-12 give the temporal locus plots to illustrate the features described above, with figure 5-13 providing greater detail of the critical region. (The first four loci are given; these illustrate the properties above and the behaviour of higher order loci is similar.) The loci may be seen to depart to $-\infty$ and subsequently emerge (apart from the locus passing through $s = -1$).

2. Second order with parametric matching.

$$s^2 + s + 1 + \exp(-sh) - \exp(-sh_0) = 0$$

with

$$C(s) = 1, \qquad G(s) = G_0(s) = \frac{1}{s(s+1)}$$

and

$$A_1 = s^2 + s + 1, \quad A_2 = 1, \quad C_1 = 1$$

In the second example of section 5.1.3 it was seen that for $h_0 = h = 0$, $s^2+s+1 = 0$ with roots $s = \frac{1}{2}(-1 \pm i\sqrt{3})$, and that for h_0 zero and $h = 3$, $s^2+s+\exp(-sh) = 0$ with two right-half plane poles.

The function W^* is given by

$$W^*(\omega, h) = \omega^4 - \omega^2 + 1 - 2(\omega^2 - 1)\cos(\omega h) - 2\omega \sin(\omega h) = 0$$

and the graph is shown in figure 5-8. There are again two critical values, $\omega_{(1)} = 0.4894$ (stabilising) and $\omega_{(2)} = 0.9312$ (destabilising), with corresponding delay values:

$h_{0(1)}$	$h_0 = 0$	1.0825			13.9211		
$h_{0(2)}$			4.0516	10.7990		17.5463	24.2936
No RHP roots	2	0	2	4	2	4	6

There is one stability window $1.0825 < h_0 < 4.0516$.

For the real axis to form part of the locus, σ must satisfy

$$\sigma^2 + \sigma + \exp(-\sigma h) \begin{cases} \epsilon(-1,0) & \text{for } \sigma \text{ real and positive} \\ \epsilon(0,\infty) & \text{for } \sigma \text{ real and negative} \end{cases}$$

There are no positive solutions as $\sigma^2+\sigma+\exp(-\sigma h) > 0$ for all $\sigma > 0$. Parametric matching implies a locus from $-\infty$ and from plots of $\exp(-\sigma h)$ and $-\sigma^2 - \sigma$ for example it may be deduced that the entire negative real axis forms part of the locus.

When $h_0 = h$, one locus passes through each of the complex conjugate roots $s = \frac{1}{2}(-1 \pm i\sqrt{3})$ with the remainder tending to $-\infty$. Figure 5-14 shows the (h, h_0) diagram with stability avenue and the line $h = 3$ marked.

As $h_0 \to \infty$, there are no left-half plane solutions. Solutions in the right-half plane will be at roots of $s^2+s+1+\exp(-sh) = 0$, of which there is a complex conjugate pair when $h = 3$. Spirals to these roots are expected, with the remainder of the loci terminating at the origin.

There will be stability independent of h_0 for small h as was discussed earlier.

The temporal loci are shown in two and three dimensions in figures 5-15 and 5-16.

3. Multiplicative parametric mismatch.

$$s + 1 + K\exp(-sh) - \exp(-sh_0) = 0 \quad K > 0$$

with

$$C(s) = 1, \quad G_0(s) = \frac{1}{s}, \quad G(s) = KG_0(s) = \frac{K}{s}$$

and

$$A_1 = s + 1, \quad A_2 = K, \quad C_1 = 1$$

At $h = h_0 = 0$, $s = -K$ and the system is stable for all $K > 0$. Keeping h_0 zero and taking positive h results in $s + K\exp(-sh) = 0$, an example considered in section 3.1.6 and shown to be critically stable for $\omega = K$, $h = (2q + 1/2)\pi/K$. Again $h = 3$ has been considered, for the particular cases $K = 1/2$, for which stability holds at $h_0 = 0$, and $K = 3/2$, for which there are two right-half plane

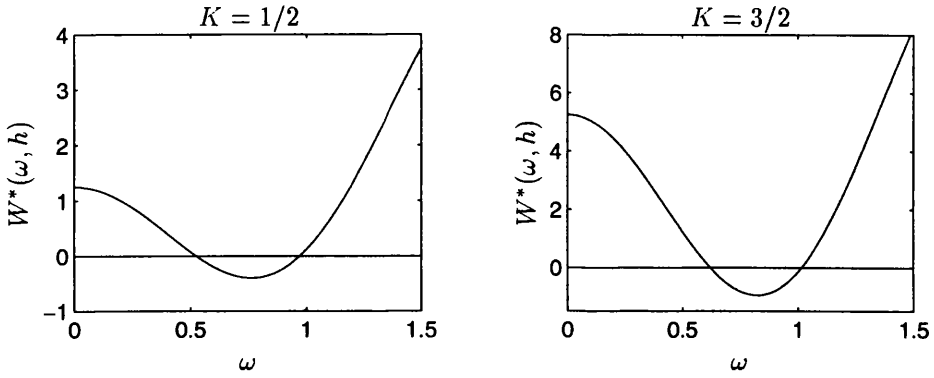


Figure 5-9: $W^*(\omega, h)$ for $s + 1 + K\exp(-3s) - \exp(-sh_0) = 0$

poles for $h_0 = 0$. When $h = 0$ the system is stable for all positive K and h_0.

For h and h_0 both positive,

$$W^*(\omega, h) = \omega^2 + K^2 + 2K(\cos(\omega h) - \omega \sin(\omega h)) = 0$$

and critical values are bounded by

$$\omega^2 = K^2 - 2K\cos(\omega(h_0 - h)) < K(K + 2)$$

Figure 5-9 shows the plot of $W^*(\omega, 3)$ against ω for $K = 1/2$ and $K = 3/2$ and critical values are found to be

$$K = \tfrac{1}{2} \begin{cases} \omega_{(1)} = 0.5238 & h_{0(1)} = 11.9503, 23.9459, \cdots & \text{(stabilising)} \\ \omega_{(2)} = 0.9713 & h_{0(2)} = 5.4061, 11.8749, 18.3437, \cdots & \text{(destabilising)} \end{cases}$$

$$K = \tfrac{3}{2} \begin{cases} \omega_{(1)} = 0.6195 & h_{0(1)} = 1.5482, 11.6898, 21.8314, \cdots & \text{(stabilising)} \\ \omega_{(2)} = 1.0149 & h_{0(2)} = 4.1351, 10.3257, 16.5164, \cdots & \text{(destabilising)} \end{cases}$$

the corresponding h_0 values being found from the original characteristic equation. For $K = 1/2$ there is one stability window $0 \leq h_0 < 5.4061$ and for $K = 3/2$ the system is stable for $1.5482 < h_0 < 4.1351$.

For real solutions for s,

$$\sigma + K\exp(-\sigma h) \begin{cases} \epsilon(-1, 0) & \text{for } \sigma \text{ real and positive} \\ \epsilon(0, \infty) & \text{for } \sigma \text{ real and negative} \end{cases}$$

For any h and K again there will be no positive solutions. It may be shown that

the entire negative real axis forms part of the locus for the values of h and K of interest here.

When $h_0 = h$, $s + 1 + (K-1)\exp(-sh) = 0$ and roots may be found by single delay methods. Given h, the left-half and right-half plane poles may be located. It is not profitable to run the single-delay program when these roots are to be found by the two-delay version, but single-delay rules will give good guidelines as to the general behaviour. For $0 < K < 2$ all roots lie in the left-half plane independent of the value of h, implying the existence of a stability avenue about $\delta = 0$ as before, but for $K > 2$ the 'avenue' terminates at $\omega = \sqrt{K(K-2)}$ and $h = (2\pi q - \tan^{-1}(\sqrt{K(K-2)}))/\sqrt{K(K-2)}$. Figures 5-17 and 5-18 show the (h, h_0) diagrams for $K = 1/2$ and $K = 3/2$ respectively; the avenues, axis behaviour and critical stability for $h = 3$ are evident. Calculations yield $\delta_c = 2.4010$ for $K = 1/2$ and $\delta_c = 1.0462$ for $K = 3/2$, the width of the avenue decreasing as K approaches 2.

As $h_0 \to \infty$, there will be no roots terminating in the left-half plane. Infinitely many will tend to the origin. Roots ending in the right-half plane will be at roots of

$$s + 1 + K\exp(-sh) = 0$$

in the right-half plane, if any. This single-delay problem is stable for zero delay and has one destabilising value of ω if $K > 1$ but is otherwise stable independent of delay. For $K = 3/2$ a locus spirals in to each of the complex conjugate roots $s = 0.0407 \pm 0.8240i$.

It may be shown that when $K = 1/2$ there exists a double negative real root at $s = -0.4762$, $h_0 = 2.0148$.

Figures 5-19 and 5-20 show the two and three dimensional temporal locus plots for $K = 1/2$. The first four loci are plotted; these illustrate the features described above and show the pattern followed by higher order loci. This example with parametric mismatch is different in that unless $h_0 = h = 0$ there is always a delay term present in the equation and hence an infinite number of loci. The $K = 3/2$ example is illustrated in figures 5-21 and 5-22. Comparison with the (h, h_0) diagrams is helpful.

