King's College London

Department of Mathematics

Course 6CCM356A

LINEAR SYSTEMS WITH CONTROL THEORY

Course 6CCM357A

INTRODUCTION TO LINEAR SYSTEMS WITH CONTROL THEORY

David Lavis

Contents

1 Differential Equations **1**

1.1 Basic Ideas . . . 1
1.2 Using MAPLE to Solve Differential Equations . . . 3
1.3 General Solution of Specific Equations . . . 4
1.3.1 First-Order Separable Equations . . . 4
1.3.2 First-Order Homogeneous Equations . . . 5
1.3.3 First-Order Linear Equations . . . 6
1.4 Equations with Constant Coefficients . . . 8
1.4.1 Finding the Complementary Function . . . 9
1.4.2 A Particular Solution . . . 11
1.5 Systems of Differential Equations . . . 15
1.5.1 MAPLE for Systems of Differential Equations . . . 18
1.6 Autonomous Systems . . . 20
1.6.1 One-Variable Autonomous Systems . . . 21
1.6.2 Digression: Matrix Algebra . . . 27
1.6.3 Linear Autonomous Systems . . . 28
1.6.4 Linearizing Non-Linear Systems . . . 31

2 Linear Transform Theory **37**

2.1 Introduction . . . 37
2.2 Some Special Functions . . . 39
2.2.1 The Gamma Function . . . 39
2.2.2 The Heaviside Function . . . 39
2.2.3 The Dirac Delta Function . . . 40
2.3 Laplace Transforms . . . 41
2.3.1 Some Particular Transforms . . . 41
2.3.2 Some General Properties . . . 43
2.3.3 Using Laplace Transforms to Solve Differential Equations . . . 45
2.4 The $\mathcal{Z}$ Transform . . . 51
2.4.1 Some Particular Transforms . . . 52
2.4.2 Some General Properties . . . 53
2.4.3 Using $\mathcal{Z}$ Transforms to Solve Difference Equations . . . 53

3 Transfer Functions and Feedback **59**
3.1 Introduction . . . 59
3.2 Systems in Cascade . . . 61
3.3 Combinations and Distribution of Inputs . . . 64
3.4 Feedback . . . 65

4 Controllability and Observability **71**
4.1 Introduction . . . 71
4.2 The Exponential Matrix . . . 72
4.3 Realizations of Systems . . . 77
4.3.1 The Companion Realization . . . 78
4.3.2 The Diagonal and Jordan Realizations . . . 85
4.4 Controllability . . . 88
4.5 Observability . . . 91
4.6 Minimal Realizations . . . 92

5 Stability **97**
5.1 Two Types of Stability . . . 97
5.2 Stability and the Transfer Function . . . 99
5.2.1 The Routh–Hurwitz Criterion . . . 100
5.3 Stability and Feedback . . . 105
5.3.1 Output Feedback . . . 105
5.3.2 State Feedback . . . 107
5.4 Discrete-Time Systems . . . 109

6 Optimal Control **113**
6.1 Digression: The Calculus of Variations . . . 113
6.1.1 The Euler-Lagrange Equations . . . 114
6.1.2 Hamilton's Principle of Least Action . . . 118
6.1.3 Constrained Problems . . . 119
6.2 The Optimal Control Problem . . . 121
6.2.1 Problems Without State-Space Variables . . . 121
6.2.2 Problems With State-Space Variables . . . 122
6.3 The Hamilton-Pontriagin Method . . . 125
6.3.1 The Connection with Dynamics: Hamilton's Equations . . . 126
6.4 Pontriagin's Principle . . . 127

7 Complex Variable Methods **141**
7.1 Introduction . . . 141
7.2 Results in Complex Variable Theory . . . 141
7.2.1 Cauchy's Residue Theorem . . . 141
7.2.2 The Argument Principle . . . 143
7.3 The Inverse Laplace Transform . . . 144
7.4 The Stability of a System with Unit Feedback . . . 146
7.4.1 The Frequency Response Function . . . 146
7.4.2 The Nyquist Criterion . . . 146

8 Non-Linear Systems **153**
8.1 Introduction . . . 153
8.2 Laplace Transforms and Transfer Functions . . . 153
8.3 Constant Control-Variable Systems . . . 154
8.3.1 The Stability of Trajectories . . . 155
8.3.2 The Lyapunov Direct Method . . . 157
8.4 The Stability of Periodic Solutions . . . 160

9 Solutions **163**
9.1 Problems 1 . . . 163
9.2 Problems 2 . . . 173
9.3 Problems 3 . . . 177
9.4 Problems 4 . . . 180
9.5 Problems 5 . . . 184
9.6 Problems 6 . . . 187
9.7 Problems 7 . . . 196
9.8 Problems 8 . . . 199

Chapter 1

Differential Equations

1.1 Basic Ideas

We start by considering a *first-order* (ordinary) differential equation of the form

$$\frac{\mathrm{d}x}{\mathrm{d}t} = F(x;t) \quad \text{with the } \textit{initial condition} \quad x(\tau) = \xi. \tag{1.1}$$

We say that $x = f(t)$ is a *solution* of (1.1) if it is satisfied identically when we substitute $f(t)$ for x. That is

$$\frac{\mathrm{d}f(t)}{\mathrm{d}t} \equiv F(f(t);t) \quad \text{and} \quad f(\tau) = \xi. \tag{1.2}$$

If we plot $(t, f(t))$ in the Cartesian t–x plane then we generate a *solution curve* of (1.1) which passes through the point (τ, ξ) determined by the initial condition. If the initial condition is varied then a *family* of solution curves is generated. This idea corresponds to thinking of the *general solution* $x = f(t, c)$ with $f(t, c_0) = f(t)$ and $f(\tau, c_0) = \xi$. The family of solutions are obtained by plotting $(t, f(t, c))$ for different values of c.

Throughout this course the variable t can be thought of as *time*. When convenient we shall use the 'dot' notation to signify differentiation with respect to t. Thus

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \dot{x}(t), \qquad \frac{\mathrm{d}^2x}{\mathrm{d}t^2} = \ddot{x}(t),$$

with (1.1) expressed in the form $\dot{x}(t) = F(x;t)$.[1] We shall also sometimes denote the solution of (1.1) simply as $x(t)$ rather than using the different letter $f(t)$.

In practice it is not possible, in most cases, to obtain a complete solution to a differential equation in terms of elementary functions. To see why this is the case consider the simple case of a *separable equation*

$$\frac{\mathrm{d}x}{\mathrm{d}t} = T(t)/X(x) \quad \text{with the } \textit{initial condition} \quad x(\tau) = \xi. \tag{1.3}$$

[1] For derivatives of higher than second order this notation becomes cumbersome and will not be used.

This can be rearranged to give

$$\int_{\tau}^{t} T(u)\mathrm{d}u = \int_{\xi}^{x} X(y)\mathrm{d}y. \tag{1.4}$$

So to complete the solution we must be able to perform both integrals and invert the solution form in x to get an explicit expression for $x = f(t)$. Both these tasks are not necessary possible. Unlike differentiation, integration is a skill rather than a science. If you know the rules and can apply them you can differentiate anything. You can solve an integral only if you can spot that it belongs to a class of easily solvable forms (by substitution, integration by parts etc.). So even a separable equation is not necessarily solvable and the problems increase for more complicated equations. It is, however, important to know whether a given equation possesses a solution and, if so, whether the solution is unique. This information is given by *Picard's Theorem*:

Theorem 1.1.1 *Consider the (square) set*

$$\mathcal{A} = \{(x,t) : |t-\tau| \le \triangle, |x-\xi| \le \triangle\} \tag{1.5}$$

and suppose that $F(x;t)$ and $\partial F/\partial x$ are continuous functions in both x and t on $\mathcal{A}$. Then the differential equation (1.1) has a unique *solution $x = f(t)$, satisfying the initial condition $\xi = f(\tau)$ on the interval $[\tau - \triangle_1, \tau + \triangle_1]$, for some $\triangle_1$ with $0 < \triangle_1 \le \triangle$.*

We shall not give a proof of this theorem, but we shall indicate an approach which could lead to a proof. The differential equation (1.1) (together with the initial condition) is equivalent to the *integral equation*

$$x(t) = \xi + \int_{\tau}^{t} F(x(u);u)\mathrm{d}u. \tag{1.6}$$

Suppose now that we define the sequence of functions $\{x^{(j)}(t)\}$ by

$$\begin{aligned} &x^{(0)}(t) = \xi, \\ &x^{(j+1)}(t) = \xi + \int_{\tau}^{t} F(x^{(j)}(u);u)\mathrm{d}u, \qquad j = 1,2,\ldots. \end{aligned} \tag{1.7}$$

The members of this sequence are known as *Picard iterates*. To prove Picard's theorem we would need to show that, under the stated conditions, the sequence of iterates converges uniformly to a limit function $x(t)$ and then prove that it is a unique solution to the differential equation. Rather than considering this general task we consider a particular case:

Example 1.1.1 Consider the differential equation

$$\dot{x}(t) = x \qquad \text{with the } \textit{initial condition} \qquad x(0) = 1. \tag{1.8}$$

Using (1.7) we can construct the sequence of Picard iterates:

$$
\begin{aligned}
x^{(0)}(t) &= 1, \\
x^{(1)}(t) &= 1 + \int_0^t 1\,\mathrm{d}u = 1 + t, \\
x^{(2)}(t) &= 1 + \int_0^t (1+u)\mathrm{d}u = 1 + t + \tfrac{1}{2!}t^2, \\
x^{(3)}(t) &= 1 + \int_0^t \left(1 + u + \tfrac{1}{2!}u^2\right)\mathrm{d}u = 1 + t + \tfrac{1}{2!}t^2 + \tfrac{1}{3!}t^3, \\
&\vdots \qquad\qquad\qquad\qquad \vdots \\
x^{(j)}(t) &= 1 + t + \tfrac{1}{2!}t^2 + \cdots + \tfrac{1}{j!}t^j.
\end{aligned}
\tag{1.9}
$$

We see that, as $j \to \infty$, $x^{(j)}(t) \to \exp(x)$, which is indeed the unique solution of (1.8).

1.2 Using MAPLE to Solve Differential Equations

The MAPLE command to differentiate a function $f(x)$ with respect to x is

```
diff(f(x),x)
```

The n-th order derivative can be obtained either by repeating the x n times or using a dollar sign. Thus the 4th derivative of $f(x)$ with respect to x is obtained by

`diff(f(x),x,x,x,x)` or `diff(f(x),x$4)`

This same notation can also be used for partial differentiation. Thus for $g(x, y, z)$, the partial derivative $\partial^4 g/\partial x^2 \partial y \partial z$ is obtained by

```
diff(g(x,y,z),x$2,y,z)
```

An n-th order differential equation is of the form

$$
\frac{\mathrm{d}^n x}{\mathrm{d}t^n} = F\left(x, \frac{\mathrm{d}x}{\mathrm{d}t}, \ldots, \frac{\mathrm{d}^{n-1}x}{\mathrm{d}t^{n-1}}; t\right). \tag{1.10}
$$

This would be coded into MAPLE as

```
diff(x(t),t$n)=F(diff(x(t),t),...,diff(x(t),t$(n-1)),t)
```

Thus the MAPLE code for the differential equation

$$
\frac{\mathrm{d}^3 x}{\mathrm{d}t^3} = x, \tag{1.11}
$$

is

```
diff(x(t),t$3)=x(t)
```

The MAPLE command for solving a differential equation is `dsolve` and the MAPLE code which obtains the *general solution* for (1.11) is

```
> dsolve(diff(x(t),t$3)=x(t));
```

$$\mathrm{x}(t) = \_C1\, e^t + \_C2\, e^{(-1/2\,t)} \sin(\frac{1}{2}\sqrt{3}\,t) + \_C3\, e^{(-1/2\,t)} \cos(\frac{1}{2}\sqrt{3}\,t)$$

Since this is a third-order differential equation the general solution contains three arbitrary constants for which MAPLE uses the notation $\_C1$, $\_C2$ and $\_C3$. Values are given to these constants if we impose three initial conditions. In MAPLE this is coded by enclosing the differential equation and the initial conditions in curly brackets and then adding a final record of the quantity required. Thus to obtain a solution to (1.11) with the initial conditions

$$x(0) = 2, \qquad \dot{x}(0) = 3, \qquad \ddot{x}(0) = 7, \tag{1.12}$$

```
> dsolve({diff(x(t),t$3)=x(t),x(0)=2,D(x)(0)=3,
> (D@@2)(x)(0)=7},x(t));
```

$$\mathrm{x}(t) = 4\, e^t - \frac{4}{3}\sqrt{3}\, e^{(-1/2\,t)} \sin(\frac{1}{2}\sqrt{3}\,t) - 2\, e^{(-1/2\,t)} \cos(\frac{1}{2}\sqrt{3}\,t)$$

Note that in the context of the initial condition the code `D(x)` is used for the first derivative of x with respect to t. The corresponding n-th order derivative is denoted by `(D@@n)(x)`.

1.3 General Solution of Specific Equations

An n-th order differential equation like (1.10) is said to be *linear* if F is a linear function of x and its derivatives; (the dependence of F on t need not be linear for the system to be linear). The equation is said to be *autonomous* if F does not depend explicitly on t. (The first-order equation of Example 1.1.1 is both linear and autonomous.) The general solution of an n-th order equation contains n arbitrary constants. These can be given specific values if we have n initial[2] conditions. These may be values for $x(t)$ at n different values of t or they may be values for x and its first $n-1$ derivatives at one value of t. We shall use C, C', $\mathsf{C}_1, \mathsf{C}_2, \ldots$ to denote arbitrary constants.

1.3.1 First-Order Separable Equations

If a differential equation is of the type of (1.3), it is said to be *separable* because it is equivalent to (1.4), where the variables have been *separated* onto opposite

[2]The terms *initial* and *boundary conditions* are both used in this context. Initial conditions have the connotation of being specified at a fixed or initial time and boundary conditions at fixed points in space at the ends or boundaries of the system.

sides of the equation. The solution is now more or less easy to find according to whether it is easy or difficult to perform the integrations.

Example 1.3.1 A simple separable equation is the *exponential growth equation* $\dot{x}(t) = \mu x$. The variable $x(t)$ could be the size of a population (e.g. of rabbits) which grows in proportion to its size. A modified version of this equation is when the rate of growth is zero when $x = \kappa$. Such an equation is

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \frac{\mu}{\kappa} x(\kappa - x). \tag{1.13}$$

This is equivalent to

$$\int \frac{\kappa \mathrm{d}x}{x(\kappa - x)} = \mu \int \mathrm{d}t + \mathsf{C}. \tag{1.14}$$

Using partial fractions

$$\begin{aligned} \mu t &= \int \frac{\mathrm{d}x}{x} + \int \frac{\mathrm{d}x}{\kappa - x} - \mathsf{C} \\ &= \ln|x| - \ln|\kappa - x| - \mathsf{C}. \end{aligned} \tag{1.15}$$

This gives

$$\frac{x}{\kappa - x} = \mathsf{C}' \exp(\mu t), \tag{1.16}$$

which can be solved for x to give

$$x(t) = \frac{\mathsf{C}' \kappa \exp(\mu t)}{1 + \mathsf{C}' \exp(\mu t)}. \tag{1.17}$$

The MAPLE code for solving this equation is

```
> dsolve(diff(x(t),t)=mu*x(t)*(kappa-x(t))/kappa);
```

$$\mathrm{x}(t) = \frac{\kappa}{1 + e^{(-\mu\, t)}\ \_C1\, \kappa}$$

It can be seen that $x(t) \to \kappa$, as $t \to \infty$.

1.3.2 First-Order Homogeneous Equations

A first-order equation of the form

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \frac{P(x,t)}{Q(x,t)}, \tag{1.18}$$

is said to be *homogeneous* if P and Q are functions such that

$$P(\lambda t, t) = t^m P(\lambda, 1), \qquad Q(\lambda t, t) = t^m Q(\lambda, 1), \tag{1.19}$$

for some m. The method of solution is to make the change of variable $y(t) = x(t)/t$. Since

$$\dot{x}(t) = y + t\,\dot{y}(t), \tag{1.20}$$

(1.18) can be re-expressed in the form

$$t\frac{\mathrm{d}y}{\mathrm{d}t} + y = \frac{P(y,1)}{Q(y,1)}, \tag{1.21}$$

which is separable and equivalent to

$$\int^{x/t} \frac{Q(y,1)\mathrm{d}y}{P(y,1) - yQ(y,1)} = \int \frac{\mathrm{d}t}{t} + \mathsf{C} = \ln|t| + \mathsf{C}. \tag{1.22}$$

Example 1.3.2 Find the general solution of

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \frac{2x^2 + t^2}{xt}. \tag{1.23}$$

In terms of the variable $y = x/t$ the equation becomes

$$t\frac{\mathrm{d}y}{\mathrm{d}t} = \frac{1 + y^2}{y}. \tag{1.24}$$

So

$$\int \frac{\mathrm{d}t}{t} = \int^{x/t} \frac{y\,\mathrm{d}y}{1 + y^2} + \mathsf{C}, \tag{1.25}$$

giving

$$\ln|t| = \tfrac{1}{2}\ln|1 + x^2/t^2| + \mathsf{C}. \tag{1.26}$$

This can be solved to give

$$x(t) = \pm\sqrt{\mathsf{C}'t^4 - t^2}. \tag{1.27}$$

1.3.3 First-Order Linear Equations

Consider the equation

$$\dot{x}(t) + f(t)x(t) = g(t). \tag{1.28}$$

Multiplying through by $\mu(t)$ (a function to be chosen later) we have

$$\mu\frac{\mathrm{d}x}{\mathrm{d}t} + \mu f x = \mu g, \tag{1.29}$$

or equivalently

$$\frac{\mathrm{d}(\mu x)}{\mathrm{d}t} - x\frac{\mathrm{d}\mu}{\mathrm{d}t} + \mu f x = \mu g. \tag{1.30}$$

Now we choose $\mu(t)$ to be a solution of

$$\frac{\mathrm{d}\mu}{\mathrm{d}t} = \mu f \tag{1.31}$$

and (1.30) becomes

$$\frac{\mathrm{d}(\mu x)}{\mathrm{d}t} = \mu g \tag{1.32}$$

which has the solution

$$\mu(t)x(t) = \int \mu(t)g(t)\mathrm{d}t + \mathsf{C}. \tag{1.33}$$

The function $\mu(t)$ given, from (1.31), by

$$\mu(t) = \exp\left(\int f(t)\mathrm{d}t\right). \tag{1.34}$$

is an *integrating factor*.

Example 1.3.3 Find the general solution of

$$\frac{\mathrm{d}x}{\mathrm{d}t} + x\cot(t) = 2\cos(t). \tag{1.35}$$

The integrating factor is

$$\exp\left(\int \cot(t)\mathrm{d}t\right) = \exp\left[\ln\sin(t)\right] = \sin(t), \tag{1.36}$$

giving

$$\begin{aligned} \sin(t)\frac{\mathrm{d}x}{\mathrm{d}t} + x\cos(t) &= 2\sin(t)\cos(t) = \sin(2t), \\ \frac{\mathrm{d}[x\sin(t)]}{\mathrm{d}t} &= \sin(2t). \end{aligned} \tag{1.37}$$

So

$$\begin{aligned} x\sin(t) &= \int \sin(2t)\mathrm{d}t + \mathsf{C} \\ &= -\tfrac{1}{2}\cos(2t) + \mathsf{C}, \end{aligned} \tag{1.38}$$

giving

$$x(t) = \frac{\mathsf{C}' - \cos^2(t)}{\sin(t)}. \tag{1.39}$$

The MAPLE code for solving this equation is

```
> dsolve(diff(x(t),t)+x(t)*cot(t)=2*cos(t));
```

$$\mathrm{x}(t) = \frac{-\frac{1}{2}\cos(2\,t) + \_C1}{\sin(t)}$$

1.4 Equations with Constant Coefficients

Consider the n-th order linear differential equation

$$\frac{\mathrm{d}^n x}{\mathrm{d}t^n} + a_{n-1}\frac{\mathrm{d}^{n-1}x}{\mathrm{d}t^{n-1}} + \cdots + a_1\frac{\mathrm{d}x}{\mathrm{d}t} + a_0 x = f(t), \tag{1.40}$$

where $a_0, a_1, \ldots, a_{n-1}$ are *real* constants. We use the $\mathcal{D}$-*operator* notation. With

$$\mathcal{D} = \frac{\mathrm{d}}{\mathrm{d}t} \tag{1.41}$$

(1.40) can be expressed in the form

$$\phi(\mathcal{D})x(t) = f(t), \tag{1.42}$$

where

$$\phi(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0. \tag{1.43}$$

Equation (1.40) (or (1.42)) is said to be *homogeneous* or *inhomogeneous* according to whether $f(t)$ is or is not identically zero. An important result for the solution of this equation is the following:

Theorem 1.4.1 *The general solution of the inhomogeneous equation (1.42) (with $f(t) \not\equiv 0$) is given by the sum of the general solution of the homogeneous equation*

$$\phi(\mathcal{D})x(t) = 0, \tag{1.44}$$

and any particular solution of (1.42).

Proof: Let $x_\mathrm{c}(t)$ be the general solution of (1.44) and let $x_\mathrm{p}(t)$ be a particular solution to (1.42). Since $\phi(\mathcal{D})x_\mathrm{c}(t) = 0$ and $\phi(\mathcal{D})x_\mathrm{p}(t) = f(t)$

$$\phi(\mathcal{D})[x_\mathrm{c}(t) + x_\mathrm{p}(t)] = f(t). \tag{1.45}$$

So

$$x(t) = x_\mathrm{c}(t) + x_\mathrm{p}(t) \tag{1.46}$$

is a solution of (1.42). That it is the *general solution* follows from the fact that, since $x_\mathrm{c}(t)$ contains n arbitrary constants, then so does $x(t)$. If

$$x'(t) = x_\mathrm{c}(t) + x'_\mathrm{p}(t) \tag{1.47}$$

were the solution obtained with a different particular solution then it is easy to see that the difference between $x(t)$ and $x'(t)$ is just a particular solution of (1.44). So going from $x(t)$ to $x'(t)$ simply involves a change in the arbitrary constants.

We divide the problem of finding the solution to (1.42) into two parts. We first describe a method for finding $x_\mathrm{c}(t)$, usually called the *complementary function*, and then we develop a method for finding a particular solution $x_\mathrm{p}(t)$.

1.4.1 Finding the Complementary Function

It is a consequence of the fundamental theorem of algebra that the n-th degree polynomial equation

$$\phi(\lambda) = 0 \tag{1.48}$$

has exactly n (possibly complex and not necessarily distinct) solutions. In this context (1.48) is called the *auxiliary equation* and we see that, since the coefficients $a_0, a_1, \ldots, a_{n-1}$ in (1.43) are all real, complex solutions of the auxiliary equation appear in conjugate complex pairs. Suppose the solutions are $\lambda_1, \lambda_2, \ldots, \lambda_n$. Then the homogeneous equation (1.44) can be expressed in the form

$$(\mathcal{D} - \lambda_1)(\mathcal{D} - \lambda_2)\cdots(\mathcal{D} - \lambda_n)x(t) = 0. \tag{1.49}$$

It follows that the solutions of the n first-order equations

$$(\mathcal{D} - \lambda_j)x(t) = 0, \qquad j = 1, 2, \ldots, n \tag{1.50}$$

are also solutions of (1.49) and hence of (1.44). The equations (1.50) are simply

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \lambda_j x, \qquad j = 1, 2, \ldots, n \tag{1.51}$$

with solutions

$$x_j(t) = \mathsf{C}_j \exp(\lambda_j t). \tag{1.52}$$

If all of the roots $\lambda_1, \lambda_2, \ldots, \lambda_n$ *are distinct* we have the complementary function given by

$$x_{\mathrm{c}}(t) = \mathsf{C}_1 \exp(\lambda_1 t) + \mathsf{C}_2 \exp(\lambda_2 t) + \cdots + \mathsf{C}_n \exp(\lambda_n t). \tag{1.53}$$

Setting aside for the moment the case of equal roots, we observe that (1.53) includes the possibility of:

(i) A zero root, when the contribution to the complementary function is just a constant term.

(ii) Pairs of complex roots. Suppose that for some j

$$\lambda_j = \alpha + \mathrm{i}\beta, \qquad \lambda_{j+1} = \alpha - \mathrm{i}\beta. \tag{1.54}$$

Then

$$\begin{aligned} \mathsf{C}_j \exp(\lambda_j t) + \mathsf{C}_{j+1} \exp(\lambda_{j+1} t) &= \exp(\alpha t)\big\{\mathsf{C}_j[\cos(\beta t) + \mathrm{i}\sin(\beta t)] \\ &\qquad + \mathsf{C}_{j+1}[\cos(\beta t) - \mathrm{i}\sin(\beta t)]\big\} \\ &= \exp(\alpha t)\left\{\mathsf{C}\cos(\beta t) + \mathsf{C}'\sin(\beta t)\right\}, \end{aligned} \tag{1.55}$$

where

$$\mathsf{C} = \mathsf{C}_j + \mathsf{C}_{j+1}, \qquad \mathsf{C}' = \mathrm{i}[\mathsf{C}_j - \mathsf{C}_{j+1}]. \tag{1.56}$$

In order to consider the case of equal roots we need the following result:

Theorem 1.4.2 *For any positive integer n and function $u(t)$*

$$\mathcal{D}^n u(t)\exp(\lambda t) = \exp(\lambda t)(\mathcal{D}+\lambda)^n u(t) \tag{1.57}$$

Proof: We prove the result by induction. For $n = 1$

$$\begin{aligned}\mathcal{D}u(t)\exp(\lambda t) &= \exp(\lambda t)\mathcal{D}u(t) + u(t)\mathcal{D}\exp(\lambda t)\\ &= \exp(\lambda t)\mathcal{D}u(t) + \exp(\lambda t)\lambda u(t)\\ &= \exp(\lambda t)(\mathcal{D}+\lambda)u(t).\end{aligned} \tag{1.58}$$

Now suppose the result is true for some n. Then

$$\begin{aligned}\mathcal{D}^{n+1}u(t)\exp(\lambda t) &= \mathcal{D}\exp(\lambda t)(\mathcal{D}+\lambda)^n u(t)\\ &= (\mathcal{D}+\lambda)^n u(t)\mathcal{D}\exp(\lambda t)\\ &\quad + \exp(\lambda t)\mathcal{D}(\mathcal{D}+\lambda)^n u(t)\\ &= \exp(\lambda t)\lambda(\mathcal{D}+\lambda)^n u(t)\\ &\quad + \exp(\lambda t)\mathcal{D}(\mathcal{D}+\lambda)^n u(t)\\ &= \exp(\lambda t)(\mathcal{D}+\lambda)^{n+1}u(t).\end{aligned} \tag{1.59}$$

and the result is established for all n.

An immediate consequence of this theorem is that, for any polynomial $\phi(\mathcal{D})$ with constant coefficients,

$$\phi(\mathcal{D})u(t)\exp(\lambda t) = \exp(\lambda t)\phi(\mathcal{D}+\lambda)u(t) \tag{1.60}$$

Suppose now that $(\mathcal{D}-\lambda')^m$ is a factor in the expansion of $\phi(\mathcal{D})$ and that all the other roots of the auxiliary equation are distinct from λ'. It is clear that one solution of the homogeneous equation (1.44) is $\mathsf{C}\exp(\lambda' t)$, but we need $m-1$ more solutions associated with this root to complete the complementary function. Suppose we try the solution $u(t)\exp(\lambda' t)$ for some polynomial $u(t)$. From (1.60)

$$\begin{aligned}(\mathcal{D}-\lambda')^m u(t)\exp(\lambda' t) &= \exp(\lambda' t)(\mathcal{D}+\lambda'-\lambda')^m u(t)\\ &= \exp(\lambda' t)\mathcal{D}^m u(t).\end{aligned} \tag{1.61}$$

The general solution of

$$\mathcal{D}^m u(t) = 0 \tag{1.62}$$

is

$$u(t) = \left[\mathsf{C}^{(0)} + \mathsf{C}^{(1)}t + \cdots + \mathsf{C}^{(m-1)}t^{m-1}\right]. \tag{1.63}$$

So the contribution to the complementary function from an m-fold degenerate root λ' of the auxiliary equation is

$$x'_{\mathrm{c}}(t) = \left[\mathsf{C}^{(0)} + \mathsf{C}^{(1)}t + \cdots + \mathsf{C}^{(m-1)}t^{m-1}\right]\exp(\lambda' t) \tag{1.64}$$

Example 1.4.1 Find the general solution of

$$\frac{\mathrm{d}^2 x}{\mathrm{d}t^2} - 3\frac{\mathrm{d}x}{\mathrm{d}t} - 4x = 0. \tag{1.65}$$

The auxiliary equation $\lambda^2 - 3\lambda - 4 = 0$ has roots $\lambda = -1, 4$. So the solution is

$$x(t) = \mathsf{C}_1 \exp(-t) + \mathsf{C}_2 \exp(4t). \tag{1.66}$$

Example 1.4.2 Find the general solution of

$$\frac{\mathrm{d}^2 x}{\mathrm{d}t^2} + 4\frac{\mathrm{d}x}{\mathrm{d}t} + 13x = 0. \tag{1.67}$$

The auxiliary equation $\lambda^2 + 4\lambda + 13 = 0$ has roots $\lambda = -2 \pm 3\mathrm{i}$. So the solution is

$$x(t) = \exp(-2t)\left[\mathsf{C}_1 \cos(3t) + \mathsf{C}_2 \sin(3t)\right]. \tag{1.68}$$

Example 1.4.3 Find the general solution of

$$\frac{\mathrm{d}^3 x}{\mathrm{d}t^3} + 3\frac{\mathrm{d}^2 x}{\mathrm{d}t^2} - 4x = 0. \tag{1.69}$$

The auxiliary equation is a cubic $\lambda^3 + 3\lambda^2 - 4 = 0$. It is easy to spot that one root is $\lambda = 1$. Once this is factorized out we have $(\lambda - 1)(\lambda^2 + 4\lambda + 4) = 0$ and the quadratic part has the two-fold degenerate root $\lambda = -2$. So the solution is

$$x(t) = \mathsf{C}_1 \exp(t) + [\mathsf{C}_2 + \mathsf{C}_3 t] \exp(-2t). \tag{1.70}$$

Of course, it is possible for a degenerate root to be complex. Then the form of that part of the solution will be a product of the appropriate polynomial in t and the form for a pair of complex conjugate roots.

1.4.2 A Particular Solution

There are a number of methods for finding a particular solution $x_{\mathrm{p}}(t)$ to the inhomogeneous equation (1.42). We shall use the method of *trial functions*. We substitute a trial function $\mathsf{T}(t)$, containing a number of arbitrary constants (A, B etc.) into the equation and then adjust the values of the constants to achieve a solution. Suppose, for example,

$$f(t) = a \exp(bt). \tag{1.71}$$

Now take the trial function $\mathsf{T}(t) = \mathsf{A} \exp(bt)$. From (1.60)

$$\phi(\mathcal{D})\mathsf{T}(t) = \mathsf{A} \exp(bt)\phi(b). \tag{1.72}$$

Equating this with $f(t)$, given by (1.71), we see that the trial function is a solution of (1.42) if

$$\mathsf{A} = \frac{a}{\phi(b)}, \tag{1.73}$$

as long as $\phi(b) \neq 0$, that is, when b is not a root of the auxiliary equation (1.48). To consider that case suppose that

$$\phi(\lambda) = \psi(\lambda)(\lambda - b)^m, \qquad \psi(b) \neq 0. \tag{1.74}$$

That is b is an m-fold root of the auxiliary equation. Now try the trial function $\mathsf{T}(t) = \mathsf{A}t^m \exp(bt)$. From (1.60)

$$\begin{aligned} \phi(\mathcal{D})\mathsf{T}(t) &= \psi(\mathcal{D})(\mathcal{D} - b)^m \mathsf{A}t^m \exp(bt) \\ &= \mathsf{A}\exp(bt)\psi(\mathcal{D} + b)[(\mathcal{D} + b) - b]^m t^m \\ &= \mathsf{A}\exp(bt)\psi(\mathcal{D} + b)\mathcal{D}^m t^m \\ &= \mathsf{A}\exp(bt)\psi(b)m! \end{aligned} \tag{1.75}$$

Equating this with $f(t)$, given by (1.71), we see that the trial function is a solution of (1.42) if

$$\mathsf{A} = \frac{a}{m!\psi(b)}. \tag{1.76}$$

Table 1.1 contains a list of trial functions to be used for different forms of $f(t)$. Trial functions when $f(t)$ is a linear combination of the forms given are simply the corresponding linear combination of the trial functions. Although there seems to be a lot of different cases it can be seen that they are all special cases of either the eighth or tenth lines. We conclude this section with two examples.

Example 1.4.4 Find the general solution of

$$\frac{\mathrm{d}^2 x}{\mathrm{d}t^2} - 4\frac{\mathrm{d}x}{\mathrm{d}t} + 3x = 6t - 11 + 8\exp(-t). \tag{1.77}$$

The auxiliary equation is

$$\lambda^2 - 4\lambda + 3 = 0, \tag{1.78}$$

with roots $\lambda = 1, 3$. So

$$x_{\mathrm{c}}(t) = \mathsf{C}_1 \exp(t) + \mathsf{C}_2 \exp(3t). \tag{1.79}$$

From Table 1.1 the trial function for

- $6t$ is $\mathrm{B}_1 t + \mathrm{B}_2$, since zero is not a root of (1.78).
- -11 is B_3, since zero is not a root of (1.78).
- $8\exp(-t)$ is $\mathsf{A}\exp(-t)$, since -1 is not a root of (1.78).

The constant B_3 can be neglected and we have

$$\mathsf{T}(t) = \mathrm{B}_1 t + \mathrm{B}_2 + \mathsf{A}\exp(-t). \tag{1.80}$$

Now

$$\phi(\mathcal{D})\mathsf{T}(t) = 3\mathrm{B}_1 t + 3\mathrm{B}_2 - 4\mathrm{B}_1 + 8\mathsf{A}\exp(-t) \tag{1.81}$$

Table 1.1: **Table of trial functions for finding a particular integral for $\phi(\mathcal{D})x = f(t)$**

$f(t)$	$\mathsf{T}(t)$	Comments
$a\exp(bt)$	$\mathsf{A}\exp(bt)$	b not a root of $\phi(\lambda) = 0$.
$a\exp(bt)$	$\mathsf{A}t^k\exp(bt)$	b a root of $\phi(\lambda) = 0$ of multiplicity k.
$a\sin(bt)$ or $a\cos(bt)$	$\mathsf{A}\sin(bt) + \mathsf{B}\cos(bt)$	$\lambda^2 + b^2$ not a factor of $\phi(\lambda)$.
$a\sin(bt)$ or $a\cos(bt)$	$t^k[\mathsf{A}\sin(bt) + \mathsf{B}\cos(bt)]$	$\lambda^2 + b^2$ a factor of $\phi(\lambda)$ of multiplicity k.
at^n	$\mathsf{A}_n t^n + \mathsf{A}_{n-1}t^{n-1} + \cdots + \mathsf{A}_0$	Zero is not a root of $\phi(\lambda) = 0$.
at^n	$t^k[\mathsf{A}_n t^n + \mathsf{A}_{n-1}t^{n-1} + \cdots + \mathsf{A}_0]$	Zero is a root of $\phi(\lambda) = 0$ of multiplicity k.
$at^n\exp(bt)$	$\exp(bt)[\mathsf{A}_n t^n + \mathsf{A}_{n-1}t^{n-1} + \cdots + \mathsf{A}_0]$	b is not a root of $\phi(\lambda) = 0$.
$at^n\exp(bt)$	$t^k\exp(bt)[\mathsf{A}_n t^n + \mathsf{A}_{n-1}t^{n-1} + \cdots + \mathsf{A}_0]$	b is a root of $\phi(\lambda) = 0$ of multiplicity k.
$at^n\sin(bt)$ or $at^n\cos(bt)$	$[\mathsf{B}_1\sin(bt) + \mathsf{B}_2\cos(bt)][t^n + \mathsf{A}_{n-1}t^{n-1} + \cdots + \mathsf{A}_0]$	$\lambda^2 + b^2$ not a factor of $\phi(\lambda)$.
$at^n\sin(bt)$ or $at^n\cos(bt)$	$t^k[\mathsf{B}_1\sin(bt) + \mathsf{B}_2\cos(bt)][t^n + \mathsf{A}_{n-1}t^{n-1} + \cdots + \mathsf{A}_0]$	$\lambda^2 + b^2$ a factor of $\phi(\lambda)$ of multiplicity k.

and comparing with $f(t)$ gives $\mathrm{B}_1 = 2$, $\mathrm{B}_2 = -1$ and $\mathsf{A} = 1$. Thus

$$x_{\mathrm{p}}(t) = 2t - 1 + \exp(-t) \tag{1.82}$$

and

$$x(t) = \mathsf{C}_1 \exp(t) + \mathsf{C}_2 \exp(3t) + 2t - 1 + \exp(-t). \tag{1.83}$$

The MAPLE code for solving this equation is

```
> dsolve(diff(x(t),t$2)-4*diff(x(t),t)+3*x(t)=6*t-11+8*exp(-t));
```

$$\mathrm{x}(t) = 2\,t - 1 + e^{(-t)} + \_C1\, e^t + \_C2\, e^{(3\,t)}$$

Example 1.4.5 Find the general solution of

$$\frac{\mathrm{d}^3 x}{\mathrm{d}t^3} + \frac{\mathrm{d}^2 x}{\mathrm{d}t^2} = 4 - 12\exp(2t). \tag{1.84}$$

The auxiliary equation is

$$\lambda^2(\lambda + 1) = 0, \tag{1.85}$$

with roots $\lambda = 0$ (twice) and $\lambda = -1$. So

$$x_{\mathrm{c}}(t) = \mathsf{C}_0 + \mathsf{C}_1 t + \mathsf{C}_3 \exp(-t). \tag{1.86}$$

From Table 1.1 the trial function for

- 4 is $\mathsf{B}t^2$, since zero is double root of (1.85).
- $-12\exp(2t)$ is $\mathsf{A}\exp(2t)$, since 2 is not a root of (1.85).

We have

$$\mathsf{T}(t) = \mathsf{B}t^2 + \mathsf{A}\exp(2t) \tag{1.87}$$

and

$$\phi(\mathcal{D})\mathsf{T}(t) = 2\mathsf{B} + 12\mathsf{A}\exp(2t). \tag{1.88}$$

Comparing with $f(t)$ gives $\mathsf{B} = 2$ and $\mathsf{A} = -1$. Thus

$$x(t) = \mathsf{C}_0 + \mathsf{C}_1 t + \mathsf{C}_3 \exp(-t) + 2t^2 - \exp(2t). \tag{1.89}$$

1.5 Systems of Differential Equations

If, for the n-th order differential equation (1.10), we define the new set of variables $x_1 = x$, $x_2 = \mathrm{d}x/\mathrm{d}t, \ldots, x_n = \mathrm{d}^{n-1}x/\mathrm{d}t^{n-1}$ then the one n-th order differential equation with independent variable t and one dependent variable x can be replaced by the system

$$\begin{aligned} \frac{\mathrm{d}x_1}{\mathrm{d}t} &= x_2(t), \\ \frac{\mathrm{d}x_2}{\mathrm{d}t} &= x_3(t), \\ \vdots \qquad & \quad \vdots \\ \frac{\mathrm{d}x_{n-1}}{\mathrm{d}t} &= x_n(t), \\ \frac{\mathrm{d}x_n}{\mathrm{d}t} &= F(x_1, x_2, \ldots, x_n; t) \end{aligned} \tag{1.90}$$

of n first-order equations with independent variable t and n dependent variables $x_1, \ldots, x_n$. In fact this is just a special case of

$$\begin{aligned} \frac{\mathrm{d}x_1}{\mathrm{d}t} &= F_1(x_1, x_2, \ldots, x_n; t), \\ \frac{\mathrm{d}x_2}{\mathrm{d}t} &= F_2(x_1, x_2, \ldots, x_n; t), \\ \vdots \qquad & \qquad\qquad \vdots \\ \frac{\mathrm{d}x_{n-1}}{\mathrm{d}t} &= F_{n-1}(x_1, x_2, \ldots, x_n; t), \\ \frac{\mathrm{d}x_n}{\mathrm{d}t} &= F_n(x_1, x_2, \ldots, x_n; t), \end{aligned} \tag{1.91}$$

where the right-hand sides of *all* the equations are now functions of the variables $x_1, x_2, \ldots, x_n$.[3] The system defined by (1.91) is called an *n-th order dynamical system*. Such a system is said to be *autonomous* if none of the functions F_ℓ is an explicit function of t.

Picard's theorem generalizes in the natural way to this n-variable case as does also the procedure for obtained approximations to a solution with Picard iterates. That is, with the initial condition $x_\ell(\tau) = \xi_\ell$, $\ell = 1, 2, \ldots, n$, we define the set of sequences $\{x_\ell^{(j)}(t)\}$, $\ell = 1, 2, \ldots, n$ with

$$\begin{aligned} x_\ell^{(0)}(t) &= \xi_\ell, \\ x_\ell^{(j+1)}(t) &= \xi_\ell + \int_\tau^t F_\ell(x_1^{(j)}(u), \ldots, x_n^{(j)}(u); u)\mathrm{d}u, \qquad j = 1, 2, \ldots \end{aligned} \tag{1.92}$$

for all $\ell = 1, 2, \ldots, n$.

[3] Of course, such a system is not, in general, equivalent to one n-th order equation.

Example 1.5.1 Consider the *simple harmonic* differential equation

$$\ddot{x}(t) = -\omega^2 x(t) \tag{1.93}$$

with the initial conditions $x(0) = 0$ and $\dot{x}(0) = \omega$.

This equation is equivalent to the system

$$\begin{aligned} &\dot{x}_1(t) = x_2(t), && \dot{x}_2(t) = -\omega^2 x_1(t), \\ &x_1(0) = 0, && x_2(0) = \omega \end{aligned} \tag{1.94}$$

which is a second-order autonomous system. From (1.92)

$$\begin{aligned} x_1^{(0)}(t) &= 0, & x_2^{(0)}(t) &= \omega, \\ x_1^{(1)}(t) &= 0 + \int_0^t \omega \mathrm{d}u, & x_2^{(1)}(t) &= \omega - \omega^2 \int_0^t 0 \mathrm{d}u \\ &= \omega t, & &= \omega, \\ x_1^{(2)}(t) &= 0 + \int_0^t \omega \mathrm{d}u, & x_2^{(2)}(t) &= \omega - \int_0^t \omega^3 u \mathrm{d}u \\ &= \omega t, & &= \omega \left\{ 1 - \tfrac{(\omega t)^2}{2!} \right\}, \\ x_1^{(3)}(t) &= 0 + \int_0^t \omega \left\{ 1 - \tfrac{(\omega t)^2}{2!} \right\} \mathrm{d}u, & x_2^{(3)}(t) &= \omega - \int_0^t \omega^3 u \mathrm{d}u \\ &= \omega t - \tfrac{(\omega t)^3}{3!}, & &= \omega \left\{ 1 - \tfrac{(\omega t)^2}{2!} \right\}, \end{aligned} \tag{1.95}$$

The pattern which is emerging is clear

$$\begin{aligned} x_1^{(2j-1)}(t) = x_1^{(2j)}(t) \quad &= \quad \omega t - \tfrac{(\omega t)^3}{3!} + \cdots + (-1)^{j+1} \tfrac{(\omega t)^{(2j-1)}}{(2j-1)!}, \\ & \qquad\qquad j = 1, 2, \ldots, \end{aligned} \tag{1.96}$$

$$\begin{aligned} x_2^{(2j)}(t) = x_1^{(2j+1)}(t) \quad &= \quad \omega \left\{ 1 - \tfrac{(\omega t)^2}{2!} + \cdots + (-1)^j \tfrac{(\omega t)^{(2j)}}{(2j)!} \right\}, \\ & \qquad\qquad j = 0, 1, \ldots. \end{aligned} \tag{1.97}$$

In the limit $j \to \infty$ (1.96) becomes the MacLaurin expansion for $\sin(\omega t)$ and (1.97) the MacLaurin expansion for $\omega \cos(\omega t)$.