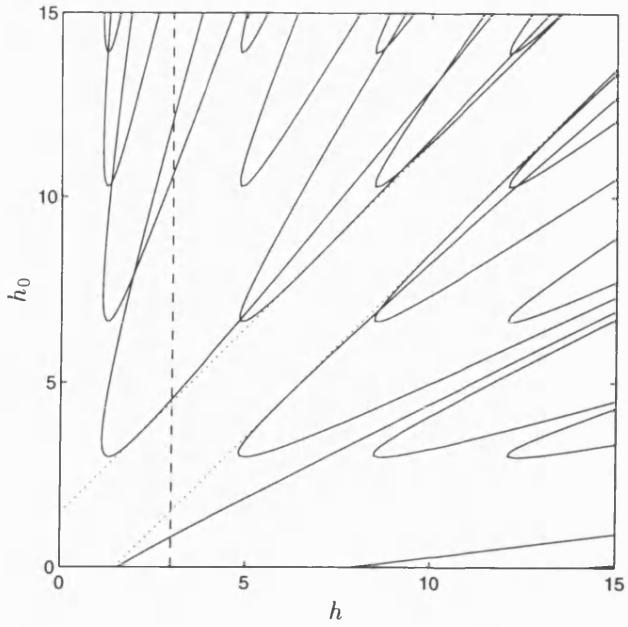


Figure 5-10: Stability Avenue for $s + 1 + \exp(-sh) - \exp(-sh_0) = 0$

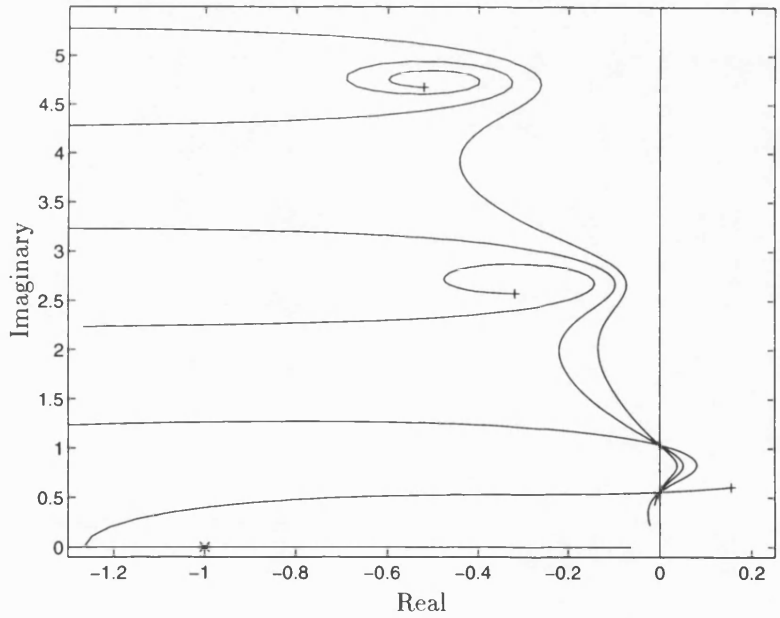


Figure 5-11: $s + 1 + \exp(-3s) - \exp(-sh_0) = 0$

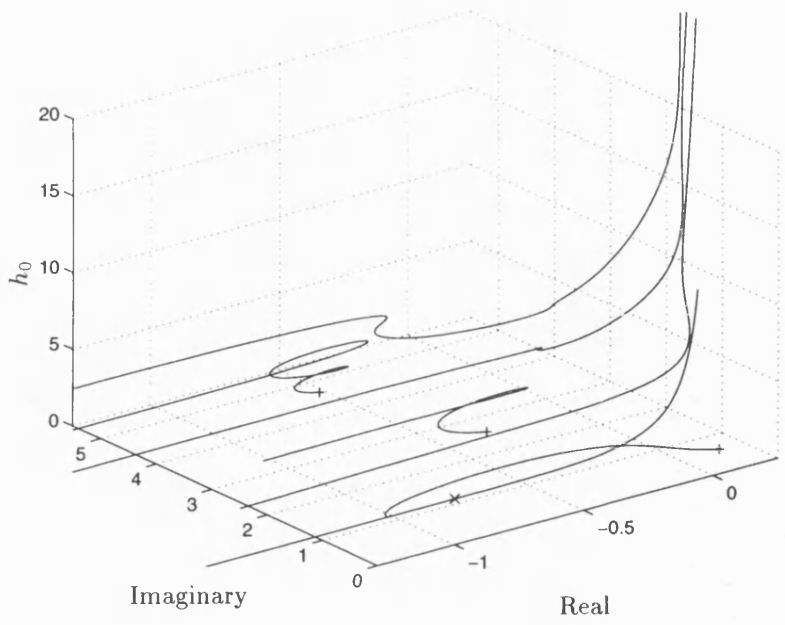


Figure 5-12: $s + 1 + \exp(-3s) - \exp(-sh_0) = 0$

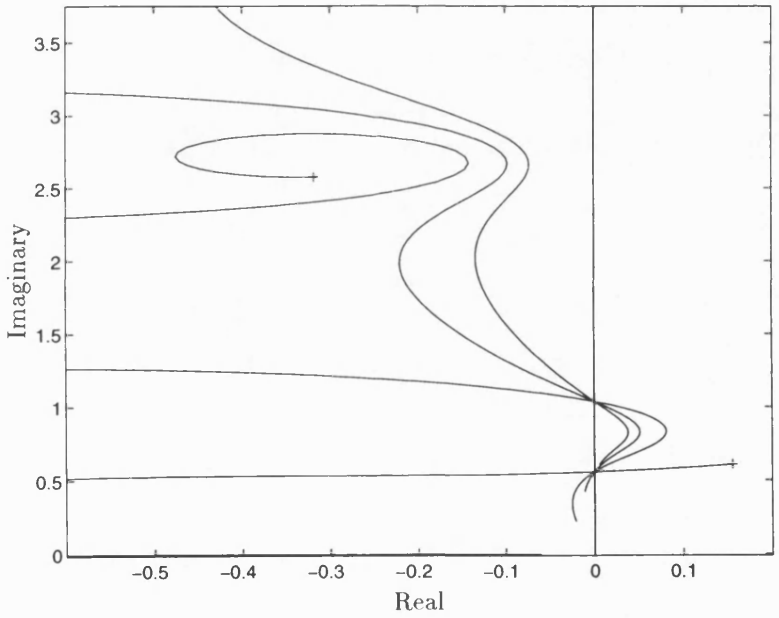


Figure 5-13: $s + 1 + \exp(-3s) - \exp(-sh_0) = 0$

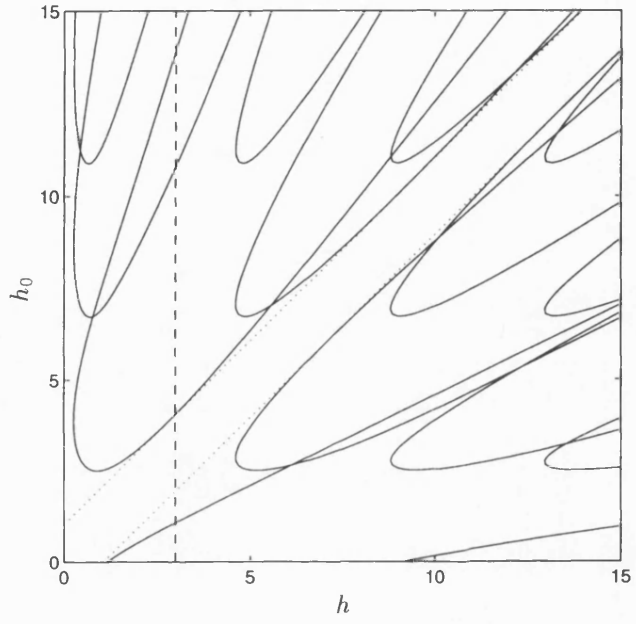


Figure 5-14: Stability Avenue for $s^2 + s + 1 + \exp(-sh) - \exp(-sh_0) = 0$

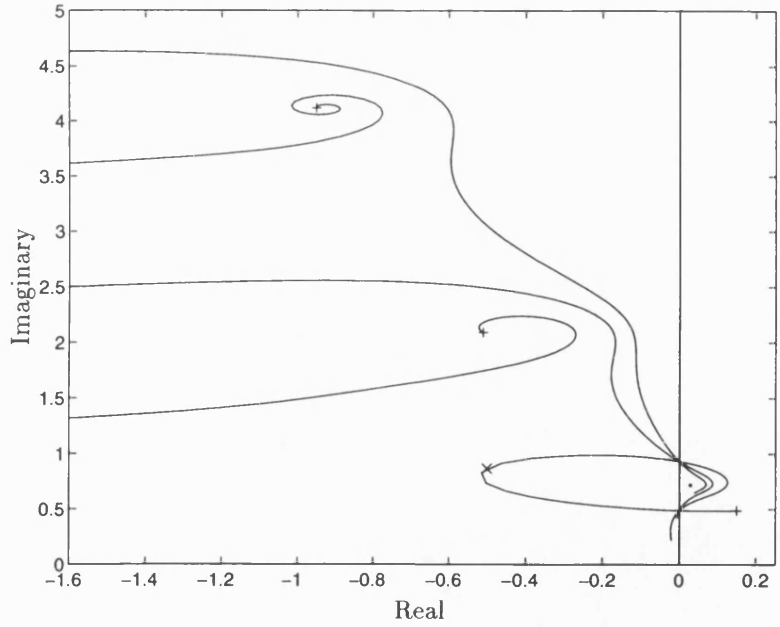


Figure 5-15: $s^2 + s + 1 + \exp(-3s) - \exp(-sh_0) = 0$

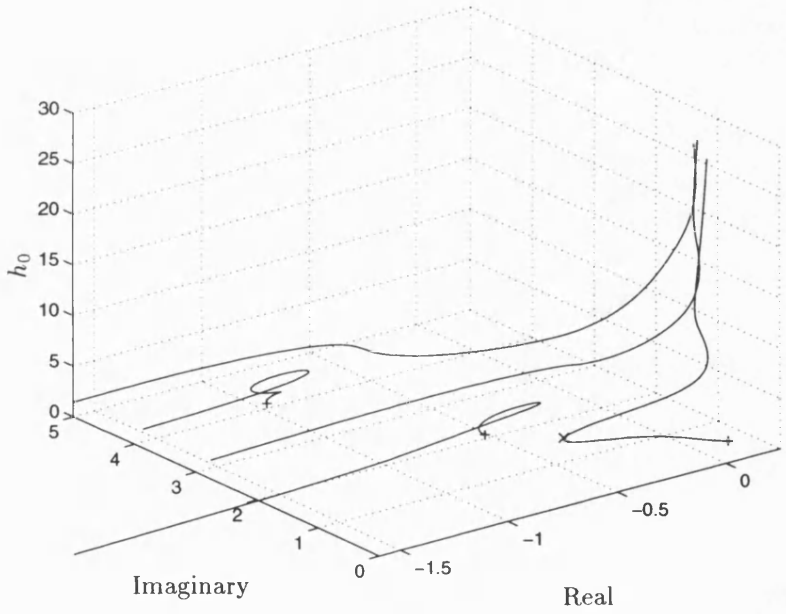


Figure 5-16: $s^2 + s + 1 + \exp(-3s) - \exp(-sh_0) = 0$

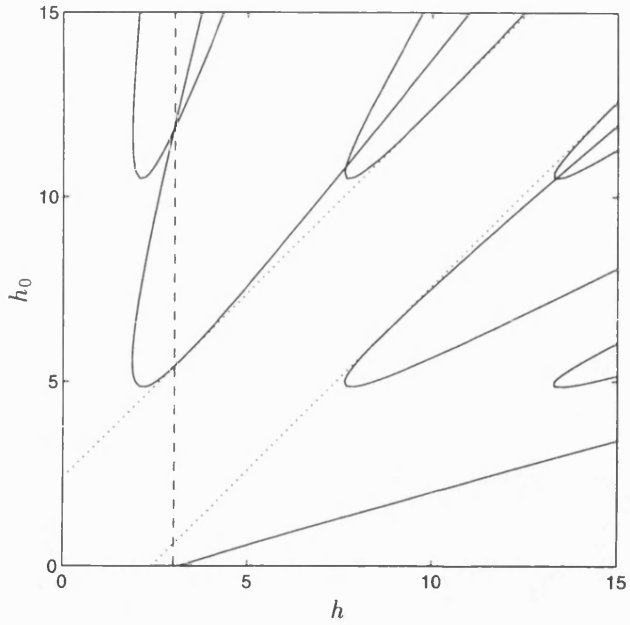


Figure 5-17: Stability Avenue for $s + 1 + K\exp(-sh) - \exp(-sh_0) = 0, K = 1/2$

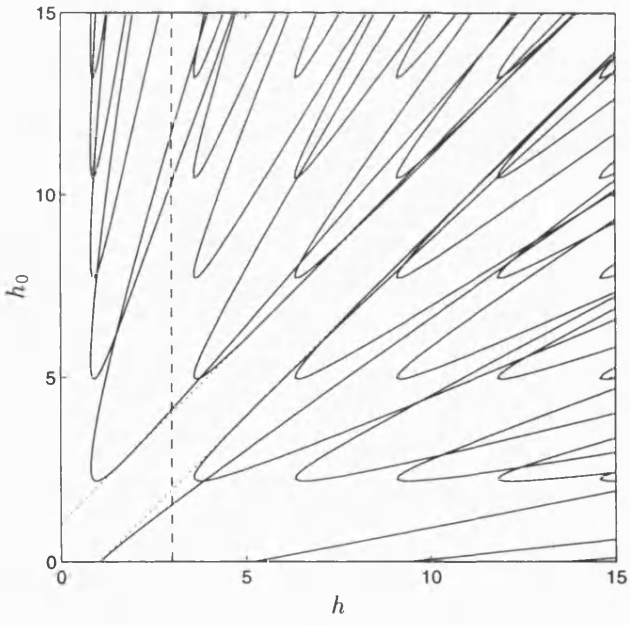


Figure 5-18: Stability Avenue for $s + 1 + K\exp(-sh) - \exp(-sh_0) = 0, K = 3/2$