The set of equations (1.91) can be written in the vector form

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{F}(\boldsymbol{x}; t), \tag{1.98}$$

where

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \qquad \boldsymbol{F}(\boldsymbol{x}; t) = \begin{pmatrix} F_1(\boldsymbol{x}; t) \\ F_2(\boldsymbol{x}; t) \\ \vdots \\ F_n(\boldsymbol{x}; t) \end{pmatrix}. \tag{1.99}$$

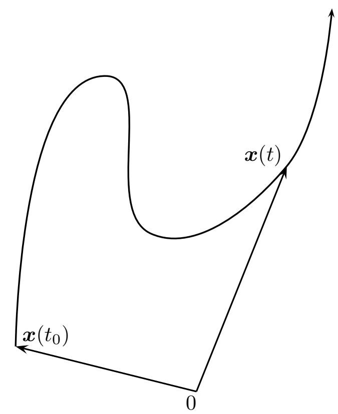


Figure 1.1: A trajectory in the phase space $\mathit{\Gamma}_n$.

As time passes the vector $\boldsymbol{x}(t)$ describes a *trajectory* in an n-dimensional *phase space* $\mathit{\Gamma}_n$ (Fig. 1.1). The trajectory is determined by the nature of the vector field $\boldsymbol{F}(\boldsymbol{x};t)$ and the location $\boldsymbol{x}(t_0)$ of the phase point at some time t_0. An important property of autonomous dynamical systems is that, if the system is at $\boldsymbol{x}^{(0)}$ at time t_0 then the state $\boldsymbol{x}^{(1)}$ of the system at t_1 is dependent on $\boldsymbol{x}^{(0)}$ and $t_1 - t_0$, but not on t_0 and t_1 individually.

A dynamical system with $2m$ degrees of freedom and variables $\{x_1, \ldots, x_m, p_1, \ldots, p_m\}$ is a *Hamiltonian system* if there exists a Hamiltonian function $H(x_1, \ldots, x_m, p_1, \ldots, p_m; t)$ in terms of which the evolution of the system is given by Hamilton's equations

$$\begin{aligned} \dot{x}_s &= \frac{\partial H}{\partial p_s}, \\ \dot{p}_s &= -\frac{\partial H}{\partial x_s}, \end{aligned} \qquad s = 1, 2, \ldots, m. \tag{1.100}$$

It follows that the rate of change of H along a trajectory is given by

$$\begin{aligned} \frac{\mathrm{d}H}{\mathrm{d}t} &= \sum_{s=1}^{m} \left\{ \frac{\partial H}{\partial x_s}\frac{\mathrm{d}x_s}{\mathrm{d}t} + \frac{\partial H}{\partial p_s}\frac{\mathrm{d}p_s}{\mathrm{d}t} \right\} + \frac{\partial H}{\partial t} \\ &= \sum_{s=1}^{m} \left\{ \frac{\partial H}{\partial x_s}\frac{\partial H}{\partial p_s} - \frac{\partial H}{\partial p_s}\frac{\partial H}{\partial x_s} \right\} + \frac{\partial H}{\partial t} \\ &= \frac{\partial H}{\partial t}. \end{aligned} \tag{1.101}$$

So if the system is autonomous ($\partial H/\partial t = 0$) the value of H does not change along a trajectory. It is said to be a *constant of motion*. In the case of many physical systems the Hamiltonian is the total energy of the system and the trajectory is a path lying on the energy surface in phase space.

As we have already seen, a system with m variables $\{x_1, x_2, \dots, x_m\}$ determined by second-order differential equations, given in vector form by

$$\ddot{\boldsymbol{x}}(t) = \boldsymbol{G}(\boldsymbol{x}; t), \tag{1.102}$$

where

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix}, \qquad \boldsymbol{G}(\boldsymbol{x}; t) = \begin{pmatrix} G_1(\boldsymbol{x}; t) \\ G_2(\boldsymbol{x}; t) \\ \vdots \\ G_m(\boldsymbol{x}; t) \end{pmatrix}, \tag{1.103}$$

is equivalent to the $2m$-th order dynamical system

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{p}(t), \qquad \dot{\boldsymbol{p}}(t) = \boldsymbol{G}(\boldsymbol{x}; t), \tag{1.104}$$

where

$$\boldsymbol{p} = \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_m \end{pmatrix} = \begin{pmatrix} \partial x_1/\partial t \\ \partial x_2/\partial t \\ \vdots \\ \partial x_m/\partial t \end{pmatrix}. \tag{1.105}$$

If there exists a scalar potential $V(\boldsymbol{x}; t)$, such that

$$\boldsymbol{G}(\boldsymbol{x}; t) = -\boldsymbol{\nabla} V(\boldsymbol{x}; t), \tag{1.106}$$

the system is said to be *conservative*. By defining

$$H(\boldsymbol{x}, \boldsymbol{p}; t) = \tfrac{1}{2}\boldsymbol{p}^2 + V(\boldsymbol{x}; t), \tag{1.107}$$

we see that a conservative system is also a Hamiltonian system. In a physical context this system can be taken to represent the motion of a set of m/d particles of unit mass moving in a space of dimension d, with position and momentum coordinates $x_1.x_2, \dots, x_m$ and $p_1, p_2, \dots, p_m$ respectively. Then $\frac{1}{2}\boldsymbol{p}^2$ and $V(\boldsymbol{x}; t)$ are respectively the kinetic and potential energies.

A rather more general case is when, for the system defined by equations (1.98) and (1.99), there exists a scalar field $U(\boldsymbol{x}; t)$ with

$$\boldsymbol{F}(\boldsymbol{x}; t) = -\boldsymbol{\nabla} U(\boldsymbol{x}; t). \tag{1.108}$$

1.5.1 MAPLE for Systems of Differential Equations

In the discussion of systems of differential equations we shall be less concerned with the analytic form of the solutions than with their qualitative structure. As we shall show below, a lot of information can be gained by finding the equilibrium points and determining their stability. It is also useful to be able to plot a trajectory with given initial conditions. MAPLE can be used for this in two (and possibly three) dimensions. Suppose we want to obtain a plot of the solution of

$$\dot{x}(t) = x(t) - y(t), \qquad \dot{y}(t) = x(t), \tag{1.109}$$

over the range $t = 0$ to $t = 10$, with initial conditions $x(0) = 1$, $y(0) = -1$. The MAPLE routine `dsolve` can be used for systems with the equations and the initial conditions enclosed in curly brackets. Unfortunately the solution is returned as a set $\{x(t) = \cdots, y(t) = \cdots\}$, which cannot be fed directly into the `plot` routine. To get round this difficulty we set the solution to some variable (`Fset` in this case) and extract $x(t)$ and $y(t)$ (renamed as $fx(t)$ and $fy(t)$) by using the MAPLE function `subs`. These functions can now be plotted parametrically. The complete MAPLE code and results are:

```
> Fset:=dsolve(
> {diff(x(t),t)=x(t)-y(t),diff(y(t),t)=x(t),x(0)=1,y(0)=-1},
> {x(t),y(t)}):
```

```
> fx:=t->subs(Fset,x(t)):
```

```
> fx(t);
```

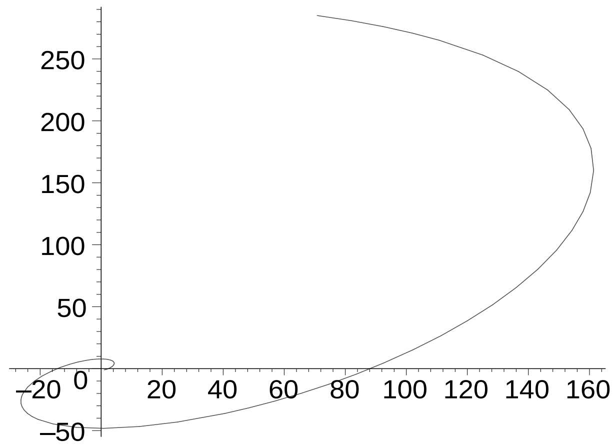
$$\frac{1}{3}\, e^{(1/2\, t)}\, (3\cos(\frac{1}{2}\, t\, \sqrt{3}) + 3\, \sqrt{3}\sin(\frac{1}{2}\, t\, \sqrt{3}))$$

```
> fy:=t->subs(Fset,y(t)):
```

```
> fy(t);
```

$$\frac{1}{3}\, e^{(1/2\, t)}\, (3\, \sqrt{3}\sin(\frac{1}{2}\, t\, \sqrt{3}) - 3\cos(\frac{1}{2}\, t\, \sqrt{3}))$$

```
> plot([fx(t),fy(t),t=0..10]);
```



1.6 Autonomous Systems

We shall now concentrate on systems of the type described by equations (1.98) and (1.99) but, where the vector field $\boldsymbol{F}$ is not an explicit function of time. These are called *autonomous systems*. In fact being autonomous is not such a severe restraint. A non-autonomous system can be made equivalent to an autonomous system by the following trick. We include the time dimension in the phase space by adding the time line Υ to Γ_n. The path in the $(n+1)$-dimensional space $\Gamma_n \times \Upsilon$ is then given by the dynamical system

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{F}(\boldsymbol{x}, x_t), \qquad \dot{x}_t(t) = 1. \tag{1.110}$$

This is called a *suspended system*.

In general the determination of the trajectories in phase space, even for autonomous systems, can be a difficult problem. However, we can often obtain a qualitative idea of the phase pattern of trajectories by considering particularly simple trajectories. The most simple of all are the *equilibrium points*.[4] These are trajectories which consist of one single point. If the phase point starts at an equilibrium point it stays there. The condition for $\boldsymbol{x}^*$ to be an equilibrium point of the autonomous system

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{F}(\boldsymbol{x}), \tag{1.111}$$

is

$$\boldsymbol{F}(\boldsymbol{x}^*) = \boldsymbol{0}. \tag{1.112}$$

[4]Also called, *fixed points*, *critical points* or *nodes*.

For the system given by (1.108) it is clear that a equilibrium point is a stationary point of $U(\boldsymbol{x})$ and for the conservative system given by (1.103)–(1.106) equilibrium points have $\boldsymbol{p} = \boldsymbol{0}$ and are stationary points of $V(\boldsymbol{x})$. An equilibrium point is useful for obtaining information about phase behaviour only if we can determine the behaviour of trajectories in its neighbourhood. This is a matter of the *stability* of the equilibrium point, which in formal terms can be defined in the following way:

The equilibrium point $\boldsymbol{x}^*$ of (1.111) is said to be ***stable (in the sense of Lyapunov)*** *if there exists, for every $\varepsilon > 0$, a $\delta(\varepsilon) > 0$, such that any solution $\boldsymbol{x}(t)$, for which $\boldsymbol{x}(t_0) = \boldsymbol{x}^{(0)}$ and*

$$|\boldsymbol{x}^* - \boldsymbol{x}^{(0)}| < \delta(\varepsilon), \tag{1.113}$$

satisfies

$$|\boldsymbol{x}^* - \boldsymbol{x}(t)| < \varepsilon, \tag{1.114}$$

for all $t \geq t_0$. If no such $\delta(\varepsilon)$ exists then $\boldsymbol{x}^*$ is said to be ***unstable (in the sense of Lyapunov)***. *If $\boldsymbol{x}^*$ is stable and*

$$\lim_{t \to \infty} |\boldsymbol{x}^* - \boldsymbol{x}(t)| = 0. \tag{1.115}$$

it is said to be ***asymptotically stable***. *If the equilibrium point is stable and (1.115) holds for every $\boldsymbol{x}^{(0)}$ then it is said to be* ***globally asymptotically stable***. *In this case $\boldsymbol{x}^*$ must be the unique equilibrium point.*

There is a warning you should note in relation to these definitions. In some texts the term ***stable*** *is used to mean what we have called 'asymptotically stable' and equilibrium points which are stable (in our sense) but not asymptotically stable are called* ***conditionally*** *or* ***marginally stable***.

An asymptotically stable equilibrium point is a type of *attractor*. Other types of attractors can exist. For example, a close (periodic) trajectory to which all neighbouring trajectories converge. These more general questions of stability will be discussed in a later chapter.

1.6.1 One-Variable Autonomous Systems

We first consider a first-order autonomous system. In general a system may contain a number of adjustable parameters $a, b, c, \ldots$ and it is of interest to consider the way in which the equilibrium points and their stability change with changes of these parameters. We consider the equation

$$\dot{x}(t) = F(a, b, c, \ldots, x), \tag{1.116}$$

where $a, b, c, \ldots$ is some a set of one or more independent parameters. An equilibrium point $x^*(a, b, c, \ldots)$ is a solution of

$$F(a, b, c, \ldots, x^*) = 0. \tag{1.117}$$

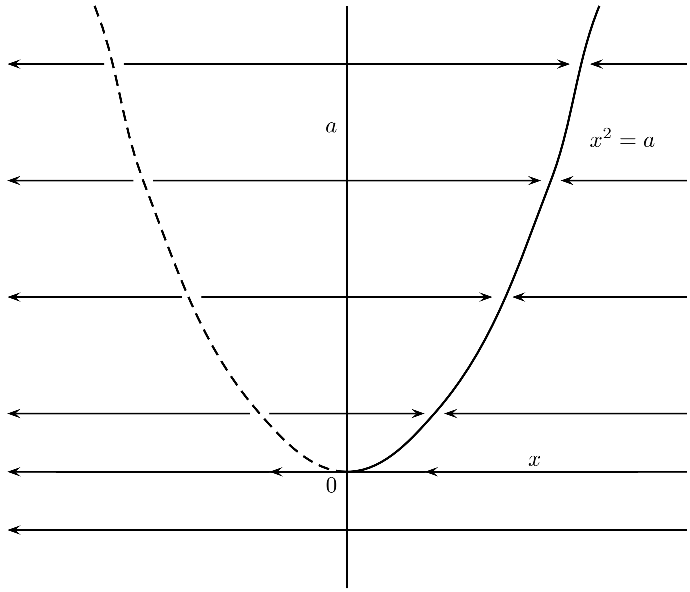


Figure 1.2: The bifurcation diagram for Example 1.6.1. The stable and unstable equilibrium solutions are shown by continuous and broken lines and the direction of the flow is shown by arrows. This is an example of a simple turning point bifurcation.

According to the Lyapunov criterion it is *stable* if, when the phase point is perturbed a small amount from x^* it remains in a neighbourhood of x^*, asymptotically stable if it converges on x^* and *unstable* if it moves away from x^*. We shall, therefore, determine the stability of equilibrium points by linearizing about the point.[5]

Example 1.6.1 Consider one-variable non-linear system given by

$$\dot{x}(t) = a - x^2. \tag{1.118}$$

The parameter a can vary over all real values and the nature of equilibrium points will vary accordingly.

The equilibrium points are given by $x = x^* = \pm\sqrt{a}$. They exist only when $a \geq 0$ and form the parabolic curve shown in Fig. 1.2. Let $x = x^* + \triangle x$ and substitute into (1.118) neglecting all but the linear terms in $\triangle x$.

$$\frac{\mathrm{d}\triangle x}{\mathrm{d}t} = a - (x^*)^2 - 2x^*\triangle x. \tag{1.119}$$

[5]A theorem establishing the formal relationship between this *linear stability* and the Lyapunov criteria will be stated below.

Since $a = (x^*)^2$ this gives

$$\frac{\mathrm{d}\triangle x}{\mathrm{d}t} = -2x^*\triangle x, \tag{1.120}$$

which has the solution

$$\triangle x = \mathrm{C}\exp(-2x^*t). \tag{1.121}$$

Thus the equilibrium point $x^* = \sqrt{a} > 0$ is asymptotically stable (denoted by a continuous line in Fig. 1.2) and the equilibrium point $x^* = -\sqrt{a} < 0$ is unstable (denoted by a broken line in Fig. 1.2). When $a \le 0$ it is clear that $\dot{x}(t) < 0$ so $x(t)$ decreases monotonically from its initial value $x(0)$. In fact for $a = 0$ equation (1.118) is easily solved:

$$\int_{x(0)}^{x} x^{-2}\mathrm{d}x = -\int_0^t \mathrm{d}t \tag{1.122}$$

gives

$$x(t) = \frac{x(0)}{1 + tx(0)} \qquad \dot{x}(t) = -\left(\frac{x(0)}{1+tx(0)}\right)^2. \tag{1.123}$$

Then

$$x(t) \to \begin{cases} 0, & \text{as } t \to \infty \text{ if } x(0) > 0, \\ -\infty, & \text{as } t \to 1/|x(0)| \text{ if } x(0) < 0. \end{cases} \tag{1.124}$$

In each case $x(t)$ decreases with increasing t. When $x(0) > 0$ it takes 'forever' to reach the origin. For $x(0) < 0$ it attains minus infinity in a finite amount of time and then 'reappears' at infinity and decreases to the origin as $t \to \infty$. The linear equation (1.120) cannot be applied to determine the stability of $x^* = 0$ as it gives $(\mathrm{d}\triangle x/\mathrm{d}t)^* = 0$. If we retain the quadratic term we have

$$\frac{\mathrm{d}\triangle x}{\mathrm{d}t} = -(\triangle x)^2. \tag{1.125}$$

So including the second degree term we see that $\mathrm{d}\triangle x/\mathrm{d}t < 0$. If $\triangle x > 0$ $x(t)$ moves towards the equilibrium point and if $\triangle x < 0$ it moves away. In the strict Lyapunov sense the equilibrium point $x^* = 0$ is unstable. But it is 'less unstable' that $x^* = -\sqrt{a}$, for $a > 0$, since there is a path of attraction. It is at the boundary between the region where there are no equilibrium points and the region where there are two equilibrium points. It is said to be *on the margin of stability*. The value $a = 0$ separates the stable range from the unstable range. Such equilibrium points are *bifurcation points*. This particular type of bifurcation is variously called a *simple turning point*, a *fold* or a *saddle-node bifurcation*. Fig.1.2 is the *bifurcation diagram*.

Example 1.6.2 The system with equation

$$\dot{x}(t) = x\{(a + c^2) - (x - c)^2\} \tag{1.126}$$

has two parameters a and c.

The equilibrium points are $x = 0$ and $x = x^* = c \pm \sqrt{a + c^2}$, which exist when $a + c^2 > 0$. Linearizing about $x = 0$ gives

$$x(t) = \mathsf{C}\exp(at) \tag{1.127}$$

The equilibrium point $x = 0$ is asymptotically stable if $a < 0$ and unstable for $a > 0$. Now let $x = x^* + \triangle x$ giving

$$\begin{aligned} \frac{\mathrm{d}\triangle x}{\mathrm{d}t} &= -2\triangle x x^*(x^* - c) \\ &= \mp 2\triangle x\sqrt{a + c^2}\left[c \pm \sqrt{a + c^2}\right]. \end{aligned} \tag{1.128}$$

This has the solution

$$\triangle x = \mathsf{C}\exp\left[\mp 2t\sqrt{a + c^2}\left(c \pm \sqrt{a + c^2}\right)\right]. \tag{1.129}$$

We consider separately the three cases:

<u>$c = 0$.</u>
Both equilibrium points $x^* = \pm\sqrt{a}$ are stable. The bifurcation diagram for this case is shown in Fig.1.3. This is an example of a *supercritical pitchfork bifurcation* with one stable equilibrium point becomes unstable and two new stable solutions emerge each side of it. The similar situation with the stability reversed is a *subcritical pitchfork bifurcation*.

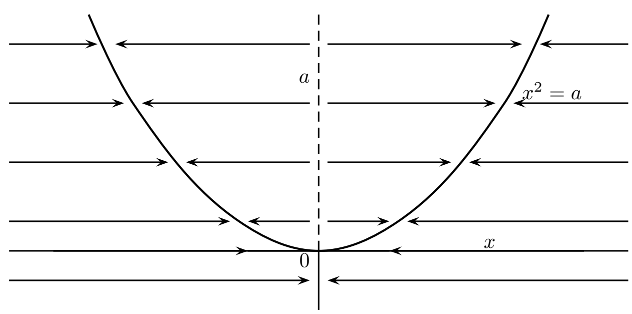


Figure 1.3: The bifurcation diagram for Example 1.6.2, $c = 0$. The stable and unstable equilibrium solutions are shown by continuous and broken lines and the direction of the flow is shown by arrows. This is an example of a supercritical pitchfork bifurcation.

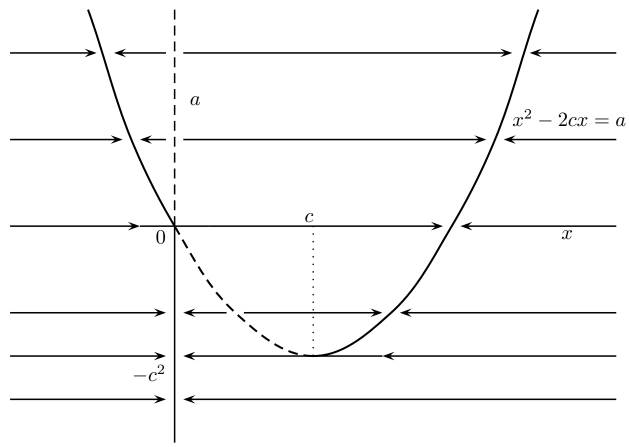


Figure 1.4: The bifurcation diagram for Example 1.6.2, $c > 0$. The stable and unstable equilibrium solutions are shown by continuous and broken lines and the direction of the flow is shown by arrows. This gives examples of both simple turning point and transcritical bifurcations.

$c > 0$.
The equilibrium point $x = c + \sqrt{a + c^2}$ is stable. The equilibrium point $x = c - \sqrt{a + c^2}$ is unstable for $a < 0$ and stable for $a > 0$. The point $x = c$, $a = -c^2$ is a simple turning point bifurcation and $x = a = 0$ is a *transcritical bifurcation*. That is the situation when the stability of two crossing lines of equilibrium points interchange. The bifurcation diagram for this example is shown in Fig.1.4.

$c < 0$.
This is the mirror image (with respect to the vertical axis) of the case $c > 0$.

Example 1.6.3

$$\dot{x}(t) = cx(b - x). \tag{1.130}$$

This is the *logistic equation*.

The equilibrium points are $x = 0$ and $x = b$. Linearizing about $x = 0$ gives

$$x(t) = \mathrm{C} \exp(cbt) \tag{1.131}$$

The equilibrium point $x = 0$ is stable or unstable according as if $cb <, > 0$. Now let $x = b + \triangle x$ giving

$$\frac{\mathrm{d}\triangle x}{\mathrm{d}t} = -cb\triangle x. \tag{1.132}$$

So the equilibrium point $x = b$ is stable or unstable according as $cb >, < 0$. Now plot the equilibrium points with the flow and stability indicated:

- In the (b, x) plane for fixed $c > 0$ and $c < 0$.
- In the (c, x) plane for fixed $b > 0$, $b = 0$ and $b < 0$.

You will see that in the (b, x) plane the bifurcation is easily identified as transcritical but in the (c, x) plane it looks rather different.

Now consider the difference equation corresponding to (1.130). Using the two-point forward derivative,

$$x_{n+1} = x_n[(\varepsilon cb + 1) - c\varepsilon x_n]. \tag{1.133}$$

Now substituting

$$x = \frac{(1 - \varepsilon cb)y + \varepsilon cb}{c\varepsilon} \tag{1.134}$$

into (1.133) gives

$$y_{n+1} = a y_n(1 - y_n), \tag{1.135}$$

where

$$a = 1 - \varepsilon cb. \tag{1.136}$$

(1.135) is the usual form of the *logistic difference equation*. The equilibrium points of (1.135), given by setting $y_{n+1} = y_n = y^*$ are

$$\begin{aligned} y^* &= 0 && \longrightarrow && x^* = b, \\ y^* &= 1 - 1/a && \longrightarrow && x^* = 0. \end{aligned} \tag{1.137}$$

Now linearize (1.135) by setting $y_n = \triangle y_n + y^*$ to give

$$\triangle y_{n+1} = a(1 - 2y^*)\triangle y_n. \tag{1.138}$$

The equilibrium point y^* is stable or unstable according as $|a(1 - 2y^*)| <, > 1$. So

- $y^* = 0$, $(x^* = b)$ is stable if $-1 < a < 1$, $(0 < \varepsilon cb < 2)$.
- $y^* = 1 - 1/a$, $(x^* = 0)$ is stable if $1 < a < 3$, $(-2 < \varepsilon cb < 0)$.

Since the differential equation corresponds to small, positive ε, these stability conditions agree with those derived for the differential equation (1.130). You may know that the whole picture for the behaviour of the difference equation (1.135) involves cycles, period doubling and chaos.[6] Here, however, we are just concerned with the situation for small ε when

$$y \simeq (c\varepsilon)x, \qquad a = 1 - (c\varepsilon)b. \tag{1.139}$$

The whole of the (b, x) plane is mapped into a small rectangle centred around $(1, 0)$ in the (a, y) plane, where a transcritical bifurcation occurs between the equilibrium points $y = 0$ and $y = 1 - 1/a$.

[6] Ian Stewart, *Does God Play Dice?*, Chapter 8, Penguin (1990)

1.6.2 Digression: Matrix Algebra

Before considering systems of more than variable we need to revise our knowledge of matrix algebra. An $n \times n$ matrix $\boldsymbol{A}$ is said to be *singular* or *non-singular* according as the determinant of $\boldsymbol{A}$, denoted by $\mathrm{Det}\{\boldsymbol{A}\}$, is zero or non-zero. The *rank* of any matrix $\boldsymbol{B}$, denoted by $\mathrm{Rank}\{\boldsymbol{B}\}$, is defined, whether the matrix is square or not, as the dimension of the largest non-singular (square) submatrix of $\boldsymbol{B}$. For the $n \times n$ matrix $\boldsymbol{A}$ the following are equivalent:

(i) The matrix $\boldsymbol{A}$ is non-singular.

(ii) The matrix $\boldsymbol{A}$ has an *inverse* denoted by $\boldsymbol{A}^{-1}$.

(iii) $\mathrm{Rank}\{\boldsymbol{A}\} = n$.

(iv) The set of n linear equations

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{c}, \tag{1.140}$$

where

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \qquad \boldsymbol{c} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}, \tag{1.141}$$

has a unique solution for the variables $x_1, x_2, \ldots, x_n$ for any numbers $c_1, c_2, \ldots, c_n$ given by

$$\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{c}. \tag{1.142}$$

(Of course, when $c_1 = c_2 = \cdots = c_n = 0$ the unique solution is the *trivial solution* $x_1 = x_2 = \cdots = x_n = 0$.)

When $\boldsymbol{A}$ is singular we form the $n \times (n+1)$ *augmented matrix* matrix $\boldsymbol{A}'$ by adding the vector $\boldsymbol{c}$ as a final column. Then the following results can be established:

(a) If

$$\mathrm{Rank}\{\boldsymbol{A}\} = \mathrm{Rank}\{\boldsymbol{A}'\} = m < n \tag{1.143}$$

then (1.140) has an infinite number of solutions corresponding to making an arbitrary choice of $n - m$ of the variables $x_1, x_2, \ldots, x_n$.

(b) If

$$\mathrm{Rank}\{\boldsymbol{A}\} < \mathrm{Rank}\{\boldsymbol{A}'\} \le n \tag{1.144}$$

then (1.140) has no solution.

Let $\boldsymbol{A}$ be a non-singular matrix. The *eigenvalues* of $\boldsymbol{A}$ are the roots of the n-degree polynomial

$$\operatorname{Det}\{\boldsymbol{A}-\lambda\boldsymbol{I}\}=0, \tag{1.145}$$

in the variable λ. Suppose that there are n *distinct* roots $\lambda^{(1)},\lambda^{(2)},\dots,\lambda^{(n)}$. Then $\operatorname{Rank}\{\boldsymbol{A}-\lambda^{(k)}\boldsymbol{I}\}=n-1$ for all $k=1,2,\dots,n$. So there is, corresponding to each eigenvalue $\lambda^{(k)}$, a *left eigenvector* $\boldsymbol{v}^{(k)}$ and a *right eigenvector* $\boldsymbol{u}^{(k)}$ which are solutions of the linear equations

$$[\boldsymbol{v}^{(k)}]^{\mathrm{T}}\boldsymbol{A}=\lambda^{(k)}[\boldsymbol{v}^{(k)}]^{\mathrm{T}}, \qquad \boldsymbol{A}\boldsymbol{u}^{(k)}=\boldsymbol{u}^{(k)}\lambda^{(k)}. \tag{1.146}$$

The eigenvectors are unique to within the choice of one arbitrary component. Or equivalently they can be thought of a unique in direction and arbitrary in length. If $\boldsymbol{A}$ is symmetric it is easy to see that the left and right eigenvectors are the same.[7] Now

$$[\boldsymbol{v}^{(k)}]^{\mathrm{T}}\boldsymbol{A}\boldsymbol{u}^{(j)}=\lambda^{(k)}[\boldsymbol{v}^{(k)}]^{\mathrm{T}}\boldsymbol{u}^{(j)}=[\boldsymbol{v}^{k)}]^{\mathrm{T}}\boldsymbol{u}^{(j)}\lambda^{(j)} \tag{1.147}$$

and since $\lambda^{(k)}\neq\lambda^{(j)}$ for $k\neq j$ the vectors $\boldsymbol{v}^{(k)}$ and $\boldsymbol{u}^{(j)}$ are orthogonal. In fact since, as we have seen, eigenvectors can always be multiplied by an arbitrary constant we can ensure that the sets $\{\boldsymbol{u}^{(k)}\}$ and $\{\boldsymbol{v}^{(k)}\}$ are *orthonormal* by dividing each for $\boldsymbol{u}^{(k)}$ and $\boldsymbol{v}^{(k)}$ by $\sqrt{\boldsymbol{u}^{(k)}.\boldsymbol{v}^{(k)}}$ for $k=1,2,\dots,n$. Thus

$$\boldsymbol{u}^{(k)}\cdot\boldsymbol{v}^{(j)}=\delta^{\mathrm{Kr}}(k-j), \tag{1.148}$$

where

$$\delta^{\mathrm{Kr}}(k-j)=\begin{cases}1, & k=j,\\ 0, & k\neq j,\end{cases} \tag{1.149}$$

is called the *Kronecker delta function.*

1.6.3 Linear Autonomous Systems

The n-th order autonomous system (1.111) is linear if

$$\boldsymbol{F}=\boldsymbol{A}\boldsymbol{x}-\boldsymbol{c}, \tag{1.150}$$

for some $n\times n$ matrix $\boldsymbol{A}$ and a vector $\boldsymbol{c}$ of constants. Thus we have

$$\dot{\boldsymbol{x}}(t)=\boldsymbol{A}\boldsymbol{x}(t)-\boldsymbol{c}, \tag{1.151}$$

An equilibrium point $\boldsymbol{x}^*$, if it exists, is a solution of

$$\boldsymbol{A}\boldsymbol{x}=\boldsymbol{c}. \tag{1.152}$$

[7] The vectors referred to in many texts simply as 'eigenvectors' are usually the right eigenvectors. But it should be remembered that non-symmetric matrices have two distinct sets of eigenvectors. The left eigenvectors of $\boldsymbol{A}$ are of course the right eigenvectors of $\boldsymbol{A}^{\mathrm{T}}$ and vice versa.

As we saw in Sect. 1.6.2 these can be either no solutions points, one solution or an infinite number of solutions. We shall concentrate on the case where $\boldsymbol{A}$ is non-singular and there is a unique solution given by

$$\boldsymbol{x}^* = \boldsymbol{A}^{-1}\boldsymbol{c}. \tag{1.153}$$

As in the case of the first-order system we consider a neighbourhood of the equilibrium point by writing

$$\boldsymbol{x} = \boldsymbol{x}^* + \triangle\boldsymbol{x}. \tag{1.154}$$

Substituting into (1.151) and using (1.153) gives

$$\frac{\mathrm{d}\triangle\boldsymbol{x}}{\mathrm{d}t} = \boldsymbol{A}\triangle\boldsymbol{x}. \tag{1.155}$$

Of course, in this case, the 'linearization' used to achieve (1.155) was exact because the original equation (1.151) was itself linear.

As in Sect. 1.6.2 we assume that all the eigenvectors of $\boldsymbol{A}$ are distinct and adopt all the notation for eigenvalues and eigenvectors defined there. The vector $\triangle\boldsymbol{x}$ can be expanded as the linear combination

$$\triangle\boldsymbol{x}(t) = w_1(t)\boldsymbol{u}^{(1)} + w_2(t)\boldsymbol{u}^{(2)} + \cdots + w_n(t)\boldsymbol{u}^{(n)}, \tag{1.156}$$

of the right eigenvectors of $\boldsymbol{A}$, where, from (1.148),

$$w_k(t) = \boldsymbol{v}^{(k)} \cdot \triangle\boldsymbol{x}(t), \qquad k = 1, 2, \ldots, n. \tag{1.157}$$

Now

$$\begin{aligned} \boldsymbol{A}\triangle\boldsymbol{x}(t) &= w_1(t)\boldsymbol{A}\boldsymbol{u}^{(1)} + w_2(t)\boldsymbol{A}\boldsymbol{u}^{(2)} + \cdots + w_n(t)\boldsymbol{A}\boldsymbol{u}^{(n)} \\ &= \lambda^{(1)}w_1(t)\boldsymbol{u}^{(1)} + \lambda^{(2)}w_2(t)\boldsymbol{u}^{(2)} + \cdots + \lambda^{(n)}w_n(t)\boldsymbol{u}^{(n)} \end{aligned} \tag{1.158}$$

and

$$\frac{\mathrm{d}\triangle\boldsymbol{x}}{\mathrm{d}t} = \dot{w}_1(t)\boldsymbol{u}^{(1)} + \dot{w}_2(t)\boldsymbol{u}^{(2)} + \cdots + \dot{w}_n(t)\boldsymbol{u}^{(n)}. \tag{1.159}$$

Substituting from (1.158) and (1.159) into (1.155) and dotting with $\boldsymbol{v}^{(k)}$ gives

$$\dot{w}_k(t) = \lambda^{(k)}w_k(t), \tag{1.160}$$

with solution

$$w_k(t) = \mathrm{C}\exp\left(\lambda^{(k)}t\right). \tag{1.161}$$

So $\triangle\boldsymbol{x}$ will grow or shrink in the direction of $\boldsymbol{u}^{(k)}$ according as $\Re\left\{\lambda^{(k)}\right\} >, < 0$. The equilibrium point will be unstable if at least one eigenvalue has a positive real part and stable otherwise. It will be asymptotically stable if the real part of every eigenvalue is (strictly) negative. Although these conclusions are based on arguments which use both eigenvalues *and* eigenvectors, it can be seen that knowledge simply of the eigenvalues is sufficient to determine stability. The eigenvectors give the directions of attraction and repulsion.

Example 1.6.4 Analyze the stability of the equilibrium points of the linear system

$$\dot{x}(t) = y(t), \qquad \dot{y}(t) = 4x(t) + 3y(t). \tag{1.162}$$

The matrix is

$$\boldsymbol{A} = \begin{pmatrix} 0 & 1 \\ 4 & 3 \end{pmatrix}, \tag{1.163}$$

with $\mathrm{Det}\{\boldsymbol{A}\} = -4$ and the unique equilibrium point is $x = y = 0$. The eigenvalues of $\boldsymbol{A}$ are $\lambda^{(1)} = -1$ and $\lambda^{(2)} = 4$. The equilibrium point is unstable because it is attractive in one direction but repulsive in the other. Such an equilibrium point is called a *saddle-point*.

For a two-variable system the matrix $\boldsymbol{A}$, obtained for a particular equilibrium point, has two eigenvalues $\lambda^{(1)}$ and $\lambda^{(2)}$. Setting aside special cases of zero or equal eigenvalues there are the following possibilities:

(i) $\lambda^{(1)}$ and $\lambda^{(2)}$ both real and (strictly) positive. $\triangle\boldsymbol{x}$ grows in all directions. This is called an *unstable node*.

(ii) $\lambda^{(1)}$ and $\lambda^{(2)}$ both real with $\lambda^{(1)} > 0$ and $\lambda^{(2)} < 0$. $\triangle\boldsymbol{x}$ grows in all directions, apart from that given by the eigenvector associated with $\lambda^{(2)}$. This, as indicated above, is called a *saddle-point*.

(iii) $\lambda^{(1)}$ and $\lambda^{(2)}$ both real and (strictly) negative. $\triangle\boldsymbol{x}$ shrinks in all directions. This is called a *stable node*.

(iv) $\lambda^{(1)}$ and $\lambda^{(2)}$ conjugate complex with $\Re\{\lambda^{(1)}\} = \Re\{\lambda^{(2)}\} > 0$. $\triangle\boldsymbol{x}$ grows in all directions, but by spiraling outward. This is called an *unstable focus*.

(v) $\lambda^{(1)} = -\lambda^{(2)}$ are purely imaginary. Close to the equilibrium point, the length of $\triangle\boldsymbol{x}$ remains approximately constant with the phase point performing a closed loop around the equilibrium point. This is called an *centre*.

(vi) $\lambda^{(1)}$ and $\lambda^{(2)}$ conjugate complex with $\Re\{\lambda^{(1)}\} = \Re\{\lambda^{(2)}\} < 0$. $\triangle\boldsymbol{x}$ shrinks in all directions, but by spiraling inwards. This is called an *stable focus*.

It is not difficult to see that the eigenvalues of the matrix for the equilibrium point $x = y = 0$ of (1.109) are $\frac{1}{2}(1 \pm \mathrm{i}\sqrt{3})$. The point is an *unstable focus* as shown by the MAPLE plot.

Example 1.6.5 Analyze the stability of the equilibrium points of the linear system

$$\dot{x}(t) = 2x(t) - 3y(t) + 4, \qquad \dot{y}(t) = -x(t) + 2y(t) - 1. \tag{1.164}$$

This can be written in the form

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{A}\boldsymbol{x}(t) - \boldsymbol{c}, \tag{1.165}$$

with

$$\boldsymbol{x} = \begin{pmatrix} x \\ y \end{pmatrix}, \qquad \boldsymbol{A} = \begin{pmatrix} 2 & -3 \\ -1 & 2 \end{pmatrix}, \qquad \boldsymbol{c} = \begin{pmatrix} -4 \\ 1 \end{pmatrix}. \tag{1.166}$$

The matrix is

$$\boldsymbol{A} = \begin{pmatrix} 2 & -3 \\ -1 & 2 \end{pmatrix}, \tag{1.167}$$

with $\mathrm{Det}\{\boldsymbol{A}\} = 1$, has inverse

$$\boldsymbol{A}^{-1} = \begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix}. \tag{1.168}$$

So the unique equilibrium point is

$$\boldsymbol{x}^* = \begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} -4 \\ 1 \end{pmatrix} = \begin{pmatrix} -5 \\ -2 \end{pmatrix}. \tag{1.169}$$

Linearizing about $\boldsymbol{x}^*$ gives an equation of the form (1.155). The eigenvalues of $\boldsymbol{A}$ are $2 \pm \sqrt{3}$. Both these numbers are positive so the equilibrium point is an *unstable node*.

1.6.4 Linearizing Non-Linear Systems

Consider now the general autonomous system (1.111) and let there by an equilibrium point given by (1.112). To investigate the stability of $\boldsymbol{x}^*$ we again make the substitution (1.154). Then for a particular member of the set of equations

$$\begin{aligned} \frac{\mathrm{d}\triangle x_\ell}{\mathrm{d}t} &= F_\ell(\boldsymbol{x}^* + \triangle\boldsymbol{x}) \\ &= \sum_{k=1}^{n} \left(\frac{\partial F_\ell}{\partial x_k}\right)^* \triangle x_k + \mathrm{O}(\triangle x_i \triangle x_j), \end{aligned} \tag{1.170}$$

where non-linear contributions in general involve all produces of pairs of the components of $\triangle\boldsymbol{x}$. Neglecting nonlinear contributions and taking all the set of equations gives

$$\frac{\mathrm{d}\triangle\boldsymbol{x}}{\mathrm{d}t} = \boldsymbol{J}^* \triangle\boldsymbol{x}, \tag{1.171}$$

where $\boldsymbol{J}^* = \boldsymbol{J}(\boldsymbol{x})$ is the *stability matrix* with

$$\boldsymbol{J}(\boldsymbol{x}) = \begin{pmatrix} \dfrac{\partial F_1}{\partial x_1} & \dfrac{\partial F_1}{\partial x_2} & \cdots & \dfrac{\partial F_1}{\partial x_m} \\ \dfrac{\partial F_2}{\partial x_1} & \dfrac{\partial F_2}{\partial x_2} & \cdots & \dfrac{\partial F_2}{\partial x_m} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial F_n}{\partial x_1} & \dfrac{\partial F_n}{\partial x_2} & \cdots & \dfrac{\partial F_n}{\partial x_m} \end{pmatrix}. \tag{1.172}$$

Analysis of the stability of the equilibrium point using the eigenvalues of $\boldsymbol{J}^*$ proceeds in exactly the same way as for the linear case. In fact it can be rigorously justified by the following theorem (also due to Lyapunov):

Theorem 1.6.1 *The equilibrium point $\boldsymbol{x}^*$ is asymptotically stable if the real parts of all the eigenvalues of the stability matrix $\boldsymbol{J}^*$ are (strictly) negative. It is unstable if they are all non-zero and at least one is positive.*

It will be see that the case where one or more eigenvalues are zero or purely imaginary is not covered by this theorem (and by linear analysis). This was the case in Example 1.6.1 at $a = 0$, where we needed the quadratic term to determine the stability.