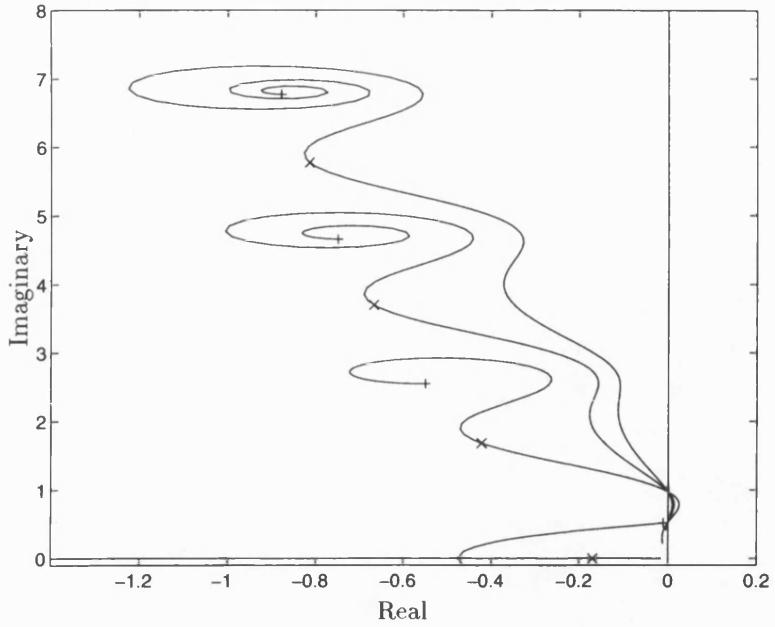


Figure 5-19: $s + 1 + K\exp(-3s) - \exp(-sh_0) = 0,\ K = 1/2$

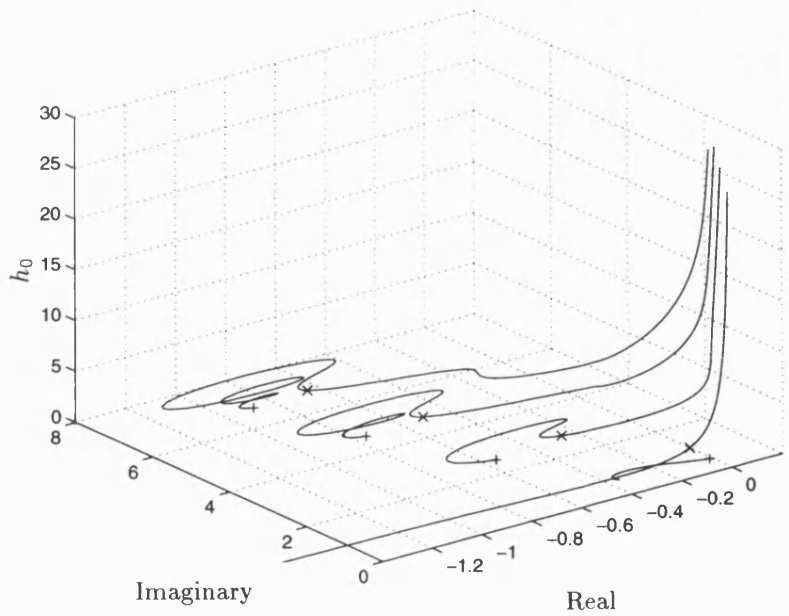


Figure 5-20: $s + 1 + K\exp(-3s) - \exp(-sh_0) = 0,\ K = 1/2$

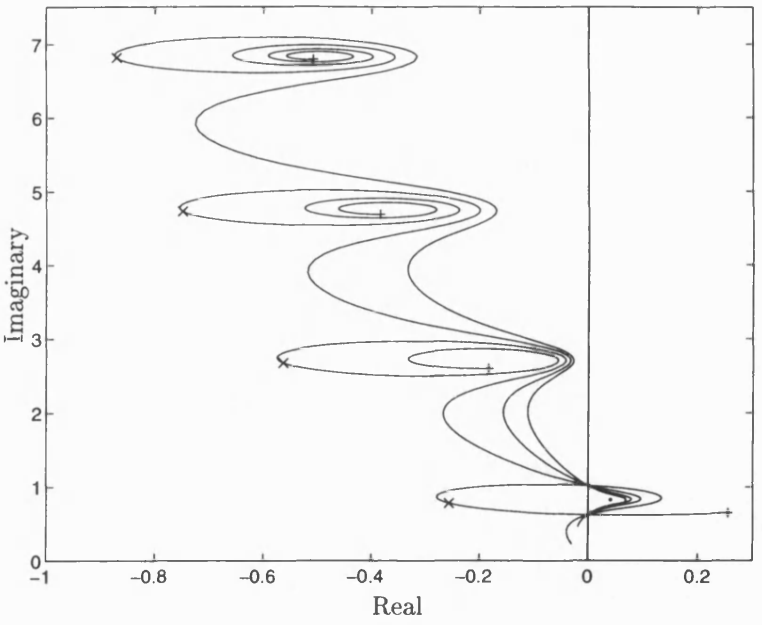


Figure 5-21: $s + 1 + K\exp(-3s) - \exp(-sh_0) = 0,\ K = 3/2$

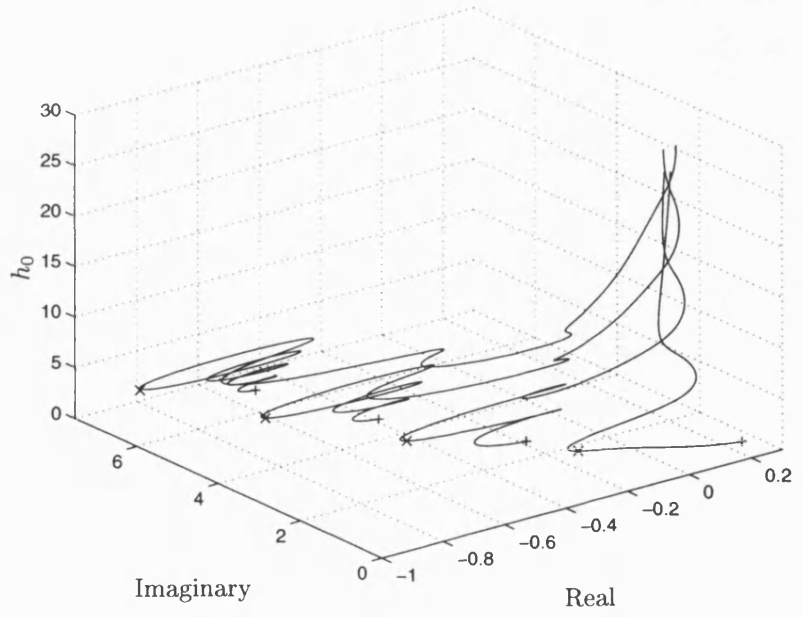


Figure 5-22: $s + 1 + K\exp(-3s) - \exp(-sh_0) = 0,\ K = 3/2$

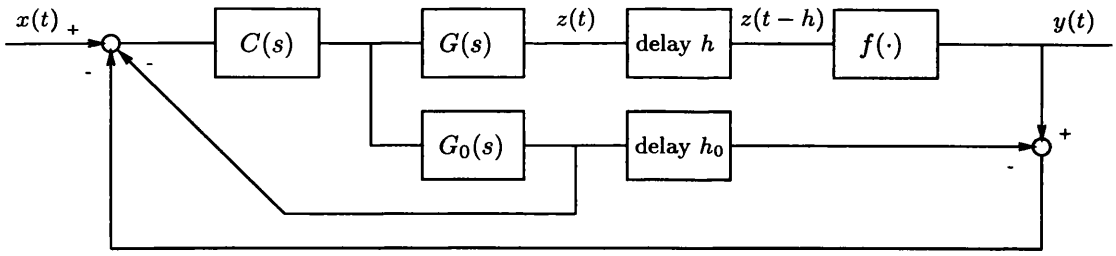


Figure 5-23: Smith Predictor with Saturation

5.2 Nonlinear Two–Delay Problems

5.2.1 Introduction

Again saturation nonlinearities and Van der Pol systems incorporating delay were considered. Smith predictor systems containing saturating elements will be discussed in section 5.2.2.

Two-delay problems of Van der Pol type exhibiting limit cycle behaviour have not been studied in great detail, given the complexity of the single-delay problem. This forms an avenue for future research. Again the vector approach is applicable for small delay values in predicting limit cycles and their approximate amplitude. Stability of the corresponding linear problem and the results and techniques of section 5.1 determine stability in the neighbourhood of the origin in the $(x, \dot{x})$ plane, yielding information on stability 'windows' of any limit cycle. One delay may be kept fixed and the other varied or a perturbation of one delay from the other considered.

5.2.2 Smith Predictor Systems containing Saturating Elements

The subject of this section is the addition of a saturating element to a Smith predictor system as in figure 5-23. $C(s)$, $G(s)$ and $G_0(s)$ are the transfer functions of the controller, plant and plant model respectively. (The Laplace domain is used for convenience and for comparison with linear methods; clearly it is not valid for the nonlinear components.) The saturating element has input-output relation as in figure 3-28 described by

$$y(t) = f(z(t-h)) = \begin{cases} -N & \text{for } z(t-h) \leq -S \\ \frac{N}{S} z(t-h) & \text{for } -S \leq z(t-h) \leq S \\ N & \text{for } z(t-h) \geq S \end{cases}$$

Specialisation is made to two cases.

Strictly a single-delay problem, the case of temporal matching and multiplicative

parametric mismatch $h = h_0$, $G = KG_0$ is treated in this chapter for convenience. Straightforward block diagram manipulation yields the single-loop feedback system shown in figure 5-24 with $w(t) = z(t)/K$ and

$$\frac{1}{H(s)} = \frac{C(s)G_0(s)}{1 + C(s)G_0(s)}$$

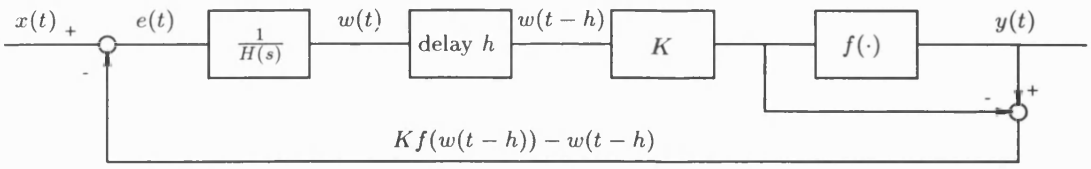
the ideal open-loop transfer function in the linear case. The nonlinear function is now

$$Kf(w(t-h)) - w(t-h) = \begin{cases} -KN - w(t-h) & \text{for} \quad w(t-h) \leq -S \\ (\frac{KN}{S} - 1)w(t-h) & \text{for} \quad -S \leq w(t-h) \leq S \\ KN - w(t-h) & \text{for} \quad w(t-h) \geq S \end{cases}$$

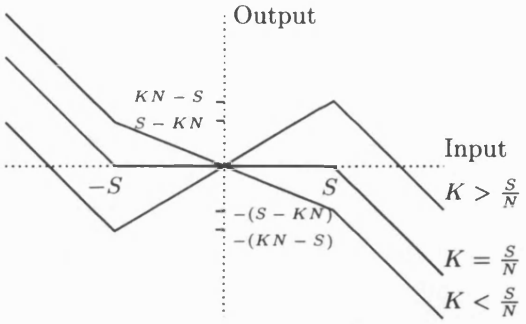
with the input-output relation also given in figure 5-24. Before saturation, feedback is $(KN/S - 1)w(t-h)$ and the system behaves linearly; should K be greater than S/N, the nonlinearity turns over at the saturation point and chaos is to be expected (see [39] for example). Solutions may be found for a given $H(s)$ by solving for $w(t)$ by the method of steps for example during the unsaturated region, finding the value of time for which saturation occurs ($w(t-h) = S$; this will depend on the values of A, K, N and S) and finally solving in the saturated region. For linear dynamics $1/H(s)$ of integrator type, wind-up is expected as in the delay-free case.

For parametric matching and temporal mismatch the system is as shown in figure 5-25. (Note that the delay and nonlinearity commute provided the stored function in the delay is suitably modified, since the delay does not affect amplitude.) A similar approach may be taken as for temporal matching and similar results are to be expected for $\delta = h_0 - h$ small.

Clearly these problems are more complex, requiring results from chaos theory together with those of sections 3.2, 3.1 and 5.1 and provide material for further study.



Equivalent System



Input-Output Relation: $Kf(\cdot) - (\cdot)$

Figure 5-24: Smith Predictor with Saturation: Multiplicative Parametric Mismatch

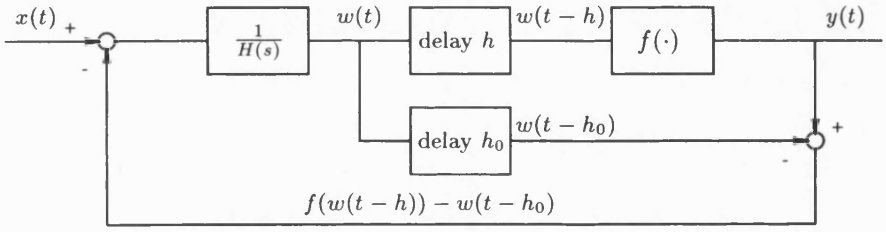


Figure 5-25: Smith Predictor with Saturation: Temporal Mismatch

Chapter 6

Conclusions

Existing methods for the stability of linear time-delay systems and nonlinear delay-free systems have been developed and extended for application to nonlinear delay-differential systems.

A review has been made of existing stability methods in both linear and nonlinear fields.

The direct method of Walton and Marshall for linear systems with delay has been of particular importance. A method for construction of the temporal loci describing stability with delay has been formulated; various construction rules are given and a program written to calculate these numerically. Temporal loci give valuable insights into the stability behaviour of time-delay systems, notably stability windows and 'stability independent of delay', the latter requiring caution as roots on successive loci approach the imaginary axis from the left-half plane.

Nonlinear delay-differential equations of Van der Pol form have been considered. The averaging method of Kryloff and Bogoliuboff found extension for small delays predicting limit cycle behaviour in the state-velocity plane. An equivalent vector approach was applicable to systems containing linear, nonlinear and delay elements. Linear time-delay analysis determined stability of the origin and enlightened stability of any limit cycle as the delay was increased. Delay in the state term gave rise to a linearised system exhibiting stability windows; these were reflected in switches in stability of the origin and limit cycle in the corresponding nonlinear system. 'Failure to alternate' brought the phenomenon of periodic orbits encircling the origin several times during a period. These were accompanied by loops and also beats when the state and velocity are plotted against time, suggesting a synchronisation effect. Increasing the delay still further revealed unsettled orbits filling bounded regions in the state-velocity plane. This unsettled behaviour is believed to be aperiodic, due to the presence of competing

oscillations, rather than chaotic. These phenomena provide an avenue for further study.

Smith predictor schemes with temporal and parametric mismatch were the subject of investigation for linear two-delay problems. Extensions of the direct method formed the basis of a temporal locus method for characteristic equations of such systems, keeping one delay fixed whilst increasing the other from zero. The effect of mismatch on stability was discussed.

This research has broadened understanding of the stability and dynamical behaviour of linear and nonlinear time-delay systems. Many interesting subjects for further investigation have been raised. These include as mentioned earlier the features associated with large delay in the state term of Van der Pol's equation, and similar problems; the addition of saturating elements to Smith predictor schemes, which is expected to give rise to chaos as described in section 5.2; and application of the extended direct method and temporal loci to Van der Pol type systems containing two non-commensurate delays.

Appendix A

Temporal Locus Algorithms and Programs

A.1 Temporal Locus by Euler-Newton Continuation

The solution of

$$F(s,h) = A(s) + C(s)\exp(-sh) = 0 \tag{A.1}$$

$s = \sigma + i\omega$ is required. Complete details of the algebra and numerics are not given here, but information on Euler-Newton Continuation may be found in standard numerical analysis texts such as [47]. Details specific to the temporal locus problem are described in what ensues.

$$F(s,h) = F(\sigma,\omega,h) = f_1(\sigma,\omega,h) + if_2(\sigma,\omega,h)$$

say, $f_1, f_2 : \mathbb{R}^3 \to \mathbb{R}$. Write

$$\boldsymbol{y} = \begin{pmatrix} \boldsymbol{x} \\ h \end{pmatrix} = \begin{pmatrix} \sigma \\ \omega \\ h \end{pmatrix}$$

($\boldsymbol{x}$ will denote the real 2 by 1 vector $\begin{pmatrix} \sigma \\ \omega \end{pmatrix}$ and s the complex scalar $\sigma + i\omega$.) Define

$$\boldsymbol{G}(\boldsymbol{x},h) := \begin{pmatrix} f_1(\boldsymbol{x},h) \\ f_2(\boldsymbol{x},h) \end{pmatrix} = \boldsymbol{0} \quad \boldsymbol{G} : \mathbb{R}^3 \to \mathbb{R}^2 \tag{A.2}$$

and let $\mathcal{L}$ be the locus of solution points of the characteristic equation:

$$\mathcal{L} = \{(\boldsymbol{x}, h)\epsilon\mathbb{R}^3 : \boldsymbol{G}(\boldsymbol{x}, h) = \mathbf{0}\}.$$

t will denote the arc length along $\mathcal{L}$ from a reference point (in this case the starting point of solution). Then it may be shown that provided

$$\text{Rank}\left(\boldsymbol{G}_{\boldsymbol{x}} \;\middle|\; \boldsymbol{G}_h \right) = 2$$

where $\boldsymbol{G}_{\boldsymbol{x}}$, $\boldsymbol{G}_h$ is the Jacobian with respect to $\boldsymbol{x}$, h respectively, the implicit function theorem holds and the system may be parameterised in terms of t:

$$\begin{pmatrix} \boldsymbol{x} \\ h \end{pmatrix} = \begin{pmatrix} \boldsymbol{x}(t) \\ h(t) \end{pmatrix}$$

with starting point

$$\begin{pmatrix} \boldsymbol{x}_0 \\ h_0 \end{pmatrix} = \begin{pmatrix} \boldsymbol{x}(t_0) \\ h(t_0) \end{pmatrix}$$

and $\mathcal{L}$ will be a unique curve in $\mathbb{R}^3$. The case when this rank condition fails is discussed later.