Example 1.6.6 Investigate the stability of the equilibrium point of

$$\dot{x}(t) = \sin[x(t)] - y(t), \qquad \dot{y}(t) = x(t). \tag{1.173}$$

The equilibrium point is $x^* = y^* = 0$. Using the McLaurin expansion of $\sin(x) = \triangle x + \mathrm{O}(\triangle x^3)$ the equations take the form (1.171), where the stability matrix is

$$\boldsymbol{J}^* = \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}. \tag{1.174}$$

This is the same stability matrix as for the linear problem (1.109) and the equilibrium point is an *unstable focus.*

Example 1.6.7

$$\dot{x}(t) = -y(t) + x(t)[a - x^2(t) - y^2(t)], \tag{1.175}$$

$$\dot{y}(t) = x(t) + y(t)[a - x^2(t) - y^2(t)]. \tag{1.176}$$

The only equilibrium point for (1.175)–(1.176) is $x = y = 0$. Linearizing about the equilibrium point gives an equation of the form (1.171) with

$$\boldsymbol{J}^* = \begin{pmatrix} a & -1 \\ 1 & a \end{pmatrix}. \tag{1.177}$$

The eigenvalues of $\boldsymbol{J}^*$ are $a \pm \mathrm{i}$. So the equilibrium point is stable or unstable according as $a < 0$ or $a > 0$. When $a = 0$ the eigenvalues are purely imaginary, so the equilibrium point is a *centre.*

We can find two integrals of (1.175)–(1.176). If (1.175) is multiplied by x and (1.176) by y and the pair are added this gives

$$x\frac{\mathrm{d}x}{\mathrm{d}t} + y\frac{\mathrm{d}y}{\mathrm{d}t} = (x^2 + y^2)(a - x^2 - y^2). \tag{1.178}$$

With $r^2 = x^2 + y^2$, if the trajectory starts with $r = r_0$ when $t = 0$,

$$2\int_0^t \mathrm{d}t = \begin{cases} \dfrac{1}{a}\displaystyle\int_{r_0}^{r} \left\{\frac{1}{a - r^2} + \frac{1}{r^2}\right\} \mathrm{d}(r^2), & a \neq 0, \\ -\displaystyle\int_{r_0}^{r} \frac{1}{r^4}\mathrm{d}(r^2), & a = 0, \end{cases} \tag{1.179}$$

giving

$$r^2(t) = \begin{cases} \dfrac{ar_0^2}{r_0^2 + \exp(-2at)\{a - r_0^2\}}, & a \neq 0, \\[2ex] \dfrac{r_0^2}{1 + 2tr_0^2}, & a = 0. \end{cases} \tag{1.180}$$

This gives

$$r(t) \longrightarrow \begin{cases} 0, & a \leq 0, \\[2ex] \sqrt{a}, & a > 0. \end{cases} \tag{1.181}$$

Now let $x = r\cos(\theta)$, $y = r\sin(\theta)$. Substituting into (1.175)–(1.176) and eliminating $\mathrm{d}r/\mathrm{d}t$ gives

$$\frac{\mathrm{d}\theta}{\mathrm{d}t} = 1. \tag{1.182}$$

If θ starts with the value $\theta(0) = \theta_0$ then

$$\theta = t + \theta_0. \tag{1.183}$$

When $a < 0$ trajectories spiral with a constant angular velocity into the origin.

When $a = 0$ linear analysis indicates that the origin is a centre. However, the full solution shows that orbits converge to the origin as $t \to \infty$, with $r(t) \simeq 1/\sqrt{2t}$, which is a slower rate of convergence than any exponential.

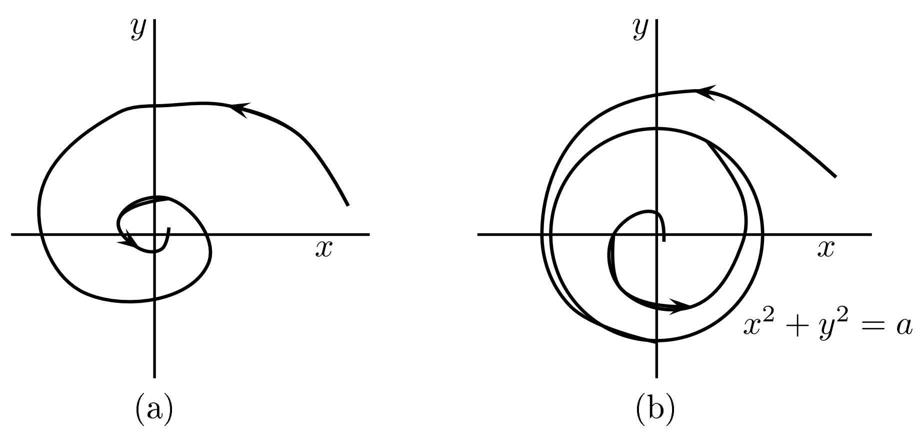


Figure 1.5: A Hopf bifurcation with (a) $a \leq 0$, (b) $a > 0$.

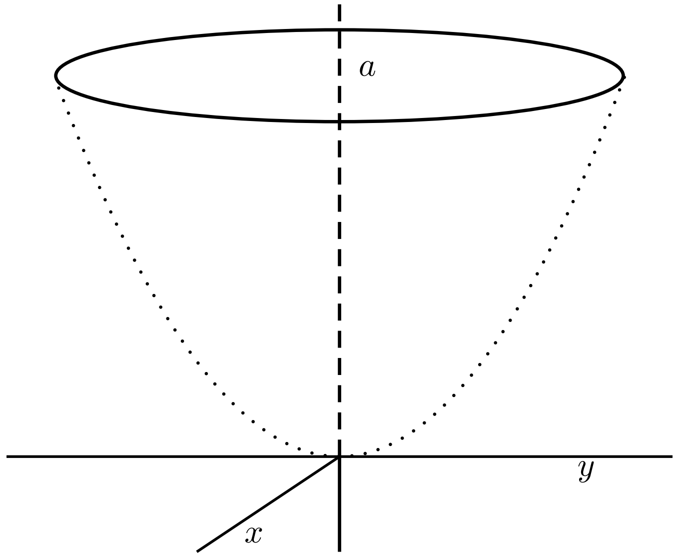


Figure 1.6: A Hopf bifurcation in the space of $\{a, x, y\}$.

When $a > 0$, if $r(0) = r_0 = \sqrt{a}$, $r(t) = \sqrt{a}$. The circle $x^2 + y^2 = a$ is invariant under the evolution of the system. The circle $x^2 + y^2 = a$ is a new type of stable solution called a *limit cycle*. Trajectories spiral, with a constant angular velocity towards the limit cycle circle, either from outside if $r_0 > \sqrt{a}$ or from inside if $r_0 < \sqrt{a}$ see Fig. 1.5. The change over in behaviour at $a = 0$ is an example of the *Hopf bifurcation*. If the behaviour is plotted in the three-dimensional space of $\{a, x, y\}$ then it resembles the supercritical pitchfork bifurcation (Fig. 1.6).

Problems 1

1) Find the general solutions of the following differential equations:

(a) $$t\,\dot{x}(t) = 2x(t)$$

(b) $$\dot{x}(t) = \frac{x(t)}{t} - \tan\left\{\frac{x(t)}{t}\right\}$$

(c) $$2t\,x(t)\,\dot{x}(t) = x^2(t) + t^2$$

(d) $$t\,\dot{x}(t) - t\,x(t) = x(t) + \exp(t)$$

(e) $$(1 - t^2)\dot{x}(t) - t\,x(t) = t$$

[There are each of one of the types described in Sects. 1.3.1-3. The first thing to do is identify the type.]

2) Find the general solutions of the following differential equations:

(a) $\ddot{x}(t) - 5\dot{x}(t) + 6x(t) = 2\exp(t) + 6t - 5$

(b) $\ddot{x}(t) + x(t) = 2\sin(t)$

(c) $\dfrac{\mathrm{d}^3x}{\mathrm{d}t^3} + 2\dfrac{\mathrm{d}^2x}{\mathrm{d}t^2} + 6\dfrac{\mathrm{d}x}{\mathrm{d}t} = 1 + 2\exp(-t)$

[These are all equations with constant coefficients as described in Sect. 1.4.]

3) Find the general solution of the differential equation

$$\ddot{x}(t) - 3\dot{x}(t) + 4x(t) = 0$$

and solve the equation

$$\ddot{x}(t) - 3\dot{x}(t) + 4x(t) = t^2\exp(t)$$

with the initial conditions $x(0) = 0$ and $\dot{x}(0) = 1$.

4) Find out as much as you can about the one-dimensional dynamic systems:

(i) $\dot{x}(t) = x(t)[a - c - ab\,x(t)]$,

(ii) $\dot{x}(t) = a\,x(t) - b\,x^2(t) + c\,x^3(t)$,

You may assume that a and b are non-zero but you can consider the case $c = 0$. You should be able to

(a) Find the equilibrium points and use linear analysis to determine their stability.

(b) Draw the bifurcation diagrams in the $\{x, a\}$–plane for the different ranges of b and c.

(c) Solve the equations explicitly.

5) Determine the nature of the equilibrium point $(0, 0)$ of the systems

(i)
$$\dot{x}(t) = x(t) + 3\,y(t),$$
$$\dot{y}(t) = 3\,x(t) + y(t)\,.$$

(ii)
$$\dot{x}(t) = 3\,x(t) + 2\,y(t),$$
$$\dot{y}(t) = x(t) + 2\,y(t).$$

6) Verify that the system

$$\dot{x}(t) = x(t) + \sin[y(t)],$$
$$\dot{y}(t) = \cos[x(t)] - 2\,y(t) - 1$$

has an equilibrium point at $x = y = 0$ and determine its type.

7) Find *all* the equilibrium points of

$$\dot{x}(t) = -x^2(t) + y(t),$$
$$\dot{y}(t) = 8\,x(t) - y^2(t)$$

and determine their type.

Chapter 2

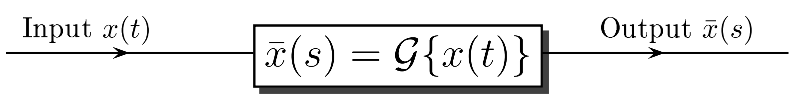
Linear Transform Theory

2.1 Introduction

Consider a function $x(t)$, where the variable t can be regarded as *time*.

Throughout the rest of the course we shall be concerned only with systems driven by autonomous differential (or difference) equations. This means that time is a relative variable and we can set, without loss of generality, the initial time to be $t = 0$.

A linear transform $\mathcal{G}$ is an operation on $x(t)$ to produce a new function $\bar{x}(s)$. It can be pictured as



The *linear property* of the transform is given by

$$\mathcal{G}\{c_1 x_1(t) + c_2 x_2(t)\} = c_1 \bar{x}_1(s) + c_2 \bar{x}_2(s), \tag{2.1}$$

for any functions $x_1(t)$ and $x_2(t)$ and constants c_1 and c_2. The variables t and s can both take a continuous range of variables or one or both of them can take a set of discrete values. In simple cases it is often the practice to use the same letter 't' for both the input and output function variables. Thus the *amplifier*

$$\bar{x}(t) = cx(t), \tag{2.2}$$

differentiator

$$\bar{x}(t) = \dot{x}(t) \tag{2.3}$$

and *integrator*

$$\bar{x}(t) = \int_0^t x(u)\mathrm{d}u + \bar{x}(0) \tag{2.4}$$

are all examples of linear transformations. Now let us examine the case of the transform given by the differential equation

$$\frac{\mathrm{d}\bar{x}(t)}{\mathrm{d}t} + \frac{\bar{x}(t)}{T} = \frac{c}{T}x(t). \tag{2.5}$$

The integrating factor (see Sect. 1.3.3) is $\exp(t/T)$ and

$$\begin{aligned}\exp(t/T)\frac{\mathrm{d}\bar{x}(t)}{\mathrm{d}t} + \exp(t/T)\frac{\bar{x}(t)}{T} &= \frac{\mathrm{d}}{\mathrm{d}t}\left[\exp(t/T)\bar{x}(t)\right] \\ &= \exp(t/T)\frac{c}{T}x(t).\end{aligned} \tag{2.6}$$

This gives

$$\bar{x}(t) = \frac{c}{T}\int_0^t \exp[-(t-u)/T]x(u)\mathrm{d}u + \bar{x}(0)\exp(-t/T). \tag{2.7}$$

In the special case of a constant input $x(t) = 1$, with the initial condition $\bar{x}(0) = 0$,

$$\bar{x}(t) = c[1 - \exp(-t/T)]. \tag{2.8}$$

A well-known example of a linear transform is the *Fourier series transformation*

$$\bar{x}(s) = \frac{1}{2\pi}\int_{-\pi}^{\pi} x(t)\exp(-\mathrm{i}st)\mathrm{d}t \tag{2.9}$$

where $x(t)$ is periodic in t with period 2π and s now takes the discrete values $s = 0, \pm 1, \pm 2, \ldots$. The inverse of this transformation is the 'usual' Fourier series

$$x(t) = \sum_{s=-\infty}^{s=\infty} \bar{x}(s)\exp(\mathrm{i}st). \tag{2.10}$$

A periodic function can be thought of as a superposition of harmonic components $\exp(\mathrm{i}st) = \cos(st) - \mathrm{i}\sin(st)$ and $\bar{x}(s)$, $s = 0, \pm 1, \pm 2, \ldots$ are just the weights or amplitudes of these components.[1]

[1] In the case of a light wave the Fourier series transformation determines the spectrum. Since different elements give off light of different frequencies the spectrum of light from a star can be used to determine the elements present on that star.

2.2 Some Special Functions

2.2.1 The Gamma Function

The *gamma function* $\Gamma(z)$ is defined by

$$\Gamma(z) = \int_0^\infty u^{z-1} \exp(-u) \mathrm{d}u, \qquad \text{for } \Re\{z\} > 0. \tag{2.11}$$

It is not difficult to show that

$$\Gamma(z+1) = z\Gamma(z), \tag{2.12}$$
$$\Gamma(1) = 1. \tag{2.13}$$

So $p! = \Gamma(p+1)$ for any integer $p \geq 0$. The gamma function is a generalization of factorial. Two other results for the gamma function are of importance

$$\Gamma(z)\Gamma(1-z) = \pi \operatorname{cosec}(\pi z), \tag{2.14}$$
$$2^{2z-1}\Gamma(z)\Gamma\left(z+\tfrac{1}{2}\right) = \Gamma(2z)\Gamma\left(\tfrac{1}{2}\right). \tag{2.15}$$

From (2.14), $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$ and this, together with (2.12), gives values for all half-integer values of z.

2.2.2 The Heaviside Function

The *Heaviside function* $\mathcal{H}(t)$ is defined by

$$\mathcal{H}(t) = \begin{cases} 1, & \text{if } t \geq 0, \\ 0, & \text{if } t < 0. \end{cases} \tag{2.16}$$

Clearly, as it stands, the Heaviside function does not have much relevance to us since it is equal to one for all times ($t \geq 0$) which interest us. However, with $t_0 > 0$,

$$\mathcal{H}(t-t_0) = \begin{cases} 1, & \text{if } t \geq t_0, \\ 0, & \text{if } t < t_0, \end{cases} \tag{2.17}$$

is very useful (see Fig. 2.1). It effectively 'turns on and off' the integrand in an integral. Thus, for $0 \leq a < b$,

$$\int_a^b \mathcal{H}(t-t_0)x(t)\mathrm{d}t = \begin{cases} 0, & \text{if } b \leq t_0, \\ \displaystyle\int_{t_0}^b x(t)\mathrm{d}t, & \text{if } b > t_0 \geq a, \\ \displaystyle\int_a^b x(t)\mathrm{d}t, & \text{if } t_0 \leq a. \end{cases} \tag{2.18}$$

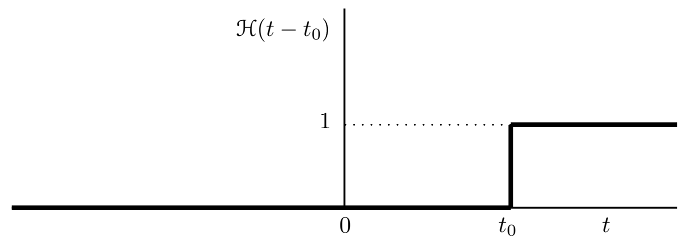


Figure 2.1: The Heaviside function $\mathcal{H}(t - t_0)$.

2.2.3 The Dirac Delta Function

The *Dirac Delta function* $\delta^{\mathrm{D}}(t)$ is defined by

$$\delta^{\mathrm{D}}(t) = \frac{\mathrm{d}\mathcal{H}(t)}{\mathrm{d}t}. \tag{2.19}$$

This function is clearly zero everywhere apart from $t = 0$. At this point it is strictly speaking undefined. However, it can be thought of as being infinite at that single point[2] leading to it often being called the *impulse function*. In spite of its peculiar nature the Dirac delta function plays an easily understood role as part of an integrand.

$$\begin{aligned}
\int_a^b \delta^{\mathrm{D}}(t - t_0)x(t)\mathrm{d}t &= \int_a^b \frac{\mathrm{d}\mathcal{H}(t-t_0)}{\mathrm{d}t} x(t)\mathrm{d}t \\
&= \Big[\mathcal{H}(t - t_0)x(t)\Big]_a^b - \int_a^b \mathcal{H}(t-t_0)\frac{\mathrm{d}x(t)}{\mathrm{d}t}\mathrm{d}t \\
&= \begin{cases} x(t_0), & \text{if } a \leq t_0 < b, \\ 0, & \text{otherwise.} \end{cases} \qquad (2.20)
\end{aligned}$$

The Dirac delta function selects a number of values from an integrand. Thus for example if $a < 0$ and $b > p$, for some positive integer p.

$$\int_a^b x(t) \sum_{j=0}^{p} \delta^{\mathrm{D}}(t - j)\mathrm{d}t = \sum_{j=0}^{p} x(j). \tag{2.21}$$

[2]More rigorous definitions can be given for this delta function. It can, for example, be defined using the limit of a normal distribution curve as the width contacts and the height increases, while maintaining the area under the curve.

2.3 Laplace Transforms

A particular example of a linear transform[3] is the *Laplace transform* defined by

$$\bar{x}(s) = \mathcal{L}\{x(t)\} = \int_0^\infty x(t)\exp(-st)\mathrm{d}t. \tag{2.22}$$

In this case s is taken to be a complex variable with $\Re\{s\} > \eta$, where η is sufficiently large to ensure the convergence of the integral for the particular function $x(t)$. For later use we record the information that the inverse of the Laplace transform is given by

$$x(t) = \frac{1}{2\pi\mathrm{i}}\int_{\alpha-\mathrm{i}\infty}^{\alpha+\mathrm{i}\infty} \bar{x}(s)\exp(st)\mathrm{d}s, \tag{2.23}$$

where $\alpha > \eta$ and the integral is along the vertical line $\Re\{s\} = \alpha$ in the complex s-plane.

It is clear that the Laplace transform satisfies the linear property (2.1). That is

$$\mathcal{L}\{c_1x_1(t) + c_2x_2(t)\} = c_1\bar{x}_1(s) + c_2\bar{x}_2(s). \tag{2.24}$$

It is also not difficult to show that

$$\mathcal{L}\{x(c\,t)\} = \frac{1}{c}\bar{x}\left(\frac{s}{c}\right). \tag{2.25}$$

We now determine the transforms of some particular function and then derive some more general properties.

2.3.1 Some Particular Transforms

A constant C.

$$\mathcal{L}\{\mathsf{C}\} = \int_0^\infty \mathsf{C}\exp(-st)\mathrm{d}t = \frac{\mathsf{C}}{s}, \qquad \Re\{s\} > 0. \tag{2.26}$$

A monomial t^p, where $p \geq 0$ is an integer. To establish this result use integration by parts

$$\begin{aligned}
\mathcal{L}\{t^p\} &= \int_0^\infty t^p\exp(-st)\mathrm{d}t \\
&= \left[-\frac{t^p\exp(-st)}{s}\right]_0^\infty + \frac{p}{s}\int_0^\infty t^{p-1}\exp(-st)\mathrm{d}t \\
&= \frac{p}{s}\int_0^\infty t^{p-1}\exp(-st)\mathrm{d}t.
\end{aligned} \tag{2.27}$$

[3]Henceforth, unless otherwise stated, we shall use $\bar{x}(s)$ to mean the Laplace transform of $x(t)$.

From (2.27) and the result (2.26) for $p = 0$, it follows by induction that

$$\mathcal{L}\{t^p\} = \frac{p!}{s^{p+1}}, \qquad \Re\{s\} > 0. \tag{2.28}$$

It is of some interest to note that this result can be generalized to t^ν, for complex ν with $\Re\{\nu\} \geq 0$. Now

$$\mathcal{L}\{t^\nu\} = \int_0^\infty t^\nu \exp(-st)\mathrm{d}t \tag{2.29}$$

and making the change of variable $st = u$

$$\mathcal{L}\{t^\nu\} = \frac{1}{s^{\nu+1}} \int_0^\infty u^\nu \exp(-u)\mathrm{d}u = \frac{\Gamma(\nu+1)}{s^{\nu+1}}, \qquad \Re\{s\} > 0, \tag{2.30}$$

where the gamma function is defined by (2.11). In fact the result is valid for all complex ν apart from at the singularities $\nu = -1, -2, \ldots$ of $\Gamma(\nu + 1)$.

The exponential function $\exp(-\alpha t)$**.**

$$\begin{aligned} \mathcal{L}\{\exp(-\alpha t)\} &= \int_0^\infty \exp[-(s+\alpha)t]\mathrm{d}t \\ &= \frac{1}{s+\alpha}, \qquad \Re\{s\} > -\Re\{\alpha\}. \end{aligned} \tag{2.31}$$

This result can now be used to obtain the Laplace transforms of the hyperbolic functions

$$\begin{aligned} \mathcal{L}\{\cosh(\alpha t)\} &= \tfrac{1}{2}\left[\mathcal{L}\{\exp(\alpha t)\} + \mathcal{L}\{\exp(-\alpha t)\}\right] \\ &= \tfrac{1}{2}\left[\frac{1}{s-\alpha} + \frac{1}{s+\alpha}\right] \\ &= \frac{s}{s^2-\alpha^2}, \qquad \Re\{s\} > |\Re\{\alpha\}|, \end{aligned} \tag{2.32}$$

and in a similar way

$$\mathcal{L}\{\sinh(\alpha t)\} = \frac{\alpha}{s^2-\alpha^2}, \qquad \Re\{s\} > |\Re\{\alpha\}|, \tag{2.33}$$

Formulae (2.32) and (2.33) can then be used to obtain the Laplace transforms of the harmonic functions

$$\mathcal{L}\{\cos(\omega t)\} = \mathcal{L}\{\cosh(\mathrm{i}\omega t)\} = \frac{s}{s^2+\omega^2}, \qquad \Re\{s\} > |\Im\{\omega\}|, \tag{2.34}$$

$$\mathcal{L}\{\sin(\omega t)\} = -\mathrm{i}\mathcal{L}\{\sinh(\mathrm{i}\omega t)\} = \frac{\omega}{s^2+\omega^2}, \qquad \Re\{s\} > |\Im\{\omega\}|. \tag{2.35}$$

The Heaviside function $\mathcal{H}(t-t_0)$ with $t_0>0$.

$$\begin{aligned}\mathcal{L}\{\mathcal{H}(t-t_0)\} &= \int_0^\infty \mathcal{H}(t-t_0)\exp(-st)\mathrm{d}t \\ &= \int_{t_0}^\infty \exp(-st)\mathrm{d}t = \frac{\exp(-st_0)}{s}, \qquad \Re\{s\}>0. \qquad (2.36)\end{aligned}$$

The Dirac delta function $\delta^{\mathrm{D}}(t-t_0)$ with $t_0>0$.

$$\mathcal{L}\{\delta^{\mathrm{D}}(t-t_0)\} = \exp(-st_0), \qquad \Re\{s\}>0. \qquad (2.37)$$

2.3.2 Some General Properties

The shift theorem.

$$\mathcal{L}\{\exp(-\alpha t)x(t)\} = \int_0^\infty x(t)\exp[-(s+\alpha)t]\mathrm{d}t \qquad (2.38)$$

$$= \bar{x}(s+\alpha), \qquad (2.39)$$

as long as $\Re\{s+\alpha\}$ is large enough to achieve the convergence of the integral. It is clear that (2.31) is a special case of this result with $x(t)=1$ and, from (2.26) $\bar{x}(s)=1/s$.

Derivatives of the Laplace transform. From (2.22)

$$\frac{\mathrm{d}^p\bar{x}(s)}{\mathrm{d}s^p} = \int_0^\infty x(t)\frac{\mathrm{d}^p\exp(-st)}{\mathrm{d}s^p}\mathrm{d}t \qquad (2.40)$$

$$= (-1)^p\int_0^\infty t^p x(t)\exp(-st)\mathrm{d}t, \qquad (2.41)$$

as long as the function is such as to allow differentiation under the integral sign. In these circumstances we have, therefore,

$$\mathcal{L}\{t^p x(t)\} = (-1)^p\frac{\mathrm{d}^p\mathcal{L}\{x(t)\}}{\mathrm{d}s^p}, \qquad \text{for integer } p\ge 0. \qquad (2.42)$$

It is clear the (2.28) is the special case of this result with $x(t)=1$.

The Laplace transform of the derivatives of a function.

$$\begin{aligned}\mathcal{L}\left\{\frac{\mathrm{d}^p x(t)}{\mathrm{d}t^p}\right\} &= \int_0^\infty \frac{\mathrm{d}^p x(t)}{\mathrm{d}t^p}\exp(-st)\mathrm{d}t \\ &= \left[\frac{\mathrm{d}^{p-1}x(t)}{\mathrm{d}t^{p-1}}\exp(-st)\right]_0^\infty + s\int_0^\infty \frac{\mathrm{d}^{p-1}x(t)}{\mathrm{d}t^{p-1}}\exp(-st)\mathrm{d}t \\ &= -\left(\frac{\mathrm{d}^{p-1}x(t)}{\mathrm{d}t^{p-1}}\right)_{t=0} + s\mathcal{L}\left\{\frac{\mathrm{d}^{p-1}x(t)}{\mathrm{d}t^{p-1}}\right\}. \qquad (2.43)\end{aligned}$$

It then follows by induction that

$$\mathcal{L}\left\{\frac{\mathrm{d}^p x(t)}{\mathrm{d}t^p}\right\} = s^p\mathcal{L}\{x(t)\} - \sum_{j=0}^{p-1} s^{p-j-1}\left(\frac{\mathrm{d}^j x(t)}{\mathrm{d}t^j}\right)_{t=0}. \tag{2.44}$$

The product of a function with the Heaviside function. With $t_0 > 0$,

$$\begin{aligned}\mathcal{L}\{x(t)\mathcal{H}(t-t_0)\} &= \int_0^\infty x(t)\mathcal{H}(t-t_0)\exp(-st)\mathrm{d}t \\ &= \int_{t_0}^\infty x(t)\exp(-st)\mathrm{d}t. \end{aligned} \tag{2.45}$$

Making the change of variable $u = t - t_0$,

$$\mathcal{L}\{x(t)\mathcal{H}(t-t_0)\} = \exp(-st_0)\mathcal{L}\{x(t+t_0)\}. \tag{2.46}$$

The product of a function with the Dirac delta function. With $t_0 > 0$,

$$\mathcal{L}\{x(t)\delta^{\mathrm{D}}(t-t_0)\} = x(t_0)\exp(-st_0), \tag{2.47}$$

The Laplace transform of a convolution. The integral

$$\int_0^t x(u)y(t-u)\mathrm{d}u \tag{2.48}$$

is called the *convolution of* $x(t)$ *and* $y(t)$. It is not difficult to see that

$$\int_0^t x(u)y(t-u)\mathrm{d}u = \int_0^t y(u)x(t-u)\mathrm{d}u. \tag{2.49}$$

So the convolution of two functions is independent of the order of the functions.

$$\mathcal{L}\left\{\int_0^t x(u)y(t-u)\mathrm{d}u\right\} = \int_0^\infty \mathrm{d}t \int_0^t \mathrm{d}u x(u)y(t-u)\exp(-st). \tag{2.50}$$

Now define

$$\mathcal{I}_\lambda(s) = \int_0^\lambda \mathrm{d}t \int_0^t \mathrm{d}u\, x(u)y(t-u)\exp(-st). \tag{2.51}$$

The region of integration is shown in Fig. 2.2. Suppose now that the functions are such that we can reverse the order of integration[4] Then (2.51) becomes

$$\mathcal{I}_\lambda(s) = \int_0^\lambda \mathrm{d}u \int_u^\lambda \mathrm{d}t\, x(u)y(t-u)\exp(-st). \tag{2.52}$$

[4]To do this it is sufficient that $x(t)$ and $y(t)$ are piecewise continuous.

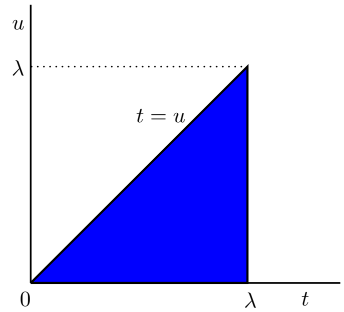


Figure 2.2: The region of integration (*shaded*) for $\mathcal{I}_\lambda(s)$.

Now make the change of variable $t = u + v$. Equation (2.52) becomes

$$\mathcal{I}_\lambda(s) = \int_0^\lambda \mathrm{d}u\, x(u) \exp(-su) \int_0^{\lambda-u} \mathrm{d}v\, y(v) \exp(-sv). \tag{2.53}$$

Now take the limit $\lambda \to \infty$ and, given that $x(t)$, $y(t)$ and s are such that the integrals converge it follows from (2.50), (2.51) and (2.53) that

$$\mathcal{L}\left\{\int_0^t x(u)y(t-u)\mathrm{d}u\right\} = \bar{x}(s)\bar{y}(s). \tag{2.54}$$

A special case of this result is when $y(t) = 1$ giving $\bar{y}(s) = 1/s$ and

$$\mathcal{L}\left\{\int_0^t x(u)\mathrm{d}u\right\} = \frac{\bar{x}(s)}{s}. \tag{2.55}$$

The results of Sects. 2.3.1 and 2.3.2 are summarized in Table 2.1

2.3.3 Using Laplace Transforms to Solve Differential Equations

Equation (2.44) suggests a method for solving differential equations by turning them into algebraic equations in s. For this method to be effective we need to be able, not only to solve the transformed equation for $\bar{x}(s)$, but to invert the Laplace transform to obtain $x(t)$. In simple cases this last step will be achieved reading Table 2.1 from right to left. In more complicated cases it will be necessary to apply the inversion formula (2.23), which often requires a knowledge of contour integration in the complex plane. We first consider a simple example.

Example 2.3.1 Consider the differential equation

$$\ddot{x}(t) + 2\xi\omega\, \dot{x}(t) + \omega^2\, x(t) = 0. \tag{2.56}$$

Table 2.1: **Table of particular Laplace transforms and their general properties.**

$\mathrm{C}t^p$	$\dfrac{\mathrm{C}p!}{s^{p+1}}$	$p \geq 0$ an integer, $\Re\{s\} > 0$.
t^ν	$\dfrac{\Gamma(\nu+1)}{s^{\nu+1}}$	$\nu \neq -1, -2, -3, \ldots$, $\Re\{s\} > 0$.
$\exp(-\alpha t)$	$\dfrac{1}{s+\alpha}$	$\Re\{s\} > \Re\{\alpha\}$.
$\cosh(\alpha t)$	$\dfrac{s}{s^2-\alpha^2}$	$\Re\{s\} > \lvert\Re\{\alpha\}\rvert$.
$\sinh(\alpha t)$	$\dfrac{\alpha}{s^2-\alpha^2}$	$\Re\{s\} > \lvert\Re\{\alpha\}\rvert$.
$\cos(\omega t)$	$\dfrac{s}{s^2+\omega^2}$	$\Re\{s\} > \lvert\Im\{\omega\}\rvert$.
$\sin(\omega t)$	$\dfrac{\omega}{s^2+\omega^2}$	$\Re\{s\} > \lvert\Im\{\omega\}\rvert$.
$c_1x_1(t) + c_2x_2(t)$	$c_1\bar{x}_1(s) + c_2\bar{x}_2(s)$	The linear property.
$x(c\,t)$	$(1/c)\bar{x}(s/c)$	
$\exp(-\alpha t)x(t)$	$\bar{x}(s+\alpha)$	The shift theorem.
$t^px(t)$	$(-1)^p\dfrac{\mathrm{d}^p\bar{x}(s)}{\mathrm{d}s^p}$	$p \geq 0$ an integer.
$\dfrac{\mathrm{d}^px(t)}{\mathrm{d}t^p}$	$s^p\bar{x}(s) - \displaystyle\sum_{j=0}^{p-1} s^{p-j-1}\left(\dfrac{\mathrm{d}^jx(t)}{\mathrm{d}t^j}\right)_{t=0}$	$p \geq 0$ an integer.
$x(t)\mathcal{H}(t-t_0)$	$\exp(-st_0)\bar{x}_1(s)$	$t_0 > 0$, where $x_1(t) = x(t+t_0)$.
$x(t)\delta^{\mathrm{D}}(t-t_0)$	$x(t_0)\exp(-st_0)$	$t_0 > 0$
$\displaystyle\int_0^t x(u)y(t-u)\mathrm{d}u$	$\bar{x}(s)\bar{y}(s)$	The convolution integral.

This is the case of a particle of unit mass moving on a line with simple harmonic oscillations of angular frequency ω, in a medium of viscosity $\xi\omega$. Suppose that $x(0) = x_0$ and $\dot{x}(0) = 0$. Then from Table 2.1 line 12

$$\begin{aligned}\mathcal{L}\{\ddot{x}(t)\} &= s^2\bar{x}(s) - sx_0,\\ \mathcal{L}\{\dot{x}(t)\} &= s\bar{x}(s) - x_0.\end{aligned} \tag{2.57}$$

So the Laplace transform of the whole of (2.56) is

$$\bar{x}(s)[s^2 + 2\xi\omega s + \omega^2] = x_0(s + 2\xi\omega). \tag{2.58}$$

Giving

$$\bar{x}(s) = \frac{x_0(s + 2\xi\omega)}{(s+\xi\omega)^2 + \omega^2(1-\xi^2)}. \tag{2.59}$$

To find the required solution we must invert the transform. Suppose that $\xi^2 < 1$ and let $\theta^2 = \omega^2(1-\xi^2)$.[5] Then (2.59) can be re-expressed in the form

$$\bar{x}(s) = x_0\left[\frac{s+\xi\omega}{(s+\xi\omega)^2+\theta^2} + \frac{\xi\omega}{(s+\xi\omega)^2+\theta^2}\right]. \tag{2.60}$$

Using Table 2.1 lines 6, 7 and 10 to invert these transforms gives

$$x(t) = x_0\exp(-\xi\omega t)\left[\cos(\theta t) + \frac{\xi\omega}{\theta}\sin(\theta t)\right]. \tag{2.61}$$

Let $\zeta = \xi\omega/\theta$ and defined ϕ such that $\tan(\phi) = \zeta$. Then (2.61) can be expressed in the form

$$x(t) = x_0\sqrt{1+\zeta^2}\exp(-\xi\omega t)\cos(\theta t - \phi). \tag{2.62}$$

This is a periodic solution with angular frequency θ subject to exponential damping. We can use MAPLE to plot $x(t)$ for particular values of ω, ξ and x_0:

```
> theta:=(omega,xi)->omega*sqrt(1-xi^2);
```

$$\theta := (\omega, \xi) \to \omega\sqrt{1-\xi^2}$$

```
> zeta:=(omega,xi)->xi*omega/theta(omega,xi);
```

$$\zeta := (\omega, \xi) \to \frac{\xi\,\omega}{\theta(\omega, \xi)}$$

```
> phi:=(omega,xi)->arcsin(zeta(omega,xi));
```

$$\phi := (\omega, \xi) \to \arcsin(\zeta(\omega, \xi))$$

[5] The case of a strong viscosity is included by taking θ imaginary.

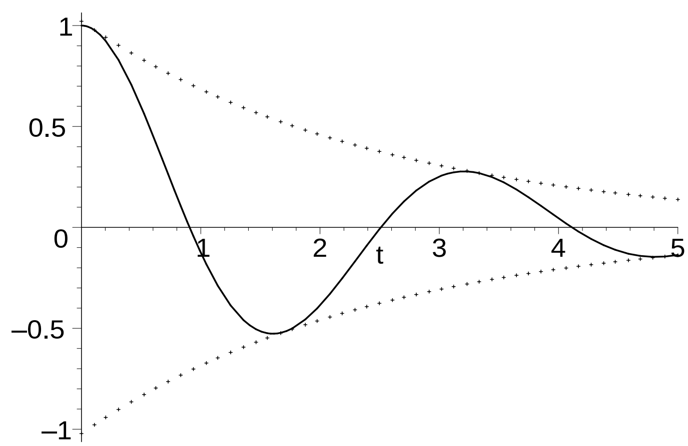
```
> y:=(t,omega,xi,x0)->x0*exp(-xi*omega*t)/sqrt(1-(zeta(omega,xi))^2);#
```

$$y := (t,\, \omega,\, \xi,\, x0) \to \frac{x0\; e^{(-\xi\,\omega\,t)}}{\sqrt{1-\zeta(\omega,\,\xi)^2}}$$

```
> x:=(t,omega,xi,x0)->y(t,omega,xi,x0)*cos(theta(omega,xi)*t-phi(omega,xi));
```

$$x := (t,\, \omega,\, \xi,\, x0) \to \mathrm{y}(t,\, \omega,\, \xi,\, x0)\cos(\theta(\omega,\, \xi)\, t - \phi(\omega,\, \xi))$$

```
> plot(
> {y(t,2,0.2,1),-y(t,2,0.2,1),x(t,2,0.2,1)},t=0..5,style=[point,point,line]);
```



Suppose that (2.56) is modified to

$$\ddot{x}(t) + 2\xi\,\omega\dot{x}(t) + \omega^2\,x(t) = f(t). \tag{2.63}$$

In physical terms the function $f(t)$ is a *forcing term* imposed on the behaviour of the oscillator. As we saw in Sect. 1.4, the general solution of (2.63) consists of the general solution of (2.56) (now called the *complementary function*) together with a particular solution of (2.63). The Laplace transform of (2.63) with the initial conditions $x(0) = x_0$ and $\dot{x}(0) = 0$ is

$$\bar{x}(s)[s^2 + 2\xi\omega s + \omega^2] = x_0(s + 2\xi\omega) + \bar{f}(s). \tag{2.64}$$

Giving

$$\bar{x}(s) = \frac{x_0(s+2\xi\omega)}{(s+\xi\omega)^2 + \omega^2(1-\xi^2)} + \frac{\bar{f}(s)}{(s+\xi\omega)^2 + \omega^2(1-\xi^2)}. \tag{2.65}$$

Comparing with (2.59), we see that the solution of (2.63) consists of the sum of the solution (2.62) of (2.56) and the inverse Laplace transform of

$$\bar{x}_{\mathrm{p}}(s) = \frac{\bar{f}(s)}{(s+\xi\omega)^2+\theta^2}. \tag{2.66}$$

From Table 2.1 lines 7, 10 and 15

$$x_{\mathrm{p}}(t) = \frac{1}{\theta}\int_0^t f(t-u)\exp(-\xi\omega u)\sin(\theta u)\mathrm{d}u. \tag{2.67}$$

So for a particular $f(t)$ we can complete the problem by solving this integral. However, this will not necessarily be the simplest approach. There are two other possibilities:

(i) Decompose $\bar{x}_{\mathrm{p}}(s)$ into a set of terms which can be individually inverse-transformed using the lines of Table 2.1 read from right to left.

(ii) Use the integral formula (2.23) for the inverse transform.

If you are familiar with the techniques of *contour integration* (ii) is often the most straightforward method. We have already used method (i) to derive (2.60) from (2.59). In more complicated cases it often involves the use of partial fractions. As an illustration of the method suppose that

$$f(t) = \mathsf{F}\exp(-\alpha t), \tag{2.68}$$

for some constant F. Then, from Table 2.1,

$$\bar{x}_{\mathrm{p}}(s) = \frac{\mathsf{F}}{(s+\alpha)[(s+\xi\omega)^2+\theta^2]}. \tag{2.69}$$

Suppose now that (2.69) is decomposed into

$$\bar{x}_{\mathrm{p}}(s) = \frac{\mathsf{A}}{(s+\alpha)} + \frac{\mathsf{B}(s+\xi\omega)+\mathsf{C}\,\theta}{(s+\xi\omega)^2+\theta^2}. \tag{2.70}$$

Then

$$x_{\mathrm{p}}(t) = \mathsf{A}\exp(-\alpha t) + \exp(-\xi\omega t)[\mathsf{B}\cos(\theta t) + \mathsf{C}\sin(\theta t)]. \tag{2.71}$$

It then remains only to determine A, B and C. This is done by recombining the terms of (2.70) into one quotient and equating the numerator with that of (2.69). This gives

$$\begin{aligned} s^2(\mathsf{A}+\mathsf{B}) + s[2\mathsf{A}\,\xi\omega + \mathsf{B}(\xi\omega+\alpha) + \mathsf{C}\,\theta]& \\ + \mathsf{A}\,\theta^2 + \mathsf{B}\xi\omega\alpha + \mathsf{C}\,\theta\alpha + \mathsf{A}\xi^2\omega^2 &= \mathsf{F}. \end{aligned} \tag{2.72}$$

Equating powers of s gives

$$\begin{aligned} \mathsf{A} = -\mathsf{B} &= -\frac{\mathsf{F}}{2\xi\omega\alpha-\alpha^2-\theta^2-\xi^2\omega^2}, \\ \mathsf{C} &= \frac{\mathsf{F}(\xi\omega-\alpha)}{\theta(2\xi\omega\alpha-\alpha^2-\theta^2-\xi^2\omega^2)}. \end{aligned} \tag{2.73}$$

In general the Laplace transform of the n-th order equation with constant coefficients (1.40) will be of the form

$$\bar{x}(s)\phi(s) - w(s) = \bar{f}(s). \tag{2.74}$$

where $\phi(s)$ is the polynomial (1.43) and $w(s)$ is some polynomial arising from the application of line 12 of Table 2.1 and the choice of initial conditions. So

$$\bar{x}(s) = \frac{w(s)}{\phi(s)} + \frac{\bar{f}(s)}{\phi(s)}. \tag{2.75}$$

since the coefficients $a_0, a_1, \ldots, a_{n-1}$ are real it has a decomposition of the form

$$\phi(s) = \left\{\prod_{j=1}^{m}(s+\alpha_j)\right\}\left\{\prod_{r=1}^{\ell}[(s+\beta_r)^2+\gamma_r^2]\right\}, \tag{2.76}$$

where all α_j, $j = 1, 2, \ldots, m$ and β_r, γ_r, $r = 1, \ldots, \ell$ are real. The terms in the first product correspond to the real factors of $\phi(s)$ and the terms in the second product correspond to conjugate pairs of complex factors. Thus $m + 2\ell = n$. When all the factors in (2.76) are distinct the method for obtaining the inverse transform of the first term on the right of (2.75) is to express it in the form

$$\frac{w(s)}{\phi(s)} = \sum_{j=1}^{m}\frac{\mathsf{A}_j}{s+\alpha_j} + \sum_{r=1}^{\ell}\frac{\mathsf{B}_r(s+\beta_r)+\mathsf{C}_r}{(s+\beta_r)^2+\gamma_r^2}, \tag{2.77}$$

Recombining the quotients to form the denominator $\phi(s)$ and comparing coefficients of s in the numerator will give all the constants A_j, B_r and C_r. If $\alpha_j = \alpha_{j+1} = \cdots = \alpha_{j+p-1}$, that is, $\phi(s)$ has a real root of degeneracy p then in place of the p terms in the first summation in (2.77) corresponding to these factors we include the terms

$$\sum_{i=1}^{p}\frac{\mathsf{A}_j^{(i)}}{(s+\alpha_j)^i}. \tag{2.78}$$

In a similar way for a p-th fold degenerate complex pair the corresponding term is

$$\sum_{i=1}^{p}\frac{\mathsf{B}_j^{(i)}(s+\beta_j)+\mathsf{C}_j^{(i)}}{[(s+\beta_j)^2+\gamma_j^2]^i}. \tag{2.79}$$

Another, often simpler, way to extract the constants $\mathsf{A}_j^{(i)}$ in (2.78) (and A_j in (2.77) as the special case $p = 1$) is to observe that

$$\begin{aligned}(s+\alpha_j)^p\frac{w(s)}{\phi(s)} &= \sum_{i=1}^{p}(s+\alpha_j)^{p-i}\mathsf{A}_j^{(i)} \\ &\quad + (s+\alpha_j)^p \times [\text{terms not involving } (s+\alpha_j)]. \end{aligned} \tag{2.80}$$

Thus

$$\mathsf{A}_j^{(i)} = \frac{1}{(p-i)!}\left[\frac{\mathrm{d}^{p-i}}{\mathrm{d}s^{p-i}}\left((s+\alpha_j)^p\frac{w(s)}{\phi(s)}\right)\right]_{s=-\alpha_j}, \qquad i=1,..,p \tag{2.81}$$

and in particular, when $p=1$ and $-\alpha_j$ is a simple root of $\phi(s)$,

$$\mathsf{A}_j = \left[(s+\alpha_j)\frac{w(s)}{\phi(s)}\right]_{s=-\alpha_j}. \tag{2.82}$$

Once the constants have been obtained it is straightforward to invert the Laplace transform. Using the shift theorem and the first line of Table 2.1

$$\mathcal{L}^{-1}\left\{\frac{1}{(s+\alpha)^i}\right\} = \frac{\exp(-\alpha t)t^{i-1}}{(i-1)!}. \tag{2.83}$$

This result is also obtainable from line 11 of Table 2.1 and the observation that

$$\frac{1}{(s+\alpha)^i} = \frac{(-1)^{i-1}}{(i-1)!}\frac{\mathrm{d}^{i-1}}{\mathrm{d}s^{i-1}}\left(\frac{1}{s+\alpha}\right). \tag{2.84}$$

The situation is somewhat more complicated for the complex quadratic factors. However, the approach exemplified by (2.84) can still be used.[6] As we saw in the example given above. The second term of the right hand side of (2.75) can be treated in the same way except that now $\bar{f}(s)$ may contribute additional factors in the denominator. Further discussion of Laplace transforms will be in the context of control theory.