A third equation is introduced which represents the equation of a plane perpendicular to the tangent vector to $\mathcal{L}$ at $\begin{pmatrix} \boldsymbol{x}_0 \\ h_0 \end{pmatrix}$ and a distance Δt away. This is as follows

$$n(\boldsymbol{x}, h, t) = \begin{pmatrix} \dot{\boldsymbol{x}}_0^T & \dot{h}_0 \end{pmatrix} \begin{pmatrix} \boldsymbol{x} - \boldsymbol{x}_0 \\ h - h_0 \end{pmatrix} - \Delta t = 0$$

Here $\begin{pmatrix} \dot{\boldsymbol{x}}_0 \\ \dot{h}_0 \end{pmatrix}$ is the unit tangent vector to the locus at the reference point. Now define

$$\boldsymbol{F}(\boldsymbol{x}, h, t) = \begin{pmatrix} \boldsymbol{G}(\boldsymbol{x}, h) \\ n(\boldsymbol{x}, h, t) \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ 0 \end{pmatrix} \tag{A.3}$$

where $\boldsymbol{F} : \mathbb{R}^3 \times \mathbb{R} \to \mathbb{R}^3$ (noting that the vector $\boldsymbol{F}$ is distinct from the scalar F of equation A.1).
The following derivatives are required for the solution.

1. $\boldsymbol{G}_{\boldsymbol{x}}(\boldsymbol{x}, h, t)$, the Jacobian of $\boldsymbol{G}$ with respect to $\boldsymbol{x}$, has a simple form. $F(\sigma, \omega, h)$

is analytic so the Cauchy-Riemann equations apply:

$$\frac{\partial f_1}{\partial \sigma} = \frac{\partial f_2}{\partial \omega} \quad \text{and} \quad \frac{\partial f_1}{\partial \omega} = -\frac{\partial f_2}{\partial \sigma}$$

Now

$$\begin{aligned}
\boldsymbol{G_x}(1,1) &= \frac{\partial f_1}{\partial \sigma} = \frac{\partial(\mathrm{Re}(F))}{\partial s} \cdot \frac{\partial s}{\partial \sigma} = \mathrm{Re}\left(\frac{\partial F}{\partial s}\right) \\
\boldsymbol{G_x}(1,2) &= \frac{\partial f_1}{\partial \omega} = -\frac{\partial f_2}{\partial \sigma} = -\mathrm{Im}\left(\frac{\partial F}{\partial s}\right) \\
\boldsymbol{G_x}(2,1) &= \frac{\partial f_2}{\partial \sigma} = -\boldsymbol{G_x}(1,2) \\
\boldsymbol{G_x}(2,2) &= \frac{\partial f_2}{\partial \omega} = \boldsymbol{G_x}(1,1)
\end{aligned}$$

So

$$\boldsymbol{G_x} = \begin{pmatrix} \mathrm{Re}\left(\frac{\partial F}{\partial s}\right) & -\mathrm{Im}\left(\frac{\partial F}{\partial s}\right) \\ \mathrm{Im}\left(\frac{\partial F}{\partial s}\right) & \mathrm{Re}\left(\frac{\partial F}{\partial s}\right) \end{pmatrix}$$

2. $\boldsymbol{G_h}(\boldsymbol{x}, h, t)$, the Jacobian of $\boldsymbol{G}$ with respect to h.

$$\boldsymbol{G_h} = \begin{pmatrix} \frac{\partial f_1}{\partial h} \\ \frac{\partial f_2}{\partial h} \end{pmatrix} = \begin{pmatrix} \mathrm{Re}\left(\frac{\partial F}{\partial h}\right) \\ \mathrm{Im}\left(\frac{\partial F}{\partial h}\right) \end{pmatrix}$$

3. $\boldsymbol{F_t}(\boldsymbol{x}, h, t)$, the Jacobian of $\boldsymbol{F}$ with respect to t.

$$\boldsymbol{F_t} = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}$$

4. $\begin{pmatrix} \dot{\boldsymbol{x}}_0 \\ \dot{h}_0 \end{pmatrix}$, the unit tangent vector to the locus, may be found by differentiating equation A.2 with respect to t and putting $t = t_0$.

$$\boldsymbol{G}^0_{\boldsymbol{x}} \dot{\boldsymbol{x}}_0 + \boldsymbol{G}^0_h \dot{h}_0 = \boldsymbol{0}$$

The superscript ^{0} denotes evaluation at t_0. Then

$$\frac{\dot{\boldsymbol{x}}_0}{\dot{h}_0} = -\boldsymbol{G}^{0\,-1}_{\boldsymbol{x}} \boldsymbol{G}^0_h$$

for $G_{\boldsymbol{x}}^0$ nonsingular. Together with

$$\left\| \begin{array}{c} \dot{\boldsymbol{x}}_0 \\ \dot{h}_0 \end{array} \right\| = 1,$$

it follows that

$$\left(\begin{array}{c} \dot{\boldsymbol{x}}_0 \\ \dot{h}_0 \end{array} \right) = \frac{1}{\left\| \left(\begin{array}{c} {G_{\boldsymbol{x}}^0}^{-1} G_h^0 \\ 1 \end{array} \right) \right\|} \left(\begin{array}{c} {G_{\boldsymbol{x}}^0}^{-1} G_h^0 \\ 1 \end{array} \right)$$

5. $\boldsymbol{F_y}(\boldsymbol{x}, h, t)$, the Jacobian of $\boldsymbol{F}$ with respect to $\boldsymbol{y}$.

$$\boldsymbol{F_y} = \left(\begin{array}{c|c} G_{\boldsymbol{x}} & G_h \\ \hline n_{\boldsymbol{x}}^T & n_h \end{array} \right) = \left(\begin{array}{c|c} G_{\boldsymbol{x}} & G_h \\ \hline \dot{\boldsymbol{x}}^T & \dot{h} \end{array} \right)$$

Algorithm for Temporal Locus by Euler-Newton Continuation

To solve

$$\boldsymbol{F}(\boldsymbol{y}, t) = \boldsymbol{0} \quad \boldsymbol{F} : \mathbb{R}^3 \times \mathbb{R} \to \mathbb{R}^3$$

with $\boldsymbol{F_y}$ nonsingular and given $(\boldsymbol{y}_0, t_0)$ satisfying

$$\boldsymbol{F}(\boldsymbol{y}_0, t_0) = \boldsymbol{0}$$

Step 1 Evaluate $\boldsymbol{F_y}(\boldsymbol{y}_0, t_0)$ and $\boldsymbol{F}_t(\boldsymbol{y}_0, t_0)$.

Step 2 Increment t: $t_1 = t_0 + \Delta t$.

Step 3 Predict the next point on the curve $\boldsymbol{y}_1 = \boldsymbol{y}(t_1)$ by Euler's method

$$\boldsymbol{y}_1^0 = \boldsymbol{y}_0 - \Delta t (\boldsymbol{F_y}(\boldsymbol{y}_0, t_0))^{-1} \boldsymbol{F}_t(\boldsymbol{y}_0, t_0).$$

The superscript ^{0} denotes a first approximation on $\boldsymbol{y}_1$.

Step 4 Solve

$$\boldsymbol{F}(\boldsymbol{y}, \Delta t) = \boldsymbol{0}$$

with starting 'guess' $\boldsymbol{y}_1^0$ to find $\boldsymbol{y}_1$ using Newton's method for systems.
For $k \geq 0$ iterate

1. Solve

$$\boldsymbol{d} = -(\boldsymbol{F_y}(\boldsymbol{y}_1^k, t_1))^{-1} \boldsymbol{F}(\boldsymbol{y}_1^k, t_1)$$

for the Newton correction $\boldsymbol{d} \epsilon \mathbb{R}^3$.

2. Update the solution

$$\boldsymbol{y}_1^{k+1} = \boldsymbol{d} + \boldsymbol{y}_1^k$$

That the root has been found to the required tolerance may be checked by examining $\|\boldsymbol{y}_1^k\|_2$.

Step 5 Set

$$(\boldsymbol{y}_0, t_0) = (\boldsymbol{y}_1, t_1)$$

and return to Step 1.

Note on Singularity

It has been assumed that Rank$\left(\boldsymbol{G_x} \mid \boldsymbol{G}_h \right) = 2$ and also that $\boldsymbol{G_x}$ and $\boldsymbol{F_y}$ are nonsingular. If $\boldsymbol{G_x}$ is nonsingular then the rank condition holds and it may be shown that $\boldsymbol{F_y}$ is nonsingular using the 'ABCD' lemma. $\boldsymbol{G_x}$ will be singular when

$$\left(\mathrm{Re}\left(\frac{\partial F}{\partial s}\right)\right)^2 + \left(\mathrm{Im}\left(\frac{\partial F}{\partial s}\right)\right)^2 = 0$$

which implies

$$\mathrm{Re}\left(\frac{\partial F}{\partial s}\right) = \mathrm{Im}\left(\frac{\partial F}{\partial s}\right) = 0 = \frac{\partial F}{\partial s}$$

This occurs at multiple roots. At such points the Jacobian vanishes completely, Rank$\left(\boldsymbol{G_x} \mid \boldsymbol{G}_h \right) = 0$ and attempts to deal with the singularity such as those in the paper [43] fail. In a neighbourhood of a multiple root the implicit function theorem no longer holds for the unique parameterisation of the curve in terms of t. $\boldsymbol{F_y}$ is invertible at all points other than these multiple roots.

The program is capable of solving the system in a very close neighbourhood of the singularity. Multiple roots may be calculated in advance and iteration performed towards them along the locus.

The plots in (σ, ω, h) space show that the points on the curves are approximately equispaced in arc length. (They may not be exactly equal after iteration from the initial Euler prediction which is a distance Δt along from the previous point.)

The MATLAB code for calculating the loci is given on the following pages. The program may be adapted straightforwardly to enable the corresponding Chu loci with varying gain to be plotted.

```
:::::::::::::::
enc.m
:::::::::::::::
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%  Euler Newton continuation method       %%
%%  for temporal root locus                %%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

clear
format long
more on
cput=cputime;

% Input the polynomials A and C and find their derivatives.
fprintf('Temporal root locus: stability with delay h of a linear time-delay')
fprintf('\nsystem with CE     A(s)+C(s)exp(-sh)=0')
fprintf('\nusing Euler-Newton Continuation.')
fprintf('\nNote that all vectors must be entered in square brackets [].\n')

a=input('Type in coeffs of A(s): ');
c=input('Type in coeffs of C(s): ');
while(length(c)>length(a)|(length(c)==length(a)&(c(1)~=0.0))),
        sprintf('Error, must have proper system; degree(A)>degree(C)\n')
        a=input('Type in coeffs of A(s): ');
        c=input('Type in coeffs of C(s): ');
end
aprime=polyder(a);
cprime=polyder(c);

more off

% Set up figures
figure(1)
        clg
        xx=[0,0];
        yy=[-10000,10000];
        % Plot imaginary axis
        line(xx,yy)
        hold on
        xlabel('Sigma')
        ylabel('Omega')
        plus='+(';
        delay=')exp(-sh)=0';
        title(sprintf('%c%c%c%c',poly2str(a,'s'),plus,poly2str(c,'s'),delay))
        axis([-5 2 -0.01 5])    % Default axes
figure(2)
        clg
        plot3(NaN,0,0,'.')      % Ensures 3-D plot
        hold on
        xlabel('Sigma')
        ylabel('Omega')
        zlabel('h')
        title(sprintf('%c%c%c%c',poly2str(a,'s'),plus,poly2str(c,'s'),delay))
        grid on

%^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
% Enter calculation loop
%^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^

% A new section of the locus may be calculated while(toggle=='y')
toggle='y';
while(toggle=='y'),

% Initialization
F=zeros(3,1);           % The system of 3 equations to be solved,
                        % comprising real and imaginary parts of the CE
                        % and the tangent plane equation
F_y=zeros(3,3);         % Jacobian with respect to y
F_t=[0;0;-1];           % Jacobian with respect to t
y=zeros(3,1);           % y=[sigma;omega;h]
yprev=zeros(3,1);       % Previous iterate of y in Newton loop
xhdot=zeros(3,1);       % Unit vector tangent to curve at previous root
t=0;                    % Current value of continuation parameter t
incr=0;                 % Number of increments in t used so far
newtcount=0;            % Number of Newton iterates for a particular
                        % step of the continuation algorithm
tot_newtcount=0;        % Total number of Newton iterates
                        % for entire continuation process

% Enter the starting point on the locus
s=input('Type in starting value of s (complex): ');
y(1)=real(s);
y(2)=imag(s);
y(3)=input('Type in starting value of h: ')
% delt and tend may be negative for reverse iteration
delt=input('Type in size of increments in continuation parameter t: ')
tend=input('Type in end value of t: ')
r=round((tend-t)/delt +0.5);    % Number of calculation points on curve
hist=zeros(r,2);                % Storage vector for locus points
hist(1,1)=y(1)+sqrt(-1)*y(2);
hist(1,2)=y(3);

fprintf('\nMay type CTRL C to stop iteration at any stage\n')
fprintf('\nStrike any key to continue\n')
pause
fprintf('\nCalculation in progress\n')

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Enter continuation loop
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

while (abs(t)<=abs(tend))
        yprev=y;
        fprintf('\nnumber of steps along curve = %g\n',incr)
        fprintf('value of t = %g\n', t)
        fprintf('Newton iterates to calculate this step= %g\n',newtcount)
        fprintf('\n************************************************\n')

        xhdot=xhdot_function(aprime,c,cprime,y);

        % Update increment counter and t
        incr=incr+1;
        t=t+delt;

        %*****************************************
        % Begin Euler prediction step
        %*****************************************

        % Evaluate Jacobian with respect to y (y is previous solution point)
        F_y=F_y_function(aprime,c,cprime,y,xhdot);

        % Euler step
        y=y-delt*((F_y)\(F_t));
        % y is now prediction of new solution point

        %*****************************************
        % End of Euler prediction step
        %*****************************************
```

```
        fprintf('\nIteration continuing: Please wait\n')

        % Newton's method to solve F(y,t)=0
        % with Euler prediction as starting guess, and t fixed.

        d=ones(3,1);    % Newton correction vector
        newtcount=0;    % Initialise newtcount

        %++++++++++++++++++++++++++++++++++++++++++
        % Enter Newton loop
        %++++++++++++++++++++++++++++++++++++++++++

        while(norm(d)>10*eps)   % Iterate until norm(d)<=10*eps

                newtcount=newtcount+1;

                % Newton correction and update
                % Assemble Jacobian with respect to y (t now fixed)
                F_y=F_y_function(aprime,c,cprime,y,xhdot);
                F=F_function(a,c,y,yprev,delt,xhdot);

                % Newton step (written to y temporarily)
                d=-(F_y)\F;
                y=y+d;

        end

        %++++++++++++++++++++++++++++++++++++++++++
        % End of Newton loop
        %++++++++++++++++++++++++++++++++++++++++++

        tot_newtcount=tot_newtcount+newtcount;  % Update tot_newtcount
        hist(incr+1,1)=y(1)+sqrt(-1)*y(2);      % Save solution point to hist
        hist(incr+1,2)=y(3);

end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% End continuation loop
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% Print out final information

fprintf('\nnumber of steps along curve = %g\n',incr)
fprintf('value of t = %3.14f\n', t)
fprintf('value of s = %3.14f+%3.14fi\n', y(1),y(2))
fprintf('value of h = %3.14f\n', y(3))
fprintf('\n\nTotal Newton iterates =  %g\n',tot_newtcount)
fprintf('\n\nTotal CPU time taken = %g\n',cputime-cput)

%$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$
% Plotting the solution
%$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$

figure(1)

        % choose plot font
        style=input('type in font style for plot, eg -  ','s');
        sigma=real(hist(:,1));
        omega=imag(hist(:,1));
        h=hist(:,2);
        plot(sigma,omega,style)
        pr=input('Print required? (y): ','s');
        if pr=='y',
                print
        end
        tog=input('Change axis? (y): ','s');
        while(tog=='y'),
                ax=input('Type in axis coords: ');
                axis([ax])
                pr=input('Print required? (y): ','s');
                if pr=='y',
                        print
                end
                tog=input('Change axis? (y): ','s');
        end

figure(2)

        plot3(sigma,omega,h,style)
        % Plot projection of curve onto (sigma,omega) plane.
        plot(sigma,omega,style)
        pr=input('Print required? (y): ','s');
        if pr=='y',
                print
        end
        tog=input('Change axis(y) or viewing angle(v)? ','s');
        while(tog=='y'|tog=='v')
                if tog=='y',
                        ax=input('Type in axis coords: ');
                        axis([ax])
                        pr=input('Print required? (y): ','s');
                        if pr=='y',
                                print
                        end
                        tog=input('Change axis(y) or viewing angle(v)? ','s');
                elseif tog=='v',
                        vi=input('Type in viewing coords [az,el]: ');
                        view([vi])
                        pr=input('Print required? (y): ','s');
                        if pr=='y',
                                print
                        end
                        tog=input('Change axis(y) or viewing angle(v)? ','s');
                end
        end

%$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$
% End plotting
%$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$

toggle=input('Continue locus? (y): ','s');
end

%^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
% End calculation loop
%^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^

figure(1); hold off
figure(2); hold off
more on

end
```

```
::::::::::::::
F_function.m
::::::::::::::
function z=F_function(a,c,y,yprev,delt,xhdot)

% System of 3 equations to be solved comprising the real and imaginary
% parts of the CE and the tangent plane equation.

z=zeros(3,1);
s=y(1)+sqrt(-1)*y(2);
fsh=polyval(a,s)+polyval(c,s)*exp(-s*y(3));
z(1)=real(fsh);
z(2)=imag(fsh);
z(3)=xhdot'*(y-yprev)-delt;

end
```

```
::::::::::::::
F_y_function.m
::::::::::::::
function F_y=F_y_function(aprime,c,cprime,y,xhdot)

% Finds the Jacobian of F with respect to y

G_x=zeros(2,2);
G_h=zeros(2,1);

[G_x,G_h]=G_xh_function(aprime,c,cprime,y);

F_y(1:2,1:2)=G_x;
F_y(1:2,3)=G_h;
F_y(3,1:3)=xhdot';

end
```

```
::::::::::::::
G_xh_function.m
::::::::::::::
function [G_x,G_h]=G_xh_function(aprime,c,cprime,y)

% Finds the Jacobian of G with respect to x and h

s=y(1)+sqrt(-1)*y(2);

% The differential of the CE with respect to s:
dfshds=polyval(aprime,s)+ ...
        exp(-s*y(3))*(polyval(cprime,s)-y(3)*polyval(c,s));

% The differential of the CE with respect to h:
dfshdh=-s*polyval(c,s)*exp(-s*y(3));

% Evaluate Jacobian of G using Cauchy-Riemann equations
G_x(1,1)=real(dfshds);
G_x(1,2)=-imag(dfshds);
G_x(2,1)=-G_x(1,2);
G_x(2,2)=G_x(1,1);
G_h(1,1)=real(dfshdh);
G_h(2,1)=imag(dfshdh);

end
```

```
::::::::::::::
xhdot_function.m
::::::::::::::
function z=xhdot_function(aprime,c,cprime,y)

% Calculates the vector [xdot',hdot]'

z=zeros(3,1);
G_x=zeros(2,2);
G_h=zeros(2,1);

[G_x,G_h]=G_xh_function(aprime,c,cprime,y);
z(1:2)=-(G_x)\(G_h);
z(3)=1;
z=z/norm(z);

end
```

A.2 Temporal and Chu Loci via Padé Approximant

The following page gives a MATLAB program for calculating starting and end points of loci and the W-polynomial and its roots. The temporal loci are also plotted using a Padé approximant; this is based on the MATLAB functions `pade` (capable of generating diagonal approximants of high order) and `series` for the transfer function of a series connection of plant and delay. Care must be exercised in the choice of order of the approximant: too low and the approximant matches the exponential function to only a few terms; too high and the calculations are subject to significant rounding errors.

As mentioned before, using the Padé approximant gives a fast initial plot of the loci and the general behaviour is evident. For low h the result will be very close to the true solution. The method gives approximate starting values for iteration by Euler-Newton continuation, particularly useful in the case of stability independent of delay when there are no imaginary axis crossing points.

It is a simple matter to adapt the program to plot Chu loci for fixed delay as the gain K varies. The Padé approximant and series connection need only be calculated once and then the roots of the characteristic equation with gain K plotted. h is fixed, and if small, the plot will be relatively accurate over the entire range of gain, whereas with the temporal loci the loci become increasingly inaccurate with increasing delay.

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%  Direct Method and Temporal Root Locus                    %%
%%  via Pade Approximant                                     %%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

clear
format long
more on

% Input the polynomials A and C
fprintf('Temporal root locus: stability with delay h of a linear time-delay')
fprintf('\nsystem with CE  A(s)+C(s)exp(-sh)=0, using Pade Approximant.')
fprintf('\nNote that all vectors must be entered in square brackets [].\n')

a=input('Type in coeffs of A(s): ');
c=input('Type in coeffs of C(s): ');
n=input('Type in order of Pade approximant: ');
while(length(c)>length(a)|(length(c)==length(a)&(c(1)~=0.0))),
        fprintf('Error, must have proper system; degree(A)>degree(C)\n')
        a=input('Type in coeffs of A(s): ');
        c=input('Type in coeffs of C(s): ');
end

% Give A, C, A+C and their roots.
% RHP roots of A and LHP roots of C are solutions as h-> infinity
% Roots of A+C are solutions for h=0
a
fprintf('Roots of A: ')
roots(a)
c
fprintf('Roots of C: ')
roots(c)
fprintf('A+C: ')
polyadd(a,c)    % Function to add polynomials of unequal length
fprintf('Roots of A+C: ')
roots(polyadd(a,c))

% Calculate the coefficients of A and C when s=i.omega
aiomega=a;      % complex polynomials in omega
ciomega=c;
for k=0:length(a)-1
        aiomega(length(a)-k)=a(length(a)-k)*(sqrt(-1)^k);
end
for k=0:length(c)-1
        ciomega(length(c)-k)=a(length(c)-k)*(sqrt(-1)^k);
end

% Calculate the W-polynomial and its roots.
% Real roots correspond to critical values of omega.
% W(omega^2) = A Abar - C Cbar          at s=i.omega
W=polyadd(conv(aiomega,conj(aiomega)),-conv(ciomega,conj(ciomega)))
fprintf('Roots of W: ')
roots(W)

more off
% Set up figure
figure(1)
        clg
        line([0,0],[-10000,10000])      % Plot imaginary axis
        hold on
        xlabel('Sigma')
        ylabel('Omega')
        plus='+(';
        delay=')exp(-sh)=0, Pade ';
        title(sprintf('%c%c%c%c%c',poly2str(a,'s'),plus,poly2str(c,'s'),
                delay,num2str(n)))
        axis([-4,2,-4,4])       % Default axes

%^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
% Enter calculation loop
%^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^

% A new section of the locus may be calculated while(toggle=='y')
toggle='y';
while(toggle=='y'),

        % Input range of delay
        hstart=input('Type in starting value of h: ');
        hend=input('Type in end value of delay: ');
        hinc=input('Type in increments of delay: ');

        % Compute and plot locus for this h range

        for h=hstart:hinc:hend,

                % Compute nth order diagonal Pade approximant for delay h
                [num,den]=pade(h,n);

                % Transfer function representing series connection of
                % delay and plant C(s)/A(s):
                [num1,den1]=series(c,a,num,den);

                % The following is necessary to avoid plotting the vector
                % against its argument, should it be purely real:
                sol=roots(num1+den1);
                plot(real(sol),imag(sol),'.')

                h

        end

        pr=input('Print required? (y): ','s');
        if pr=='y',
                print
        end
        tog=input('Change axis? (y): ','s');
        while(tog=='y'),
                ax=input('Type in axis coords: ');
                axis([ax])
                pr=input('Print required? (y): ','s');
                if pr=='y',
                        print
                end
                tog=input('Change axis? (y): ','s');
        end
        toggle=input('Continue locus? (y): ','s');

end

%^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
% End calculation loop
%^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^