2.4 The $\mathcal{Z}$ Transform

Equations (2.9) and (2.10), which define the Fourier series transform, are an example of a transform from a function of a continuous variable to a function of a discrete variable. The $\mathcal{Z}$ transform is similar to this except that we normally think of it as a transformation from a *sequence*, that is a function $x(k)$ of a discrete time variable $k=0,1,2,\ldots$, to its transform $\tilde{x}(z)$, which is a function of the continuous variable z. The definition of the transform is

$$\tilde{x}(z) = \mathcal{Z}\{x(k)\} = x(0) + x(1)z^{-1} + x(2)z^{-2} + \cdots, \tag{2.85}$$

where conditions are applied to z to ensure convergence of the series. Again, for later reference, we record the fact that, as a consequence of Cauchy's theorem, the inverse of the transform is

$$x(k) = \frac{1}{2\mathrm{i}\pi}\oint_C z^{k-1}\tilde{x}(z)\mathrm{d}z, \tag{2.86}$$

[6] We just have to be more careful because differentiation throws up linear terms in s in the numerator.

where the integration is anticlockwise around a simple closed contour C enclosing the singularities of $\tilde{x}(z)$. It is clear that this transform satisfies the linearity property (2.1) since

$$\begin{aligned}\mathcal{Z}\{c_1 x_1(k) + c_2 x_2(k)\} &= [c_1 x(0) + c_2 x_2(0)] + [c_1 x(1) + c_2 x_2(1)]z^{-1} \\ &\quad + [c_1 x(2) + c_2 x_2(2)]z^{-2} + \cdots \\ &= c_1 \tilde{x}_1(z) + c_2 \tilde{x}_2(z). \qquad (2.87)\end{aligned}$$

Just as in the case of the Laplace transform we determine the $\mathcal{Z}$ transform of some particular sequences and derive some more general properties.

2.4.1 Some Particular Transforms

A number of particular transforms can be derived from

$$\mathcal{Z}\{a^k\} = 1 + az^{-1} + a^2 z^{-2} + a^3 z^{-3} + \cdots = \frac{z}{z-a}, \qquad |z| > |a|. \quad (2.88)$$

Thus

$$\begin{aligned}\mathcal{Z}\{\mathrm{C}\} &= \frac{\mathrm{C}\,z}{z-1}, & |z| > 1, \qquad (2.89)\\ \mathcal{Z}\{\exp(-\alpha k)\} &= \frac{z}{z-\exp(-\alpha)}, & |z| > \exp\left(-\Re\{\alpha\}\right). \qquad (2.90)\end{aligned}$$

Also

$$\begin{aligned}\mathcal{Z}\{\cosh(\alpha k)\} &= \frac{1}{2}\left[\mathcal{Z}\{\exp(\alpha k)\} + \mathcal{Z}\{\exp(-\alpha k)\}\right] \\ &= \frac{z[z-\cosh(\alpha)]}{z^2 - 2z\cosh(\alpha) + 1}, \qquad |z| > \exp\left(|\Re\{\alpha\}|\right) \qquad (2.91)\end{aligned}$$

and in a similar way

$$\mathcal{Z}\{\sinh(\alpha k)\} = \frac{z\sinh(\alpha)}{z^2 - 2z\cosh(\alpha) + 1}, \qquad |z| > \exp\left(|\Re\{\alpha\}|\right), \qquad (2.92)$$

$$\mathcal{Z}\{\cos(\omega k)\} = \frac{z[z-\cos(\omega)]}{z^2 - 2z\cos(\omega) + 1}, \qquad |z| > \exp\left(|\Im\{\omega\}|\right), \qquad (2.93)$$

$$\mathcal{Z}\{\sin(\omega k)\} = \frac{z\sin(\omega)}{z^2 - 2z\cos(\omega) + 1}, \qquad |z| > \exp\left(|\Im\{\omega\}|\right). \qquad (2.94)$$

Other important results can be derived from these using the general properties derived below. The *Kronecker delta function* $\delta^{\mathrm{Kr}}(k)$ is defined in (1.149). With $m \geq 0$, the terms of the sequence $\delta^{\mathrm{Kr}}(k-m)$, $k = 0, 1, \ldots$ are all zero except that for which $k = m$, which is unity. Thus

$$\mathcal{Z}\{\delta^{\mathrm{Kr}}(k-m)\} = \frac{1}{z^m}, \qquad m \geq 0. \qquad (2.95)$$

2.4.2 Some General Properties

For $p > 0$

$$\begin{aligned}\mathcal{Z}\{x(k+p)\} &= x(p) + x(p+1)z^{-1} + x(p+2)z^{-2} + \cdots \\ &= z^p\tilde{x}(z) - \sum_{j=0}^{p-1} x(j)z^{p-j}. \end{aligned} \tag{2.96}$$

If we continue the sequence to negative indices by defining $x(k) = 0$ if $k < 0$ then

$$\begin{aligned}\mathcal{Z}\{x(k-p)\} &= x(0)z^{-p} + x(1)z^{-p-1} + x(2)z^{-p-2} + \cdots \\ &= z^{-p}\tilde{x}(z). \end{aligned} \tag{2.97}$$

As a generalization of (2.88)

$$\begin{aligned}\mathcal{Z}\{a^k x(k)\} &= x(0) + ax(1)z^{-1} + a^2x(2)z^{-2} + \cdots \\ &= x(0) + x(1)(z/a)^{-1} + x(2)(z/a)^{-2} + \cdots \\ &= \tilde{x}(z/a). \end{aligned} \tag{2.98}$$

With the result

$$\begin{aligned}\mathcal{Z}\{kx(k)\} &= x(1)z^{-1} + 2x(2)z^{-2} + 3x(3)z^{-3} + \cdots \\ &= -z\frac{\mathrm{d}\tilde{x}(z)}{\mathrm{d}z}, \end{aligned} \tag{2.99}$$

formulae for any derivative of the $\mathcal{Z}$ transform can be derived. Consider now

$$\tilde{x}(z) = x(0) + x(1)z^{-1} + x(2)z^{-2} + x(3)z^{-3} + \cdots, \tag{2.100}$$

$$\tilde{y}(z) = y(0) + y(1)z^{-1} + y(2)z^{-2} + y(3)z^{-3} + \cdots. \tag{2.101}$$

The coefficient of z^{-k} in the product $\tilde{x}(z)\tilde{y}(z)$ is

$$\begin{aligned}&x(0)y(k) + x(1)y(k-1) + x(2)y(k-2) + \cdots \\ &\qquad + x(k-1)y(1) + x(k)y(0). \end{aligned} \tag{2.102}$$

So

$$\mathcal{Z}\left\{\sum_{j=0}^{k} x(j)y(k-j)\right\} = \tilde{x}(z)\tilde{y}(z). \tag{2.103}$$

This is the $\mathcal{Z}$ transform analogue of the convolution formula (2.55). The results of Sects. 2.4.1 and 2.4.2 are summarized in Table 2.2.

2.4.3 Using $\mathcal{Z}$ Transforms to Solve Difference Equations

This is most easily illustrated by giving some examples

Table 2.2: **Table of particular $\mathscr{Z}$ transforms and their general properties.**

$\mathrm{C}a^k$	$\frac{\mathrm{C}z}{z-a}$	$\|z\| > \|a\|$.
$\exp(-\alpha k)$	$\frac{z}{z-\exp(-\alpha)}$	$\|z\| > \exp(-\Re\{\alpha\})$.
$\cosh(\alpha k)$	$\frac{z[z-\cosh(\alpha)]}{z^2-2z\cosh(\alpha)+1}$	$\|z\| > \exp(\|\Re\{\alpha\}\|)$.
$\sinh(\alpha k)$	$\frac{z\sinh(\alpha)}{z^2-2z\cosh(\alpha)+1}$	$\|z\| > \exp(\|\Re\{\alpha\}\|)$.
$\cos(\omega k)$	$\frac{z[z-\cos(\omega)]}{z^2-2z\cos(\omega)+1}$	$\|z\| > \exp(\|\Im\{\omega\}\|)$.
$\sin(\omega k)$	$\frac{z\sin(\omega)}{z^2-2z\cos(\omega)+1}$	$\|z\| > \exp(\|\Im\{\omega\}\|)$.
$\delta^{\mathrm{Kr}}(k-m)$	$\frac{1}{z^m}$	$m \geq 0$.
$c_1x_1(k)+c_2x_2(k)$	$c_1\tilde{x}_1(z)+c_2\tilde{x}_2(z)$	The linear property.
$a^kx(k)$	$\tilde{x}(z/a)$	
$x(k+p)$	$z^p\tilde{x}(z)-\sum_{j=0}^{p-1}x(j)z^{(p-j)}$	$p>0$ an integer.
$x(k-p)$ $(x(k)=0,\ k<0.)$	$z^{-p}\tilde{x}(z)$	$p \geq 0$ an integer.
$kx(k)$	$-z\frac{\mathrm{d}\tilde{x}(z)}{\mathrm{d}z}$	
$\sum_{j=0}^{k}x(j)y(k-j)$	$\tilde{x}(z)\tilde{y}(z)$	The convolution formula.

Example 2.4.1 Solve the difference equation

$$x(k+2)+5x(k+1)+6x(k)=0, \tag{2.104}$$

subject to the conditions $x(0)=\alpha$, $x(1)=\beta$.

From (2.96)

$$\begin{aligned}\mathcal{Z}\{x(k+2)\} &= z^2\tilde{x}(z)-\alpha z^2-\beta z,\\ \mathcal{Z}\{x(k+1)\} &= z\tilde{x}(z)-\alpha z,\end{aligned} \tag{2.105}$$

So transforming (2.104) gives

$$\tilde{x}(z)[z^2+5z+6]=\alpha z^2+(\beta+5\alpha)z \tag{2.106}$$

and thus

$$\tilde{x}(z)=z\left[\frac{\alpha z+(\beta+5\alpha)}{(z+2)(z+3)}\right] \tag{2.107}$$

giving

$$\tilde{x}(z)=\frac{z(3\alpha+\beta)}{z+2}-\frac{z(2\alpha+\beta)}{z+3}. \tag{2.108}$$

Inverting the transform using line 1 of Table 2.2

$$x(k)=[3\alpha+\beta](-2)^k-[2\alpha+\beta](-3)^k. \tag{2.109}$$

The method can also be used to study systems of difference equations. This is an example based on a simple model for the buffalo population in the American West starting in the year 1830.[7]

Example 2.4.2 Let $x(k)$ and $y(k)$ be the number of female and male buffalo at the start of any one year ($k=0$ is 1830). Five percent of adults die each year. Buffalo reach maturity at two years and the number of new females alive at the beginning of year $k+2$, taking into account infant mortality, is 12% of $x(k)$. More male calves than female are born and the corresponding figure is 14% of $x(k)$. Show that in the limit $k\to\infty$ the population grows by 6.3% per year.

The difference equations are

$$\begin{aligned}x(k+2)&=0.95x(k+1)+0.12x(k),\\ y(k+2)&=0.95y(k+1)+0.14x(k).\end{aligned} \tag{2.110}$$

Applying the $\mathcal{Z}$ transform to these equations and using $x(0)=x_0$, $x(1)=x_1$, $y(0)=y_0$ and $y(1)=y_1$

$$\begin{aligned}z^2\tilde{x}(z)-x_0z^2-x_1z&=0.95[z\tilde{x}(z)-zx_0]+0.12\tilde{x}(z),\\ z^2\tilde{y}(z)-y_0z^2-y_1z&=0.95[z\tilde{y}(z)-zy_0]+0.14\tilde{x}(z)\end{aligned} \tag{2.111}$$

[7] Taken from Barnett and Cameron(1985) p. 21.

Since we are interested only in the long-time behaviour of the total population $p(k) = x(k) + y(k)$ we need extract simply the formula for $\tilde{p}(z)$ from these equations. With $p_0 = x_0 + y_0$, $p_1 = x_1 + y_1$

$$\tilde{p}(z) = z\left[\frac{(p_0 z + p_1 - 0.95p_0)}{z(z-0.95)} + \frac{0.26(x_0 z + x_1 - 0.95x_0)}{z(z-0.95)(z^2-0.95z-0.12)}\right]. \tag{2.112}$$

The reason for retaining the factor z in the numerators (with a consequent z in the denominators) can be seen by looking at the first line of Table 2.2. Factors of the form $z/(z-a)$ are easier to handle than $1/(z-a)$. We now resolve the contents of the brackets into partial fractions. You can if you like use MAPLE to do this. The code is

```
> convert(0.26*(x0*z+x1-0.95*x0)/(z*(z-0.95)*(z^2-0.95*z-0.12)),parfrac,z);
```

$$\frac{1.958625701\,x0 - 1.842720774\,x1}{z + .1128988008} - \frac{-2.280701754\,x1 + 2.166666667\,x0}{z} - \frac{.1058808501\,10^{-8}\,x0 + 2.280701750\,x1}{z - .9500000000} + \frac{.2080409669\,x0 + 1.842720770\,x1}{z - 1.062898801}$$

```
> convert((p0*z+p1-0.95*p0)/(z*(z-0.95)),parfrac,z);
```

$$1.052631579\,\frac{p1}{z - .9500000000} + \frac{-1.052631579\,p1 + p0}{z}$$

Thus we have

$$\begin{aligned}\tilde{p}(z) &= (p_0 - 1.053p_1 - 2.167x_0 + 2.281x_1) + \frac{z(1.053p_1 - 2.281x_1)}{z-0.95}\\ &\quad + \frac{z(1.959x_0 - 1.843x_1)}{z+0.113} + \frac{z(0.208x_0 + 1.843x_1)}{z-1.063}\end{aligned} \tag{2.113}$$

and inverting the transform using lines 1 and 7 of Table 2.2 gives

$$\begin{aligned}p(k) &= (p_0 - 1.053p_1 - 2.167x_0 + 2.281x_1)\delta^{\mathrm{Kr}}(0)\\ &\quad + (1.053p_1 - 2.281x_1)(0.95)^k\\ &\quad + (1.959x_0 - 1.843x_1)(-0.113)^k\\ &\quad + (0.208x_0 + 1.843x_1)(1.063)^k.\end{aligned} \tag{2.114}$$

In the limit of large k this expression is dominated by the last term

$$p(k) \simeq (0.208x_0 + 1.843x_1)(1.063)^k. \tag{2.115}$$

The percentage yearly increase is

$$\frac{p(k+1) - p(k)}{p(k)} \times 100 = 6.3. \tag{2.116}$$

Problems 2

1) Show, using the standard results in the table of Laplace transforms, that:

(i) $\mathcal{L}\{t\sin(\omega t)\} = \dfrac{2\omega s}{(\omega^2+s^2)^2}$,

(ii) $\mathcal{L}\{\sin(\omega t) - \omega t\cos(\omega t)\} = \dfrac{2\omega^3}{(\omega^2+s^2)^2}$.

Hence solve the differential equation

$$\ddot{x}(t) + \omega^2 x(t) = \sin(\omega t),$$

when $x(0) = \dot{x}(0) = 0$.

2) Use Laplace transforms to solve the differential equation

$$\frac{\mathrm{d}^3 x(t)}{\mathrm{d}t^3} + x(t) = 1,$$

where $x(0) = \dot{x}(0) = \ddot{x}(0) = 0$.

3) Given that $x(t) = -t$ and

$$\bar{y}(s) = \frac{\bar{x}(s)}{(s-1)^2},$$

find $y(t)$.

4) Show using your notes that

$$\mathcal{L}\{t^{-\frac{1}{2}}\} = \frac{\sqrt{\pi}}{s^{\frac{1}{2}}}.$$

The *error function* $\mathrm{Erf}(z)$ is defined by

$$\mathrm{Erf}(z) = \frac{2}{\sqrt{\pi}}\int_0^z \exp(-u^2)\mathrm{d}u.$$

Show that

$$\mathcal{L}\left\{\mathrm{Erf}(t^{\frac{1}{2}})\right\} = \frac{1}{s(s+1)^{\frac{1}{2}}}.$$

5) Find the sequences $x(0), x(1), x(2), \ldots$ for which $\mathcal{Z}\{x(k)\} = \tilde{x}(z)$ are:

(i) $\dfrac{z}{(z-1)(z-2)}$,

(ii) $\dfrac{z}{z^2+a^2}$,

(iii) $\dfrac{z^3 + 2z^2 + 1}{z^3}$.

6) Use the $\mathcal{Z}$ transform to solve the following difference equations for $k \geq 0$:

(i) $8x(k+2) - 6x(k+1) + x(k) = 9$, where $x(0) = 1$ and $x(1) = \frac{3}{2}$,

(ii) $x(k+2) + 2x(k) = 0$, where $x(0) = 1$ and $x(1) = \sqrt{2}$.

7) A person's capital at the beginning of year k is $x(k)$ and their expenditure during year k is $y(k)$. Given that these satisfy the difference equations

$$x(k+1) = 1.5x(k) - y(k),$$
$$y(k+1) = 0.21x(k) + 0.5y(k)$$

Show that in the long time limit the person's capital changes at a rate of 20% per annum.

Chapter 3

Transfer Functions and Feedback

3.1 Introduction

Linear control theory deals with a linear time-invariant system having a set of inputs $\{u_1(t), u_2(t), \ldots\}$ and outputs $\{x_1(t), x_2(t), \ldots\}$. The input functions are controlled by the experimenter, that is, they are *known* functions. The aim of control theory is to

(i) Construct a model relating inputs to outputs. (Usually differential equations for continuous time and difference equations for discrete time.) The time invariant nature of the system implies that the equations are autonomous.

(ii) Devise a strategy for choosing the input functions and possibly changing the design of the system (and hence the equations) so that the output have some specific required form. If the aim is to produce outputs as close as possible to some *reference functions* $\{\rho_1(t), \rho_2(t), \ldots\}$ then the system is called a *servomechanism*. If each of the reference functions is constant the system is a *regulator*.

Consider, for example, the simple case of one input function $u(t)$ and one output function $x(t)$ related by the differential equation

$$\frac{\mathrm{d}^n x}{\mathrm{d}t^n} + a_{n-1}\frac{\mathrm{d}^{n-1} x}{\mathrm{d}t^{n-1}} + \cdots + a_1\frac{\mathrm{d}x}{\mathrm{d}t} + a_0 x = b_m\frac{\mathrm{d}^m u}{\mathrm{d}t^m} + b_{m-1}\frac{\mathrm{d}^{m-1} u}{\mathrm{d}t^{m-1}} + \cdots + b_1\frac{\mathrm{d}u}{\mathrm{d}t} + b_0 u, \tag{3.1}$$

and with

$$\begin{aligned} \left(\frac{\mathrm{d}^i x}{\mathrm{d}t^i}\right)_{t=0} &= 0, \qquad i = 0, 1, \ldots, n-1, \\ \left(\frac{\mathrm{d}^j u}{\mathrm{d}t^j}\right)_{t=0} &= 0, \qquad j = 0, 1, \ldots, m-1. \end{aligned} \tag{3.2}$$

Now taking the Laplace transform of (3.1) gives

$$\phi(s)\bar{x}(s) = \psi(s)\bar{u}(s), \tag{3.3}$$

where

$$\begin{aligned} \phi(s) &= s^n + a_{n-1}s^{n-1} + \cdots + a_1 s + a_0, \\ \psi(s) &= b_m s^m + b_{m-1}s^{m-1} + \cdots + b_1 s + b_0. \end{aligned} \tag{3.4}$$

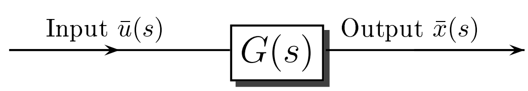
(Cf (1.43).) Equation (3.3) can be written

$$\bar{x}(s) = G(s)\bar{u}(s), \tag{3.5}$$

where

$$G(s) = \frac{\psi(s)}{\phi(s)} \tag{3.6}$$

is called the *transfer function*. This system can be represented in block diagrammatic form as



Three simple examples of transfer functions are:

(i) Proportional control when

$$x(t) = \mathsf{K}\, u(t), \tag{3.7}$$

where K is a constant. This gives

$$G(s) = \mathsf{K}. \tag{3.8}$$

(ii) Integral control when

$$x(t) = \int_0^t \mathsf{K}\, u(\tau)\mathrm{d}\tau, \tag{3.9}$$

where K is a constant. This gives[1]

$$G(s) = \frac{\mathsf{K}}{s}. \tag{3.10}$$

[1]From the last line of Table 2.1 with $y(t) = 1$, $\bar{y}(s) = 1/s$.

(iii) Differential control when

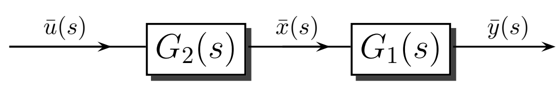
$$x(t) = \mathsf{K}\,\frac{\mathrm{d}u(t)}{\mathrm{d}t}, \tag{3.11}$$

where K is a constant and $u(0) = 0$. This gives

$$G(s) = \mathsf{K}\,s. \tag{3.12}$$

3.2 Systems in Cascade

If the output variable (or variables) are used as control variables for a second system then the two systems are said to be in *cascade*. For the one control function/one output function case the block diagram takes the form



with equations

$$\begin{aligned} \bar{x}(s) &= G_2(s)\bar{u}(s), \\ \bar{y}(s) &= G_1(s)\bar{x}(s), \end{aligned} \tag{3.13}$$

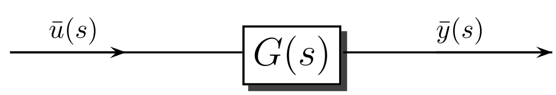
where $G_1(s)$ is called the *process transfer function* and $G_2(s)$ is called the *controller transfer function*. Combining these equations gives

$$\bar{y}(s) = G(s)\bar{u}(s), \tag{3.14}$$

where

$$G(s) = G_1(s)G_2(s) \tag{3.15}$$

and the block diagram is



For many-variable systems the x-variables may be large in number or difficult to handle. The y-variables may then represent a smaller or more easily accessible set. Thus in Example 2.4.2 the two x-variables could be the numbers of males and females. I don't know a lot about buffalo, but in practice it may be difficult to count these and the sum of the two (the whole population) may be easier to count. Thus the second stage would be simply to sum the populations of the two sexes.

Example 3.2.1 We want to construct a model for the temperature control of an oven.

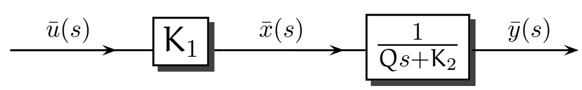
Let the control variable $u(t)$ be the position on the heating dial of the oven and suppose that this is directly proportional to the heat $x(t)$ supplied to the oven by the heat source. This is the situation of proportional control with $x(t) = \mathsf{K}_1 u(t)$. Let the output variable $y(t)$ be the temperature difference at time t between the oven and its surroundings. Some of the heat supplied to the oven will be lost by radiation; this will be proportional to $y(t)$. So the heat used to raise the temperature of the oven is $x(t) - \mathsf{K}_2 y(t)$. According to the laws of thermodynamics this is $\mathsf{Q}\dot{y}(t)$, where Q is the heat capacity of the oven. Then we have

$$\begin{aligned} x(t) &= \mathsf{K}_1 u(t), && (3.16)\\ \mathsf{Q}\dot{y}(t) + \mathsf{K}_2 y(t) &= x(t). && (3.17) \end{aligned}$$

Suppose that the oven is at the temperature of its surroundings at $t = 0$. Then the Laplace transforms of (3.16)–(3.17) are

$$\begin{aligned} \bar{x}(s) &= \mathsf{K}_1 \bar{u}(s), && (3.18)\\ (s\mathsf{Q} + \mathsf{K}_2)\bar{y}(s) &= \bar{x}(s) && (3.19) \end{aligned}$$

with block diagram



From (3.18)–(3.19)

$$\bar{y}(s) = \frac{\mathsf{K}_1}{\mathsf{Q}s + \mathsf{K}_2}\,\bar{u}(s). \tag{3.20}$$

Suppose the dial is turned from zero to a value u_0 at $t = 0$. Then, remembering that we always assume $t \geq 0$, $u(t) = u_0$ and $\bar{u}(s) = u_0/s$. So

$$\bar{y}(s) = \frac{u_0 \mathsf{K}_1}{s(\mathsf{Q}s + \mathsf{K}_2)}. \tag{3.21}$$

Using partial fractions this gives

$$\bar{y}(s) = u_0 \frac{\mathsf{K}_1}{\mathsf{K}_2}\left(\frac{1}{s} - \frac{1}{\mathsf{T}^{-1} + s}\right), \tag{3.22}$$

where $\mathsf{T} = \mathsf{Q}/\mathsf{K}_2$. Inverting the Laplace transform gives

$$y(t) = u_0 \frac{\mathsf{K}_1}{\mathsf{K}_2}[1 - \exp(-t/\mathsf{T})]. \tag{3.23}$$

Suppose now that the proportional control condition (3.16) is replaced by the integral control condition

$$x(t) = \mathsf{K}_1 \int_0^t u(\tau)\mathrm{d}\tau. \tag{3.24}$$

giving in place of (3.18)

$$\bar{x}(s) = \frac{\mathsf{K}_1}{s}\,\bar{u}(s). \tag{3.25}$$

The formula (3.20) is now replaced by

$$\bar{y}(s) = \frac{\mathsf{K}_1}{s(\mathsf{Q}s + \mathsf{K}_2)}\,\bar{u}(s). \tag{3.26}$$

If we now use the form $u(t) = u_0$ this is in fact equivalent to $x(t) = \mathsf{K}_1\, u_0 t$ which implies a linear buildup of heat input over the time interval $[0, t]$. Formula (3.21) is replaced by

$$\bar{y}(s) = \frac{u_0 \mathsf{K}_1}{s^2(\mathsf{Q}s + \mathsf{K}_2)}. \tag{3.27}$$

Resolving into partial fractions gives

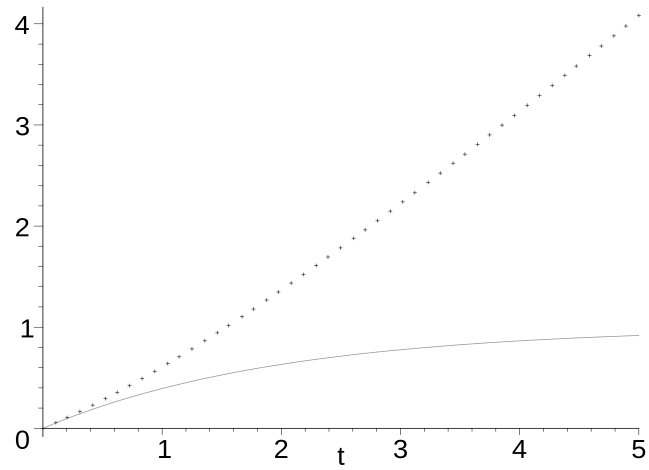
$$\bar{y}(s) = u_0 \frac{\mathsf{K}_1}{\mathsf{K}_2}\left(\frac{1}{s^2} - \frac{\mathsf{T}}{s} + \frac{\mathsf{T}}{\mathsf{T}^{-1} + s}\right), \tag{3.28}$$

where, as before, $\mathsf{T} = \mathsf{Q}/\mathsf{K}_2$. Inverting the Laplace transform gives

$$y(t) = u_0 \frac{\mathsf{T}\mathsf{K}_1}{\mathsf{K}_2}\left[\frac{t}{\mathsf{T}} - 1 + \exp(-t/\mathsf{T})\right]. \tag{3.29}$$

We now use MAPLE to compare the results of (3.23) and (3.29) (with $u_0\mathsf{K}_1/\mathsf{K}_2 = 1$, $\mathsf{T} = 2$).

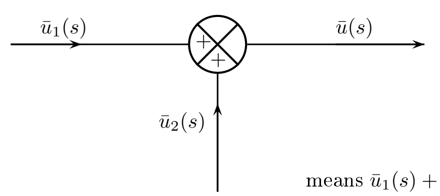
```
> plot({t-2+2*exp(-t/2),1-exp(-t/2)
> },t=0..5,style=[point,line]);
```



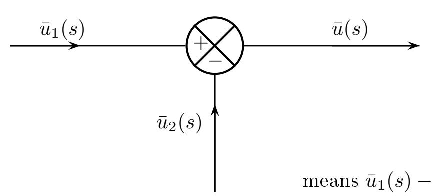
It will be see that with proportional control the temperature of the oven reaches a steady state, whereas (if it were allowed to do so) it would rise steadily with time for the case of integral control.

3.3 Combinations and Distribution of Inputs

In some cases, particularly in relation to feedback, we need to handle sums or differences of inputs. To represent these on block diagrams the following notation is convenient:

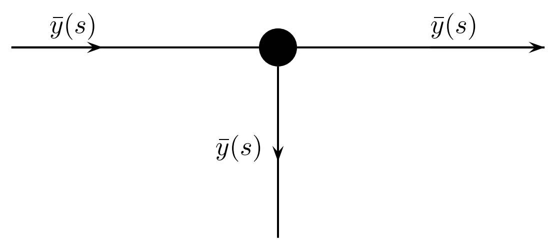


means $\bar{u}_1(s) + \bar{u}_2(s) = \bar{u}(s)$.



means $\bar{u}_1(s) - \bar{u}_2(s) = \bar{u}(s)$.

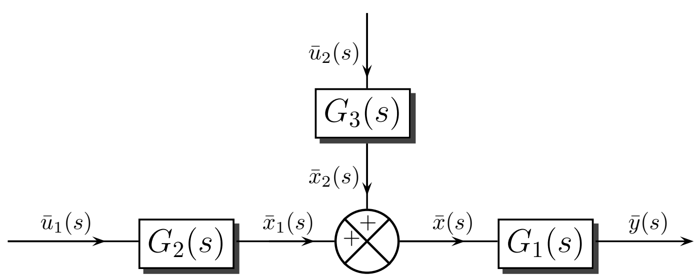
We shall also need to represent a device which receives an input and transmits it unchanged in two (or more) directions. This will be represented by



A simple example of the use of this formalism is the case where equations (3.13) are modified to

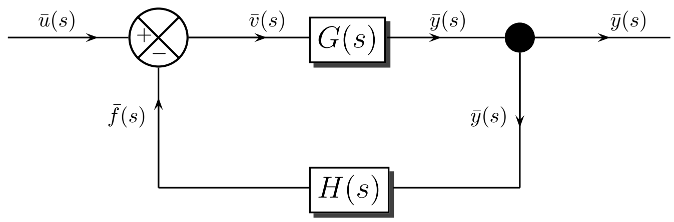
$$
\begin{aligned}
\bar{x}_1(s) &= G_2(s)\bar{u}_1(s), \\
\bar{x}_2(s) &= G_3(s)\bar{u}_2(s), \\
\bar{x}(s) &= \bar{x}_1(s) + \bar{x}_2(s), \\
\bar{y}(s) &= G_1(s)\bar{x}(s),
\end{aligned}
\tag{3.30}
$$

The block diagram is then



3.4 Feedback

Feedback is present in a system when the output is fed, usually via some *feedback transfer function*, back into the input to the system. The classical linear control system with feedback can be represented by the block diagram



The equations relating parts of the system are

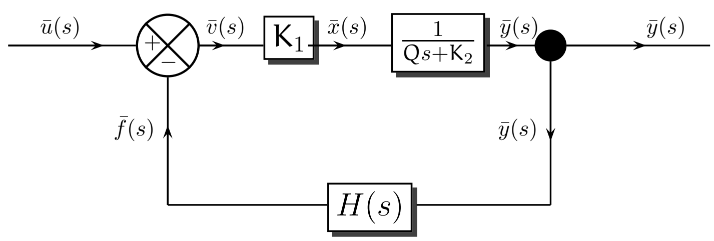
$$\begin{aligned}\bar{f}(s) &= H(s)\bar{y}(s),\\ \bar{v}(s) &= \bar{u}(s) - \bar{f}(s), \qquad (3.31)\\ \bar{y}(s) &= G(s)\bar{v}(s).\end{aligned}$$

Eliminating $\bar{f}(s)$ and $\bar{v}(s)$ gives

$$\bar{y}(s) = \frac{G(s)}{1 + G(s)H(s)}\,\bar{u}(s). \tag{3.32}$$

Example 3.4.1 We modify the model of Example 3.2.1 by introducing a feedback.

The block diagram for this problem is obtained by introducing feedback into the block diagram of Example 3.2.1.



From (3.20)

$$\bar{y}(s) = \frac{\mathsf{K}_1}{\mathsf{Q}s + \mathsf{K}_2}\,\bar{v}(s) \tag{3.33}$$

and

$$\bar{v}(s) = \bar{u}(s) - H(s)\bar{y}(s). \tag{3.34}$$

Giving

$$\bar{y}(s) = \frac{\mathsf{K}_1}{\mathsf{Q}\,s + \mathsf{K}_2 + H(s)\mathsf{K}_1}\,\bar{u}(s). \tag{3.35}$$

To complete this problem we need to make some assumptions about the nature of the feedback and we must also know the form of $u(t)$. Suppose as in Example 3.2.1 $u(t) = u_0$, giving $\bar{u}(s) = u_0/s$ and assume a proportional feedback. That is $H(s) = \mathsf{H}$, a constant. Then

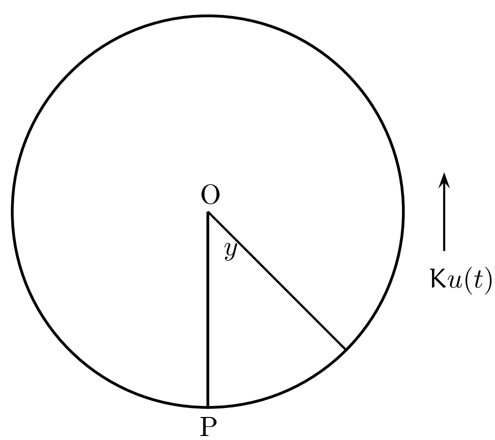
$$\bar{y}(s) = \frac{u_0 \mathsf{K}_1}{s[\mathsf{Q}\, s + (\mathsf{K}_2 + \mathsf{H}\mathsf{K}_1)]}. \tag{3.36}$$

Comparing with equations (3.21)–(3.23) we see that the effect of the feedback is to replace K_2 by $\mathsf{K}_2 + \mathsf{H}\mathsf{K}_1$. The solution is therefore

$$y(t) = u_0 \frac{\mathsf{K}_1}{\mathsf{K}_2 + \mathsf{H}\mathsf{K}_1}[1 - \exp(-t/\mathsf{T}')], \tag{3.37}$$

where $\mathsf{T}' = \mathsf{Q}/(\mathsf{K}_2 + \mathsf{H}\mathsf{K}_1)$. With $\mathsf{H}\mathsf{K}_1 > 0$, $\mathsf{T}' < \mathsf{T}$, so the feedback promotes a faster response of the output to the input. A typical case is that of *unitary feedback* where $H(s) = 1$.

Example 3.4.2 We have a heavy flywheel, centre O, of moment of inertia I. Suppose that P designates a point on the circumference with the flywheel initially at rest and P vertically below O. We need to devise a system such that, by applying a torque $\mathsf{K}u(t)$ to the wheel, in the long-time limit OP subtends an angle y^* with the downward vertical.



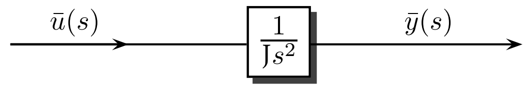
The equation of motion of the wheel is

$$\mathrm{I}\frac{\mathrm{d}^2 y}{\mathrm{d}t^2} = \mathsf{K}u(t). \tag{3.38}$$

Let $\mathrm{J} = \mathrm{I}/\mathsf{K}$ and take the Laplace transform of (3.38). Remembering that $y(0) = \dot{y}(0) = 0$

$$\bar{y}(s) = \frac{1}{\mathrm{J}s^2}\bar{u}(s). \tag{3.39}$$

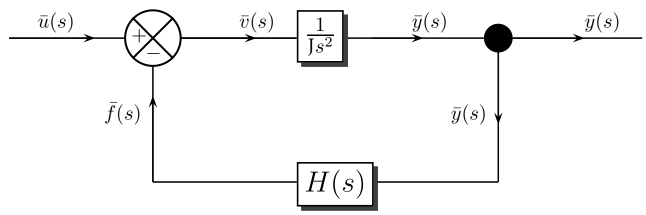
and the block diagram is



Suppose that the torque is $u(t) = u_0$. Then $\bar{u}(s) = u_0/s$ and inverting the transform

$$y(t) = \frac{u_0}{2J}t^2. \tag{3.40}$$

The angle grows, with increasing angular velocity, which does not achieve the required result. Suppose now that we introduce feedback $H(s)$. The block diagram is modified to



With a unitary feedback ($H(s) = 1$)

$$Js^2\bar{y}(s) = \bar{u}(s) - \bar{y}(s), \tag{3.41}$$

Again with $u(t) = u_0$ this gives

$$\bar{y}(s) = \frac{u_0}{s(Js^2+1)} = \frac{u_0}{s} - \frac{u_0 s}{s^2 + J^{-1}}. \tag{3.42}$$

Inverting the Laplace transform this gives

$$y(t) = u_0[1 - \cos(\omega_0 t)], \tag{3.43}$$

where $\omega_0 = 1/\sqrt{J}$. Again the objective is not achieved since $y(t)$ oscillates about u_0. Suppose we now modify the feedback to

$$H(s) = 1 + as. \tag{3.44}$$

Then

$$Js^2\bar{y}(s) = \bar{u}(s) - \bar{y}(s)(1 + as) \tag{3.45}$$

and with $u(t) = u_0$ this gives

$$\bar{y}(s) = \frac{u_0}{s(Js^2 + as + 1)} = \frac{u_0}{s} - u_0\frac{(s + \frac{1}{2}a\omega_0^2) + \frac{1}{2}a\omega_0^2}{(s + \frac{1}{2}a\omega_0^2)^2 + \omega^2}, \tag{3.46}$$

where $\omega^2 = \omega_0^2 - \frac{1}{4}a^2\omega_0^4$. Inverting the Laplace transform this gives

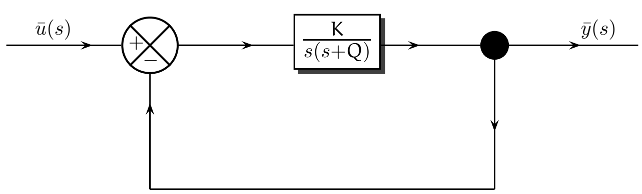
$$y(t) = u_0\left\{1 - \exp\left(-\tfrac{1}{2}at\omega_0^2\right)\left[\cos(\omega t) + \frac{a\omega_0^2}{2\omega}\sin(\omega t)\right]\right\}. \tag{3.47}$$

As $t \to \infty$ $y(t) \to u_0$.[2] So by setting $u_0 = y^*$ we have achieve the required objective. From (3.44), the feedback is

$$f(t) = a\frac{\mathrm{d}y(t)}{\mathrm{d}t} + y(t). \tag{3.48}$$

Problems 3

1) For the system with block diagram



show that

$$\bar{y}(s) = \frac{\mathsf{K}\bar{u}(s)}{s^2 + \mathsf{Q}s + \mathsf{K}}.$$

Given that $u(t) = u_0$, where u_0 is constant, show that

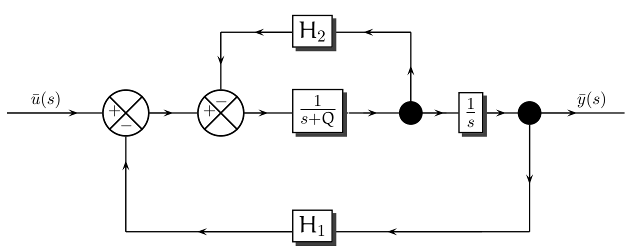
(i) when $\mathsf{K} - \frac{1}{4}\mathsf{Q}^2 = \omega^2 > 0$,

$$y(t) = u_0\left[1 - \exp\left(-\tfrac{1}{2}\mathsf{Q}t\right)\left\{\cos(\omega t) + \frac{\mathsf{Q}}{2\omega}\sin(\omega t)\right\}\right],$$

(ii) when $\frac{1}{4}\mathsf{Q}^2 - \mathsf{K} = \zeta^2 > 0$,

$$y(t) = u_0\left[1 - \frac{1}{2\zeta}\exp\left(-\tfrac{1}{2}\mathsf{Q}t\right)\left\{\left[\tfrac{1}{2}\mathsf{Q} + \zeta\right]\exp(\zeta t) - \left[\tfrac{1}{2}\mathsf{Q} - \zeta\right]\exp(-\zeta t)\right\}\right].$$

2) For the system with block diagram



[2]The term involving $\exp\left(-\frac{1}{2}at\omega_0^2\right)$ is known as the *transient contribution*.

show that

$$\bar{y}(s) = \frac{\bar{u}(s)}{s^2 + s(\mathsf{H}_2 + \mathsf{Q}) + \mathsf{H}_1}.$$

(*Hint: Put in all the intermediate variables, write down the equations associated with each box and switch and eliminate to find the relationship between* $\bar{u}(s)$ *and* $\bar{y}(s)$.)

3) Discrete time systems, where input $u(k)$ is related to output (response) $y(k)$ by a difference equation can be solved by using the $\mathcal{Z}$ transform to obtain a formula of the type $\tilde{y}(z) = G(z)\tilde{u}(z)$, where $G(z)$ is the discrete time version of the transfer function. For the following two cases find the transfer function

   (i) $y(k) - 2y(k-1) = u(k-1)$.

   (ii) $y(k) + 5y(k-1) + 6y(k-2) = u(k-1) + u(k-2)$.

   Obtain $y(k)$, in each case when $u(k) = 1$ for all k.

Chapter 4

Controllability and Observability

4.1 Introduction

In Sect. 3.1 we introduced a system with inputs $\{u_1(t), u_2(t), \ldots\}$ and outputs $\{x_1(t), x_2(t), \ldots\}$. In Sect. 3.2 this was modified to take into account the fact that the number of x-variables may be large or be difficult to handle. They may also include unnecessary details about the system. If the x-variables are now renamed *state-space variables* a new, possibly smaller, set $\{y_1(t), y_2(t), \ldots\}$ of *output variables* can be introduced. In terms of transfer functions this situation is represented by systems in cascade with one set of equations relating input variables to state-space variables and a second set relating state-space variables to output variables. The Laplace transformed equations for one variable of each type are (3.13). In general if there are m input variables n state-space variables and r output variables we need a matrix formulation to specify the system. The most general form for a linear system with continuous time is

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{A}(t)\boldsymbol{x}(t) + \boldsymbol{B}(t)\boldsymbol{u}(t), \tag{4.1}$$

$$\boldsymbol{y}(t) = \boldsymbol{C}(t)\boldsymbol{x}(t), \tag{4.2}$$

where $\boldsymbol{u}(t)$ is the m-dimensional column vector of input variables, $\boldsymbol{x}(t)$ is the n-dimensional column vector of input variables and $\boldsymbol{y}(t)$ is the r-dimensional column vector of output variables. The matrices $\boldsymbol{A}(t)$, $\boldsymbol{B}(t)$ and $\boldsymbol{C}(t)$ are respectively $n \times n$, $n \times m$ and $r \times n$. Equation (4.1) is a system of n first-order differential equations. They will be equations with constant coefficients if the elements of the matrix $\boldsymbol{A}$ are not time-dependent. The control system represented by the pair of equations (4.1) and (4.2) is said to be *a constant system* if none of the matrices $\boldsymbol{A}$, $\boldsymbol{B}$ and $\boldsymbol{C}$ is time-dependent.

The concentration on first-order differential equations is not a serious restriction. As we saw in Sect. 1.5, a single n-th order differential equation can be replaced by an equivalent set of n first-order equations. In a similar way a number of higher-order differential equations could be replaced by a larger number of first-order equations.

A simple case would be the problem of Example 3.4.2. Let $x_1(t) = y(t)$ and $x_2(t) = \dot{y}(t)$. Then,

$$y(t) = \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}(t) \tag{4.3}$$

and, from (3.48),

$$f(t) = \boldsymbol{a}^{\mathrm{T}}\boldsymbol{x}(t), \tag{4.4}$$

where

$$\boldsymbol{x}(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix}, \qquad \boldsymbol{a} = \begin{pmatrix} 1 \\ a \end{pmatrix}, \qquad \boldsymbol{c} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \tag{4.5}$$

With $v(t) = u(t) - f(t)$ replacing $u(t)$ in (3.38)

$$\begin{aligned} \dot{x}_1(t) &= x_2(t), \\ \dot{x}_2(t) &= \mathrm{J}^{-1}v(t) = \mathrm{J}^{-1}[u(t) - f(t)], \end{aligned} \tag{4.6}$$

giving

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{b}u(t), \tag{4.7}$$

where

$$\boldsymbol{b} = \begin{pmatrix} 0 \\ \mathrm{J}^{-1} \end{pmatrix}, \qquad \boldsymbol{X} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \tag{4.8}$$

$$\boldsymbol{A} = \boldsymbol{X} - \boldsymbol{b}\boldsymbol{a}^{\mathrm{T}} = \begin{pmatrix} 0 & 1 \\ -\mathrm{J}^{-1} & -a\mathrm{J}^{-1} \end{pmatrix}. \tag{4.9}$$

Equations (4.3) and (4.7) are a particular case of (4.1) and (4.2) with $m = r = 1$ and $n = 2$.

4.2 The Exponential Matrix

We shall now concentrate on the solution of (4.1) for a constant system. That is

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{B}\boldsymbol{u}(t). \tag{4.10}$$

We can *define* the *exponential matrix* $\exp(\boldsymbol{A}t)$ by the expansion

$$\exp(\boldsymbol{A}t) = \boldsymbol{I} + t\boldsymbol{A} + \frac{1}{2!}t^2\boldsymbol{A}^2 + \cdots + \frac{1}{k!}t^k\boldsymbol{A}^k + \cdots, \tag{4.11}$$

and it then follows, as is the case for a scalar (1×1 matrix) $\boldsymbol{A}$, that

$$\frac{\mathrm{d}\exp(\boldsymbol{A}t)}{\mathrm{d}t} = \exp(\boldsymbol{A}t)\boldsymbol{A}. \tag{4.12}$$

Then (4.10) can be transformed by premultiplication by $\exp(-\boldsymbol{A}t)$ into the form

$$\frac{\mathrm{d}\exp(-\boldsymbol{A}t)\boldsymbol{x}(t)}{\mathrm{d}t} = \exp(-\boldsymbol{A}t)\boldsymbol{B}\boldsymbol{u}(t). \tag{4.13}$$

This is, of course, equivalent to $\exp(-\boldsymbol{A}t)$ being the integrating factor for (4.10) (see Sect. 1.3.3), except that here, since we are dealing with matrices, the order of the terms in any product is important. Integrating (4.13) gives

$$\boldsymbol{x}(t) = \exp(\boldsymbol{A}t)\left[\boldsymbol{x}(0) + \int_0^t \exp(-\boldsymbol{A}\tau)\boldsymbol{B}\boldsymbol{u}(\tau)\mathrm{d}\tau\right]. \tag{4.14}$$

Of course, this formula is of practical use for determining $\boldsymbol{x}(t)$ only if we have a closed-form expression for $\exp(\boldsymbol{A}t)$. One case of this kind would be when $\boldsymbol{A}$ has a set of n distinct eigenvalues $\lambda^{(j)}$, $j = 1, 2, \ldots, n$ and we are able to calculate all the left eigenvectors $\boldsymbol{p}^{(j)}$ and right eigenvectors $\boldsymbol{q}^{(j)}$, which satisfy the orthonormality condition (1.148).[1] Then the matrix $\boldsymbol{P}$ formed by having the vectors $\left[\boldsymbol{p}^{(j)}\right]^{\mathrm{T}}$ (in order) as rows and the matrix $\boldsymbol{Q}$ formed by having the vectors $\boldsymbol{q}^{(j)}$ (in order) as columns satisfy the conditions

$$\boldsymbol{P} = \boldsymbol{Q}^{-1}, \tag{4.15}$$
$$\boldsymbol{P}\boldsymbol{A}\boldsymbol{Q} = \boldsymbol{\Lambda}, \tag{4.16}$$

where $\boldsymbol{\Lambda}$ is the diagonal matrix with the eigenvalues (in order) as diagonal elements. Then

$$\begin{aligned}\exp(\boldsymbol{A}t) &= \boldsymbol{I} + t\boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{P} + \frac{1}{2!}t^2(\boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{P})^2 + \cdots + \frac{1}{k!}t^k(\boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{P})^k + \cdots \\ &= \boldsymbol{Q}[\boldsymbol{I} + t\boldsymbol{\Lambda} + \frac{1}{2!}t^2\boldsymbol{\Lambda}^2 + \cdots + \frac{1}{k!}t^k\boldsymbol{\Lambda}^k + \cdots]\boldsymbol{P} \\ &= \boldsymbol{Q}\exp(\boldsymbol{\Lambda}t)\boldsymbol{P},\end{aligned} \tag{4.17}$$

where $\exp(\boldsymbol{\Lambda}t)$ is the diagonal matrix with diagonal elements $\exp(\lambda^{(j)}t)$, $j = 1, 2, \ldots n$ (in order).

The problem with this method is that it involves calculating (or using MAPLE to calculate) all the eigenvalues *and* the left and right eigenvectors. It is also valid only when all the eigenvalues are distinct, so that the left and right eigenvectors are orthonormalizible. We now develop a method of obtaining $\exp(\boldsymbol{A}t)$ as a polynomial in $\boldsymbol{A}$ which depends only on deriving the eigenvalues and which is valid even if some are degenerate. The characteristic equation of $\boldsymbol{A}$ is

$$\Delta(\lambda) = 0, \tag{4.18}$$

[1] We have changed the letters from those used in Sect. 1.6.2 to avoid a clash with present usage.

where

$$\Delta(\lambda) = \mathrm{Det}\{\lambda \boldsymbol{I} - \boldsymbol{A}\} = \lambda^n + \alpha_{n-1}\lambda^{n-1} + \cdots + \alpha_1\lambda + \alpha_0 \tag{4.19}$$

is called the *characteristic polynomial*. The zeros of $\Delta(\lambda)$ are the eigenvalues of $\boldsymbol{A}$. Suppose that the eigenvalues are $\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(m)}$, where $\lambda^{(j)}$ is $\mu^{(j)}$-fold degenerate. Then

$$\sum_{j=1}^{m} \mu^{(j)} = n \tag{4.20}$$

and

$$\Delta(\lambda^{(j)}) = \left(\frac{\mathrm{d}\Delta(\lambda)}{\mathrm{d}\lambda}\right)_{\lambda=\lambda^{(j)}} = \cdots = \left(\frac{\mathrm{d}^{\mu^{(j)}-1}\Delta(\lambda)}{\mathrm{d}\lambda^{\mu^{(j)}-1}}\right)_{\lambda=\lambda^{(j)}} = 0, \qquad j = 1, 2, \ldots, m \tag{4.21}$$

An important result of linear algebra is the *Cayley-Hamilton Theorem*. This asserts that $\boldsymbol{A}$ satisfies its own characteristic equation. That is

$$\Delta(\boldsymbol{A}) = \boldsymbol{A}^n + \alpha_{n-1}\boldsymbol{A}^{n-1} + \cdots + \alpha_1\boldsymbol{A} + \alpha_0\boldsymbol{I} = 0. \tag{4.22}$$

This is easily proved when (4.15) and (4.16) are valid (no degenerate eigenvalues, $\mu_j = 1$, for $j = 1, 2, \ldots, m$). Then

$$\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{P}, \tag{4.23}$$

$$\boldsymbol{A}^s = \boldsymbol{Q}\boldsymbol{\Lambda}^s\boldsymbol{P}, \qquad s = 1, 2, \ldots \tag{4.24}$$

and

$$\Delta(\boldsymbol{A}) = \boldsymbol{Q}[\boldsymbol{\Lambda}^n + \alpha_{n-1}\boldsymbol{\Lambda}^{n-1} + \cdots + \alpha_1\boldsymbol{\Lambda} + \alpha_0\boldsymbol{I}]\boldsymbol{P}. \tag{4.25}$$

Since the eigenvalues satisfy the characteristic equation, the matrix obtained by summing all the terms in the square brackets has *every* element zero, which establishes the theorem. The result still holds for repeated eigenvalues but the proof is a little longer. An important result for our discussion is the following:

Theorem 4.2.1 *The power series for* $\exp(zt)$ *can be decomposed in the form*

$$\exp(zt) = \mathrm{D}(z;t) + \Delta(z)f(z) \tag{4.26}$$

where

(i) $$\mathrm{D}(z;t) = \beta_0(t) + \beta_1(t)z + \cdots + \beta_{n-1}(t)z^{n-1} \tag{4.27}$$

(ii) $\Delta(z)$ *is the characteristic polynomial of an* $n \times n$ *matrix* $\boldsymbol{A}$.

(iii) $f(z)$ *is a regular function of* z.

This theorem is established by considering the Laurent expansion of

$$g(z) = \frac{\exp(zt)}{\Delta(z)} \tag{4.28}$$

about each of its poles $z = \lambda^{(1)}, \lambda^{(2)}, \dots, \lambda^{(m)}$ in order. Thus, for $\lambda^{(1)}$,

$$\begin{aligned} g(z) &= \sum_{i=1}^{\mu^{(1)}} \frac{\gamma_i}{\left[z - \lambda^{(1)}\right]^i} + \sum_{k=0}^{\infty} \rho_k \left[z - \lambda^{(1)}\right]^k \\ &= \frac{P_1(z)}{\left[z - \lambda^{(1)}\right]^{\mu^{(1)}}} + g_1(z), \end{aligned} \tag{4.29}$$

where $P_1(z)$ is a polynomial in z of degree $\mu^{(1)} - 1$ and $g_1(z)$ is regular at $z = \lambda^{(1)}$, but with the same poles as $g(z)$ at all the other eigenvalues. Repeating this process gives

$$g(z) = \sum_{j=1}^{m} \frac{P_j(z)}{\left[z - \lambda^{(j)}\right]^{\mu^{(j)}}} + f(z), \tag{4.30}$$

Multiplying through by $\Delta(z)$ gives

$$\exp(zt) = \sum_{j=1}^{m} \frac{\Delta(z)P_j(z)}{\left[z - \lambda^{(j)}\right]^{\mu^{(j)}}} + \Delta(z)f(z). \tag{4.31}$$

Each of the terms in the summation is a polynomial of degree $n-1$; so together they form a polynomial $\mathrm{D}(z;t)$ of degree $n-1$ and given by (4.27). To determine the coefficients $\beta_0(t), \dots, \beta_{n-1}(t)$ we use (4.21) and (4.26). For the eigenvalue $\lambda^{(j)}$ we have $\mu^{(j)}$ linear equations for these coefficients given by

$$\begin{aligned} &\exp\left(\lambda^{(j)}t\right) = \mathrm{D}\left(\lambda^{(j)};t\right), \\ &t^i \exp\left(\lambda^{(j)}t\right) = \left(\frac{\mathrm{d}^i \mathrm{D}(z;t)}{\mathrm{d}z^i}\right)_{z=\lambda^{(j)}}, \qquad i = 1, 2, \dots, \mu^{(j)} - 1. \end{aligned} \tag{4.32}$$

This then gives in all n linear equations in n unknowns. It is not difficult to show that they are independent and will thus yield the coefficients.

It now follows from the Cayley-Hamilton result (4.22) and (4.26) and (4.27) that

$$\exp(\boldsymbol{A}t) = \beta_0(t)\boldsymbol{I} + \beta_1(t)\boldsymbol{A} + \cdots + \beta_{n-1}(t)\boldsymbol{A}^{n-1}. \tag{4.33}$$

Example 4.2.1 Consider the matrix

$$\boldsymbol{A} = \begin{pmatrix} 0 & 1 \\ 6 & 1 \end{pmatrix} \tag{4.34}$$

Then

$$\Delta(\lambda) = \lambda^2 - \lambda - 6, \tag{4.35}$$

with eigenvalues $\lambda^{(1)} = 3$, $\lambda^{(2)} = -2$. Now let

$$\exp(zt) = \beta_0(t) + \beta_1(t)z + \Delta(z)f(z), \tag{4.36}$$

giving

$$\begin{aligned} \exp(3t) &= \beta_0(t) + 3\beta_1(t), \\ \exp(-2t) &= \beta_0(t) - 2\beta_1(t). \end{aligned} \tag{4.37}$$

Thus

$$\begin{aligned} \exp(\boldsymbol{A}t) &= \tfrac{1}{5}\boldsymbol{I}[2\exp(3t) + 3\exp(-2t)] + \tfrac{1}{5}\boldsymbol{A}[\exp(3t) - \exp(-2t)] \\ &= \frac{1}{5}\exp(3t)\begin{pmatrix} 2 & 1 \\ 6 & 3 \end{pmatrix} + \frac{1}{5}\exp(-2t)\begin{pmatrix} 3 & -1 \\ -6 & 2 \end{pmatrix}. \end{aligned} \tag{4.38}$$

An alternative approach to deriving a compact form for $\exp(\boldsymbol{A}t)$ is to use a Laplace transform. Taking the Laplace transform of (4.10) gives

$$s\bar{\boldsymbol{x}}(s) - \boldsymbol{x}(0) = \boldsymbol{A}\bar{\boldsymbol{x}}(s) + \boldsymbol{B}\bar{\boldsymbol{u}}(s). \tag{4.39}$$

Writing this equation in the form

$$\bar{\boldsymbol{x}}(s) = (s\boldsymbol{I} - \boldsymbol{A})^{-1}\boldsymbol{x}(0) + (s\boldsymbol{I} - \boldsymbol{A})^{-1}\boldsymbol{B}\bar{\boldsymbol{u}}(s) \tag{4.40}$$

and inverting the Laplace transform it follows, on comparison with (4.14) that

$$\mathcal{L}^{-1}\left\{(s\boldsymbol{I} - \boldsymbol{A})^{-1}\right\} = \exp(\boldsymbol{A}t). \tag{4.41}$$

Now we apply this method to Example 4.2.1.

$$s\boldsymbol{I} - \boldsymbol{A} = \begin{pmatrix} s & -1 \\ -6 & s-1 \end{pmatrix} \tag{4.42}$$

and

$$(s\boldsymbol{I} - \boldsymbol{A})^{-1} = \frac{1}{\Delta(s)}\begin{pmatrix} s-1 & 1 \\ 6 & s \end{pmatrix}, \tag{4.43}$$

where $\Delta(s)$ is given by (4.35). Inverting this Laplace transformed matrix element by element gives (4.38).

4.3 Realizations of Systems

For simplicity we shall, in most of the work which follows, consider a constant system with one (scalar) input $u(t)$ and one (scalar) output $y(t)$. Thus we have

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{b}\,u(t), \tag{4.44}$$

$$y(t) = \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}(t), \tag{4.45}$$

where the n-dimensional column vector $\boldsymbol{b}$ replaces the matrix $\boldsymbol{B}$ and the n-dimensional row vector $\boldsymbol{c}^{\mathrm{T}}$ replaces the matrix $\boldsymbol{C}$. Let $\boldsymbol{T}$ be an $n \times n$ non-singular matrix. Equations (4.44) and (4.45) can be expressed in the form

$$\frac{\mathrm{d}\boldsymbol{T}\boldsymbol{x}(t)}{\mathrm{d}t} = \boldsymbol{T}\boldsymbol{A}\boldsymbol{T}^{-1}\boldsymbol{T}\boldsymbol{x}(t) + \boldsymbol{T}\boldsymbol{b}\,u(t), \tag{4.46}$$

$$y(t) = \boldsymbol{c}^{\mathrm{T}}\boldsymbol{T}^{-1}\boldsymbol{T}\boldsymbol{x}(t). \tag{4.47}$$

With

$$\begin{aligned} &\boldsymbol{x}'(t) = \boldsymbol{T}\boldsymbol{x}(t), \qquad \boldsymbol{T}\boldsymbol{A}\boldsymbol{T}^{-1} = \boldsymbol{A}', \\ &\boldsymbol{T}\boldsymbol{b} = \boldsymbol{b}', \qquad \boldsymbol{c}^{\mathrm{T}}\boldsymbol{T}^{-1} = (\boldsymbol{c}')^{\mathrm{T}}, \end{aligned} \tag{4.48}$$

this gives

$$\frac{\mathrm{d}\boldsymbol{x}'(t)}{\mathrm{d}t} = \boldsymbol{A}'\boldsymbol{x}'(t) + \boldsymbol{b}'\,u(t), \tag{4.49}$$

$$y(t) = (\boldsymbol{c}')^{\mathrm{T}}\boldsymbol{x}'(t). \tag{4.50}$$

These equations still have the same input and output variables, but the state variables and matrices have been transformed. These equivalent expressions of the problem are called *realizations* and are denoted by $[\boldsymbol{A}, \boldsymbol{b}, \boldsymbol{c}^{\mathrm{T}}]$ and $[\boldsymbol{A}', \boldsymbol{b}', (\boldsymbol{c}')^{\mathrm{T}}]$. The Laplace transforms of (4.44) and (4.45) are

$$s\bar{\boldsymbol{x}}(s) - \boldsymbol{x}(0) = \boldsymbol{A}\bar{\boldsymbol{x}}(s) + \boldsymbol{b}\,\bar{u}(s), \tag{4.51}$$

$$\bar{y}(s) = \boldsymbol{c}^{\mathrm{T}}\bar{\boldsymbol{x}}(s). \tag{4.52}$$

With $\boldsymbol{x}(0) = \boldsymbol{0}$ these give

$$\bar{y}(s) = G(s)\bar{u}(s), \tag{4.53}$$

where

$$G(s) = \boldsymbol{c}^{\mathrm{T}}(s\boldsymbol{I} - \boldsymbol{A})^{-1}\boldsymbol{b} \tag{4.54}$$

is the *transfer function*. It is clear that it is invariant under the change of realization given by (4.48). So these are said to be *different realizations of the* *transfer function*. From the definition of the inverse of a matrix and (4.19), equation (4.54) can also be written in the form[2]

$$G(s) = \frac{\boldsymbol{c}^{\mathrm{T}}\mathrm{Adj}\{s\boldsymbol{I}-\boldsymbol{A}\}\boldsymbol{b}}{\Delta(s)} \tag{4.55}$$

It follows that the poles of the transfer function are the eigenvalues of $\boldsymbol{A}$, which are, of course, invariant under the transformation $\boldsymbol{TAT}^{-1} = \boldsymbol{A}'$.

Since the inverse of the transpose of a matrix is equal to the transpose of the inverse it is clear that (4.54) can be rewritten in the form

$$G(s) = \boldsymbol{b}^{\mathrm{T}}(s\boldsymbol{I}-\boldsymbol{A}^{\mathrm{T}})^{-1}\boldsymbol{c}. \tag{4.56}$$

It follows that $[\boldsymbol{A}^{\mathrm{T}}, \boldsymbol{c}, \boldsymbol{b}^{\mathrm{T}}]$, that is to say,

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{x}(t) + \boldsymbol{c}\,u(t), \tag{4.57}$$

$$y(t) = \boldsymbol{b}^{\mathrm{T}}\boldsymbol{x}(t), \tag{4.58}$$

is also a realization of the system. It is called the *dual realization*.

4.3.1 The Companion Realization

Suppose we have a system with one input variable $u(t)$ and one output variable $y(t)$ related by

$$\frac{\mathrm{d}^n y}{\mathrm{d}t^n} + a_{n-1}\frac{\mathrm{d}^{n-1} y}{\mathrm{d}t^{n-1}} + \cdots + a_1\frac{\mathrm{d}y}{\mathrm{d}t} + a_0 y = b_m\frac{\mathrm{d}^m u}{\mathrm{d}t^m} + b_{m-1}\frac{\mathrm{d}^{m-1} u}{\mathrm{d}t^{m-1}} + \cdots + b_1\frac{\mathrm{d}u}{\mathrm{d}t} + b_0 u, \tag{4.59}$$

with $n > m$ and initial conditions

$$\begin{aligned} \left(\frac{\mathrm{d}^i y}{\mathrm{d}t^i}\right)_{t=0} &= 0, \qquad i = 0, 1, \ldots, n-1, \\ \left(\frac{\mathrm{d}^j u}{\mathrm{d}t^j}\right)_{t=0} &= 0, \qquad j = 0, 1, \ldots, m-1. \end{aligned} \tag{4.60}$$

Then

$$\bar{y}(s) = G(s)\bar{u}(s), \tag{4.61}$$

where the transfer function

$$G(s) = \frac{\psi(s)}{\phi(s)}, \tag{4.62}$$

with

$$\begin{aligned} \phi(s) &= s^n + a_{n-1}s^{n-1} + \cdots + a_1 s + a_0, \\ \psi(s) &= b_m s^m + b_{m-1}s^{m-1} + \cdots + b_1 s + b_0. \end{aligned} \tag{4.63}$$

[2] The *adjoint* $\mathrm{Adj}\{\boldsymbol{X}\}$ of a square matrix $\boldsymbol{X}$ is the transpose of the matrix of cofactors.

We now construct a particular realization of the system by defining a set of input variables $x_1(t), \ldots, x_n(t)$. Let $x_1(t)$ be a solution of

$$y(t) = b_m \frac{\mathrm{d}^m x_1}{\mathrm{d}t^m} + b_{m-1} \frac{\mathrm{d}^{m-1} x_1}{\mathrm{d}t^{m-1}} + \cdots + b_1 \frac{\mathrm{d}x_1}{\mathrm{d}t} + b_0 x_1, \tag{4.64}$$

and define

$$\begin{aligned} x_2(t) &= \dot{x}_1(t), \\ x_3(t) &= \dot{x}_2(t) = \frac{\mathrm{d}^2 x_1}{\mathrm{d}t^2}, \\ &\vdots \qquad\qquad \vdots \\ x_n(t) &= \dot{x}_{n-1}(t) = \frac{\mathrm{d}^{n-1} x_1}{\mathrm{d}t^{n-1}}. \end{aligned} \tag{4.65}$$

Then (4.64) becomes

$$y(t) = b_m x_{m+1}(t) + b_{m-1} x_m(t) + \cdots + b_0 x_1(t), \tag{4.66}$$

which can be expressed in the form (4.50) with

$$\boldsymbol{c}^{\mathrm{T}} = \begin{pmatrix} b_0 & \cdots & b_m & 0 & \cdots & 0 \end{pmatrix}, \qquad \boldsymbol{x}(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{pmatrix} \tag{4.67}$$

Let

$$\dot{x}_n(t) = \kappa u(t) + \sum_{k=1}^{n} \gamma_{k-1} x_k(t) \tag{4.68}$$

which is equivalent to

$$u(t) = \frac{1}{\kappa} \frac{\mathrm{d}^n x_1}{\mathrm{d}t^n} - \frac{\gamma_{n-1}}{\kappa} \frac{\mathrm{d}^{n-1} x_1}{\mathrm{d}t^{n-1}} - \cdots - \frac{\gamma_1}{\kappa} \frac{\mathrm{d}x_1}{\mathrm{d}t} - \frac{\gamma_0}{\kappa} x_1(t). \tag{4.69}$$

We now substitute into (4.59) from (4.64) and (4.69). This gives

$$\sum_{j=0}^{m} b_j \frac{\mathrm{d}^j w}{\mathrm{d}t^j} = 0, \tag{4.70}$$

where

$$w(t) = \left(1 - \frac{1}{\kappa}\right) \frac{\mathrm{d}^n x_1}{\mathrm{d}t^n} + \sum_{i=0}^{n-1} \left(a_i + \frac{\gamma_i}{\kappa}\right) \frac{\mathrm{d}^i x_1}{\mathrm{d}t^i}. \tag{4.71}$$

Thus, by choosing

$$\kappa = 1, \qquad \gamma_i = -a_i, \qquad i = 1, 2, \ldots, n-1, \tag{4.72}$$

(4.70), and hence (4.59), are identically satisfied and (4.68) and (4.65) can be combined in the matrix form (4.44) with

$$\boldsymbol{A} = \begin{pmatrix} 0 & 1 & 0 & \cdots & \cdots & \cdots & 0 & 0 \\ 0 & 0 & 1 & 0 & \cdots & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \cdots & 0 & 1 \\ -a_0 & -a_1 & \cdots & \cdots & \cdots & \cdots & \cdots & -a_{n-1} \end{pmatrix} \tag{4.73}$$

$$\boldsymbol{b} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}. \tag{4.74}$$

When $\boldsymbol{A}$, $\boldsymbol{b}$ and $\boldsymbol{c}^{\mathrm{T}}$ are of this type they are said to be in *companion form*. It is not difficult to show that the characteristic function for $\boldsymbol{A}$ is

$$\Delta(\lambda) = \phi(\lambda), \tag{4.75}$$

again establishing that the poles of the transfer function are the eigenvalues of $\boldsymbol{A}$.

For any realization, the $n \times n$ matrix

$$\boldsymbol{U} = \begin{pmatrix} \boldsymbol{b} & \boldsymbol{A}\boldsymbol{b} & \boldsymbol{A}^2\boldsymbol{b} & \cdots & \boldsymbol{A}^{n-1}\boldsymbol{b} \end{pmatrix}, \tag{4.76}$$

that is the matrix with j-th column $\boldsymbol{A}^{j-1}\boldsymbol{b}$, is[3] called the *controllability matrix*. The $n \times n$ matrix

$$\boldsymbol{V} = \begin{pmatrix} \boldsymbol{c}^{\mathrm{T}} \\ \boldsymbol{c}^{\mathrm{T}}\boldsymbol{A} \\ \boldsymbol{c}^{\mathrm{T}}\boldsymbol{A}^2 \\ \vdots \\ \boldsymbol{c}^{\mathrm{T}}\boldsymbol{A}^{n-1} \end{pmatrix}, \tag{4.77}$$

[3] For reasons which we shall see below.

that is the matrix with j-th row $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{A}^{j-1}$, is[4] called the *observability matrix.*

For a companion realization it is not difficult to see that the controllability matrix has determinant of plus or minus one, depending on its order. We now prove an important theorem:

Theorem 4.3.1 *A realization given by (4.44) and (4.45) can be transformed by (4.48) into a companion realization if and only if the controllability matrix is non-singular.*

Proof:

Sufficiency. If $\boldsymbol{U}$ is non-singular $\boldsymbol{U}^{-1}$ exists. Let $\boldsymbol{\xi}^{\mathrm{T}}$ be the n-th row of $\boldsymbol{U}^{-1}$. That is

$$\boldsymbol{\xi}^{\mathrm{T}} = (\boldsymbol{b}')^{\mathrm{T}}\boldsymbol{U}^{-1}, \tag{4.78}$$

where

$$(\boldsymbol{b}')^{\mathrm{T}} = (0\,0\,\cdots\,0\,1) \tag{4.79}$$

and define the matrix

$$\boldsymbol{T} = \begin{pmatrix} \boldsymbol{\xi}^{\mathrm{T}} \\ \boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{A} \\ \vdots \\ \boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{A}^{n-1} \end{pmatrix} \tag{4.80}$$

We need to show that $\boldsymbol{T}^{-1}$ exists, which is the case when $\boldsymbol{T}$ is non-singular, that is when the rows of the matrix are independent. Suppose the contrary, that there exist $\gamma_1, \ldots, \gamma_n$, not all zero, such that

$$\gamma_1\boldsymbol{\xi}^{\mathrm{T}} + \gamma_2\boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{A} + \cdots + \gamma_n\boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{A}^{n-1} = 0. \tag{4.81}$$

Multiplying on the right by $\boldsymbol{A}^k\boldsymbol{b}$ the coefficient of γ_j is $\boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{A}^{j-1+k}\boldsymbol{b}$. The vector $\boldsymbol{A}^{j-1+k}\boldsymbol{b}$ is the $(j+k)$-th column of $\boldsymbol{U}$ so this coefficient is non-zero (and equal to unity) if and only if $j+k=n$. So by varying k from 0 to $n-1$ we establish that $\gamma_1 = \gamma_2 = \cdots = \gamma_n = 0$ and thus that the matrix $\boldsymbol{T}$ is non-singular. Now

$$\boldsymbol{T}\boldsymbol{b} = \begin{pmatrix} \boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{b} \\ \boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{b} \\ \vdots \\ \boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{A}^{n-1}\boldsymbol{b} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} = \boldsymbol{b}'. \tag{4.82}$$

[4] Again for reasons which we shall see below.

Let the columns of $\boldsymbol{T}^{-1}$ be denoted by $\boldsymbol{e}_j$, $j = 1, 2, \dots, n$. Then

$$\boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{A}^k\boldsymbol{e}_j = \delta^{\mathrm{Kr}}(k+1-j), \qquad k = 0, 1, \dots, n-1. \tag{4.83}$$

Now the matrix $\boldsymbol{A}'$ given in (4.48) has k–j-th element given by

$$A'_{kj} = \boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{A}^{k-1}\boldsymbol{A}\boldsymbol{e}_j = \delta^{\mathrm{Kr}}(k+1-j), \qquad k = 1, 2, \dots, n-1. \tag{4.84}$$

So in each row of $\boldsymbol{A}'$ apart from the last there is one non-zero element equal to unity. In the k-th row the element is the $k+1$-th which is exactly the form of (4.73) if we define

$$a_j = -\boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{A}^n\boldsymbol{e}_{j+1}, \qquad j = 0, 1, \dots, n-1. \tag{4.85}$$

Finally we note that

$$(\boldsymbol{c}')^{\mathrm{T}} = \begin{pmatrix} \boldsymbol{c}^{\mathrm{T}}\boldsymbol{e}_1 & \boldsymbol{c}^{\mathrm{T}}\boldsymbol{e}_2 & \cdots & \boldsymbol{c}^{\mathrm{T}}\boldsymbol{e}_n \end{pmatrix} \tag{4.86}$$

Necessity. Suppose that a transformation by a non-singular matrix $\boldsymbol{T}$ exists. We know that in this companion realization the controllablity matrix

$$\boldsymbol{U}' = \begin{pmatrix} \boldsymbol{b}' & \boldsymbol{A}'\boldsymbol{b}' & (\boldsymbol{A}')^2\boldsymbol{b}' & \cdots & (\boldsymbol{A}')^{n-1}\boldsymbol{b}' \end{pmatrix}, \tag{4.87}$$

is non-singular ($|\mathrm{Det}\{\boldsymbol{U}'\}| = 1$). But

$$\boldsymbol{U}' = \boldsymbol{T}\boldsymbol{U}, \tag{4.88}$$

So $\boldsymbol{U}$ is also non-singular.

From (4.77) and (4.76)

$$\boldsymbol{V}^{\mathrm{T}} = \begin{pmatrix} \boldsymbol{c} & \boldsymbol{A}^{\mathrm{T}}\boldsymbol{c} & (\boldsymbol{A}^{\mathrm{T}})^2\boldsymbol{c} & \cdots & (\boldsymbol{A}^{\mathrm{T}})^{n-1}\boldsymbol{c} \end{pmatrix}, \tag{4.89}$$

and

$$\boldsymbol{U}^{\mathrm{T}} = \begin{pmatrix} \boldsymbol{b}^{\mathrm{T}} \\ \boldsymbol{b}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}} \\ \boldsymbol{b}^{\mathrm{T}}(\boldsymbol{A}^{\mathrm{T}})^2 \\ \vdots \\ \boldsymbol{b}^{\mathrm{T}}(\boldsymbol{A}^{\mathrm{T}})^{n-1} \end{pmatrix}. \tag{4.90}$$

It follows that the controllability and observability matrices of a realization (4.44)–(4.45) are the transposes of the observability and controllability matrices of it dual (4.57)–(4.58). This is, of course, a symmetric relationship because the transpose of the transpose of a matrix is the matrix itself and the dual of the dual of a realization is the realization itself. This idea can be developed further by defining an *alternative companion realization*. This is one where the matrix $\boldsymbol{A}$ has a form like the transpose of (4.73) and $\boldsymbol{c}$ replaces $\boldsymbol{b}$ in (4.74). We then have the theorem:

Theorem 4.3.2 *A realization given by (4.44) and (4.45) can be transformed by (4.48) into an alternative companion realization if and only if the observability matrix is non-singular.*

Proof: We first take the dual $[\boldsymbol{A}^{\mathrm{T}}, \boldsymbol{c}, \boldsymbol{b}^{\mathrm{T}}]$ of $[\boldsymbol{A}, \boldsymbol{b}, \boldsymbol{c}^{\mathrm{T}}]$. Since $\boldsymbol{V}^{\mathrm{T}}$ is the controllability matrix of $[\boldsymbol{A}^{\mathrm{T}}, \boldsymbol{c}, \boldsymbol{b}^{\mathrm{T}}]$, there exists a matrix $\boldsymbol{T}$ giving a companion realization $[\boldsymbol{T}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{T}^{-1}, \boldsymbol{T}\boldsymbol{c}, \boldsymbol{b}^{\mathrm{T}}\boldsymbol{T}^{-1}]$ if and only if $\boldsymbol{V}$ is nonsingular. Taking the dual of this realization gives an alternative companion realization $[\boldsymbol{T}'\boldsymbol{A}\boldsymbol{T}'^{-1}, \boldsymbol{T}'\boldsymbol{b}, \boldsymbol{c}^{\mathrm{T}}(\boldsymbol{T}')^{-1}]$, where $\boldsymbol{T}' = (\boldsymbol{T}^{\mathrm{T}})^{-1} = (\boldsymbol{T}^{-1})^{\mathrm{T}}$.

Example 4.3.1 Consider the system with a realization $[\boldsymbol{A}, \boldsymbol{b}, \boldsymbol{c}^{\mathrm{T}}]$ given by

$$\boldsymbol{A} = \begin{pmatrix} -1 & 0 & 1 \\ -1 & -2 & -1 \\ -2 & -2 & -3 \end{pmatrix}, \quad \boldsymbol{b} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad \boldsymbol{c} = \begin{pmatrix} -2 \\ -5 \\ 5 \end{pmatrix}. \tag{4.91}$$

From (4.76) and (4.91) the controllability matrix is

$$\boldsymbol{U} = \begin{pmatrix} 1 & 0 & -5 \\ 0 & -2 & 9 \\ 1 & -5 & 19 \end{pmatrix} \tag{4.92}$$

Since $\mathrm{Det}\{\boldsymbol{U}\} = -3$ a transformation to a companion realization exists. Inverting the controllability matrix

$$\boldsymbol{U}^{-1} = \begin{pmatrix} -\frac{7}{3} & -\frac{25}{3} & \frac{10}{3} \\ -3 & -8 & 3 \\ -\frac{2}{3} & -\frac{5}{3} & \frac{2}{3} \end{pmatrix} \tag{4.93}$$

and, from (4.78) and (4.80),

$$\boldsymbol{\xi}^{\mathrm{T}} = \begin{pmatrix} -\frac{2}{3} & -\frac{5}{3} & \frac{2}{3} \end{pmatrix}, \qquad \boldsymbol{T} = \begin{pmatrix} -\frac{2}{3} & -\frac{5}{3} & \frac{2}{3} \\ 1 & 2 & -1 \\ -1 & -2 & 2 \end{pmatrix}, \tag{4.94}$$

with

$$\mathrm{Det}\{\boldsymbol{T}\} = \frac{1}{3}, \qquad \boldsymbol{T}^{-1} = \begin{pmatrix} 6 & 6 & 1 \\ -3 & -2 & 0 \\ 0 & 1 & 1 \end{pmatrix}. \tag{4.95}$$

Then the companion realization is given by

$$\begin{aligned} \boldsymbol{A}' = \boldsymbol{T}\boldsymbol{A}\boldsymbol{T}^{-1} &= \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -6 & -11 & -6 \end{pmatrix}, \\ \boldsymbol{b}' = \boldsymbol{T}\boldsymbol{b} &= \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \qquad (\boldsymbol{c}')^{\mathrm{T}} = \boldsymbol{c}^{\mathrm{T}}\boldsymbol{T}^{-1} = \begin{pmatrix} 3 & 3 & 3 \end{pmatrix}. \end{aligned} \tag{4.96}$$

From (4.77) and (4.91) the observability matrix is

$$\boldsymbol{V} = \begin{pmatrix} -2 & -5 & 5 \\ -3 & 0 & -12 \\ 27 & 24 & 33 \end{pmatrix} \tag{4.97}$$

Since $\mathrm{Det}\{\boldsymbol{V}\} = 189$ a transformation to an alternative companion realization exists. Inverting the observability matrix

$$\boldsymbol{V}^{-1} = \begin{pmatrix} \frac{32}{21} & \frac{95}{63} & \frac{20}{63} \\ -\frac{25}{21} & -\frac{67}{63} & -\frac{13}{63} \\ -\frac{8}{21} & -\frac{29}{63} & -\frac{5}{63} \end{pmatrix}. \tag{4.98}$$

Take the last column of this matrix an denote it by $\boldsymbol{\chi}$. Then the matrix $\boldsymbol{T}_{\mathrm{A}}$ required to obtain alternative companion realization is given by

$$\begin{aligned} (\boldsymbol{T}_{\mathrm{A}})^{-1} &= \begin{pmatrix} \boldsymbol{\chi} & \boldsymbol{A}\boldsymbol{\chi} & \boldsymbol{A}^2\boldsymbol{\chi} \end{pmatrix} \\ &= \begin{pmatrix} \frac{20}{63} & -\frac{25}{63} & \frac{26}{63} \\ -\frac{13}{63} & \frac{11}{63} & \frac{2}{63} \\ -\frac{5}{63} & \frac{1}{63} & \frac{25}{63} \end{pmatrix} \end{aligned} \tag{4.99}$$

and

$$\boldsymbol{T}_{\mathrm{A}} = \begin{pmatrix} -13 & -31 & 16 \\ -15 & -30 & 18 \\ -2 & -5 & 5 \end{pmatrix}. \tag{4.100}$$

Then the alternative companion realization is given by

$$\begin{aligned} \boldsymbol{A}'' = \boldsymbol{T}_{\mathrm{A}}\boldsymbol{A}(\boldsymbol{T}_{\mathrm{A}})^{-1} &= \begin{pmatrix} 0 & 0 & -6 \\ 1 & 0 & -11 \\ 0 & 1 & -6 \end{pmatrix}, \\ \boldsymbol{b}'' = \boldsymbol{T}_{\mathrm{A}}\boldsymbol{b} &= \begin{pmatrix} 3 \\ 3 \\ 3 \end{pmatrix}, \qquad (\boldsymbol{c}'')^{\mathrm{T}} = \boldsymbol{c}^{\mathrm{T}}(\boldsymbol{T}_{\mathrm{A}})^{-1} = \begin{pmatrix} 0 & 0 & 1 \end{pmatrix}. \end{aligned} \tag{4.101}$$

4.3.2 The Diagonal and Jordan Realizations

We already know that the poles of the transfer function are the eigenvalues of the matrix $\boldsymbol{A}$ for any realization of the system. We also know that when $\boldsymbol{A}$ has distinct eigenvalues it can be diagonalized to $\boldsymbol{\Lambda}$ by the matrix $\boldsymbol{P}$ according to equations (4.15) and (4.16). The realization $[\boldsymbol{\Lambda}, \boldsymbol{b}'', (\boldsymbol{c}'')^{\mathrm{T}}]$, where $\boldsymbol{b}'' = \boldsymbol{P}\boldsymbol{b}$ and $(\boldsymbol{c}'')^{\mathrm{T}} = \boldsymbol{c}^{\mathrm{T}}\boldsymbol{P}^{-1}$ is called a *diagonal realization*. An alternative way of obtaining a diagonal realization is from the transfer function.