hold off
more on

end
```

Appendix B

Finite Horizon Cost Functionals

Application of Temporal Locus Calculations to Finite Horizon Cost Functionals – Tail Approximation

There are various criteria for comparing the 'cost' of different controllers, the Integral of Squared Error (ISE) criterion of performance being the most widely implemented of these for linear systems.
ISE:

$$\min\left\{J = \int_0^\infty e^2(t)\,dt\right\}$$

Here the error $e(t)$ is usually defined to be the difference in output from the desired response. The references [35] and [70] give an extensive treatment of the evaluation of cost functionals for linear continuous delayed and all-pass systems and sampled-data systems. This is based on Parseval's theorem of the Laplace transform and complex variable methods. (Particular reference is made to Smith-type predictors where it is shown that improvement to the cost may be made by mismatch of the delays.)

This section is concerned with the evaluation of the cost over an interval in time of finite duration, that is

$$J_T = \int_0^T e^2(t)\,dt$$

In general, an analytic expression for J_T will not be available and work by Dr. K. Walton and Dr. J. E. Marshall has been based on two alternative approximation schemes.

1.

$$\int_0^T e^2(t)\,dt = \int_0^\infty e^2(t)\,dt - \int_T^\infty e^2(t)\,dt$$

It is assumed that J_∞ is known; the second integral on the right-hand side of the above equation is approximated. (This is preferable to approximating $\int_0^T e^2(t)\,dt$

as the error tends exponentially to zero for a stable system.)

2.

$$\int_0^T e^2(t)\,dt = \int_0^\infty [H(t) - H(t-T)]e^2(t)\,dt$$

$[H(t) - H(t-T)]$ is approximated in such a way that the integral of its product with $e^2(t)$ is known.

The temporal locus finds application in approximating the integral in scheme 1. The greatest contribution to the cost for small delay will be made by the primary locus. As critical stability is reached, the cost becomes infinite; the primary locus still dominates even though remote loci are approaching and making a greater contribution.

Examples

1. An integrator in series with delay was considered in section 3.1.9. Here

$$E(s) = \frac{1}{s + \exp(-sh)}$$

Then by the inversion theorem of the Laplace transform

$$\begin{aligned} e(t) &= \sum_{\text{LHP}} \text{res} \frac{\exp(st)}{s + \exp(-sh)} \\ &= \sum_n \frac{\exp(s_n t)}{1 - h\exp(-s_n h)} = \sum_n \frac{\exp(s_n t)}{1 + hs_n} \end{aligned}$$

where $s_n + \exp(-s_n h) = 0$ and s_n distinct. (The case of the double root follows.) The temporal locus for $s + \exp(-sh)$ was the first example considered in section 3.1.10. The primary locus consists of two negative real roots for $0 < h < \exp(-1)$, a pair of real roots at $s = -\exp(1)$, $h = \exp(-1)$, and a complex conjugate pair with negative real part for $h > \exp(-1)$. Critical stability is reached at $h = \pi/2$. The primary locus roots are denoted s_1 and s_2. Then approximating,

$$e(t) \approx \frac{\exp(s_1 t)}{1 + hs_1} + \frac{\exp(s_2 t)}{1 + hs_2}$$

So

$$\begin{aligned} \int_T^\infty e^2(t)\,dt &\approx \int_T^\infty \left(\frac{\exp(s_1 t)}{1 + hs_1} + \frac{\exp(s_2 t)}{1 + hs_2}\right)^2 dt \\ &= -\frac{\exp(2s_1 T)}{2s_1(1 + hs_1)^2} - \frac{2\exp((s_1 + s_2)T)}{(s_1 + s_2)(1 + hs_1)(1 + hs_2)} - \frac{\exp(2s_2 T)}{2s_2(1 + hs_2)^2} \end{aligned}$$

[35] gives

$$J_\infty = \frac{\cos(h)}{2(1-\sin(h))}$$

So

$$\begin{aligned} J_T &\approx \frac{\cos(h)}{2(1-\sin(h))} + \frac{\exp(2s_1T)}{2s_1(1+hs_1)^2} + \frac{\exp(2s_2T)}{2s_2(1+hs_2)^2} \\ &\quad + \frac{2\exp((s_1+s_2)T)}{(s_1+s_2)(1+hs_1)(1+hs_2)} \end{aligned} \tag{B.1}$$

For the real double root,

$$\begin{aligned} e(t) &\approx \lim_{s\to -\exp(1)} \frac{d}{ds}\left(\frac{(s+\exp(1))^2\exp(st)}{s+\exp(-sh)}\right) \\ &= \frac{2(3\exp(1)t+1)}{3\exp(\exp(1)t)} = -2\left(st-\frac{1}{3}\right)\exp(st) \end{aligned}$$

and

$$\int_T^\infty e^2(t)\,dt \approx -\frac{2}{s}\exp(2sT)\left(\left(sT-\frac{1}{3}\right)^2 - \left(sT-\frac{1}{3}\right)+\frac{1}{2}\right)$$

Finally

$$J_T \approx \frac{\cos(h)}{2(1-\sin(h))} + \frac{2}{s}\exp(2sT)\left(\left(sT-\frac{1}{3}\right)^2 - \left(sT-\frac{1}{3}\right)+\frac{1}{2}\right) \tag{B.2}$$

for $h = \exp(-1)$, $s = -\exp(1)$. The accuracy of the approximation may be examined by letting $T = mh$, $m\epsilon\mathbb{N}$ and comparing with the result with the analytical solution via the method of steps (in general, such a solution is not available). Clearly, the larger the interval T the better the result as the tail is smaller. The expressions for $e(t)$ found in section 3.1.9 may be used to obtain

$$\begin{aligned} J_h &= h \\ J_{2h} &= 2h - h^2 + \frac{h^3}{3} \\ J_{3h} &= 3h - 4h^2 + 3h^3 - \frac{7h^4}{12} + \frac{h^5}{20} \end{aligned}$$

and these are compared with result B.1 for various h (or result B.2 in the case of the double root). The temporal locus method is used to find pairs s_1 and s_2 and corresponding h on the primary locus. The results may of course be improved by including other loci in the calculation.

Figure B-1 gives graphs of the cost for $0 \le h < \pi/2$ for the cases $T = h$, $3h$,

together with plots of $(\tilde{J}_T - J_T)/J_T$, the relative error in the approximation. ($\tilde{J}_T$ denotes the approximation to J_T.) It may be seen that for $T = 3h$ the true and approximate solutions are indistinguishable over this range of h.

2.

$$E(s) = \frac{s + 1 - \exp(-sh)}{s(s + 1 + \exp(-sh))}$$

This system exhibits stability independent of delay. In the previous example the cost functional was considered over the range $0 \leq h < \pi/2$ ('fairly low h'). At $h = \pi/2$ the system becomes unstable and the cost infinite. Here if high values of h are considered, it is expected that approximations to J_T using the primary locus only will be inaccurate since successive remote loci are approaching the origin and their contribution is no longer negligible.

The primary locus of

$$s + 1 + \exp(-sh) = 0$$

consists of a real root from $s = -2$ meeting a locus from $-\infty$ at $s = -4.5911$, $h = 0.2785$, bifurcating to a pair of complex conjugate roots and tending to the origin as $h \to \infty$. Remote loci similarly tend to the origin as $h \to \infty$. By standard methods it may be shown that

$$J_\infty = \int_0^\infty e^2(t)\, dt = \frac{1}{4} + \frac{3h}{4} + \frac{h^2}{4} + \frac{h^3}{12}$$

and by the method of steps that

$$\begin{aligned}
J_h &= h \\
J_{2h} &= -2 + 2h + 4\exp(-h) - 2\exp(-2h) \\
J_{3h} &= -5 + 3h + (10 + 4h)\exp(-h) - (9 + 6h + 2h^2)\exp(-2h) \\
&\quad +(6 + 4h)\exp(-3h) - 2\exp(-4h)
\end{aligned}$$

By similar analysis to the first example, $\tilde{J}_T$ is found by considering primary locus roots only. From the inversion theorem of the Laplace transform

$$\begin{aligned}
e(t) &= \sum_{\text{LHP}} \text{res} \frac{(s + 1 - \exp(-sh))\exp(st)}{s(s + 1 + \exp(-sh))} \\
&= \sum_n \frac{(s_n + 1 - \exp(-s_n h))\exp(s_n t)}{s_n(1 - h\exp(-s_n h))} = \sum_n \frac{2(s_n + 1)\exp(s_n t)}{s_n(1 + h(s_n + 1))}
\end{aligned}$$

where $s_n + 1 + \exp(-s_n h) = 0$. (Note that $s = 0$ is a root of the numerator of

$E(s)$, so that this root is not included in the residue calculations.) Restricting attention to the primary roots and evaluating the integral $\int_T^\infty e^2(t)\,dt$ yields

$$\begin{aligned} J_T \approx\ & \frac{1}{4}+\frac{3h}{4}+\frac{h^2}{4}+\frac{h^3}{12} \\ & +\frac{2(s_1+1)^2\exp(2s_1T)}{{s_1}^3(1+h(s_1+1))^2}+\frac{2(s_2+1)^2\exp(2s_2T)}{{s_2}^3(1+h(s_2+1))^2} \\ & +\frac{8(s_1+1)(s_2+1)\exp((s_1+s_2)T)}{(s_1+s_2)s_1s_2(1+h(s_1+1))(1+h(s_2+1))} \end{aligned}$$

Figure B-2 shows cost and relative error plots for $T = h, 3h$. As the delay increases the accuracy of the approximation deteriorates.

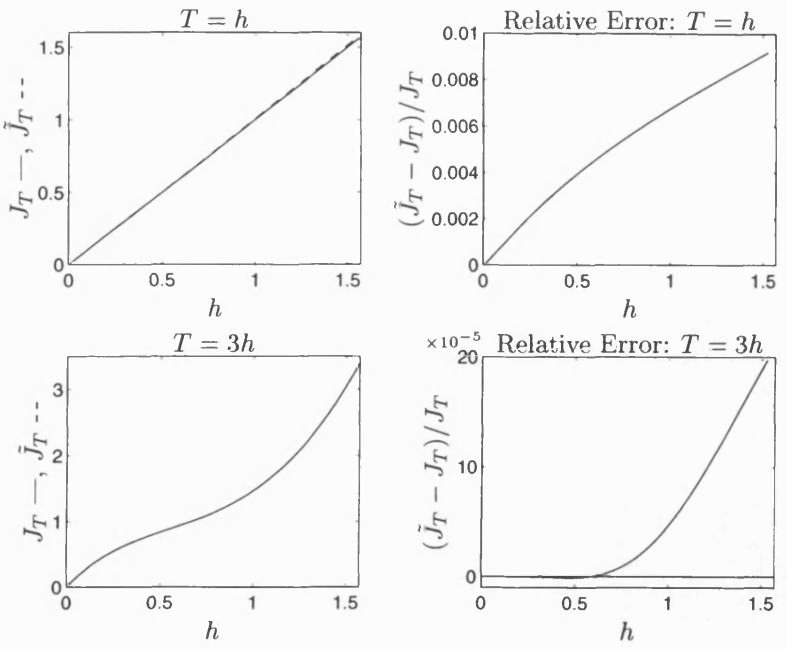


Figure B-1: Finite Horizon Tail Approximation for $E(s) = \frac{1}{s+\exp(-sh)}$

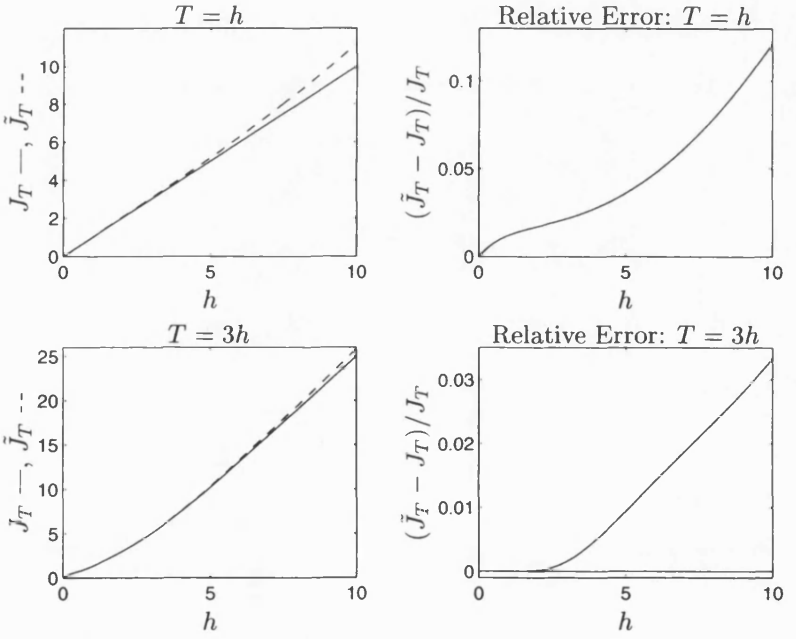


Figure B-2: Finite Horizon Tail Approximation for $E(s) = \frac{s+1-\exp(-sh)}{s(s+1+\exp(-sh))}$

Appendix C

Delayed Van der Pol Systems

Algorithms and Programs for Numerical Solution

This appendix describes the numerical solution employed for equations of form 4.1, 4.2 and 4.3, and which generated the results and figures of chapter 4. The program written may easily be adapted for other delayed equations of second or higher order. The MATLAB programming environment has been used for both the calculation and plotting routines. The algorithm is described in detail for the extended Van der Pol system with delay in the velocity term 4.3; adaptation for the delayed systems 4.1 and 4.2 is straightforward. Complete descriptions of the numerical analysis is not given here but may be found in standard texts such as [6].

Writing the differential equation in system form,

$$\begin{pmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{pmatrix} = \begin{pmatrix} x_2(t) \\ (\mu_1 - \mu_0 x_1^2(t))x_2(t-h) - x_1(t) \end{pmatrix}$$

or

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{x}(t-h))$$

where $x_1 = x$, $x_2 = \dot{x}$ and $\boldsymbol{x}$ is the 2 by 1 vector $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$. Solving at mesh points separated by a stepsize Δt in time, the delay necessitates that the stepsize should divide the delay h exactly so that values of the solution h seconds earlier are available. This means that a numerical method with *fixed* step size must be used. Similarly a Runge-Kutta method, even with fixed stepsize, is unsuitable, involving as it does evaluation of the function $\boldsymbol{f}$ at points in between the steps ($\boldsymbol{x}(t-h)$ is not available at such points). Hence an Adams fourth order predictor-corrector method has been employed. The stepsize is fixed and $\boldsymbol{f}$ need only be evaluated at the actual step points.

Note that the delay means that solution points h seconds previously and the four preceding ones to that are needed. The entire solution is stored so that these results may be accessed and that the solution may subsequently be plotted. Initially, if the stepsize Δt is chosen such that $4\Delta t < h$ then four solution points may be obtained before the delay takes effect in the equation; these may be found by a Runge-Kutta method for example. (The equation 4.1 does admit an analytical solution for $0 \leq t < h$, but for the program to be valid for more general systems this is not exploited and a Runge-Kutta method implemented to obtain the first solution points.) Note that the 'usual' predictor-corrector step will not apply until the delay has fully taken effect in the equation, that is when $t = h + 4\Delta t$.

The stored function in the delay has been taken to be zero so that $x(t-h) = 0$ for $0 \leq t < h$. Clearly this has consequences for the time taken for the solution to approach steady state behaviour, and may cause the solution to become unbounded before the delay has fully taken effect. An improvement may be made either using the predicted results from the extension of Kryloff and Bogoliuboff's method for x and $\dot{x}$ should only the steady-state and not transient behaviour be of interest, or by running the program once with zero stored function in the delay and using this solution for starting values in a subsequent program run. However the latter does not deal with the unboundedness problem.

The complete algorithm is as follows.

Algorithm for Solution of Modified Van der Pol Equations with Delay

To solve

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{x}(t-h))$$

for $0 = t_0 \leq t \leq t_e$, given h, $\boldsymbol{x}_0 = \boldsymbol{x}(0)$ and n, the total number of steps in each time interval of duration h.

Step 1:
Set $m = nt_e/h$, the total number of steps, and $\Delta t = h/n$, the stepsize.
(If $h = 0$, $n = m$ becomes the total number of steps and $\Delta t = t_e/m$.)

Step 2: Runge-Kutta Method for Starting Values
For $i = 1, 2, 3$, iterate:

$$\begin{aligned} \boldsymbol{K}_1 &= \Delta t \boldsymbol{f}(\boldsymbol{x}_{i-1}, \boldsymbol{0}) \\ \boldsymbol{K}_2 &= \Delta t \boldsymbol{f}(\boldsymbol{x}_{i-1} + \tfrac{1}{2}\boldsymbol{K}_1, \boldsymbol{0}) \\ \boldsymbol{K}_3 &= \Delta t \boldsymbol{f}(\boldsymbol{x}_{i-1} + \tfrac{1}{2}\boldsymbol{K}_2, \boldsymbol{0}) \end{aligned}$$

$$
\begin{aligned}
\boldsymbol{K}_4 &= \Delta t \boldsymbol{f}(\boldsymbol{x}_{i-1} + \boldsymbol{K}_3, \boldsymbol{0}) \\
\boldsymbol{x}_i &= \boldsymbol{x}_{i-1} + \tfrac{1}{6}(\boldsymbol{K}_1 + 2\boldsymbol{K}_2 + 2\boldsymbol{K}_3 + \boldsymbol{K}_4) \\
t_i &= i\Delta t
\end{aligned}
$$

Step 3: Predictor-Corrector Method

For $i = 4, \ldots, m$, iterate:

$$t_i = i\Delta t$$

Predict $\boldsymbol{x}_i$

$$
\begin{aligned}
\boldsymbol{x}_i = \boldsymbol{x}_{i-1} + \frac{\Delta t}{24}(55\boldsymbol{f}(\boldsymbol{x}_{i-1}, \boldsymbol{x}_{i-1-n}) - 59\boldsymbol{f}(\boldsymbol{x}_{i-2}, \boldsymbol{x}_{i-2-n}) \\
+37\boldsymbol{f}(\boldsymbol{x}_{i-3}, \boldsymbol{x}_{i-3-n}) - 9\boldsymbol{f}(\boldsymbol{x}_{i-4}, \boldsymbol{x}_{i-4-n}))
\end{aligned}
$$

Correct $\boldsymbol{x}_i$

$$
\begin{aligned}
\boldsymbol{x}_i = \boldsymbol{x}_{i-1} + \frac{\Delta t}{24}(9\boldsymbol{f}(\boldsymbol{x}_i, \boldsymbol{x}_{i-n}) + 19\boldsymbol{f}(\boldsymbol{x}_{i-1}, \boldsymbol{x}_{i-1-n}) \\
-5\boldsymbol{f}(\boldsymbol{x}_{i-2}, \boldsymbol{x}_{i-2-n}) + \boldsymbol{f}(\boldsymbol{x}_{i-3}, \boldsymbol{x}_{i-3-n}))
\end{aligned}
$$

where in each case if $i - j - n < 0$ then $\boldsymbol{x}_{i-j-n} = \boldsymbol{0}$ for $0 \leq j \leq 4$.

The program listing is given on the following pages. (Note that the index i has been increased by 1 since arrays are numbered starting from index 1 which corresponds to the starting values at $t = 0$.) The program `w_function.m` solves the implicit equation for ω arising from the simultaneous equations of Kryloff and Bogoliuboff's method or vector approach given a starting 'guess' for ω by Newton's method. The starting point is readily obtainable from a plot of this implicit equation against ω. The solution is not necessarily unique. From this value of the frequency the approximate amplitude may be calculated for comparison with the numerical solution. `f_function.m` evaluates the function $\boldsymbol{f}$ at given values $\boldsymbol{x}(t)$ and $\boldsymbol{x}(t-h)$.

An alternative method of solution is via SIMULINK, a program for simulation dynamical systems (see [38]). This extension of MATLAB enables block diagram representation of system models and then simulation of the behaviour via numerical methods. Various integration routines are available. The results obtained from both methods were found to be consistent.