Example 4.3.2 Determine a diagonal realization for the system given by $\bar{y}(s) = G(s)\bar{u}(s)$, where

$$G(s) = \frac{3(s^2+s+1)}{(s+1)(s+2)(s+3)}. \tag{4.102}$$

First express the transfer function in partial fractions so that

$$\bar{y}(s) = \frac{3\bar{u}(s)}{2(s+1)} - \frac{9\bar{u}(s)}{(s+2)} + \frac{21\bar{u}(s)}{2(s+3)}. \tag{4.103}$$

Then define state variables $x_1(t)$, $x_2(t)$ and $x_3(t)$ with Laplace transforms related to the input variable by

$$\bar{x}_1(s) = \frac{\bar{u}(s)}{s+1}, \qquad \bar{x}_2(s) = \frac{\bar{u}(s)}{s+2}, \qquad \bar{x}_3(s) = \frac{\bar{u}(s)}{s+3}, \tag{4.104}$$

giving

$$\bar{y}(s) = \tfrac{3}{2}\bar{x}_1(s) - 9\bar{x}_2(s) + \tfrac{21}{2}\bar{x}_3(s). \tag{4.105}$$

Inverting the Laplace transforms and expressing the results in matrix form gives (4.44) and (4.45) with

$$\boldsymbol{A} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -3 \end{pmatrix}, \quad \boldsymbol{b} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \boldsymbol{c} = \begin{pmatrix} \frac{3}{2} \\ -9 \\ \frac{21}{2} \end{pmatrix}. \tag{4.106}$$

Example 4.3.3 Try applying the procedure of Example 4.3.2 when

$$G(s) = \frac{1}{(s+1)^3(s+2)(s+3)}. \tag{4.107}$$

This case corresponds to a 3-fold degenerate eigenvalue -1, so we do not expect to be able to obtain a diagonal realization. Resolving the transfer functions into partial fractions gives

$$\bar{y}(s) = \frac{\bar{u}(s)}{2(s+1)^3} - \frac{3\bar{u}(s)}{4(s+1)^2} + \frac{7\bar{u}(s)}{8(s+1)} - \frac{\bar{u}(s)}{(s+2)} + \frac{\bar{u}(s)}{8(s+3)}. \tag{4.108}$$

Let

$$\bar{x}_3(s) = \frac{\bar{u}(s)}{s+1}, \qquad \bar{x}_2(s) = \frac{\bar{x}_3(s)}{s+1} = \frac{\bar{u}(s)}{(s+1)^2},$$
$$\bar{x}_1(s) = \frac{\bar{x}_2(s)}{s+1} = \frac{\bar{u}(s)}{(s+1)^3}, \quad \bar{x}_4(s) = \frac{\bar{u}(s)}{s+2}, \quad \bar{x}_5(s) = \frac{\bar{u}(s)}{s+3} \tag{4.109}$$

Then

$$\begin{aligned} s\bar{x}_1(s) &= -\bar{x}_1(s) + \bar{x}_2(s), \\ s\bar{x}_2(s) &= -\bar{x}_2(s) + \bar{x}_3(s), \\ s\bar{x}_3(s) &= -\bar{x}_3(s) + \bar{u}(s), \\ s\bar{x}_4(s) &= -2\bar{x}_4(s) + \bar{u}(s), \\ s\bar{x}_5(s) &= -3\bar{x}_5(s) + \bar{u}(s), \end{aligned} \tag{4.110}$$

$$\bar{y}(s) = \tfrac{1}{2}\bar{x}_1(s) - \tfrac{3}{4}\bar{x}_2(s) + \tfrac{7}{8}\bar{x}_3(s) - \bar{x}_4(s) + \tfrac{1}{8}\bar{x}_5(s). \tag{4.111}$$

Inverting the Laplace transforms and expressing the results in matrix form gives (4.44) and (4.45) with

$$\boldsymbol{A} = \begin{pmatrix} -1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 0 & -3 \end{pmatrix}, \qquad \boldsymbol{b} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \tag{4.112}$$

$$\boldsymbol{c}^{\mathrm{T}} = \begin{pmatrix} \frac{1}{2} & -\frac{3}{4} & \frac{7}{8} & -1 & \frac{1}{8} \end{pmatrix}.$$

In this case $\boldsymbol{A}$ is diagonal in the rows and columns containing the non-degenerate eigenvalues -2 and -3 but has a 3×3 block corresponding to the 3-fold degenerate eigenvalue -1. The matrix is said to be in *Jordan canonical form* and this is an example of a *Jordan realization*.

We have seen that a system defined by its transfer function can lead to different realization of the same dimensions related by a transformation (4.48). In fact the realizations need not be of the same dimension.

Example 4.3.4 Consider the system with a realization for which

$$\boldsymbol{A} = \begin{pmatrix} -1 & 0 & 1 \\ -1 & -2 & -1 \\ -2 & -2 & -3 \end{pmatrix}, \qquad \boldsymbol{b} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \qquad \boldsymbol{c} = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}. \tag{4.113}$$

We can now find the transfer function using (4.54). The code to do the calculation in MAPLE is

```
> with(linalg):

Warning, new definition for norm

Warning, new definition for trace
> A:=array([[-1,0,1],[-1,-2,-1],[-2,-2,-3]]);
```

$$A := \begin{bmatrix} -1 & 0 & 1 \\ -1 & -2 & -1 \\ -2 & -2 & -3 \end{bmatrix}$$

```
> b:=array([[1],[0],[1]]);
```

$$b := \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

```
> ct:=array([[1,1,-1]]);
```

$$ct := \begin{bmatrix} 1 & 1 & -1 \end{bmatrix}$$

```
> II:=array([[1,0,0],[0,1,0],[0,0,1]]);
```

$$II := \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

```
> W:=s->inverse(s*II-A);
```

$$\mathrm{W} := s \to \mathrm{inverse}(s\, II - A)$$

```
> W(s);
```

$$\begin{bmatrix} \dfrac{s+4}{s^2+5\,s+6} & -2\,\dfrac{1}{s^3+6\,s^2+11\,s+6} & \dfrac{1}{s^2+4\,s+3} \\ -\dfrac{1}{s^2+5\,s+6} & \dfrac{s^2+4\,s+5}{s^3+6\,s^2+11\,s+6} & -\dfrac{1}{s^2+4\,s+3} \\ -2\,\dfrac{1}{s^2+5\,s+6} & -2\,\dfrac{1}{s^2+5\,s+6} & \dfrac{1}{s+3} \end{bmatrix}$$

```
> G:=s->simplify(multiply(ct,W(s),b));
```

$$G := s \to \text{simplify}(\text{multiply}(ct,\ \text{W}(s),\ b))$$

```
> G(s);
```

$$\left[\ 3\,\frac{1}{(s+3)\,(s+2)}\ \right]$$

It is not difficult to show that this system, with the transfer function

$$G(s) = \frac{3}{(s+2)(s+3)}, \tag{4.114}$$

also has the two-dimensional diagonal realization

$$\boldsymbol{A} = \begin{pmatrix} -2 & 0 \\ 0 & -3 \end{pmatrix}, \quad \boldsymbol{b} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \boldsymbol{c} = \begin{pmatrix} 3 \\ -3 \end{pmatrix}. \tag{4.115}$$

4.4 Controllability

As indicated above an essential step in dealing with many control problems is to determine whether a desired outcome can be achieved by manipulating the input (control) variable. The outcome is determined in terms of a particular set of state variables and the controllability is that of the realization rather than the system. In fact the outcome $y(t)$ does not play a role and the equation of interest is (4.44).

> *The realization given by (4.44) is* ***controllable*** *if, given a finite time interval $[t_0, t_{\mathrm{f}}]$ and state vectors $\boldsymbol{x}_0$ and $\boldsymbol{x}_{\mathrm{f}}$, we can find an input $u(t)$ over the interval $[t_0, t_{\mathrm{f}}]$ such that $\boldsymbol{x}(t_0) = \boldsymbol{x}_0$ and $\boldsymbol{x}(t_{\mathrm{f}}) = \boldsymbol{x}_{\mathrm{f}}$.*

Since we are concerned with a constant system we can, without loss of generality, set $t_0 = 0$. Then from (4.14),

$$\boldsymbol{x}_{\mathrm{f}} = \exp\left(\boldsymbol{A}t_{\mathrm{f}}\right)\left[\boldsymbol{x}_0 + \int_0^{t_{\mathrm{f}}} \exp(-\boldsymbol{A}\tau)\boldsymbol{b}u(\tau)\mathrm{d}\tau\right]. \tag{4.116}$$

This equation can be rewritten in the form

$$\exp\left(-\boldsymbol{A}t_{\mathrm{f}}\right)\boldsymbol{x}_{\mathrm{f}} - \boldsymbol{x}_0 = \int_0^{t_{\mathrm{f}}} \exp(-\boldsymbol{A}\tau)\boldsymbol{b}u(\tau)\mathrm{d}\tau. \tag{4.117}$$

Since $\boldsymbol{x}_{\mathrm{f}}$, $\boldsymbol{x}_0$ and t_{f} are all arbitrary it is both sufficient and necessary for controllability that for any state vector $\boldsymbol{x}^*$ and time interval t_{f} we can find an input $u(t)$ to satisfy

$$\boldsymbol{x}^* = \int_0^{t_{\mathrm{f}}} \exp(-\boldsymbol{A}\tau)\boldsymbol{b}u(\tau)\mathrm{d}\tau. \tag{4.118}$$

Theorem 4.4.1 (The controllability theorem.) *A realization is controllable if and only if the controllability matrix* $\boldsymbol{U}$ *given by (4.76) is non-singular.*

Proof:

We use the polynomial formula (4.33) for the exponential matrix and thus, from (4.118)

$$\begin{aligned} \boldsymbol{x}^* &= \sum_{k=0}^{n-1} \boldsymbol{A}^k \boldsymbol{b} \int_0^{t_\mathrm{f}} \beta_k(-\tau) u(\tau) \mathrm{d}\tau \\ &= \boldsymbol{U}\boldsymbol{\Omega} \end{aligned} \tag{4.119}$$

where

$$\boldsymbol{\Omega}^\mathrm{T} = \begin{pmatrix} \Omega_0 & \Omega_1 & \cdots & \Omega_{n-1} \end{pmatrix} \tag{4.120}$$

and

$$\Omega_k = \int_0^{t_\mathrm{f}} \beta_k(-\tau) u(\tau) \mathrm{d}\tau. \tag{4.121}$$

Sufficiency. If $\boldsymbol{U}$ is non-singular it has an inverse and the values of Ω_k, $k = 0, 1, \ldots, n-1$ are uniquely given by $\boldsymbol{\Omega} = \boldsymbol{U}^{-1}\boldsymbol{x}^*$. By using an input with n adjustable parameters these values can be satisfied by the integrals.

Necessity. If $\boldsymbol{U}$ is singular then the set of linear equations in $q_1, q_2, \ldots, q_n$ given by $\boldsymbol{q}^\mathrm{T}\boldsymbol{U} = 0$ has a non-trivial solution. Thus $\boldsymbol{q}^\mathrm{T}\boldsymbol{x}^* = \boldsymbol{q}^\mathrm{T}\boldsymbol{U}\boldsymbol{\Omega} = 0$. Which is true only if $\boldsymbol{x}^*$ is orthogonal to $\boldsymbol{q}$. So (4.118) cannot be satisfied for arbitrary $\boldsymbol{x}^*$.

Example 4.4.1 Investigate the controllability of the realization

$$\boldsymbol{A} = \begin{pmatrix} -2 & 1 \\ 0 & -2 \end{pmatrix}, \qquad \boldsymbol{b} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{4.122}$$

The controllability matrix is

$$\boldsymbol{U} = \begin{pmatrix} 0 & 1 \\ 1 & -2 \end{pmatrix} \tag{4.123}$$

with $\mathrm{Det}\{\boldsymbol{U}\} = -1$; so the realization is controllable and

$$\boldsymbol{U}^{-1} = \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix}. \tag{4.124}$$

The eigenvalues of $\boldsymbol{A}$ are both -2 so, using the method of Sect. 4.2,

$$\exp(\boldsymbol{A}t) = \beta_0(t)\boldsymbol{I} + \beta_1(t)\boldsymbol{A}, \tag{4.125}$$

where $\beta_0(t)$ and $\beta_1(t)$ are given by

$$\begin{aligned}\exp(-2t) &= \beta_0(t) - 2\beta_1(t),\\ t\exp(-2t) &= \beta_1(t).\end{aligned} \tag{4.126}$$

Thus

$$\begin{aligned}\exp(\boldsymbol{A}t) &= (1+2t)\exp(-2t)\boldsymbol{I} + t\exp(-2t)\boldsymbol{A}\\ &= \exp(-2t)\begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}\end{aligned} \tag{4.127}$$

and

$$\exp(-\boldsymbol{A}t) = \exp(2t)\begin{pmatrix} 1 & -t \\ 0 & 1 \end{pmatrix}. \tag{4.128}$$

From (4.121), (4.125) and (4.127)

$$\begin{aligned}\Omega_0 &= \int_0^{t_\mathrm{f}} [1-2\tau]\exp(2\tau)u(\tau)\mathrm{d}\tau,\\ \Omega_1 &= -\int_0^{t_\mathrm{f}} \tau\exp(2\tau)u(\tau)\mathrm{d}\tau.\end{aligned} \tag{4.129}$$

From (4.119) and (4.123)

$$x_1^* = \Omega_1, \qquad x_2^* = \Omega_0 - 2\Omega_1. \tag{4.130}$$

We *know* because the system is controllable that for any given t_f, x_1^* and x_2^* we can find a form for $u(t)$ which will satisfy (4.129) and (4.130). Suppose we try the form

$$u(t) = (\mathsf{A}_0 + \mathsf{A}_1 t)\exp(-2t). \tag{4.131}$$

Then by substituting into (4.129) and then (4.130)

$$\begin{aligned}x_1^* &= -\frac{1}{2}\mathsf{A}_0 t_\mathrm{f}^2 - \frac{1}{3}\mathsf{A}_1 t_\mathrm{f}^3,\\ x_2^* &= \mathsf{A}_0 t_\mathrm{f} + \frac{1}{2}\mathsf{A}_1 t_\mathrm{f}^2.\end{aligned} \tag{4.132}$$

The determinant of this pair of linear equations is $\frac{1}{12}t_\mathrm{f}^4$, so they have a unique solution for any t_f. In particular suppose we want to control the system to produce $x_1^* = x_2^* = 1$ in time $t_\mathrm{f} = 1$, then we choose

$$u(t) = (10 - 18t)\exp(-2t). \tag{4.133}$$

4.5 Observability

Allied to the concept of controllability is that of *observability*. The essence of this concept is that by observing the the input and output of a system over a period of time the state at the beginning of that time can be inferred. More formally:

> *The realization given by (4.44) and (4.45) is* **observable** *if, given a time t_0, there exists a time interval $[t_0, t_\mathrm{f}]$ such that the initial state $\boldsymbol{x}(t_0) = \boldsymbol{x}_0$ is determined by the input function $u(t)$ and the output function $y(t)$ over $[t_0, t_\mathrm{f}]$.*

As for controllability, since we are concerned with a constant system, we can set $t_0 = 0$. From (4.14) and (4.45)

$$y(t) = \boldsymbol{c}^{\mathrm{T}} \exp\left(\boldsymbol{A}t\right) \left[\boldsymbol{x}_0 + \int_0^t \exp(-\boldsymbol{A}\tau)\boldsymbol{b}u(\tau)\mathrm{d}\tau\right], \quad 0 \le t \le t_\mathrm{f}. \tag{4.134}$$

Let

$$y^*(t) = y(t) - \boldsymbol{c}^{\mathrm{T}} \exp\left(\boldsymbol{A}t\right) \int_0^t \exp(-\boldsymbol{A}\tau)\boldsymbol{b}u(\tau)\mathrm{d}\tau. \tag{4.135}$$

Then (4.134) is equivalent to

$$y^*(t) = \boldsymbol{c}^{\mathrm{T}} \exp\left(\boldsymbol{A}t\right) \boldsymbol{x}_0, \quad 0 \le t \le t_\mathrm{f} \tag{4.136}$$

and thus a realization is observable if $\boldsymbol{x}_0$ can be obtained from (4.136) for any arbitrary function $y^*(t)$.

Theorem 4.5.1 (The observability theorem.) *A realization is observable if and only if the observability matrix $\boldsymbol{V}$ given by (4.77) is non-singular.*

Proof:

We use the polynomial form (4.33) for the exponential matrix and, from (4.136)

$$\begin{aligned} y^*(t) &= \sum_{k=0}^{n-1} \beta_k(t)\boldsymbol{c}^{\mathrm{T}}\boldsymbol{A}^k\boldsymbol{x}_0 \\ &= \boldsymbol{\beta}^{\mathrm{T}}(t)\boldsymbol{V}\boldsymbol{x}_0, \end{aligned} \tag{4.137}$$

where

$$\boldsymbol{\beta}^{\mathrm{T}}(t) = \begin{pmatrix} \beta_0(t) & \beta_1(t) & \cdots & \beta_{n-1}(t) \end{pmatrix} \tag{4.138}$$

Sufficiency. Suppose $\boldsymbol{V}$ is non-singular and suppose that two states $\boldsymbol{x}_0$ and $\boldsymbol{x}_0'$ both satisfy (4.137). Then

$$\boldsymbol{\beta}^{\mathrm{T}}(t)\boldsymbol{V}\triangle\boldsymbol{x}_0 = 0, \tag{4.139}$$

where $\triangle\boldsymbol{x}_0 = \boldsymbol{x}_0 - \boldsymbol{x}_0'$. Since (4.139) holds for all t in the interval $[0, t_{\mathrm{f}}]$, it must be the case that

$$\boldsymbol{V}\triangle\boldsymbol{x}_0 = \boldsymbol{0}, \tag{4.140}$$

Since $\boldsymbol{V}$ is non-singular (4.140) has the unique solution $\triangle\boldsymbol{x}_0 = \boldsymbol{0}$. That is $\boldsymbol{x}_0 = \boldsymbol{x}_0'$, which means that the realization is observable.

Necessity. Suppose that $\boldsymbol{V}$ is singular. Then there exists a vector $\boldsymbol{p}$ such the $\boldsymbol{V}\boldsymbol{p} = \boldsymbol{0}$. Then if $\boldsymbol{x}_0$ satisfies (4.137) so does $\boldsymbol{x}_0 + \mu\boldsymbol{p}$ for any μ and the realization is not observable.

It is clear from these results that a realization is controllable if and only if its dual realization is observable.

4.6 Minimal Realizations

We have seen that for a system with transfer function $G(s)$ any realization $[\boldsymbol{A}, \boldsymbol{b}, \boldsymbol{c}^{\mathrm{T}}]$ must satisfy (4.54) and equivalently (4.56). We have also seen by means of example 4.3.4 that the dimension n need not be the same for all realizations of a system. Let n_{min} be the least possible dimension for any realization of a system. Then any realization which has this dimension is called *minimal*. We state without proof an important theorem.

Theorem 4.6.1 (The minimality theorem.) *A realization is minimal if and only if it is both controllable and observable.*

An obvious corollary to this theorem which scarcely needs proof is:

Theorem 4.6.2 *If a realization of dimension n is*

(i) controllable and observable then all realizations of dimension n are controllable and observable.

(ii) not both controllable and observable then no realization of dimension n is both controllable and observable.

Example 4.6.1 Compare the realizations given by (4.91) and (4.113).

They differ only in the vector $\boldsymbol{c}$. We showed in Example 4.3.1 that (4.91) was both controllable and observable and it is therefore minimal, with $n_{\mathrm{min}} = 3$. In Example 4.3.4 we derived the transfer function (4.114) corresponding to (4.113) and showed that it also had a realization of dimension 2. This means, of course, that (4.113) cannot be minimal. Since it is controllable (in this respect it is identical to (4.91)) it cannot be observable. We confirm this by working out the observability matrix. From (4.77) and (4.113)

$$\boldsymbol{V} = \begin{pmatrix} 1 & 1 & -1 \\ 0 & 0 & 3 \\ -6 & -6 & -9 \end{pmatrix} \tag{4.141}$$

which is clearly singular. So this realization is controllable but not observable. What about the two-dimensional realization of the same system given by (4.115)? From this equation and (4.76) and (4.77)

$$\boldsymbol{U} = \begin{pmatrix} 1 & -2 \\ 1 & -3 \end{pmatrix}, \qquad \boldsymbol{V} = \begin{pmatrix} 3 & -3 \\ -6 & 9 \end{pmatrix}. \tag{4.142}$$

Since neither of these matrices is singular the realization is both controllable and observable and therefore minimal with $n_{\min} = 2$. All other two-dimensional realizations will also be both controllable and observable and no realization of larger dimension can be both.

Example 4.6.2 Now consider the system with transfer function

$$G(s) = \frac{(s+c)}{(s+1)(s+2)^2}, \tag{4.143}$$

where c is some constant.

Resolving the transfer function into partial fractions gives

$$\bar{y}(s) = \frac{\bar{u}(s)(c-1)}{(s+1)} + \frac{\bar{u}(s)(2-c)}{(s+2)^2} + \frac{\bar{u}(s)(1-c)}{(s+2)}. \tag{4.144}$$

Let

$$\bar{x}_1(s) = \frac{\bar{u}(s)}{(s+1)}, \qquad \bar{x}_3(s) = \frac{\bar{u}(s)}{s+2}, \qquad \bar{x}_2(s) = \frac{\bar{x}_3(s)}{s+2} = \frac{\bar{u}(s)}{(s+2)^2}. \tag{4.145}$$

Then

$$\bar{y}(s) = (c-1)\bar{x}_1(s) + (2-c)\bar{x}_2(s) + (1-c)\bar{x}_3(s). \tag{4.146}$$

Inverting the Laplace transforms and expressing the results in matrix form gives (4.44) and (4.45) with

$$\boldsymbol{A} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -2 & 1 \\ 0 & 0 & -2 \end{pmatrix}, \quad \boldsymbol{b} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \tag{4.147}$$

$$\boldsymbol{c}^{\mathrm{T}} = \begin{pmatrix} c-1 & 2-c & 1-c \end{pmatrix}.$$

Since $\boldsymbol{A}$ and $\boldsymbol{b}$ are not functions of c it is clear that the controllability or uncontrollability of the system is not affected by the value of c. From (4.76) and (4.147)

$$\boldsymbol{U} = \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & -4 \\ 1 & -2 & 4 \end{pmatrix}. \tag{4.148}$$

Since $\mathrm{Det}\{\boldsymbol{U}\} = -1$ the realization is controllable. Now from (4.77) and (4.147)

$$\boldsymbol{V} = \begin{pmatrix} c-1 & 2-c & 1-c \\ 1-c & 2c-4 & c \\ c-1 & 8-4c & -4 \end{pmatrix}. \tag{4.149}$$

Since $\mathrm{Det}\{\boldsymbol{V}\} = (c-1)(c-2)^2$ the realization is observable unless $c = 1$ or $c = 2$. These are precisely the cases where there is cancellation in the transfer function leading to a quadratic denominator. The procedure for deriving a realization used here will now lead to two-dimensional realizations. You can easily check that

<u>For $c = 1$</u>

$$\boldsymbol{A} = \begin{pmatrix} -2 & 1 \\ 0 & -2 \end{pmatrix}, \qquad \boldsymbol{b} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tag{4.150}$$

$$\boldsymbol{c}^{\mathrm{T}} = \begin{pmatrix} 1 & 0 \end{pmatrix}.$$

with

$$\boldsymbol{U} = \begin{pmatrix} 0 & 1 \\ 1 & -2 \end{pmatrix}, \qquad \boldsymbol{V} = \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix}. \tag{4.151}$$

<u>For $c = 2$</u>

$$\boldsymbol{A} = \begin{pmatrix} -1 & 0 \\ 0 & -2 \end{pmatrix}, \qquad \boldsymbol{b} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \tag{4.152}$$

$$\boldsymbol{c}^{\mathrm{T}} = \begin{pmatrix} 1 & -1 \end{pmatrix}.$$

with

$$\boldsymbol{U} = \begin{pmatrix} 1 & -1 \\ 1 & -2 \end{pmatrix}, \qquad \boldsymbol{V} = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}. \tag{4.153}$$

In each case the realization is both controllable and observable and this minimal with $n_{\min} = 2$.

It is clear, from this example, that the dimension of a realization, that is the number of state variables, corresponds, when it is derived from a transfer function $G(s)$, to the number of partial fractions of $G(s)$. This in term is simply the degree of the polynomial denominator of $G(s)$. In the cases we have observed, the realization was minimal unless there was some cancellation in factors between the numerator and denominator of $G(s)$. The following theorem, therefore, comes as no surprise.

Theorem 4.6.3 *Suppose that a system is given by (4.61)–(4.63) where $m < n$. Then $\phi(s)$ and $\psi(s)$ have no common factors, the degree n of the denominator $\phi(s)$ is the dimension of minimal realizations of the system.*

Problems 4

1) Let $\boldsymbol{A}$ be a 2×2 matrix with eigenvalues λ and μ. Show that, when $\lambda \neq \mu$,

$$\exp(\boldsymbol{A}t) = \frac{[\lambda \exp(\mu t) - \mu \exp(\lambda t)]\boldsymbol{I} + [\exp(\lambda t) - \exp(\mu t)]\boldsymbol{A}}{\lambda - \mu}.$$

What is the corresponding result when $\lambda = \mu$?

2) Consider the system with realization

$$\begin{aligned} \dot{\boldsymbol{x}}(t) &= \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2 & 1 & -2 \end{pmatrix} \boldsymbol{x}(t) + \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} u(t), \\ y(t) &= \begin{pmatrix} 1 & 2 & 0 \end{pmatrix} \boldsymbol{x}(t). \end{aligned}$$

Use the exponential matrix to find the output $y(t)$, for $u(t) = \mathsf{K}t$, where K is a constant and $\boldsymbol{x}(0) = \boldsymbol{0}$.

3) For the system with realization

$$\begin{aligned} \dot{\boldsymbol{x}}(t) &= \begin{pmatrix} -\frac{3}{4} & -\frac{1}{4} \\ -\frac{1}{2} & -\frac{1}{2} \end{pmatrix} \boldsymbol{x}(t) + \begin{pmatrix} 1 \\ 1 \end{pmatrix} u(t), \\ y(t) &= \begin{pmatrix} 4 & 2 \end{pmatrix} \boldsymbol{x}(t) \end{aligned}$$

calculate the transfer function $G(s)$, the controllability matrix $\boldsymbol{U}$ and the observability matrix $\boldsymbol{V}$.

4) For the system with realization

$$\begin{aligned}\dot{\boldsymbol{x}}(t) &= \begin{pmatrix} -1 & -1 \\ 2 & -4 \end{pmatrix} \boldsymbol{x}(t) + \begin{pmatrix} 1 \\ 3 \end{pmatrix} u(t), \\ y(t) &= \begin{pmatrix} -1 & 1 \end{pmatrix} \boldsymbol{x}(t)\end{aligned}$$

verify that the controllability matrix is non-singular and find the matrix $\boldsymbol{T}$ which transforms it to the companion realization.

5) Obtain Jordan representations for the systems which have transfer functions:

(i) $G(s) = \dfrac{s^2 + s + 1}{(s+1)^3}$,

(ii) $G(s) = \dfrac{4}{(s+1)^2(s+3)}$.

Chapter 5

Stability

5.1 Two Types of Stability

Given an $m \times n$ matrix $\boldsymbol{Q}$ with elements Q_{ij}, the *Euclidean norm* $||\boldsymbol{Q}||$ of $\boldsymbol{Q}$ is defined by

$$||\boldsymbol{Q}|| = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}|Q_{ij}|^2}. \tag{5.1}$$

The Euclidean norm of a vector of real elements is, of course, the 'usual' modulus of the vector. For any two matrices $\boldsymbol{P}$ and $\boldsymbol{Q}$ (with appropriate dimensions) and any scalar μ the following properties hold:

(i) $||\boldsymbol{Q}|| > 0$ unless $\boldsymbol{Q} = \boldsymbol{0}$.

(ii) $||\mu\boldsymbol{Q}|| = |\mu|\,||\boldsymbol{Q}||$.

(iii) $||\boldsymbol{P}+\boldsymbol{Q}|| \le ||\boldsymbol{P}|| + ||\boldsymbol{Q}||$.

(iv) $||\boldsymbol{P}\boldsymbol{Q}|| \le ||\boldsymbol{P}||\,||\boldsymbol{Q}||$.

*The $n \times n$ matrix $\boldsymbol{A}$ is called a **stability matrix** if each of its eigenvalues has a (strictly) negative real part.*

Theorem 5.1.1 *If the $n \times n$ matrix $\boldsymbol{A}$ is a stability matrix then there exist positive constants K and k such that*

$$||\exp(\boldsymbol{A}t)|| \le \mathsf{K}\exp(-kt) \tag{5.2}$$

for all $t \ge 0$ and hence

$$\exp(\boldsymbol{A}t) \to \boldsymbol{0}, \qquad as \qquad t \to \infty. \tag{5.3}$$

If $\boldsymbol{A}$ is not a stability matrix (5.3) does not hold.

In Sect. 1.6.1 we discussed the dynamic system

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{F}(\boldsymbol{x}) \tag{5.4}$$

and defined what it meant to say that the equilibrium point $\boldsymbol{x}^*$, which satisfies

$$\boldsymbol{F}(\boldsymbol{x}^*) = \boldsymbol{0}, \tag{5.5}$$

is *asymptotically stable* (in the sense of Lyapunov). For a linear autonomous system (5.4) takes the form

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{A}\boldsymbol{x}(t), \tag{5.6}$$

where $\boldsymbol{A}$ is a constant matrix. In this case the only equilibrium point is $\boldsymbol{x}^* = \boldsymbol{0}$ and we have the following result:

Theorem 5.1.2 (Cf. Thm. 1.6.1.) $\boldsymbol{x}^* = \boldsymbol{0}$ *is an asymptotically stable equilibrium point of (5.6) if and only if* $\boldsymbol{A}$ *is a stability matrix.*

Proof:
According to the definition, $\boldsymbol{x}^* = \boldsymbol{0}$ is an asymptotically stable equilibrium point of (5.6) if $\boldsymbol{x}(t) \to \boldsymbol{0}$ as $t \to \infty$. The solution of (5.6) is

$$\boldsymbol{x}(t) = \exp[\boldsymbol{A}(t - t_0)]\boldsymbol{x}(t_0). \tag{5.7}$$

and thus $\boldsymbol{x}(t) \to \boldsymbol{0}$ as $t \to \infty$ if and only if $\exp[\boldsymbol{A}(t - t_0)] \to \boldsymbol{0}$ as $t \to \infty$. It follows from Thm. 5.1.1 that $\boldsymbol{x}^* = \boldsymbol{0}$ is an asymptotically stable equilibrium point if and only if $\boldsymbol{A}$ is a stability matrix. This result provides most of a rather belated proof of Thm. 1.6.1.

We now need to to generalize the discussion of stability to the case of a system with input $u(t)$, output $y(t)$ and a realization

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{b}\,u(t), \tag{5.8}$$

$$y(t) = \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}(t). \tag{5.9}$$

The system with realization (5.8)–(5.9) is said to be ***bounded input–bounded output stable*** *for* $t \geq t_0$ *if any bounded input produces a bounded output. That is, given, that there exists a positive constant* $\mathcal{B}_1$*, such that*

$$|u(t)| < \mathcal{B}_1, \quad \textit{for all} \quad t \geq t_0, \tag{5.10}$$

then there exists a positive constant $\mathcal{B}_2$*, such that*

$$|y(t)| < \mathcal{B}_2, \quad \textit{for all} \quad t \geq t_0, \tag{5.11}$$

regardless of the initial state $\boldsymbol{x}(t_0)$*.*

Two important theorems relate these two types of stability.

Theorem 5.1.3 *If $\boldsymbol{x}^* = \mathbf{0}$ is an asymptotically stable equilibrium point of (5.6), then the system given by (5.8)–(5.9) is bounded input–bounded output stable for all t_0.*

Proof: From (5.9) and (4.134), with the intial time $t = t_0$ replacing $t = 0$,

$$y(t) = \boldsymbol{c}^{\mathrm{T}} \exp[\boldsymbol{A}(t-t_0)]\boldsymbol{x}(t_0) + \boldsymbol{c}^{\mathrm{T}} \int_{t_0}^{t} \exp[\boldsymbol{A}(t-\tau)]\boldsymbol{b}u(\tau)\mathrm{d}\tau. \tag{5.12}$$

Since $\boldsymbol{A}$ is a stability matrix, it follows from Thm. 5.1.1 that there exist positive constants K and k such that

$$\begin{aligned}
||y(t)|| &\leq ||\boldsymbol{c}||\,||\exp[\boldsymbol{A}(t-t_0)]||\,||\boldsymbol{x}(t_0)|| + ||\boldsymbol{c}|| \int_{t_0}^{t} ||\exp[\boldsymbol{A}(t-\tau)]||\,||\boldsymbol{b}||\,|u(\tau)|\mathrm{d}\tau \\
&\leq \mathsf{K}||\boldsymbol{c}|| \left[||\boldsymbol{x}(t_0)|| \exp[k(t_0-t)] + ||\boldsymbol{b}||\mathcal{B}_1 \int_{t_0}^{t} \exp[k(\tau-t)]\mathrm{d}\tau\right] \\
&= \mathsf{K}||\boldsymbol{c}|| \left\{||\boldsymbol{x}(t_0)|| \exp[k(t_0-t)] + k^{-1}||\boldsymbol{b}||\mathcal{B}_1[1-\exp[k(t_0-t)]]\right\} \\
&\leq \mathsf{K}||\boldsymbol{c}|| \left\{||\boldsymbol{x}(t_0)|| + k^{-1}||\boldsymbol{b}||\mathcal{B}_1\right\}.
\end{aligned} \tag{5.13}$$

So the system is bounded input–bounded output stable.

Theorem 5.1.4 *If the system with a minimal realization (5.8)–(5.9) is bounded input–bounded output stable for all t_0 then $\boldsymbol{x}^* = \mathbf{0}$ is an asymptotically stable equilibrium point of (5.6).*

The proof of this theorem is somewhat more complicated and will be omitted. We can see on the basis of these two theorems that the asymptotic stability of $\boldsymbol{x}^* = \mathbf{0}$ is a stronger condition than bounded input–bounded output stability and can be deduced from the bounded input–bounded output stability only for minimal[1] realizations.

5.2 Stability and the Transfer Function

In the development of the companion realization in Sect. 4.3.1 we Laplace transformed (4.59), assuming, with (4.60), that all the initial conditions were zero. If this assumption is not made (4.61) is replaced by

$$\bar{y}(s) = G(s)\bar{u}(s) + G_0(s), \tag{5.14}$$

[1] That is, from Thm. 4.6.1, controllable and observable.

where

$$G(s) = \frac{\psi(s)}{\phi(s)}, \qquad G_0(s) = \frac{\theta(s)}{\phi(s)} \tag{5.15}$$

are respectively the transfer function, as before, and the contribution from the initial conditions, with

$$\begin{aligned} \phi(s) &= s^n + a_{n-1}s^{n-1} + \cdots + a_1 s + a_0, \\ \psi(s) &= b_m s^m + b_{m-1}s^{m-1} + \cdots + b_1 s + b_0, \\ \theta(s) &= c_r s^r + c_{r-1}s^{r-1} + \cdots + c_1 s + c_0. \end{aligned} \tag{5.16}$$

Both m and r are less than n. Suppose now that $\phi(s)$ and $\psi(s)$ have no common factors. It then follows from Thm. 4.6.3 that any realization derived from (5.14) will be minimal with the characteristic polynomial of $\boldsymbol{A}$ being $\phi(\lambda)$. Thus, from Thms. 5.1.2–5.1.4,

Theorem 5.2.1 *If (5.8)–(5.9) is a realization derived from the transfer function, $G(s) = \psi(s)/\phi(s)$, where $\psi(s)$ and $\phi(s)$ have no common factors, then it will be bounded input–bounded output stable, and the corresponding system (5.6) will have $\boldsymbol{x}^* = \mathbf{0}$ as an asymptotically stable equilibrium point, if all the zeros of $\phi(s)$ have (strictly) negative real parts.*

So far we have used the phrase *asymptotically stable* qualified by reference to the origin for the system (5.6). *We shall henceforth use it freely for the system itself.* We now know that the criterion for asymptotic stability is that the poles of the transfer function or equivalently the eigenvalues of the matrix $\boldsymbol{A}$ must all have (strictly) negative real parts. In each case there is a qualification. For the transfer function it is that the denominator and numerator of $G(s)$ do not have a common factor and equivalently for the matrix $\boldsymbol{A}$ the realization must be minimal. It is easy to see that these conditions are important. Otherwise we could remove the stability of the system by multiplying the transfer function top and bottom by $(s - \omega)$ for some ω with $\Re\{\omega\} > 0$. In the corresponding realization derived from this there would be a new eigenvalue ω, but, of course the realization would no longer be minimal. When at least one of the poles of the transfer function has (strictly) positive part the system (like the equilibrium point at the origin for (5.6) is called *unstable*. The intermediate case, where the system is not unstable but some of the poles have zero real part is sometimes called *stable* and sometimes *conditionally* or *marginally stable*.

5.2.1 The Routh–Hurwitz Criterion

We now concentrate on using Thm. 5.2.1 to determine the stability of a system from the location of the zeros of the polynomial[2]

$$\phi(s) = a_n s^n + a_{n-1}s^{n-1} + \cdots + a_1 s + a_0. \tag{5.17}$$

[2]For convenience we include a leading coefficient of a_n which can always, as in (5.16) be set equal to one.

We define the $n \times n$ determinant

$$\Phi_n = \begin{vmatrix} a_{n-1} & a_n & 0 & 0 & 0 & 0 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ a_{n-3} & a_{n-2} & a_{n-1} & a_n & 0 & 0 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ a_{n-5} & a_{n-4} & a_{n-3} & a_{n-2} & a_{n-1} & a_n & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 & a_0 & a_1 & a_2 & a_3 & a_4 \\ 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 & 0 & 0 & a_0 & a_1 & a_2 \\ 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 & 0 & 0 & 0 & 0 & a_0 \end{vmatrix} \tag{5.18}$$

The way to build this determinant is as follows:

(i) Extend the range of the definition of the coefficients over all integer values by defining $a_\ell = 0$ if $\ell > n$ or $\ell < 0$.

(ii) The i–j-th element is a_{n-2i+j}.

Thus

(a) For $n = 2$

$$\phi(s) = a_2 s^2 + a_1 s + a_0, \tag{5.19}$$

$$\Phi_2 = \begin{vmatrix} a_1 & a_2 \\ a_{-1} & a_0 \end{vmatrix} = \begin{vmatrix} a_1 & a_2 \\ 0 & a_0 \end{vmatrix}. \tag{5.20}$$

(b) For $n = 3$

$$\phi(s) = a_3 s^3 + a_2 s^2 + a_1 s + a_0, \tag{5.21}$$

$$\Phi_3 = \begin{vmatrix} a_2 & a_3 & a_4 \\ a_0 & a_1 & a_2 \\ a_{-2} & a_{-1} & a_0 \end{vmatrix} = \begin{vmatrix} a_2 & a_3 & 0 \\ a_0 & a_1 & a_2 \\ 0 & 0 & a_0 \end{vmatrix}. \tag{5.22}$$

(c) For $n = 4$

$$\phi(s) = a_4 s^4 + a_3 s^3 + a_2 s^2 + a_1 s + a_0, \tag{5.23}$$

$$\Phi_4 = \begin{vmatrix} a_3 & a_4 & a_5 & a_6 \\ a_1 & a_2 & a_3 & a_4 \\ a_{-1} & a_0 & a_1 & a_2 \\ a_{-3} & a_{-2} & a_{-1} & a_0 \end{vmatrix} = \begin{vmatrix} a_3 & a_4 & 0 & 0 \\ a_1 & a_2 & a_3 & a_4 \\ 0 & a_0 & a_1 & a_2 \\ 0 & 0 & 0 & a_0 \end{vmatrix}. \tag{5.24}$$

For any n we now define a hierarchy of subdeterminants $\Phi_n^{(k)}$, $k = 1, 2, \ldots, n-1$,[3] where $\Phi_n^{(k)}$ is the $(n-k) \times (n-k)$ determinant obtained by deleting the last k rows and columns from Φ_n. Thus, for example,

$$\Phi_4^{(2)} = \begin{vmatrix} a_3 & a_4 \\ a_1 & a_2 \end{vmatrix}. \tag{5.25}$$

Then the following theorem holds:

Theorem 5.2.2 (The Routh–Hurwitz criterion.) *If $\phi(s)$, given by (5.17), has $a_n > 0$ then the roots of $\phi(s) = 0$ all have (strictly) negative real parts if and only if*

$$\Phi_n^{(k)} > 0, \qquad \textit{for all} \quad k = 0, 1, \ldots, n-1. \tag{5.26}$$

From (5.18)

$$\Phi_n^{(n-1)} = a_{n-1}, \tag{5.27}$$

$$\Phi_n = \Phi_n^{(0)} = a_0 \Phi_n^{(1)}. \tag{5.28}$$

These results mean that (5.26) can be confined to the range $k = 1, 2, \ldots, n-2$ with the additional conditions $a_{n-1} > 0$ and $a_0 > 0$.

Example 5.2.1 Investigate the roots of the cubic equation

$$s^3 + \kappa s^2 + 3s + 2 = 0 \tag{5.29}$$

as κ varies.

The conditions $a_3 = 1 > 0$ and $a_0 = 2 > 0$ are satisfied and another necessary condition for all the roots to have negative real part is $a_2 = \kappa > 0$. The only other condition for sufficiency is, from (5.22), given using

$$\Phi_3^{(1)} = \begin{vmatrix} \kappa & 1 \\ 2 & 3 \end{vmatrix} = 3\kappa - 2. \tag{5.30}$$

So according to the Routh-Hurwitz criterion the roots all have a negative real part if $\kappa > 2/3$. We can check this out using MAPLE .

```
> phi:=(s,k)->s^3+k*s^2+3*s+2;
```

$$\phi := (s, k) \to s^3 + k\, s^2 + 3\, s + 2$$

```
> fsolve(phi(s,1)=0,s,complex);
```

[3] We shall also, for convenience, use $\Phi_n^{(0)} = \Phi_n$.

$$-.7152252384,\ -.1423873808 - 1.666147574\,I,\ -.1423873808 + 1.666147574\,I$$

```
> fsolve(phi(s,3/4)=0,s,complex);
```

$$-.6777314603,\ -.03613426987 - 1.717473625\,I,\ -.03613426987 + 1.717473625\,I$$

```
> fsolve(phi(s,2/3)=0,s,complex);
```

$$-.6666666667,\ -1.732050808\,I,\ 1.732050808\,I$$

```
> fsolve(phi(s,1/2)=0,s,complex);
```

$$-.6462972136,\ .07314860682 - 1.757612233\,I,\ .07314860682 + 1.757612233\,I$$

```
> fsolve(phi(s,0)=0,s,complex);
```

$$-.5960716380,\ .2980358190 - 1.807339494\,I,\ .2980358190 + 1.807339494\,I$$

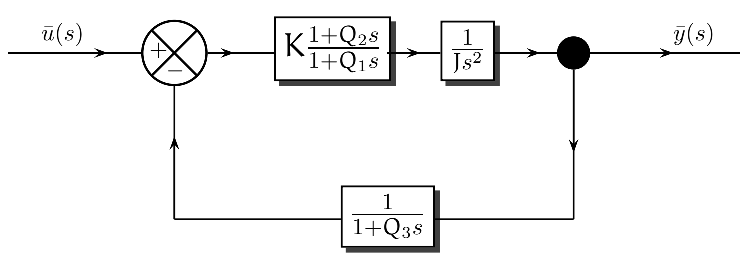
Example 5.2.2 Show, using the Routh-Hurwitz criterion, that, when $a_0, \dots, a_4$ are all real and greater than zero, the roots of the equation

$$\phi(s) = a_4 s^4 + a_3 s^3 + a_2 s^2 + a_1 s + a_0 = 0, \tag{5.31}$$

all have (strictly) negative real parts if and only if

$$a_1(a_2 a_3 - a_1 a_4) > a_0 a_3^2. \tag{5.32}$$

Hence show that the system given by this block diagram



where $\mathsf{Q}_1, \mathsf{Q}_2, \mathsf{Q}_3, \mathsf{J}, \mathsf{K}$ are all positive, can be stable only if

$$\mathsf{Q}_2 > \mathsf{Q}_1 + \mathsf{Q}_3. \tag{5.33}$$

With this condition satisfied, find the maximum value of K for stability.