```
::::::::::::::
MDVDP.m
::::::::::::::
% Program for numerical solution of the delayed Van der Pol oscillator
% (delay in the velocity term):
%
%       xddot(t)-(mu1-mu0*x^2(t))*xdot(t-h)+x(t)=0
%
% by Predictor-Corrector method, with Runge-Kutta to start solution.
% In addition, the approximate values of amplitude a and frequency w
% (and hence period T) are found by solving the two simultaneous equations
% (arising from Kryloff and Bogoliuboff or vector method) using Newton's
% method. Iteration of the solution may then be performed for an integer
% number of periods.

clear
format long
more on

% Input the parameters required for solution: data file contains
% h,mu0,mu1,n,tend,revs,x0
% n is number of increments for each h, or total number of steps if h==0
% tend= time end; revs=number of periods; one of these must be zero.
% x0=[x(0),xdot(0)] starting point

dat=input('Type in input data file: ','s');
eval(['load ' dat])
res=input('Type in file to which t,x,a,w to be saved: ','s');
h,mu0,mu1,n,tend,revs,x0

% Kryloff and Bogoliuboff Method
% Newton's method for the simultaneous solution of a and phi equations.
% Given h and mu1 may find omega. From this may calculate a(w).
[w,T,a]=w_function(h,mu0,mu1);
a,w,T

% Predictor-Corrector Method for solution calculation

fprintf('\nPress any key to continue\n');
pause
more off
if tend==0
        tend=revs*T;
end
if h==0
        m=n;
        n=0;
        delt=tend/m;
        k1=zeros(1,2);
        k2=zeros(1,2);
        k3=zeros(1,2);
        k4=zeros(1,2);
else
        m=round(n*tend/h +.5);
%       m is the total number of steps (hence m+1 plots since have starting
%       value also).
        delt=h/n;
end

x=zeros(m+1,2);
f=zeros(1,2);
t=zeros(m+1,1);
x(1,:)=x0;
```

```
if h==0
% First four iterations obtained by Runge-Kutta
        for i=2:4,
                t(i)=(i-1)*delt;
                t(i)
                k1=delt*f_function(x(i-1,:),x(i-1,:),mu0,mu1);
                k2=delt*f_function(x(i-1,:)+.5*k1(1,:), ...
                        x(i-1,:)+.5*k1(1,:),mu0,mu1);
                k3=delt*f_function(x(i-1,:)+.5*k2(1,:), ...
                        x(i-1,:)+.5*k2(1,:),mu0,mu1);
                k4=delt*f_function(x(i-1,:)+k3(1,:), ...
                        x(i-1,:)+k3(1,:),mu0,mu1);
                x(i,:)=x(i-1,:)+(1/6)*(k1+2*k2+2*k3+k4);
        end
else
% Again first four iterations obtained by Runge-Kutta
        for i=2:4,
                t(i)=(i-1)*delt;
                t(i)
                k1=delt*f_function(x(i-1,:),[0,0],mu0,mu1);
                k2=delt*f_function(x(i-1,:)+.5*k1(1,:),[0,0],mu0,mu1);
                k3=delt*f_function(x(i-1,:)+.5*k2(1,:),[0,0],mu0,mu1);
                k4=delt*f_function(x(i-1,:)+k3(1,:),[0,0],mu0,mu1);
                x(i,:)=x(i-1,:)+(1/6)*(k1+2*k2+2*k3+k4);
        end
% No output from the delay until t=h => second argument of f_function = 0
% (assuming zero stored function in delay).
        for i=5:n,
                t(i)=(i-1)*delt;
                t(i)
                x(i,:)=x(i-1,:)+(delt/24)*  ...
                        (55*f_function(x(i-1,:),[0,0],mu0,mu1) ...
                        -59*f_function(x(i-2,:),[0,0],mu0,mu1) ...
                        +37*f_function(x(i-3,:),[0,0],mu0,mu1) ...
                        -9*f_function(x(i-4,:),[0,0],mu0,mu1));
                x(i,:)=x(i-1,:)+(delt/24)* ...
                        (9*f_function(x(i,:),[0,0],mu0,mu1) ...
                        +19*f_function(x(i-1,:),[0,0],mu0,mu1) ...
                        -5*f_function(x(i-2,:),[0,0],mu0,mu1) ...
                        +f_function(x(i-3,:),[0,0],mu0,mu1));
        end
% In the next 4 steps along the solution curve the delay term is starting to
% take effect - each successive one has the delayed term becoming non-zero.
        i=n+1;
                t(i)=(i-1)*delt;
                t(i)
                x(i,:)=x(i-1,:)+(delt/24)*  ...
                        (55*f_function(x(i-1,:),[0,0],mu0,mu1) ...
                        -59*f_function(x(i-2,:),[0,0],mu0,mu1) ...
                        +37*f_function(x(i-3,:),[0,0],mu0,mu1) ...
                        -9*f_function(x(i-4,:),[0,0],mu0,mu1));
                x(i,:)=x(i-1,:)+(delt/24)* ...
                        (9*f_function(x(i,:),x(i-n,:),mu0,mu1) ...
                        +19*f_function(x(i-1,:),[0,0],mu0,mu1) ...
                        -5*f_function(x(i-2,:),[0,0],mu0,mu1) ...
                        +f_function(x(i-3,:),[0,0],mu0,mu1));
        i=n+2;
                t(i)=(i-1)*delt;
                t(i)
                x(i,:)=x(i-1,:)+(delt/24)*  ...
                        (55*f_function(x(i-1,:),x(i-1-n,:),mu0,mu1) ...
                        -59*f_function(x(i-2,:),[0,0],mu0,mu1) ...
                        +37*f_function(x(i-3,:),[0,0],mu0,mu1) ...
                        -9*f_function(x(i-4,:),[0,0],mu0,mu1));
```

```
        x(i,:)=x(i-1,:)+(delt/24)* ...
                (9*f_function(x(i,:),x(i-n,:),mu0,mu1) ...
                +19*f_function(x(i-1,:),x(i-1-n,:),mu0,mu1) ...
                -5*f_function(x(i-2,:),[0,0],mu0,mu1) ...
                +f_function(x(i-3,:),[0,0],mu0,mu1));
    i=n+3;
        t(i)=(i-1)*delt;
        t(i)
        x(i,:)=x(i-1,:)+(delt/24)*  ...
                (55*f_function(x(i-1,:),x(i-1-n,:),mu0,mu1) ...
                -59*f_function(x(i-2,:),x(i-2-n,:),mu0,mu1) ...
                +37*f_function(x(i-3,:),[0,0],mu0,mu1) ...
                -9*f_function(x(i-4,:),[0,0],mu0,mu1));
        x(i,:)=x(i-1,:)+(delt/24)* ...
                (9*f_function(x(i,:),x(i-n,:),mu0,mu1) ...
                +19*f_function(x(i-1,:),x(i-1-n,:),mu0,mu1) ...
                -5*f_function(x(i-2,:),x(i-2-n,:),mu0,mu1) ...
                +f_function(x(i-3,:),[0,0],mu0,mu1));
    i=n+4;
        t(i)=(i-1)*delt;
        t(i)
        x(i,:)=x(i-1,:)+(delt/24)*  ...
                (55*f_function(x(i-1,:),x(i-1-n,:),mu0,mu1) ...
                -59*f_function(x(i-2,:),x(i-2-n,:),mu0,mu1) ...
                +37*f_function(x(i-3,:),x(i-3-n,:),mu0,mu1) ...
                -9*f_function(x(i-4,:),[0,0],mu0,mu1));
        x(i,:)=x(i-1,:)+(delt/24)* ...
                (9*f_function(x(i,:),x(i-n,:),mu0,mu1) ...
                +19*f_function(x(i-1,:),x(i-1-n,:),mu0,mu1) ...
                -5*f_function(x(i-2,:),x(i-2-n,:),mu0,mu1) ...
                +f_function(x(i-3,:),x(i-3-n,:),mu0,mu1));
end
% Usual Predictor-Corrector step now that x(i-4-n,:) is available (i>n+4).
for i=n+5:m+1,
    t(i)=(i-1)*delt;
    t(i)
    x(i,:)=x(i-1,:)+(delt/24)*  ...
            (55*f_function(x(i-1,:),x(i-1-n,:),mu0,mu1) ...
            -59*f_function(x(i-2,:),x(i-2-n,:),mu0,mu1) ...
            +37*f_function(x(i-3,:),x(i-3-n,:),mu0,mu1) ...
            -9*f_function(x(i-4,:),x(i-4-n,:),mu0,mu1));
    x(i,:)=x(i-1,:)+(delt/24)* ...
            (9*f_function(x(i,:),x(i-n,:),mu0,mu1) ...
            +19*f_function(x(i-1,:),x(i-1-n,:),mu0,mu1) ...
            -5*f_function(x(i-2,:),x(i-2-n,:),mu0,mu1) ...
            +f_function(x(i-3,:),x(i-3-n,:),mu0,mu1));
end
% Save results to file
eval(['save ' res ' t x a w'])
more on
end
```



```
::::::::::::::
w_function.m
::::::::::::::
function [w,T,a]=w_function(h,mu0,mu1)

% Newton's method for the simultaneous solution of a and phi equations
% for
% x1dot=x2(t)
% x2dot=(mu1-mu0*x1(t)^2)*x2(t-h)-x1(t)
% Given h and mu1 may find frequency omega.
% From this the amplitude a(w) may be calculated.

w=input('type starting value (guess) of omega:   ');
fprintf('\nh=%2.10f, mu0=%2.5f, mu1=%2.5f, wstart=%2.10f\n',h,mu0,mu1,w)

% Enter Newton loop
% Iterate until an accuracy of 10*eps is achieved
d=1;
icount=0;
while norm(d)>10*eps)
        % Update iteration counter
        icount = icount + 1;
        % Set up the value of F and the derivative at the current iterate
        F=w^2-2*w*mu1*sin(w*h)-1;
        JAC=2*w-2*mu1*(sin(w*h)+w*h*cos(w*h));
        % Compute Newton correction
        d=-F/JAC;
        % Print out information on latest iterate
        fprintf('\niterate %g, w=%2.10f,\nNorm of correction %10.10g \n',
                icount,w,abs(d))
        % Update solution
        w=w+d;
end
% End Newton loop

T=2*pi/w;
fprintf('\nThe computed root is \n')
fprintf('\nh=%3.14f, mu0=%2.5f, mu1=%2.5f, w=%3.14f\n\n',h,mu0,mu1,w)
a=sqrt(4*mu1/mu0);
fprintf('a=%3.14f, T=%3.14f\n\n',a,T)

end
```

```
::::::::::::::
f_function.m
::::::::::::::
function z=f_function(y,yn,mu0,mu1)

% x1dot=x2
% x2dot=(mu1-mu0*x1(t)^2)*x2(t-h)-x1(t)

z=zeros(1,2);
z(1,1)=y(1,2);
z(1,2)=(mu1-mu0*(y(1,1)^2))*yn(1,2)-y(1,1);

end
```

Bibliography

[1] D. K. Arrowsmith and C. M. Place. *Ordinary Differential Equations.* Chapman and Hall, 1982.

[2] D. P. Atherton. *Stability of Nonlinear Systems.* Research Studies Press, 1981.

[3] G. A. Baker, Jr. and P. Graves-Morris. *Padé Approximants*, volume 13 and 14 of *Encyclopedia of Mathematics and its Applications: Mathematics of Physics.* Addison Wesley, 1981.

[4] S. P. Banks. *Control Systems Engineering.* Prentice Hall, 1986.

[5] N. F. Britton. Spatial Structures and Periodic Travelling Waves in an Integro-Differential Reaction-Diffusion Population Model. *SIAM Journal on Applied Mathematics*, 50(6):1663–1688, 1990.

[6] R. L. Burden and J. D. Faires. *Numerical Analysis.* Prindle, Weber and Schmidt, 5th edition, 1993.

[7] N. H. Chosky. Time-Lag Systems. In R. H. Macmillan, editor, *Progress in Control Engineering Volume I*, pages 19–38. Heywood, 1962.

[8] Y. Chu. Feedback Control Systems with Dead-Time Lag or Distributed Lag by Root Locus Method. *Transactions of the AIEE, Part 2, Applications and Industry*, 71:291–296, 1952.

[9] Y. Chu. Synthesis of Feedback Control System by Phase-Angle Loci. *Transactions of the AIEE, Part 2, Applications and Industry*, 71:330–339, 1952.

[10] P. A. Cook. *Nonlinear Dynamical Systems.* Prentice Hall, 2nd edition, 1994.

[11] J. L. Douce. *An Introduction to the Mathematics of Servomechanisms.* English Universities Press, 1963.

[12] W. R. Evans. Control System Synthesis by Root Locus Method. *AIEE Transactions*, 69:66–69, 1950.

[13] W. R. Evans. *Control System Dynamics.* McGraw-Hill, 1954.

[14] G. F. Franklin, J. D. Powell, and A. Emami-Naeini. *Feedback Control of Dynamic Systems.* Addison Wesley, 2nd edition, 1991.

[15] B. Garland. Sensitivity and Stability of Time-Delay Systems. Master's thesis, School of Mathematical Sciences, University of Bath, 1974.

[16] K. Gopalsamy. *Stability and Oscillations in Delay-Differential Equations of Population Dynamics.* Kluwer Academic, 1992.

[17] R. B. Grafton. Periodic Solutions of certain Liénard Equations with Delay. *Journal of Differential Equations*, 11:519–527, 1972.

[18] R. B. Grafton. Periodic Solutions of Liénard Equations with Delay: Some Theoretical and Numerical Results. In K. Schmitt, editor, *Delay and Functional Differential Equations and their Applications*, pages 321–334. Academic Press, 1972.

[19] P. R. Graves-Morris, editor. *Padé Approximants.* The Institute of Physics, 1973. Lectures delivered at a summer school held at the University of Kent, July 1972.

[20] J. Haag. *Oscillatory Motions.* Wadsworth, 1962. Translated by R. M. Rosenberg.

[21] J. K. Hale. *Theory of Functional Differential Equations.* Springer-Verlag, 1977.

[22] J. K. Hale and S. M. Lunel. *Introduction to Functional Differential Equations.* Springer-Verlag, 1993.

[23] I. D. Huntley and R. M. Johnson. *Linear and Nonlinear Differential Equations.* Ellis Horwood, 1983.

[24] A. Hurwitz. On the Conditions under which an Equation has only Roots with Negative Real Parts. In R. Bellman and R. Kalaba, editors, *Selected Papers on Mathematical Trends in Control Theory*, chapter 4, pages 70–82. Dover, 1964. Originally in German: Über die Bedingungen, unter welchen eine Gleichung nur Wurzeln mit negativen reellen Theilen besitzt. *Mathematische Annalen,* 46:273-284, 1895.

[25] P. E. Hydon, J. E. Marshall, and K. Walton. All-Pass Feedback Systems. In N. K. Nichols and D. H. Owens, editors, *The Mathematics of Control Theory*, pages

371–380. Oxford University Press, 1992. Proceedings of the 5th IMA Conference on Control Theory, University of Strathclyde, Glasgow, September 1988.

[26] O. L. R. Jacobs. *Introduction to Control Theory.* Oxford University Press, 1974.

[27] D. W. Jordan and P. Smith. *Nonlinear Ordinary Differential Equations.* Oxford University Press, 2nd edition, 1987.

[28] R. E. Kalman and J. E. Bertram. Control System Analysis and Design via the "Second Method" of Lyapunov. *Transactions of the ASME, Journal of Basic Engineering*, 82:371–400, June 1960.

[29] N. Kryloff and N. Bogoliuboff. *Introduction to Nonlinear Mechanics.* Princeton University Press, 1943.

[30] J. D. Lawrence. *A Catalog of Special Plane Curves.* Dover, 1972.

[31] S. Lefschetz. *Differential Equations: Geometric Theory.* Dover, 2nd edition, 1977.

[32] J. E. Marshall. *Digital Control of Systems with Time-Delay Elements.* PhD thesis, School of Mathematical Sciences, University of Bath, 1971.

[33] J. E. Marshall. *Control of Time-Delay Systems.* Peter Peregrinus, 1979.

[34] J. E. Marshall. Smith Predictors. In M. G. Singh, editor, *Systems and Control Encyclopaedia: Theory, Technology, Applications*, pages 4457–4463. Pergamon Press, 1987.

[35] J. E. Marshall, H. Górecki, K. Walton, and A. Korytowski. *Time-Delay Systems: Stability and Performance Criteria with Applications.* Ellis Horwood, 1992.

[36] J. E. Marshall and K. Walton. Temporal Mismatch: Stability Avenues. *IMA International Conference on Control, UMIST*, 1992. Preprint.

[37] The MathWorks, Inc. *MATLAB Reference Guide*, August 1992.

[38] The MathWorks, Inc. *SIMULINK User's Guide*, December 1993.

[39] A. I. Mees. *Dynamics of Feedback Systems.* John Wiley, 1981.

[40] A. I. Mees and C. T. Sparrow. Chaos. *IEE Proceedings*, 128 Part D(5):201–205, September 1981.

[41] A. V. Mikhailov. Harmonic Analysis in the Theory of Automatic Control. *Automatika i Telemekhanika*, (3), 1938. Originally in Russian: Metod garmonicheskogo analiza v teorii regulirovaniia.

[42] N. Minorsky. *Introduction to Nonlinear Mechanics.* Edwards Brothers, 1947.

[43] G. Moore. The Numerical Treatment of Nontrivial Bifurcation Points. *Numerical Function Analysis and Optimization*, 2(6):441–472, 1980.

[44] I. J. Nagrath and M. Gopal. *Control Systems Engineering.* John Wiley, 2nd edition, 1982.

[45] K. S. Narendra and J. H. Taylor. *Frequency Domain Criteria for Absolute Stability.* Academic Press, 1973.

[46] H. Nyquist. Regeneration Theory. In R. Bellman and R. Kalaba, editors, *Selected Papers on Mathematical Trends in Control Theory*, chapter 5, pages 83–105. Dover, 1964. Originally in *Bell System Technical Journal,* 11:126-147, 1932.

[47] J. M. Ortega and W. C. Rheinboldt. *Iterative Solution of Nonlinear Equations in Several Variables.* Academic Press, 1970.

[48] H. Padé. Mémoire sur les Développements en Fractions Continues de la Fonction Exponentielle. *Annales Scientifiques de l'École Normale Supérieure*, 16:395–426, 1899. In French.

[49] P. C. Parks. A. M. Lyapunov's Stability Theory - 100 Years On. *IMA Jornal of Mathematical Control and Information*, 9(4):275–303, 1992.

[50] P. C. Parks and V. Hahn. *Stability Theory.* Prentice Hall, 1983. Translated from the German 1981.

[51] L. A. Pipes. *Applied Mathematics for Engineers and Physicists.* McGraw-Hill, 2nd edition, 1958.

[52] E. P. Popov. *The Dynamics of Automatic Control Systems.* Pergamon Press, 1962.

[53] C. E. Rohrs, J. L. Melsa, and D. G. Schultz. *Linear Control Systems.* McGraw-Hill, 1993.

[54] E. J. Routh. *Advanced Dynamics of a System of Rigid Bodies.* Dover, 1960. Originally Macmillan, 1905.

[55] J. A. Sanders and F. Verhulst. *Averaging Methods in Nonlinear Dynamical Systems.* Springer-Verlag, 1985.

[56] M. J. Satché. Stability of Linear Oscillating Systems with Constant Time Lag. *Journal of Applied Mechanics*, 16:419–420, December 1949.

[57] S. M. Shinners. *Modern Control System Theory and Design.* John Wiley, 1992.

[58] J. Shohat. On Van der Pol's and Related Nonlinear Differential Equations. *Journal of Applied Physics*, 15:568–574, July 1944.

[59] M. G. Singh, editor. *Systems and Control Encyclopaedia: Theory, Technology, Applications*, volume 5. Pergamon Press, 1987.

[60] O. J. M. Smith. *Feedback Control Systems.* McGraw-Hill, 1958.

[61] O. J. M. Smith. A Controller to Overcome Dead Time. *ISA Journal*, 6(2):28–33, 1959.

[62] G. Stépán. Delay as a Bifurcation Parameter. *International Series of Numerical Mathematics*, 79:298–305, 1986.

[63] G. Stépán. *Retarded Dynamical Systems: Stability and Characteristic Functions.* Longman Scientific and Technical, 1989.

[64] J. M. T. Thompson and H. B. Stewart. *Nonlinear Dynamics and Chaos.* John Wiley, 1986.

[65] M. R. Tomblin. Temporal Root Locus. University of Bath, Final Year Project, 1993.

[66] J. G. Truxal. *Automatic Feedback Control System Synthesis.* McGraw-Hill, 1955.

[67] J. G. Truxal. *Introductory Systems Engineering.* McGraw-Hill, 1972.

[68] B. Van der Pol. Forced Oscillations in a Circuit with Nonlinear Resistance (Reception with Reactive Triode). In R. Bellman and R. Kalaba, editors, *Selected Papers on Mathematical Trends in Control Theory*, chapter 7, pages 124–140. Dover, 1964. Originally in *The London, Edinburgh and Dublin Philosophical Magazine and Journal of Science,* 3:65-80, 1927.

[69] F. Verhulst. *Nonlinear Differential Equations and Dynamical Systems.* Springer-Verlag, 1990. Originally in Dutch 1985.

[70] K. Walton and J. E. Marshall. Closed Form Solution for Time-Delay Systems' Cost Functionals. *International Journal of Control*, 39(5):1063–1071, 1984.

[71] K. Walton and J. E. Marshall. Direct Method for TDS Stability Analysis. *IEE Proceedings*, 134 Part D(2):101–107, March 1987.

[72] J. C. West. The Describing Function, Technique and Application. In R. H. Macmillan, editor, *Nonlinear Control Systems Analysis*, chapter 3 and 4. Pergamon Press, 1962.

[73] C. H. Wilts. *Principles of Feedback Control.* Addison Wesley, 1960.