In this example we are given that $a_n = a_4$, $a_{n-1} = a_3$ and a_0 are all positive. The only remaining conditions for the zeros to have negative real parts are

$$\Phi_4^{(1)} = \begin{vmatrix} a_3 & a_4 & 0 \\ a_1 & a_2 & a_3 \\ 0 & a_0 & a_1 \end{vmatrix} = a_1(a_2a_3 - a_1a_4) - a_0a_3^2 > 0, \tag{5.34}$$

$$\Phi_4^{(2)} = \begin{vmatrix} a_3 & a_4 \\ a_1 & a_2 \end{vmatrix} = a_2a_3 - a_1a_4 > 0, \tag{5.35}$$

Condition (5.34) is equivalent to (5.32) and if it is satisfied, then because all the coefficients are positive, (5.35) is automatically satisfied. If we use the same intermediate variables for this block diagram as we did in Example 3.4.1

$$\begin{aligned} \bar{u}(s) &= \bar{v}(s) + \bar{f}(s), \\ \bar{f}(s) &= \frac{\bar{y}(s)}{1 + \mathsf{Q}_3 s}, \\ \bar{x}(s) &= \mathsf{K}\frac{1 + \mathsf{Q}_2 s}{1 + \mathsf{Q}_1 s}\bar{v}(s), \\ \bar{y}(s) &= \frac{\bar{x}(s)}{\mathsf{J}s^2}. \end{aligned} \tag{5.36}$$

Eliminating the intermediate variables, the transfer function is

$$G(s) = \frac{\mathsf{K}(1 + \mathsf{Q}_2 s)(1 + \mathsf{Q}_3 s)}{(1 + \mathsf{Q}_1 s)(1 + \mathsf{Q}_3 s)\mathsf{J}s^2 + \mathsf{K}(1 + \mathsf{Q}_2 s)}. \tag{5.37}$$

The stability condition is deduced from the denominator

$$\phi(s) = \mathsf{J}\mathsf{Q}_1\mathsf{Q}_3\, s^4 + \mathsf{J}(\mathsf{Q}_1 + \mathsf{Q}_3)\, s^3 + \mathsf{J}\, s^2 + \mathsf{K}\mathsf{Q}_2\, s + \mathsf{K}. \tag{5.38}$$

Substituting into (5.32) gives

$$\mathsf{Q}_2[\mathsf{J}(\mathsf{Q}_1 + \mathsf{Q}_3) - \mathsf{K}\mathsf{Q}_1\mathsf{Q}_2\mathsf{Q}_3] > \mathsf{J}(\mathsf{Q}_1 + \mathsf{Q}_3)^2. \tag{5.39}$$

This simplifies to

$$\mathsf{J}(\mathsf{Q}_1 + \mathsf{Q}_3)(\mathsf{Q}_2 - \mathsf{Q}_1 - \mathsf{Q}_3) > \mathsf{K}\mathsf{Q}_1\mathsf{Q}_2^2\mathsf{Q}_3, \tag{5.40}$$

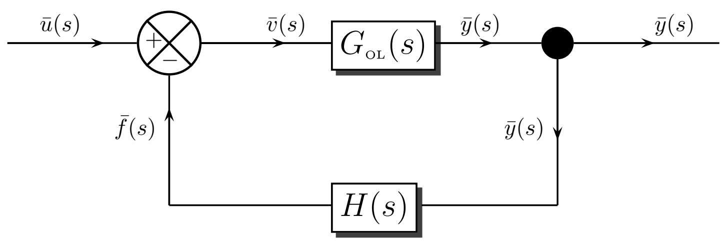
which shows that (5.33) is a necessary condition for stability. The maximum value of K is given, from (5.40) by

$$\frac{\mathsf{J}(\mathsf{Q}_1 + \mathsf{Q}_3)(\mathsf{Q}_2 - \mathsf{Q}_1 - \mathsf{Q}_3)}{\mathsf{Q}_1\mathsf{Q}_2^2\mathsf{Q}_3} > \mathsf{K}. \tag{5.41}$$

5.3 Stability and Feedback

5.3.1 Output Feedback

In Sect. 3.4 we discussed the introduction of a feedback $H(s)$ into a system giving a block diagram of the form:



The transfer function in the absence of feedback would be $G_{\text{OL}}(s)$ and in the presence of feedback it becomes

$$G_{\text{CL}}(s) = \frac{G_{\text{OL}}(s)}{1 + G_{\text{OL}}(s)H(s)}. \tag{5.42}$$

$G_{\text{OL}}(s)$ and $G_{\text{CL}}(s)$ are often referred to as *open-loop* and *closed-loop* transfer functions (hence the notation) and the signal

$$\bar{v}(s) = \bar{u}(s) - \bar{f}(s) = \bar{u}(s) - H(s)\bar{y}(s) \tag{5.43}$$

is called the *error*. In the present context this type of feedback will be called *output feedback*. For a system which is unstable it is sometimes possible, as we saw in Example 3.4.2, to achieve stability by altering parameters in the feedback. The way this happens can be seen if we let

$$G_{\text{CL}}(s) = \frac{\psi(s)}{\phi(s)}, \tag{5.44}$$

with

$$\begin{aligned} \phi(s) &= s^n + a_{n-1}s^{n-1} + \cdots + a_1 s + a_0, \\ \psi(s) &= b_m s^m + b_{m-1}s^{m-1} + \cdots + b_1 s + b_0. \end{aligned} \tag{5.45}$$

Now suppose that the change in feedback leads to the modification of the closed-loop transfer function to

$$G'_{\text{CL}}(s) = \frac{\psi(s)}{\phi(s) + \chi(s)}, \tag{5.46}$$

where

$$\chi(s) = h_{n-1}s^{n-1} + \cdots + h_1 s + h_0. \tag{5.47}$$

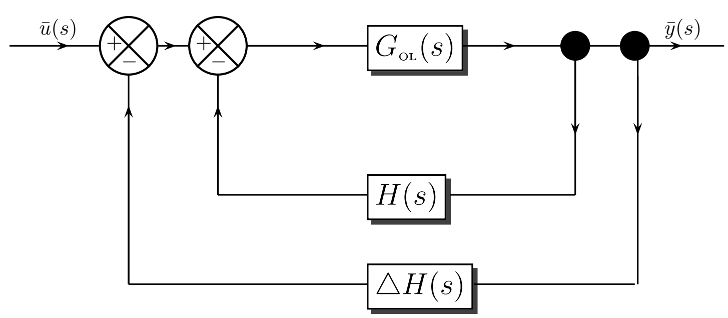
We know that the system is stable if the zeros of the denominator of the closed-loop transfer function all have negative real parts and that this can be ascertained using the Routh-Hurwitz criterion. If the system is unstable with $\phi(s)$ alone in the denominator, then the introduction of the extra factor $\chi(s)$ can be used to stabilize it. Now suppose that the introduction of $\chi(s)$ corresponds to replacing the feedback $H(s)$ in the block diagram by $H(s) + \triangle H(s)$. Then

$$\frac{G_{\mathrm{OL}}(s)}{1 + G_{\mathrm{OL}}(s)[H(s) + \triangle H(s)]} = \frac{\psi(s)}{\phi(s) + \chi(s)}. \tag{5.48}$$

From (5.42), (5.44) and (5.48)

$$\triangle H(s) = \frac{\chi(s)}{\psi(s)}. \tag{5.49}$$

Although we have considered $\triangle H(s)$ simply to be an additive factor in the feedback, it could also be applied as an additional feedback on the system. Consider the block diagram



From (5.42)

$$G'_{\mathrm{CL}}(s) = \frac{G_{\mathrm{CL}}(s)}{1 + G_{\mathrm{CL}}(s)\triangle H(s)} = \frac{G_{\mathrm{OL}}(s)}{1 + G_{\mathrm{OL}}(s)[H(s) + \triangle H(s)]}. \tag{5.50}$$

This result does, of course, illustrate the general point that feedback can be added either as elements of one feedback loop or as a succession of separate loops as shown in this block diagram.

5.3.2 State Feedback

State variables normally arise in the context of a realization. Suppose that (5.8)–(5.9) is a minimal realization. It is, therefore, controllable and we can assume that it is in companion form. The matrix $\boldsymbol{A}$ is given by (4.73), where the elements of the last row are the negatives of coefficients of the denominator of the transfer function, or equivalently of the characteristic function of $\boldsymbol{A}$. *State feedback* is introduced by replacing $u(t)$ in (5.8) by $u(t) - \boldsymbol{h}^{\mathrm{T}}\boldsymbol{x}(t)$, where

$$\boldsymbol{h}^{\mathrm{T}} = \begin{pmatrix} h_0 & h_1 & \cdots & h_{n-1} \end{pmatrix} \tag{5.51}$$

thus giving

$$\dot{\boldsymbol{x}}(t) = [\boldsymbol{A} - \boldsymbol{b}\,\boldsymbol{h}^{\mathrm{T}}]\boldsymbol{x}(t) + \boldsymbol{b}u(t). \tag{5.52}$$

The vector $\boldsymbol{b}$ is given by (4.74) and so

$$\boldsymbol{b}\,\boldsymbol{h}^{\mathrm{T}} = \begin{pmatrix} 0 & 0 & 0 & \cdots & \cdots & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \cdots & 0 & 0 \\ h_0 & h_1 & \cdots & \cdots & \cdots & \cdots & \cdots & h_{n-1} \end{pmatrix} \tag{5.53}$$

The net effect of this change is to replace the coefficients a_j in the characteristic polynomial $\Delta(\lambda)$, or equivalently in $\phi(s)$ by $a_j + h_j$. Thus $\phi(s)$ is replaced by $\phi(s) + \chi(s)$ as was the case in going from the transfer function of (5.44) to that of (5.46). It is clear that the state feedback can be used to ensure that $\boldsymbol{A}$ is a stability matrix, or equivalently, that all the zeros of $\phi(s) + \chi(s)$ have negative real parts.

Example 5.3.1 Show that companion realization

$$\boldsymbol{A} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 6 & -11 & 6 \end{pmatrix}, \qquad \boldsymbol{b} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \tag{5.54}$$

$$\boldsymbol{c}^{\mathrm{T}} = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix}.$$

is minimal but that the system is unstable. By the use of a state feedback obtain a stable realization with eigenvalues -1, -2 and -3. Determine the transfer function of the original system and the form of $\triangle H(s)$ required to produce the stabilization.

From (4.76) and (4.77) the controllability and observability matrices are respectively

$$\boldsymbol{U} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 6 \\ 1 & 6 & 25 \end{pmatrix}, \qquad \boldsymbol{V} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{5.55}$$

Since neither of these matrices is singular the realization is both controllable and observable and, therefore, minimal. The characteristic equation for $\boldsymbol{A}$ is

$$\Delta(\lambda) = \lambda^3 - 6\lambda^2 + 11\lambda - 6 = 0. \tag{5.56}$$

It is easy to see that this equation has one root $\lambda = 1$ and thus to extract the remaining roots $\lambda = 2$ and $\lambda = 3$. It follows that the system is unstable. The polynomial with the required roots is

$$(\lambda + 1)(\lambda + 2)(\lambda + 3) = \lambda^3 + 6\lambda^2 + 11\lambda + 6. \tag{5.57}$$

Subtracting (5.56) and (5.57) we see that the coefficients of the state feedback are $h_0 = 12$, $h_1 = 0$ and $h_2 = 12$ with

$$\chi(s) = 12s^2 + 12. \tag{5.58}$$

We now calculate the transfer function from this realization using (4.55). It is not difficult to show that

$$\psi(s) = \boldsymbol{c}^{\mathrm{T}} \mathrm{Adj}\{s\boldsymbol{I} - \boldsymbol{A}\}\boldsymbol{b} = 1. \tag{5.59}$$

and, of course,

$$\phi(s) = \Delta(s) = s^3 - 6s^2 + 11s - 6, \tag{5.60}$$

Giving

$$G(s) = \frac{1}{s^3 - 6s^2 + 11s - 6}. \tag{5.61}$$

The transfer function modified to produce stability is

$$G'(s) = \frac{1}{\phi(s) + \chi(s)} = \frac{1}{s^3 + 6s^2 + 11s + 6}. \tag{5.62}$$

From (5.49) this can be interpreted as an output feedback

$$\triangle H(s) = 12s^2 + 12. \tag{5.63}$$

5.4 Discrete-Time Systems

In Problems, question 3, we used the $\mathcal{Z}$ transform to derive the transfer function for two discrete-time systems. In each case the difference equation is a particular case of

$$y(k) + a_1 y(k-1) + a_0 y(k-2) = b_1 u(k-1) + b_0 u(k-2) \tag{5.64}$$

and applying the $\mathcal{Z}$ transform we have

$$\tilde{y}(z) = \frac{b_1 z + b_0}{z^2 + a_1 z + a_0}\tilde{u}(z). \tag{5.65}$$

Now rewrite (5.64) as

$$y(k+2) + a_1 y(k+1) + a_0 y(k) = b_1 u(k+1) + b_0 u(k). \tag{5.66}$$

By treating $y(k+j)$ and $u(k+j)$ like the j-th derivatives of $y(k)$ and $u(k)$ respectively we can now mirror the derivation of the companion realization in Sect. 4.3.1 to obtain

$$\begin{aligned} \boldsymbol{x}(k+1) &= \boldsymbol{A}\boldsymbol{x}(k) + \boldsymbol{b}\,u(k), \\ y(k) &= \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}(k) \end{aligned} \tag{5.67}$$

where

$$\begin{aligned} \boldsymbol{x}(k) &= \begin{pmatrix} x_1(k) \\ x_2(k) \end{pmatrix}, & \boldsymbol{A} &= \begin{pmatrix} 0 & 1 \\ -a_0 & -a_1 \end{pmatrix}, \\ \boldsymbol{b} &= \begin{pmatrix} 0 \\ 1 \end{pmatrix}, & \boldsymbol{c}^{\mathrm{T}} &= \begin{pmatrix} b_0 & b_1 \end{pmatrix}. \end{aligned} \tag{5.68}$$

For simplicity we have considered a two-dimensional case, but as for the continuous-time case the analysis applies for an n-dimensional realization When $u(k) = 0$ for all k the solution of the first of equations (5.68) is

$$\boldsymbol{x}(k) = \boldsymbol{A}^k \boldsymbol{x}(0). \tag{5.69}$$

Equation (5.69) is the discrete-time equivalent of (5.7) so we might expect a result similar to Thm. 5.1.2 to be true here. This is given by the definition

> *The $n \times n$ matrix $\boldsymbol{A}$ is called a* **convergent matrix** *if each of its eigenvalues is of magnitude (strictly) less than one.*

and the theorem

Theorem 5.4.1 $\boldsymbol{x}^* = \boldsymbol{0}$ *is an asymptotically stable equilibrium point of*

$$\boldsymbol{x}(k+1) = \boldsymbol{A}\boldsymbol{x}(k) \tag{5.70}$$

if and only if $\boldsymbol{A}$ is a convergent matrix.

This change from a condition on the real part of the eigenvalues to their magnitudes is easily understood when when realize that we have changed from the exponential matrix to the matrix itself.[4] The definition of bounded input–bounded output stability of Sect. 5.1 carries over to the discrete-time case with the index k replacing t. Thm. 5.1.3, relating the asymptotic stability of $\boldsymbol{x} = \boldsymbol{0}$ and bounded input–bounded output stability is also valid and can be proved in a similar way from (5.67). A slightly different approach, which we shall outline for the two-dimensional case, is to note that the characteristic function of $\boldsymbol{A}$ is

$$\Delta(\lambda) = \lambda^2 + a_1\lambda + a_0. \tag{5.71}$$

With a change of variable this is the denominator in (5.65), which on applying partial fractions can be written in the form

$$\tilde{y}(z) = \frac{\mathsf{C}_1 z\tilde{u}(z)}{z-\lambda_1} + \frac{\mathsf{C}_2 z\tilde{u}(z)}{z-\lambda_2}, \tag{5.72}$$

for constants C_1, C_2 and the eigenvalues λ_1 and λ_2 of $\boldsymbol{A}$. Using the last line of Table 2.2 to invert the $\mathcal{Z}$ transform gives

$$y(k) = \mathsf{C}_1 \sum_{j=0}^{k} \lambda_1^j u(k-j) + \mathsf{C}_2 \sum_{j=0}^{k} \lambda_2^j u(k-j). \tag{5.73}$$

If $|u(k)| < \mathcal{B}_1$ then

$$|y(k)| \le |\mathsf{C}_1|\,\mathcal{B}_1 \sum_{j=0}^{k} |\lambda_1|^j + |\mathsf{C}_2|\,\mathcal{B}_1 \sum_{j=0}^{k} |\lambda_2|^j. \tag{5.74}$$

If $\boldsymbol{A}$ is a convergent matrix each sum is less that the infinite binomial series and

$$|y(k)| \le \mathcal{B}_1 \left\{ \frac{|\mathsf{C}_1|}{1-|\lambda_1|} + \frac{|\mathsf{C}_2|}{1-|\lambda_2|} \right\}. \tag{5.75}$$

The system is bounded input–bounded output stable. Equation (5.65) is a particular case of the discrete time analogue

$$\tilde{y}(z) = G(z)\tilde{u}(z) \tag{5.76}$$

of (4.53), with $G(z)$ being the discrete-time transfer function. As in the case of continuous time we have seen that the poles of the transfer function are the eigenvalues of the matrix $\boldsymbol{A}$ when a realization is derived from it. Again we shall use *asymptotically stable* and *unstable* as descriptions of the system itself (with *stable* or *conditionally stable* or *marginally stable* denoting the intermediate cases). The only difference is that the criterion for stability is now whether the magnitudes of all the eigenvalues are less than one.

Example 5.4.1 Consider the stability of the female and male buffalo populations discussed in Example 2.4.2.

[4] Note that $|\exp(\zeta)| < 1$ if and only if $\Re\{\zeta\} < 0$.

We define the state variables

$$\begin{aligned} x_1(k) &= x(k), & x_2(k) &= x(k+1), \\ x_3(k) &= y(k), & x_4(k) &= y(k+1). \end{aligned} \tag{5.77}$$

Then (2.110) can be expressed in the form (5.70) with

$$\boldsymbol{A} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0.12 & 0.95 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0.14 & 0 & 0 & 0.95 \end{pmatrix}. \tag{5.78}$$

It is easy to show that the characteristic function for this matrix is

$$\Delta(\lambda) = \lambda(\lambda - 0.95)(\lambda - 1.063)(\lambda + 0.113). \tag{5.79}$$

In fact, of course these are precisely the numbers which appear in our solution (2.114). $\boldsymbol{A}$ is not a convergent matrix and the zero-population state is not asymptotically stable. According to this simplified model, as we saw in Example 2.4.2 the buffalo population would grow by 6.3% a year. In fact, due to indiscriminate slaughter,[5] the population of buffalo fell from 60 million in 1830 to 200 in 1887. Attempts are currently being made to reintroduce buffalo in the plains of South Dakota. Even then of course it may ultimately be necessary and economically desirable to implement a policy of culling.[6] Suppose that γ% of females and ξ% of males are culled each year. Then

$$\begin{aligned} x_2(k+1) &= 0.12x_1(k) + (0.95 - 0.01\gamma)x_2(k), \\ x_4(k+1) &= 0.14x_1(k) + (0.95 - 0.01\xi)x_4(k), \end{aligned} \tag{5.80}$$

and $\boldsymbol{A}$ is replaced by

$$\boldsymbol{A}' = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0.12 & 0.95 - 0.01\gamma & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0.14 & 0 & 0 & 0.95 - 0.01\xi \end{pmatrix}, \tag{5.81}$$

with characteristic equation

$$\Delta'(\lambda) = \lambda[\lambda - (0.95 - 0.01\xi)][\lambda - \omega^{(+)}(\gamma)][\lambda - \omega^{(-)}(\gamma)], \tag{5.82}$$

where

$$\omega^{(\pm)}(\gamma) = \frac{1}{200}\left[95 - \gamma \pm \sqrt{13825 - 190\gamma + \gamma^2}\right]. \tag{5.83}$$

[5] Encouraged by the United States Government.

[6] A single buffalo carcass will provide about 250 kg. of meat, enough for 10 people for year.

It will be observed that the culling of male has no affect on the long-term population. This is because we have assumed, perhaps unrealistically, that the number of calves born is proportional to the number of female adults, irrespective of the number of males. The root of $\Delta'(\lambda)$ which leads to the explosion in population is $\omega^{(+)}(\gamma)$. This is a decreasing function of γ with $\omega^{(+)}(7) = 1$. So a 7% cull of females will have the effect of stabilizing the population at a fixed value. The size of this population and the proportions of males and females, given particular initial populations can be calculated as in Example 2.4.2. This is an example of stabilization by linear feedback.

Problems 5

1) Show that the system with state-space equations

$$\begin{aligned} \dot{x}_1(t) &= -2x_1(t) + 4x_2(t), \\ \dot{x}_2(t) &= 4x_1(t) - 4x_2(t) + u(t), \end{aligned} \qquad y(t) = x_1(t)$$

is unstable. Derive the transfer function. Now investigate the situation when the input is changed to $u(t) - \gamma x_1(t)$ for some constant γ and interpret this change as an output feedback $\triangle H(s)$. Show that the system is asymptotically stable if $\gamma > 2$ and find $y(t)$ given that $\gamma = 5$, $u(t) = u_0$ and $x_1(0) = x_2(0) = 0$.

2) Consider the system with block diagram given on page 105 of the notes and

$$G_{\text{OL}}(s) = \frac{\mathsf{K}(\alpha + \beta s)}{s(1 + 2s)^2}, \qquad H(s) = 1,$$

where α and β are positive. Determine the closed loop transfer function and using the Routh-Hurwitz stability criterion show that

(i) If $\beta < \alpha$ then the system is asymptotically stable for $0 < \mathsf{K} < (\alpha - \beta)^{-1}$.

(ii) If $\alpha < \beta$ the system is asymptotically stable for all $\mathsf{K} > 0$.

Find a minimal realization when $\alpha = 1$, $\beta = 2$, $\mathsf{K} = -6$.

3) Consider the system with block diagram given on page 105 of the notes and

$$G_{\text{OL}}(s) = \frac{1}{s^3 + s^2 + s + 1}, \qquad H(s) = \gamma.$$

Determine the closed loop transfer function and show that the system is unstable for all $\gamma > 0$. Show by including the output feedback $\triangle H(s) = \alpha s^2 + \beta s$ with suitable values of α β and γ the system can be stabilized with poles of the closed-loop transfer function at $s = -1, -2, -3$. With these values of α, β and γ determine the output when $u(t) = u_0$. (The initial values of $y(t)$ and any derivatives may be assumed zero.)

Chapter 6

Optimal Control

6.1 Digression: The Calculus of Variations

The calculus of variations developed from a problem posed by the Swiss mathematician Johann Bernouilli (1667–1748). Suppose a wire lies in a vertical plane and stretches between two points A and B, with A higher than B. Given that a bead is able to move under gravity without friction on the wire and that it is released from rest at A. What form should the wire take in order that the time taken in going from A to B is a minimum?

This is called the *brachistochrone problem.*[1] Its solution is non-trivial[2] and to get some idea of how it might be tackled we will now generalize. Suppose that

$$\gamma = \{\boldsymbol{x}(\tau) : \tau_{\mathsf{A}} \le \tau \le \tau_{\mathsf{B}}\} \tag{6.1}$$

is a curve parameterized by τ in the phase space Γ_n of the vector $\boldsymbol{x}$ between the points A and B. At any point on the path the tangent vector is in the direction of $\dot{\boldsymbol{x}}(\tau)$.[3] Now for any function $f(\boldsymbol{x}(\tau), \dot{\boldsymbol{x}}(\tau); \tau)$ we define

$$\mathcal{I}[\boldsymbol{x}] = \int_{\tau_{\mathsf{A}}}^{\tau_{\mathsf{B}}} f(\boldsymbol{x}(\tau), \dot{\boldsymbol{x}}(\tau); \tau)\, \mathrm{d}\tau. \tag{6.2}$$

For fixed τ_{A} and τ_{B} the curve γ, including in general the points A and B will vary with the functional form of $\boldsymbol{x}(\tau)$, as will $\mathcal{I}$, which known as a *functional.* The technique for finding a form for $\boldsymbol{x}(\tau)$ which, for a specific $f(\boldsymbol{x}(\tau), \dot{\boldsymbol{x}}(\tau); \tau)$ and designated constraints on A and B, gives an extreme value for $\mathcal{I}[\boldsymbol{x}]$ is the *calculus of variations.*

Example 6.1.1 (The brachistochrone problem.) Consider cartesian axes with the y–axis vertically upwards and the x–axis horizontal. The bead on the

[1] From the Greek: *brachist* meaning shortest and *chronos* meaning time.

[2] Unless B is vertically below A the answer is not the straight line AB.

[3] We now extend the use of the dot notation to include differentiation with respect to τ.

wire descends from one end of the wire at $x = y = 0$ to the other end for which $x = x_{\mathsf{B}}$, $y = y_{\mathsf{B}}$. If the particle has mass m and the acceleration due to gravity is g then the total energy is conserved with

$$\tfrac{1}{2}m\left\{\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2\right\} + mgy = 0. \tag{6.3}$$

Now suppose the equation of the path is $x = w(y)$ with $w(0) = 0$. Then

$$\left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2 \left\{1 + [w'(y)]^2\right\} = -2gy. \tag{6.4}$$

This equation can be integrated to find the total time

$$\mathcal{T}[w] = \frac{1}{\sqrt{2g}}\int_0^{y_{\mathsf{B}}} \sqrt{\frac{1 + [w'(y)]^2}{-y}}\,\mathrm{d}y \tag{6.5}$$

that the bead takes for the path $x = w(y)$. The problem now is to find the function $w(y)$, describing the shape of the wire, which minimizes $\mathcal{T}[w]$ subject to whatever constraint we impose on x_{B}.

6.1.1 The Euler-Lagrange Equations

We suppose that $\boldsymbol{x}^*(\tau)$ is the functional form for $\boldsymbol{x}(\tau)$ with gives an extremum for $\mathcal{I}[\boldsymbol{x}]$. Let

$$\boldsymbol{x}(\tau) = \boldsymbol{x}^*(\tau) + \varepsilon\,\boldsymbol{\xi}(\tau), \tag{6.6}$$

where $\boldsymbol{\xi}(\tau)$ is a continuous, differentiable function of τ. Thus ε parameterizes a family of curves over the parameter interval $\tau_{\mathsf{A}} \le \tau \le \tau_{\mathsf{B}}$. The variation of $\mathcal{I}$ over the members of the family is given by

$$\frac{\mathrm{d}\mathcal{I}}{\mathrm{d}\varepsilon} = \int_{\tau_{\mathsf{A}}}^{\tau_{\mathsf{B}}} \frac{\mathrm{d}f}{\mathrm{d}\varepsilon}\,\mathrm{d}\tau. \tag{6.7}$$

Since

$$\frac{\mathrm{d}\boldsymbol{x}(\tau)}{\mathrm{d}\varepsilon} = \boldsymbol{\xi}(\tau), \qquad \frac{\mathrm{d}\dot{\boldsymbol{x}}(\tau)}{\mathrm{d}\varepsilon} = \dot{\boldsymbol{\xi}}(\tau), \tag{6.8}$$

$$\begin{aligned}\frac{\mathrm{d}f}{\mathrm{d}\varepsilon} &= \nabla_{\boldsymbol{x}} f[\boldsymbol{x}^*(\tau) + \varepsilon\,\boldsymbol{\xi}(\tau), \dot{\boldsymbol{x}}^*(\tau) + \varepsilon\,\dot{\boldsymbol{\xi}}(\tau);\tau]\cdot\boldsymbol{\xi}(\tau) \\ &\quad + \nabla_{\dot{\boldsymbol{x}}} f[\boldsymbol{x}^*(\tau) + \varepsilon\,\boldsymbol{\xi}(\tau), \dot{\boldsymbol{x}}^*(\tau) + \varepsilon\,\dot{\boldsymbol{\xi}}(\tau);\tau]\cdot\dot{\boldsymbol{\xi}}(\tau).\end{aligned} \tag{6.9}$$

Now we need to substitute from (6.9) into (6.7). In doing so we apply integration by parts to the second term

$$\int_{\tau_{\mathsf{A}}}^{\tau_{\mathsf{B}}} \nabla_{\dot{\boldsymbol{x}}} f\cdot\dot{\boldsymbol{\xi}}(\tau)\mathrm{d}\tau = \Big[\nabla_{\dot{\boldsymbol{x}}} f\cdot\boldsymbol{\xi}(\tau)\Big]_{\tau_{\mathsf{A}}}^{\tau_{\mathsf{B}}} - \int_{\tau_{\mathsf{A}}}^{\tau_{\mathsf{B}}} \frac{\mathrm{d}(\nabla_{\dot{\boldsymbol{x}}} f)}{\mathrm{d}\tau}\cdot\boldsymbol{\xi}(\tau)\mathrm{d}\tau. \tag{6.10}$$

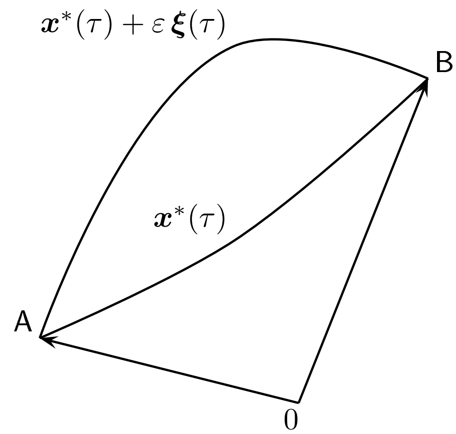


Figure 6.1: Two paths (one the extremum) from A to B in the phase space Γ_n.

For a stationary value

$$\left(\frac{\mathrm{d}\mathcal{I}}{\mathrm{d}\varepsilon}\right)_{\varepsilon=0} = 0 \tag{6.11}$$

and, from (6.7), (6.9) and (6.10),

$$\left(\frac{\mathrm{d}\mathcal{I}}{\mathrm{d}\varepsilon}\right)_{\varepsilon=0} = \Big[\nabla_{\dot{\boldsymbol{x}}} f \cdot \boldsymbol{\xi}(\tau)\Big]_{\tau_{\mathsf{A}}}^{\tau_{\mathsf{B}}} + \int_{\tau_{\mathsf{A}}}^{\tau_{\mathsf{B}}} \left[\nabla_{\boldsymbol{x}} f - \frac{\mathrm{d}(\nabla_{\dot{\boldsymbol{x}}} f)}{\mathrm{d}\tau}\right] \cdot \boldsymbol{\xi}(\tau)\,\mathrm{d}\tau, \tag{6.12}$$

where now the gradients of f with respect to $\boldsymbol{x}(\tau)$ and $\dot{\boldsymbol{x}}(\tau)$ are evaluated along the extremum curve $\boldsymbol{x}^*(\tau)$. We now apply this analysis to two cases:

(i) Both A and B are fixed points with vector locations $\boldsymbol{x}_{\mathsf{A}}$ and $\boldsymbol{x}_{\mathsf{B}}$ respectively.

In this case all paths pass through the same end-points (see Fig. 6.1) and so

$$\boldsymbol{\xi}(\tau_{\mathsf{A}}) = \boldsymbol{\xi}(\tau_{\mathsf{B}}) = \mathbf{0}. \tag{6.13}$$

The first term on the right-hand side of (6.12) is zero and for (6.11) and (6.12) to be satisfied for any $\boldsymbol{\xi}(\tau)$ satisfying (6.13) we must have

$$\frac{\mathrm{d}(\nabla_{\dot{\boldsymbol{x}}} f)}{\mathrm{d}\tau} - \nabla_{\boldsymbol{x}} f = \mathbf{0}. \tag{6.14}$$

In scalar form this equation is

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\left(\frac{\partial f}{\partial \dot{x}_j}\right) - \frac{\partial f}{\partial x_j} = 0, \qquad j = 1, 2, \ldots, n. \tag{6.15}$$

These are two forms of the *Euler-Lagrange equations.*

(ii) A is a fixed point but the location of the final point on the path is allowed to vary.
In this case the Euler-Lagrange equations must still be satisfied but we must also have

$$\nabla_{\dot{\boldsymbol{x}}} f(\boldsymbol{x}(\tau_{\mathsf{B}}), \dot{\boldsymbol{x}}(\tau_{\mathsf{B}}); \tau_{\mathsf{B}}) = \mathbf{0}, \tag{6.16}$$

or in scalar form

$$\left(\frac{\partial f}{\partial \dot{x}_j}\right)_{\tau=\tau_{\mathsf{B}}} = 0, \qquad j = 1, 2, \ldots, n. \tag{6.17}$$

For reasons which will be explained below (6.16) and (6.17) are known as *transversality conditions*.

From (6.14)

$$\dot{\boldsymbol{x}}(\tau)\cdot\frac{\mathrm{d}(\nabla_{\dot{\boldsymbol{x}}} f)}{\mathrm{d}\tau} - \dot{\boldsymbol{x}}(\tau)\cdot\nabla_{\boldsymbol{x}} f = \mathbf{0}, \tag{6.18}$$

giving

$$\frac{\mathrm{d}[\dot{\boldsymbol{x}}(\tau)\cdot\nabla_{\dot{\boldsymbol{x}}} f]}{\mathrm{d}\tau} - \ddot{\boldsymbol{x}}(\tau)\cdot\nabla_{\dot{\boldsymbol{x}}} f - \dot{\boldsymbol{x}}(\tau)\cdot\nabla_{\boldsymbol{x}} f = \mathbf{0}. \tag{6.19}$$

Now

$$\frac{\mathrm{d}f}{\mathrm{d}\tau} = \ddot{\boldsymbol{x}}(\tau)\cdot\nabla_{\dot{\boldsymbol{x}}} f + \dot{\boldsymbol{x}}(\tau)\cdot\nabla_{\boldsymbol{x}} f + \frac{\partial f}{\partial \tau}, \tag{6.20}$$

so we have the alternative form

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\left[\dot{\boldsymbol{x}}(\tau)\cdot\nabla_{\dot{\boldsymbol{x}}} f - f\right] + \frac{\partial f}{\partial \tau} = 0, \tag{6.21}$$

for the Euler-Lagrange equations.

In whatever form they are represented and given a particular function $f(\boldsymbol{x}, \dot{\boldsymbol{x}}, \tau)$, the Euler-Lagrange equations are a set of n second-order differential equations for the variables $x_1(\tau), x_2(\tau), \ldots, x_n(\tau)$. In two special cases first-integrals can be derived immediately:

(a) When f is not a function of x_j for some j it follows from (6.15) that

$$\frac{\partial f}{\partial \dot{x}_j} = \text{constant}. \tag{6.22}$$

(b) When f is not an explicit function of τ it follows from (6.21)

$$\dot{\boldsymbol{x}}(\tau)\cdot\nabla_{\dot{\boldsymbol{x}}} f(\boldsymbol{x}, \dot{\boldsymbol{x}}) - f(\boldsymbol{x}, \dot{\boldsymbol{x}}) = \text{constant}. \tag{6.23}$$

Strictly speaking condition (6.11) gives a stationary value which could be a maximum rather than a minimum. However, in most cases there are good physical reasons for supposing that the result gives a minimum.

It is also the case that in many problems the parameter τ is in fact t the time. However, this is not always so as we see in the next example.

Example 6.1.2 (The brachistochrone problem.) In this case τ is the variable y and

$$f(x(y),\dot{x}(y);y)=\sqrt{\frac{1+[\dot{x}(y)]^2}{-y}}. \tag{6.24}$$

Since f is not an explicit function of $x(y)$ condition (6.22) is applicable giving

$$\frac{\partial f}{\partial \dot{x}}=\frac{\dot{x}(y)}{\sqrt{-y(1+[\dot{x}(y)]^2)}}=\frac{1}{\mathsf{C}}. \tag{6.25}$$

We first note that if x_{B} is allowed to vary along the line $y=y_{\mathsf{B}}$ then the transversality condition (6.17) applies giving $\mathsf{C}=\infty$, and thus $\dot{x}(y)=0$. The wire is vertically downwards with $x_{\mathsf{B}}=0$, as would be expected. We now suppose that both ends of the wire are fixed. Let $\dot{x}(y)=\cot(\theta)$. Then, from (6.25),

$$y=-\tfrac{1}{2}\mathsf{C}^2[1+\cos(2\theta)]. \tag{6.26}$$

Thus

$$\frac{\mathrm{d}x}{\mathrm{d}\theta}=-\mathsf{C}^2[1+\cos(2\theta)]. \tag{6.27}$$

When $y=0$, $\theta=\frac{1}{2}\pi$, so integrating (6.27) and choosing the constant of integration so that $x=0$ when $\theta=\frac{1}{2}\pi$ gives

$$x=\tfrac{1}{2}\mathsf{C}^2[2\theta-\pi+\sin(2\theta)]. \tag{6.28}$$

These are the parametric equations of a *cycloid*. A plot for $\mathsf{C}=1$ is given by

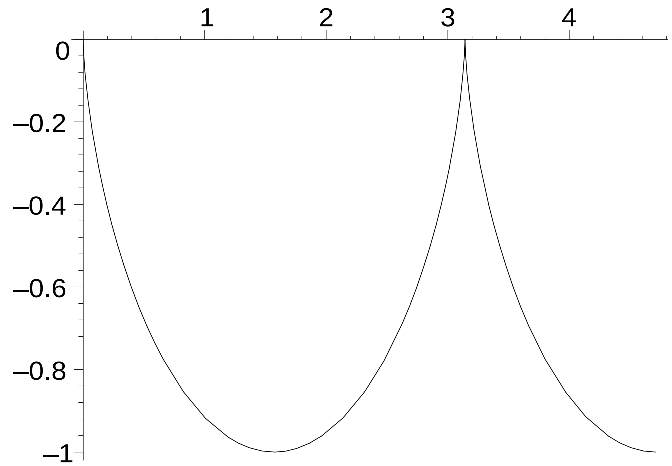
```
> x:=(theta,c)->c^2*(2*theta-Pi+sin(2*theta))/2;
```

$$x:=(\theta,\,c)\to\frac{1}{2}\,c^2\,(2\,\theta-\pi+\sin(2\,\theta))$$

```
> y:=(theta,c)->-c^2*(1+cos(2*theta))/2;
```

$$y:=(\theta,\,c)\to-\frac{1}{2}\,c^2\,(1+\cos(2\,\theta))$$

```
> plot([x(theta,1),y(theta,1),theta=Pi/2..2*Pi]);
```



Given particular values for x_{B} and y_{B} the equations can be solved to obtain C and θ_{B}.

6.1.2 Hamilton's Principle of Least Action

Suppose we have a collection of ν particles of mass m moving in three-dimensional space. The the location of all the particles at time t is given by a vector $\boldsymbol{x}(t) = (x_1(t), x_2(t), \ldots, x_n(t))^{\mathrm{T}}$ in phase space $\mathit{\Gamma}_n$, where $n = 3\nu$. If the forces acting on the particles are $\boldsymbol{G}(\boldsymbol{x}; t) = (G_1(\boldsymbol{x}; t), G_2(\boldsymbol{x}; t), \ldots, G_n(\boldsymbol{x}; t))^{\mathrm{T}}$ then the usual starting point for mechanics is *Newton's second law*

$$m\ddot{\boldsymbol{x}}(t) = \boldsymbol{G}(\boldsymbol{x}; t), \tag{6.29}$$

(see (1.102)). Now suppose that there exists a potential function $V(\boldsymbol{x}; t)$ related to $\boldsymbol{G}(\boldsymbol{x}; t)$ by (1.106). Then

$$m\ddot{\boldsymbol{x}}(t) = -\nabla V(\boldsymbol{x}; t), \tag{6.30}$$

An alternative axiomatic approach to mechanics is to define the *Lagrangian*

$$L(\boldsymbol{x}(t), \dot{\boldsymbol{x}}(t); t) = \tfrac{1}{2}m\dot{\boldsymbol{x}}^2 - V(\boldsymbol{x}; t), \tag{6.31}$$

and the *action*

$$\mathcal{I}[\boldsymbol{x}] = \int_{t_{\mathsf{A}}}^{t_{\mathsf{B}}} L(\boldsymbol{x}(t), \dot{\boldsymbol{x}}(t); t)\mathrm{d}t \tag{6.32}$$

Then *Hamilton's principle of least action* states that the path in $\mathit{\Gamma}_n$ which represents the configuration of the particles in a time interval $[t_{\mathsf{A}}, t_{\mathsf{B}}]$ from a fixed initial configuration $\boldsymbol{x}_{\mathsf{A}}$ to a fixed final configuration $\boldsymbol{x}_{\mathsf{B}}$ is that which minimizes the action. Thus we have from the Euler-Lagrange equations (6.15)

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\dot{x}_j}\right) - \frac{\partial L}{\partial x_j} = 0, \qquad j = 1, 2, \ldots, n. \tag{6.33}$$

These are call *Lagrange's equations.* From (6.31) and (6.33)

$$m\ddot{x}_j(t) = -\frac{\partial V}{\partial x_j}, \qquad j = 1, 2, \ldots, n, \tag{6.34}$$

which is just the scalar form of (6.30).

6.1.3 Constrained Problems

Suppose that we wish to find an extremum of $\mathcal{I}[\boldsymbol{x}]$ given by (6.2) but that now the paths are subject to a constrain. This can be of two forms

Integral Constraints In this case the constraint is of the form

$$\mathcal{J}[\boldsymbol{x}] = \int_{\tau_{\mathsf{A}}}^{\tau_{\mathsf{B}}} g(\boldsymbol{x}(\tau), \dot{\boldsymbol{x}}(\tau); \tau)\,\mathrm{d}\tau = \mathrm{J}, \tag{6.35}$$

where J is some constant. To solve this problem we use the method of Lagrange's undetermined multipliers. We find an extremum for

$$\mathcal{I}[\boldsymbol{x}] + p\mathcal{J}[\boldsymbol{x}] = \int_{\tau_{\mathsf{A}}}^{\tau_{\mathsf{B}}} [f(\boldsymbol{x}(\tau), \dot{\boldsymbol{x}}(\tau); \tau) + p\, g(\boldsymbol{x}(\tau), \dot{\boldsymbol{x}}(\tau); \tau)]\,\mathrm{d}\tau, \tag{6.36}$$

for some constant p. This replaces the Euler-Lagrange equations by

$$\frac{\mathrm{d}[\nabla_{\dot{\boldsymbol{x}}}(f + p\,g)]}{\mathrm{d}\tau} - \nabla_{\boldsymbol{x}}(f + p\,g) = \mathbf{0}. \tag{6.37}$$

Once the extremum function $\boldsymbol{x}^*(t)$ has been found, p is determined by substituting into (6.35).

Example 6.1.3 Consider all the curves in the x–y plane between $(0, 0)$ and $(2, 0)$ which are of length π. Find the equation of the one which encloses the maximum area between it and the x–axis.

The area enclosed is

$$\mathcal{A}[y(x)] = \int_0^2 y(x)\mathrm{d}x. \tag{6.38}$$

Now the length of an element of the curve $y = y(x)$ is

$$\sqrt{(\mathrm{d}x)^2 + (\mathrm{d}y)^2} = \sqrt{1 + [\dot{y}(x)]^2}\mathrm{d}x. \tag{6.39}$$

So the length of the curve is

$$\mathcal{L}[y(x)] = \int_0^2 \sqrt{1+[\dot{y}(x)]^2}\mathrm{d}x. \tag{6.40}$$

and the constraint is

$$\mathcal{L}[y(x)] = \pi. \tag{6.41}$$

From (6.36)

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[\frac{p\,\dot{y}(x)}{\sqrt{1+[\dot{y}(x)]^2}}\right] - 1 = 0, \tag{6.42}$$

giving

$$\frac{p\,\dot{y}(x)}{\sqrt{1+[\dot{y}(x)]^2}} = \mathsf{C} + x. \tag{6.43}$$

Now let

$$\dot{y}(x) = \cot(\theta), \tag{6.44}$$

with the range $[\theta_0, \theta_1]$ for θ corresponding to the range $[0, 2]$ for x. Substituting into (6.43) gives

$$x = p\,[\cos(\theta) - \cos(\theta_0)] \tag{6.45}$$

and then from (6.44)

$$\frac{\mathrm{d}y}{\mathrm{d}\theta} = -p\,\cos(\theta) \tag{6.46}$$

giving

$$y = p\,[\sin(\theta_0) - \sin(\theta)] \tag{6.47}$$

From the constraint condition given by (6.40) and (6.42)

$$\pi = \int_{\theta_0}^{\theta_1} \mathrm{cosec}(\theta)\frac{\mathrm{d}x}{\mathrm{d}\theta}\mathrm{d}\theta = -p\,(\theta_1 - \theta_0). \tag{6.48}$$

and from (6.45) and (6.47)

$$\begin{aligned} &2 = -p\,[\cos(\theta_0) - \cos(\theta_1)] \\ &\sin(\theta_0) = \sin(\theta_1) \end{aligned} \tag{6.49}$$

These equations are satisfied by $\theta_0 = 0,\ \theta_1 = \pi$ and $p = -1$. The curve is give by

$$\left.\begin{aligned} &x = 1 - \cos(\theta), \\ &y = \sin(\theta), \end{aligned}\right. \qquad 0 \le \theta \le \pi, \tag{6.50}$$

which is the upper semicircle of the circle radius one centre $(1,0)$. This is an example where we have needed a maximum rather than a minimum stationary value. As is usually the case with variational problems the evidence that this result does indeed correspond to the enclosing of a maximum area is not difficult to find. You can reduce the area to almost zero with a smooth curve of length π in the first quadrant from $(0,0)$ to $(2,0)$.

Non-Integral Constraints In this case the constraint is of the form

$$g(\boldsymbol{x}(\tau), \dot{\boldsymbol{x}}(\tau); \tau) = 0. \tag{6.51}$$

Again we use the method of Lagrange's undetermined multipliers and find an extremum for

$$\mathcal{I}_p[\boldsymbol{x}] = \int_{\tau_{\mathsf{A}}}^{\tau_{\mathsf{B}}} [f(\boldsymbol{x}(\tau), \dot{\boldsymbol{x}}(\tau); \tau) + p(\tau) g(\boldsymbol{x}(\tau), \dot{\boldsymbol{x}}(\tau); \tau)] \, \mathrm{d}\tau, \tag{6.52}$$

except there here p is a function of τ. The form of the Euler-Lagrange equations is again (6.37) but now we must not forget the derivative of p with respect to τ which arises from the first term.

The case of both integral and non-integral constraints can be easily generalized to a number of constraints.

6.2 The Optimal Control Problem

The optimal control problem is concerned with developing quantitative criteria for the efficiency of control systems and obtaining the form for the input $\boldsymbol{u}(t)$ which best satisfies the criteria. Since we are normally interested in the performance of the system over some period of time $[t_{\mathrm{I}}, t_{\mathrm{F}}]$ the measure of efficiency will be a time integral which must be minimized relative to input, output and, in general, state space variables. This quantity is called the *cost functional*. The ability to reduce this quantity is a *performance indicator* and the challenge is to devise the input function which will produce a minimum value.

6.2.1 Problems Without State-Space Variables

In this case the cost functional is of the form

$$\mathcal{I}[u, y] = \int_{t_{\mathrm{I}}}^{t_{\mathrm{F}}} f(u(t), \dot{u}(t), y(t), \dot{y}(t)) \mathrm{d}t. \tag{6.53}$$

Example 6.2.1 Suppose we have a system with the block diagram shown at the beginning of Sect. 5.3.1 with

$$G_{\mathrm{OL}}(s) = \frac{1}{\mathsf{Q}s}, \qquad H(s) = 1. \tag{6.54}$$

The aim it to run the system so that $y(t_{\mathrm{I}}) = y_{\mathrm{I}}$ and $y(t_{\mathrm{F}}) = y_{\mathrm{F}}$ while at the same time minimizing the time average of the *square error* $[v(t)]^2$, where, from (5.43)

$$v(t) = u(t) - y(t). \tag{6.55}$$

Thus we have

$$\mathcal{I}[u, y] = \int_{t_{\mathrm{I}}}^{t_{\mathrm{F}}} [u(t) - y(t)]^2 \mathrm{d}t. \tag{6.56}$$

The standard formula $\bar{y}(s) = G(s)\bar{u}(s)$ with, in this case,

$$G(s) = \frac{1}{1 + \mathrm{Q}s}, \tag{6.57}$$

gives

$$\mathrm{Q}\dot{y}(t) = u(t) - y(t). \tag{6.58}$$

This relationship could be used to give a constraint on the minimization of $\mathcal{I}[u, y]$. However, in this particular case, it is simpler just to substitute from (6.58) into (6.56) to give

$$\mathcal{I}[u, y] = \mathrm{Q}^2 \int_{t_{\mathrm{I}}}^{t_{\mathrm{F}}} [\dot{y}(t)]^2 \mathrm{d}t. \tag{6.59}$$

The Euler-Lagrange equation then gives an extremum for $\mathcal{I}[u, y]$ when

$$\ddot{y}(t) = 0. \tag{6.60}$$

Using the initial and final conditions on $y(t)$ gives

$$y(t) = \frac{(y_{\mathrm{I}} - y_{\mathrm{F}})t + (y_{\mathrm{F}} t_{\mathrm{I}} - y_{\mathrm{I}} t_{\mathrm{F}})}{t_{\mathrm{I}} - t_{\mathrm{F}}}, \tag{6.61}$$

and from (6.58)

$$u(t) = \frac{(y_{\mathrm{I}} - y_{\mathrm{F}})(t + \mathrm{Q}) + (y_{\mathrm{F}} t_{\mathrm{I}} - y_{\mathrm{I}} t_{\mathrm{F}})}{t_{\mathrm{I}} - t_{\mathrm{F}}}. \tag{6.62}$$

6.2.2 Problems With State-Space Variables

In this case the cost functional is of the form

$$\mathcal{I}[u, \boldsymbol{x}] = \int_{t_{\mathrm{I}}}^{t_{\mathrm{F}}} f(u(t), \boldsymbol{x}(t), \dot{\boldsymbol{x}}(t); t)\, \mathrm{d}t. \tag{6.63}$$

with the constraints

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{X}(u(t), \boldsymbol{x}(t); t) \tag{6.64}$$

and the output given by

$$y(t) = Y(u(t), \boldsymbol{x}(t); t). \tag{6.65}$$

For simplicity we have excluded possible dependence of the integrand of the cost functional on $\dot{u}(t)$ and $\dot{y}(t)$. It is also convenient to remove any dependence on $y(t)$ by substituting from (6.65). In the linear problems, which we have so far considered, (6.64) and (6.65) take the usual forms

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{b}u(t), \tag{6.66}$$

$$y(t) = \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}(t). \tag{6.67}$$

respectively, although neither (6.65) nor its linear form (6.67) play any role in our discussion. The essence of the task is to determine forms for $\boldsymbol{x}(t)$ and $u(t)$ which give a minimum for $\mathcal{I}[u, \boldsymbol{x}]$ subject to the constraints (6.64)

Example 6.2.2 Suppose we have a system with

$$\dot{x}(t) = \alpha x(t) + u(t), \tag{6.68}$$

The aim is to minimize

$$\mathcal{I}[u, x] = \int_{t_{\mathrm{I}}}^{t_{\mathrm{F}}} \left\{[u(t)]^2 + \beta[x(t)]^2\right\} \mathrm{d}t. \tag{6.69}$$

subject to the constraint (6.68) with $x(t_{\mathrm{I}}) = x_{\mathrm{I}}$ but $x(t_{\mathrm{F}})$ unrestricted.

The simplest way to do the problem is to use (6.68) to replace $u(t)$ in (6.69). Alternatively (6.68) can be included by using an undetermined multiplier. As an exercise we shall try the second method. Thus (6.68) can be treated as a constraint. From (6.52)

$$\mathcal{I}_p\,[u, x] = \int_{t_{\mathrm{I}}}^{t_{\mathrm{F}}} \left\{[u(t)]^2 + \beta[x(t)]^2 + p(t)[\dot{x}(t) - \alpha x(t) - u(t)]\right\} \mathrm{d}t \tag{6.70}$$

and the Euler-Lagrange equations give

$$\begin{aligned} &2u(t) - p(t) = 0, \\ &\dot{p}(t) - 2\beta\, x(t) + p(t)\alpha = 0 \end{aligned} \tag{6.71}$$

Since $x(t)$ is not constrained we have, from the transversality condition (6.17),

$$p(t_{\mathrm{F}}) = 0. \tag{6.72}$$

From (6.71)

$$\dot{u}(t) = \beta\, x(t) - \alpha\, u(t). \tag{6.73}$$

A simple way to solve the pair of differential equations (6.68) and (6.73) is by Laplace transforming. We have

$$\begin{pmatrix} s-\alpha & -1 \\ -\beta & s+\alpha \end{pmatrix} \begin{pmatrix} \bar{x}(s) \\ \bar{u}(s) \end{pmatrix} = \begin{pmatrix} x_{\mathrm{I}} \\ u_{\mathrm{I}} \end{pmatrix}. \tag{6.74}$$

Inverting the matrix and extracting $\bar{u}(s)$ gives

$$\bar{u}(s) = \frac{(s-\alpha)u_{\mathrm{I}} + \beta x_{\mathrm{I}}}{s^2 - (\alpha^2+\beta)}. \tag{6.75}$$

Assuming that $\alpha^2 + \beta = \omega^2 > 0$ gives

$$u(t) = u_{\mathrm{I}} \cosh(\omega t) + \omega^{-1}(\beta x_{\mathrm{I}} - \alpha u_{\mathrm{I}}) \sinh(\omega t). \tag{6.76}$$

Now using the transversality condition $u(t_{\mathrm{F}}) = 0$ to eliminate u_{I} gives

$$u(t) = \frac{\beta x_{\mathrm{I}} \sinh[\omega(t_{\mathrm{F}} - t)]}{\alpha \sinh(\omega t_{\mathrm{F}}) - \omega \cosh(\omega t_{\mathrm{F}})}. \tag{6.77}$$

Example 6.2.3 Suppose we have a system with

$$\ddot{x}(t) = -\dot{x}(t) + u(t), \tag{6.78}$$

The aim is to minimize

$$\mathcal{I}[u,x] = \int_0^\infty \left\{[x(t)]^2 + 4[u(t)]^2\right\} \mathrm{d}t. \tag{6.79}$$

subject to the constraint (6.78) with $x(0) = 0$, $\dot{x}(0) = 1$ and $x(t) \to 0$ and $\dot{x}(t) \to 0$ as $t \to \infty$.

By writing $x_1(t) = x(t)$ and $x_2(t) = \dot{x}(t)$ (6.78) can be written

$$\begin{aligned} \dot{x}_1(t) &= x_2(t), \\ \dot{x}_2(t) &= -x_2(t) + u(t) \end{aligned} \tag{6.80}$$

Thus we have two constraints. From (6.52)

$$\begin{aligned} \mathcal{I}_p[u,x_1,x_2] \quad = \quad & \int_0^\infty \Big\{[x_1(t)]^2 + 4[u(t)]^2 + p_1(t)\,[\dot{x}_1(t) - x_2(t)] \\ & + p_2(t)\,[\dot{x}_2(t) - u(t) + x_2(t)]\Big\}\mathrm{d}t \end{aligned} \tag{6.81}$$

and the Euler-Lagrange equations give

$$\begin{aligned} \dot{p}_1(t) - 2x_1(t) &= 0, \\ \dot{p}_2(t) + p_1(t) - p_2(t) &= 0, \\ 8u(t) - p_2(t) &= 0. \end{aligned} \tag{6.82}$$

Eliminating $p_1(t)$ and $p_2(t)$ gives

$$4\ddot{u}(t) = 4\dot{u}(t) - x(t). \tag{6.83}$$

We could now proceed by Laplace transforming (6.78) and (6.83). In this case it is probably easier to eliminate $u(t)$ to give

$$4\frac{\mathrm{d}^4 x(t)}{\mathrm{d}t^4} - 4\frac{\mathrm{d}^2 x(t)}{\mathrm{d}t^2} + x(t) = 0. \tag{6.84}$$

The auxiliary equation is

$$4\lambda^4 - 4\lambda^2 + 1 = (2\lambda^2 - 1)^2 = 0. \tag{6.85}$$

So the general solution is

$$x(t) = \mathsf{A}\exp(-t/\sqrt{2}) + \mathsf{B}t\exp(-t/\sqrt{2}) + \mathsf{C}\exp(t/\sqrt{2}) + \mathsf{D}t\exp(t/\sqrt{2}). \tag{6.86}$$

Since both $x(t)$ and $\dot{x}(t)$ tend to zero as $t \to \infty$ $\mathsf{C} = \mathsf{D} = 0$. Using the initial conditions, $\mathsf{A} = 0$ and $\mathsf{B} = 1$. Then substituting into (6.78)

$$u(t) = (1 - \sqrt{2})\left(1 + \tfrac{1}{2}t\right)\exp(-t/\sqrt{2}). \tag{6.87}$$

6.3 The Hamilton-Pontriagin Method

In proposing the form (6.63) for the cost functional we remarked that any dependence of the integrand on the output $y(t)$ could be removed by substitution from (6.65). In the *Hamilton-Pontriagin method* the process is taken a step further by removing explicit dependence on $\dot{\boldsymbol{x}}(t)$ by substituting from the constraint conditions (6.64).[4] Thus we have the cost functional

$$\mathcal{I}[u, \boldsymbol{x}] = \int_{t_{\mathrm{I}}}^{t_{\mathrm{F}}} f(u(t), \boldsymbol{x}(t); t)\,\mathrm{d}t. \tag{6.88}$$

with the constraints

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{X}(u(t), \boldsymbol{x}(t); t) \tag{6.89}$$

For these n constraints we introduce n undetermined (time dependent) multipliers $p_j(t)$, $j = 1, 2, \dots, n$. The variables $x_j(t)$ and $p_j(t)$ are said to be *conjugate*. With $\boldsymbol{p}(t) = (p_1(t), p_2(t), \dots, p_n(t))^{\mathrm{T}}$ we have

$$F(u(t), \boldsymbol{x}(t), \boldsymbol{p}(t); t) = f(u(t), \boldsymbol{x}(t); t) + \boldsymbol{p}(t) \cdot [\dot{\boldsymbol{x}}(t) - \boldsymbol{X}(u(t), \boldsymbol{x}(t); t)] \tag{6.90}$$

and

$$\mathcal{I}_p[u, \boldsymbol{x}] = \int_{t_{\mathrm{I}}}^{t_{\mathrm{F}}} F(u(t), \boldsymbol{x}(t), \boldsymbol{p}(t); t)\,\mathrm{d}t. \tag{6.91}$$

The task is now to find an extremum for $\mathcal{I}_p[u, \boldsymbol{x}]$. Before deriving the equations for this we reformulate the problem slightly by defining the *Hamiltonian*

$$H(u(t), \boldsymbol{x}(t), \boldsymbol{p}(t); t) = \boldsymbol{p}(t) \cdot \boldsymbol{X}(u(t), \boldsymbol{x}(t); t) - f(u(t), \boldsymbol{x}(t); t). \tag{6.92}$$

Then (6.91) becomes

$$\mathcal{I}_p[u, \boldsymbol{x}] = \int_{t_{\mathrm{I}}}^{t_{\mathrm{F}}} [\boldsymbol{p}(t) \cdot \dot{\boldsymbol{x}}(t) - H(u(t), \boldsymbol{x}(t), \boldsymbol{p}(t); t)]\,\mathrm{d}t. \tag{6.93}$$

[4] If in particular problems time derivatives appear in the integrand which do not correspond to constraints then they can be removed by 'inventing' new variables, rather in the way that we treated the second-order time derivative in Example 6.2.3.

Since $\nabla_{\dot{\boldsymbol{x}}}[\boldsymbol{p}(t)\cdot\dot{\boldsymbol{x}}(t)] = \boldsymbol{p}(t)$ the Euler-Lagrange equations are

$$\dot{\boldsymbol{p}}(t) = -\nabla_{\boldsymbol{x}} H, \tag{6.94}$$

$$\frac{\partial H}{\partial u} = 0. \tag{6.95}$$

In this context these are often referred to as the *Hamilton-Pontriagin equations.* It is also interesting (but not particularly useful) to note that (6.89) can be rewritten as

$$\dot{\boldsymbol{x}}(t) = \nabla_{\boldsymbol{p}} H. \tag{6.96}$$

In scalar form (6.94) and (6.96) are

$$\dot{p}_j(t) = -\frac{\partial H}{\partial x_j}, \tag{6.97}$$

$$\dot{x}_j(t) = \frac{\partial H}{\partial p_j}, \tag{6.98}$$

for $j = 1, 2, \ldots, n$. Examples 6.2.2 and 6.2.3 can both be formulated in the way described here. In fact in these cases, since the integrand of the cost functional does not involve time derivatives, no substitutions from the constraint conditions are needed. For Example 6.2.2 the Hamiltonian is

$$H(u(t), x(t), p(t)) = p(t)[u(t) + \alpha x(t)] - [u(t)]^2 - \beta[x(t)]^2 \tag{6.99}$$

and (6.95) and (6.98) yield (6.71). For Example 6.2.3 the Hamiltonian is

$$H(u(t), x_1(t), x_2(t), p_1(t), p_2(t)) = p_1(t)x_2(t) + p_2(t)[u(t) - x_2(t)] - 4[u(t)]^2 - [x_1(t)]^2 \tag{6.100}$$

and (6.95) and (6.98) yield (6.82).

In cases like Example 6.2.2 where the location of the final point on the path is unrestricted we need transversality conditions. From (6.90) and (6.91) these are given by

$$p_j(t_{\mathrm{F}}) = 0, \quad \text{if } x_j(t_{\mathrm{F}}) \text{ is unrestricted for } j = 1, \ldots, n. \tag{6.101}$$

in agreement with the transversality condition (6.72) of Example 6.2.2.

6.3.1 The Connection with Dynamics: Hamilton's Equations

For the dynamic system described in Sect. 6.1.2 we need first to express the Lagrangian without using time-derivative variables. To do this we introduce the new variables

$$\dot{x}_j(t) = v_j(t), \qquad j = 1, 2, \ldots, n, \tag{6.102}$$

or in vector form

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{v}(t). \tag{6.103}$$

Then

$$L(\boldsymbol{x}(t), \boldsymbol{v}(t); t) = \tfrac{1}{2} m \boldsymbol{v}^2 - V(\boldsymbol{x}; t) \tag{6.104}$$

and from (6.92) the Hamiltonian is

$$\begin{aligned} H(\boldsymbol{x}(t), \boldsymbol{v}(t), \boldsymbol{p}(t); t) &= \boldsymbol{p}(t) \cdot \boldsymbol{v}(t) - L(\boldsymbol{x}(t), \boldsymbol{v}(t); t), & (6.105) \\ &= \boldsymbol{p}(t) \cdot \boldsymbol{v}(t) - \tfrac{1}{2} m \boldsymbol{v}^2 + V(\boldsymbol{x}; t) & (6.106) \end{aligned}$$

giving

$$F(\boldsymbol{x}(t), \dot{\boldsymbol{x}}(t), \boldsymbol{v}(t); t) = \boldsymbol{p}(t) \cdot \dot{\boldsymbol{x}}(t) - H(\boldsymbol{x}(t), \boldsymbol{p}(t), \boldsymbol{v}(t); t). \tag{6.107}$$

The Euler-Lagrange equation using this expression for the state-space variables $\boldsymbol{x}(t)$ simply give (6.94) or its scalar form (6.97). However, we also have an equation using $\boldsymbol{v}$. Since $\dot{\boldsymbol{v}}(t)$ is not present this is simply

$$\nabla_{\boldsymbol{v}} F = \boldsymbol{p}(t) - m\boldsymbol{v}(t) = \boldsymbol{0}. \tag{6.108}$$

Substituting into (6.106) gives the expression

$$H(\boldsymbol{x}(t), \boldsymbol{p}(t); t) = \frac{1}{2m} \boldsymbol{p}^2 + V(\boldsymbol{x}; t), \tag{6.109}$$

for the Hamiltonian. Equations (6.97) and (6.98) (or their vector forms (6.94) and (6.96)) are in this context called *Hamilton's equations* [see (1.100) and the following discussion].

6.4 Pontriagin's Principle

In real-life problems the input control variable $u(t)$ is usually subject to some constraint on its magnitude. Typically this is of the form

$$u_{\mathrm{L}} \le u(t) \le u_{\mathrm{U}}. \tag{6.110}$$

Such a constraint will, of course, affect the derivation of an optimal control if the unconstrained optimum value for $u(t)$ given by (6.95) lies outside the range given by (6.110). We now need to obtain the 'best' value subject to the constraint. This will depend on whether we are looking for a maximum or minimum extremum for $\mathcal{I}[u, \boldsymbol{x}]$ given by (6.88). It follows from (6.92) that a minimum for $f(u(t), \boldsymbol{x}(t); t)$ corresponds to a maximum for $H(u(t), \boldsymbol{x}(t), \boldsymbol{p}(t); t)$ and vice-versa. If you are uncomfortable with this crossover you can change the sign in the definition of the Hamiltonian by replacing H by $-H$ in (6.92). This has the disadvantage of giving the wrong sign for the mechanics Hamiltonian (6.109).[5] We shall for the sake of definiteness assume that we are looking for a

[5]Because of this choice of sign in the problem Pontriagin's principle is variously referred to as *Pontriagin's minimum principle* and *Pontriagin's maximum principle* in the literature.

minimum of $f(u(t),\boldsymbol{x}(t);t)$ and thus a maximum of $H(u(t),\boldsymbol{x}(t),\boldsymbol{p}(t);t)$, which we assume to be a continuous differentiable function of u over the range (6.110). Then the condition for an optimum $u^*(t)$ for $u(t)$ is that

$$H(u^*(t),\boldsymbol{x}(t),\boldsymbol{p}(t);t) > H(u^*(t)+\delta u,\boldsymbol{x}(t),\boldsymbol{p}(t);t), \tag{6.111}$$

for all δu compatible with (6.110). Expanding in powers of δu we have

$$\left(\frac{\partial H}{\partial u}\right)^*\delta u + \frac{1}{2}\left(\frac{\partial^2 H}{\partial u^2}\right)^*(\delta u)^2 + \mathrm{O}\left([\delta u]^3\right) < 0. \tag{6.112}$$

If there is an extremum of H in the allowed range (6.112) gives the usual conditions for a maximum. Otherwise (6.112) can be approximated by its leading term to give

$$\left(\frac{\partial H}{\partial u}\right)^*\delta u < 0. \tag{6.113}$$

In these circumstances H is a monotonic function of u. If it is increasing then $\delta u < 0$, which means that $u^*(t) = u_{\mathrm{U}}$, if it is decreasing then $\delta u > 0$ which means that $u^*(t) = u_{\mathrm{L}}$. The optimum value of u will be at one of the boundaries of the range. Since the sign of $\partial H/\partial u$ will depend on the other variables $\boldsymbol{x}(t)$ and $\boldsymbol{p}(t)$ there may, as t varies, be a sudden change of $u^*(t)$ between u_{L} and u_{U}. This sudden change is called *bang-bang control* and

$$\mathcal{S}(t) = \frac{\partial H}{\partial u} \tag{6.114}$$

is called the *switching function*.

Example 6.4.1 The motor-racing problem. A vehicle of mass m moves in a straight line Ox under an engine thrust $mu(t)$, where the control variable $u(t)$ is constrained by the condition $|u(t)| \le u_{\mathbf{B}}$. Assuming that friction is absent find the control strategy which takes the vehicle from rest at $x = 0$ to rest at $x = x_{\mathrm{F}}$ in the least time.

Let $x_1(t) = x(t)$ and $x_2(t) = \dot{x}(t)$. Then

$$\begin{aligned} \dot{x}_1(t) &= x_2(t), \\ \dot{x}_2(t) &= u(t). \end{aligned} \tag{6.115}$$

If t_{F} is the time of the drive

$$\mathcal{I}[u,x_1,x_2] = t_{\mathrm{F}} = \int_0^{t_{\mathrm{F}}} \mathrm{d}t. \tag{6.116}$$

This gives $f(u(t),x_1(t),x_2(t);t) = 1$ and, from (6.92),

$$H(u(t),x_1(t),x_2(t),p_1(t),p_2(t)) = p_1(t)x_2(t) + p_2(t)u(t) - 1. \tag{6.117}$$

Then, from (6.97)

$$\begin{aligned} \dot{p}_1(t) &= 0, \\ \dot{p}_2(t) &= -p_1(t) \end{aligned} \tag{6.118}$$

and (6.98) gives, of course, (6.115). From (6.118)

$$\begin{aligned} p_1(t) &= \mathsf{A}, \\ p_2(t) &= \mathsf{B} - \mathsf{A}t \end{aligned} \tag{6.119}$$

and

$$\frac{\partial H}{\partial u} = p_2(t) = \mathsf{B} - \mathsf{A}t. \tag{6.120}$$

Thus H is a monotonic strictly increasing or strictly decreasing function of $u(t)$ for all t except at $t = \mathsf{B}/\mathsf{A}$ if this lies in the interval of time of the journey. Since the vehicle starts from rest at $t = 0$ and comes to rest at $t = t_{\mathrm{F}}$ it must be the case that $\ddot{x}(0) = u(0) > 0$ and $\ddot{x}(t_{\mathrm{F}}) = u(t_{\mathrm{F}}) < 0$. So in the early part of the journey $u(t) = u_{\mathbf{B}}$ and in the later part of the journey $u(t) = -u_{\mathbf{B}}$. The switch over occurs when $p_2(t)$ changes sign. So $p_2(t)$ is the switching function $\mathcal{S}(t)$. For the first part of the journey

$$\begin{aligned} \ddot{x}(t) &= u_{\mathbf{B}}, \\ \dot{x}(t) &= u_{\mathbf{B}}t, \\ x(t) &= \tfrac{1}{2}u_{\mathbf{B}}t^2. \end{aligned} \tag{6.121}$$

For the second part of the journey

$$\begin{aligned} \ddot{x}(t) &= -u_{\mathbf{B}}, \\ \dot{x}(t) &= u_{\mathbf{B}}(t_{\mathrm{F}} - t), \\ x(t) &= x_{\mathrm{F}} - \tfrac{1}{2}u_{\mathbf{B}}(t_{\mathrm{F}} - t)^2. \end{aligned} \tag{6.122}$$

Since both $\dot{x}(t)$ and $x(t)$ are continuous over the whole journey the switch occurs at

$$t = t_{\mathrm{S}} = t_{\mathrm{F}}/2, \tag{6.123}$$

with

$$\begin{aligned} t_{\mathrm{F}} &= 2\sqrt{x_{\mathrm{F}}/u_{\mathbf{B}}}, \\ x(t_{\mathrm{S}}) &= x_{\mathrm{F}}/2. \end{aligned} \tag{6.124}$$

These results give us the strategy for completing the journey in the minimum time. Suppose alternatively we chose

$$u(t) = u_{\mathbf{B}} - \mu t \tag{6.125}$$

throughout the whole journey. Then

$$\dot{x}(t) = u_{\mathrm{B}}t - \tfrac{1}{2}\mu t^2, \qquad x(t) = \tfrac{1}{2}u_{\mathrm{B}}t^2 - \tfrac{1}{6}\mu t^3. \tag{6.126}$$

The conditions that the journey ends with zero velocity and $x = x_{\mathrm{F}}$ gives

$$\mu = \sqrt{2u_{\mathrm{B}}^3/3x_{\mathrm{F}}}, \qquad t_{\mathrm{F}} = \sqrt{6x_{\mathrm{F}}/u_{\mathrm{B}}}. \tag{6.127}$$

Comparing this result with (6.124) we see that this procedure yield a journey time $\sqrt{3/2}$ longer. Plots of velocity against distance can be obtained using MAPLE :

```
> tF1:=(uB,xF)->2*sqrt(xF/uB);
```

$$tF1 := (uB,\, xF) \to 2\sqrt{\frac{xF}{uB}}$$

```
> v1:=(uB,xF,t)->uB*t;
```

$$v1 := (uB,\, xF,\, t) \to uB\, t$$

```
> x1:=(uB,xF,t)->uB*t^2/2;
```

$$x1 := (uB,\, xF,\, t) \to \frac{1}{2}\, uB\, t^2$$

```
> v2:=(uB,xF,t)->uB*(tF1(uB,xF)-t);
```

$$v2 := (uB,\, xF,\, t) \to uB\, (\mathrm{tF1}(uB,\, xF) - t)$$

```
> x2:=(uB,xF,t)->xF-uB*(tF1(uB,xF)-t)^2/2;
```

$$x2 := (uB,\, xF,\, t) \to xF - \frac{1}{2}\, uB\, (\mathrm{tF1}(uB,\, xF) - t)^2$$

```
> tF2:=(uB,xF)->sqrt(6*xF/uB);
```

$$tF2 := (uB,\, xF) \to \sqrt{6\,\frac{xF}{uB}}$$

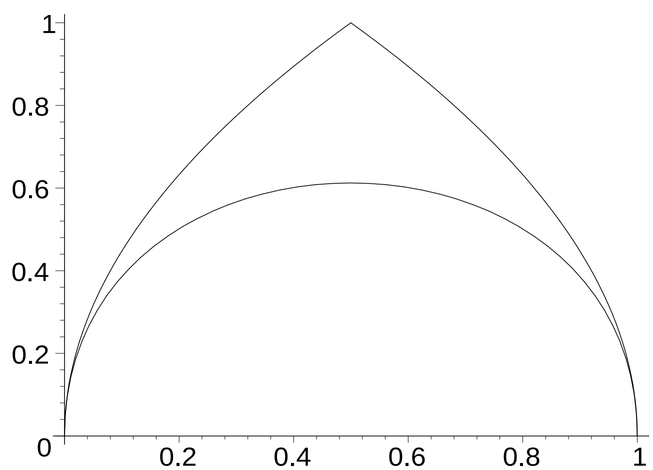
```
> v3:=(uB,xF,t)->t*uB*(tF2(uB,xF)-t)/tF2(uB,xF);
```

$$v3 := (uB,\, xF,\, t) \to \frac{t\, uB\, (\mathrm{tF2}(uB,\, xF) - t)}{\mathrm{tF2}(uB,\, xF)}$$

```
> x3:=(uB,xF,t)->t^2*uB*(3*tF2(uB,xF)-2*t)/(6*tF2(uB,xF));
```

$$x3 := (uB,\, xF,\, t) \to \frac{1}{6}\, \frac{t^2\, uB\, (3\, \mathrm{tF2}(uB,\, xF) - 2\, t)}{\mathrm{tF2}(uB,\, xF)}$$

```
> plot(
> {[x1(1,1,t),v1(1,1,t),t=0..tF1(1,1)/2],[x2(1,1,t),v2(1,1,t),
> t=tF1(1,1)/2..tF1(1,1)],[x3(1,1,t),v3(1,1,t),t=0..tF2(1,1)]
> },linestyle=1);
```



The upper curve corresponds to the optimum control and the lower to the control with linearly decreasing thrust. Since the area under the plot of the reciprocal of the velocity against distance would give the time of the journey the larger area in this plot corresponds to a shorter journey time. The latter part of the upper curve and its extension, which are given parametrically by the second and third of equations (6.122) is called the *switching curve*. It represents the points in the velocity-position space from which the final destination can be reached, arriving with zero velocity, by applying maximum deceleration. So the point where this curve is crossed by the first branch of the journey is the point when switching from acceleration to deceleration must occur.

Example 6.4.2 The soft-landing problem. A space vehicle of mass m is released at a height x_{I} above the surface of a planet with an upward velocity v_{I}. The engine exerts a downward thrust $mu(t)$, where $|u(t)| \leq u_{\mathsf{B}}$ and $v_{\mathrm{I}}^2 < 2x_{\mathrm{I}}(u_{\mathsf{B}} - g)$, g being the acceleration due to gravity. It is required to reach the surface in minimum time arriving there with zero velocity.

Let $x_1(t) = x(t)$ and $x_2(t) = \dot{x}(t)$. Then

$$\begin{aligned} \dot{x}_1(t) &= x_2(t), \\ \dot{x}_2(t) &= -[u(t) + g]. \end{aligned} \tag{6.128}$$

If t_{F} is the time taken to land

$$\mathcal{I}[u, x_1, x_2] = t_{\mathrm{F}} = \int_0^{t_{\mathrm{F}}} \mathrm{d}t. \tag{6.129}$$

This gives $f(u(t), x_1(t), x_2(t); t) = 1$ and, from (6.92),

$$H(u(t), x_1(t), x_2(t), p_1(t), p_2(t)) \quad = \quad p_1(t)x_2(t) - p_2(t)[u(t) + g] - 1. \tag{6.130}$$

Then, from (6.97)

$$\begin{aligned} \dot{p}_1(t) &= 0, \\ \dot{p}_2(t) &= -p_1(t) \end{aligned} \tag{6.131}$$

giving (6.118)

$$\begin{aligned} p_1(t) &= \mathsf{A}, \\ p_2(t) &= \mathsf{B} - \mathsf{A}t \end{aligned} \tag{6.132}$$

and

$$\frac{\partial H}{\partial u} = -p_2(t) = -\mathsf{B} + \mathsf{A}t. \tag{6.133}$$

Now $-p_2(t)$ is the switching function. H is a monotonic strictly increasing or strictly decreasing function of $u(t)$ for all t except at $t = \mathsf{B}/\mathsf{A}$ if this lies in the interval of time of the journey. Since the vehicle begins the landing process with an upward velocity its engine thrust must be initially downwards. (Otherwise the initial total downward thrust would be $m(g - u_{\mathsf{B}})$ which is negative.) There is one switch to an upward engine thrust to give a soft landing. For the first part of the journey

$$\begin{aligned} \ddot{x}(t) &= -(u_{\mathsf{B}} + g), \\ \dot{x}(t) &= v_{\mathrm{I}} - t(u_{\mathsf{B}} + g), \\ x(t) &= x_{\mathrm{I}} + v_{\mathrm{I}}t - \tfrac{1}{2}t^2(u_{\mathsf{B}} + g). \end{aligned} \tag{6.134}$$

For the second part of the journey

$$\begin{aligned}
\ddot{x}(t) &= (u_{\mathsf{B}} - g), \\
\dot{x}(t) &= (u_{\mathsf{B}} - g)(t - t_{\mathrm{F}}), \\
x(t) &= \tfrac{1}{2}(u_{\mathsf{B}} - g)(t - t_{\mathrm{F}})^2.
\end{aligned} \tag{6.135}$$

Since both $\dot{x}(t)$ and $x(t)$ are continuous over the whole journey the switch occurs at

$$t = t_{\mathrm{S}} = \frac{t_{\mathrm{F}}(u_{\mathsf{B}} - g) + v_{\mathrm{I}}}{2u_{\mathsf{B}}}, \tag{6.136}$$

with

$$t_{\mathrm{F}} = \frac{1}{u_{\mathsf{B}} + g}\left\{v_{\mathrm{I}} + \sqrt{\frac{2u_{\mathsf{B}}[v_{\mathrm{I}}^2 + 2x_{\mathrm{I}}(u_{\mathsf{B}} + g)]}{u_{\mathsf{B}} - g}}\right\}. \tag{6.137}$$

These results give us the strategy for completing the journey in the minimum time. The plot of velocity against distance is given by:

```
> v1:=(uB,xI,vI,g,t)->vI-t*(uB+g);
```

$$v1 := (uB,\, xI,\, vI,\, g,\, t) \to vI - t\,(uB + g)$$

```
> x1:=(uB,xI,vI,g,t)->xI+vI*t-t^2*(uB+g)/2;
```

$$x1 := (uB,\, xI,\, vI,\, g,\, t) \to xI + vI\,t - \frac{1}{2}\,t^2\,(uB + g)$$

```
> tF:=(uB,xI,vI,g)->(vI+sqrt(2*uB*(vI^2+2*xI*(uB+g))/(uB-g)))/(uB+g);
```

$$tF := (uB,\, xI,\, vI,\, g) \to \frac{vI + \sqrt{2\,\dfrac{uB\,(vI^2 + 2\,xI\,(uB + g))}{uB - g}}}{uB + g}$$

```
> v2:=(uB,xI,vI,g,t)->(uB-g)*(t-tF(uB,xI,vI,g));
```

$$v2 := (uB,\, xI,\, vI,\, g,\, t) \to (uB - g)\,(t - \mathrm{tF}(uB,\, xI,\, vI,\, g))$$

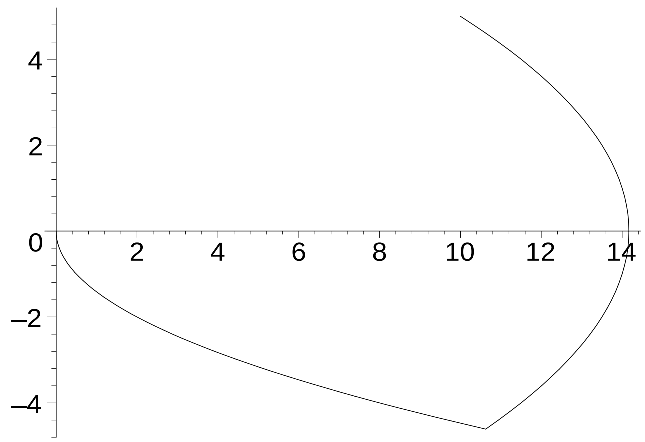
```
> x2:=(uB,xI,vI,g,t)->(uB-g)*(t-tF(uB,xI,vI,g))^2/2;
```

$$x2 := (uB,\, xI,\, vI,\, g,\, t) \to \frac{1}{2}\,(uB - g)\,(t - \mathrm{tF}(uB,\, xI,\, vI,\, g))^2$$

```
> tS:=(uB,xI,vI,g)->(tF(uB,xI,vI,g)*(uB-g)+vI)/(2*uB);
```

$$tS := (uB,\, xI,\, vI,\, g) \to \frac{1}{2}\,\frac{\mathrm{tF}(uB,\, xI,\, vI,\, g)\,(uB - g) + vI}{uB}$$

```
> plot(
> {[x1(2,10,5,1,t),v1(2,10,5,1,t),t=0..tS(2,10,5,1)],

> [x2(2,10,5,1,t),v2(2,10,5,1,t),t=tS(2,10,5,1)..tF(2,10,5,1)]
> });
```



The lower branch of the plot is the switching curve.

Example 6.4.3 The flywheel problem. The equation of motion of a flywheel with friction is

$$\ddot{\theta}(t) + 2\dot{\theta}(t) = u(t),$$

where the input variable u is restricted by $|u(t)| \leq 2$. It is required to bring the flywheel from $\theta(0) = 0$, $\dot{\theta}(0) = 0$ to $\theta = \pi$ with $\dot{\theta} = 0$ in minimum time. Use the *Hamilton-Pontriagin method* with $x_1 = \theta$ and $x_2 = \dot{\theta}$ to show that during the motion either $u(t) = 2$ or $u(t) = -2$, and that the variables $p_1(t)$ and $p_2(t)$ conjugate to $x_1(t)$ and $x_2(t)$ are given by

$$p_1(t) = \mathsf{C}, \qquad p_2(t) = \tfrac{1}{2}\mathsf{C} + \mathrm{B}\exp(2t),$$

where C and B are constants. Deduce that there is exactly one switch between $u(t) = 2$ and $u(t) = -2$.

Given that the total time for the motion of the flywheel is $t_{\rm f}$ and that the switch occurs when $t = t_{\rm S}$, show that

$$t_{\rm S} = \tfrac{1}{2}(\pi + t_{\rm f}) \qquad \text{with} \qquad t_{\rm f} = \operatorname{arccosh}[\exp(\pi)].$$

The two constraints are

$$\dot{x}_1(t) = x_2(t),$$
$$\dot{x}_2(t) = u(t) - 2\,x_2(t)$$

Since we are looking for an extremum of time $f(u(t), x_1(t), x_2(t); t) = 1$ and the Hamiltonian is

$$H(u(t), x_1(t), x_2(t), p_1(t), p_2(t)) = p_1(t)x_2(t) + p_2(t)[u(t) - 2\,x_2(t)] - 1.$$

Then the Hamiltonian-Pontriagin equations are

$$\dot{p}_1(t) = 0, \qquad \dot{p}_2(t) = 2\,p_2(t) - p_1(t),$$

with

$$\frac{\partial H}{\partial u} = p_2(t).$$

Thus

$$p_1(t) = \mathsf{C}$$

and

$$\frac{\mathrm{d}}{\mathrm{d}t}\,[p_2(t)\exp(-2t)] = -\mathsf{C}\exp(-2t),$$

giving

$$p_2(t) = \tfrac{1}{2}\mathsf{C} + \mathrm{B}\exp(2t).$$

This is the switching function. Since it is a monotonically increasing function it has at most one zero in the range $0 \le t \le t_{\rm F}$, so there will be at most one switch between $u(t) = 2$ and $u(t) = -2$. So there must be exactly one switch in order for the wheel to begin from rest and return to rest.

For the first part of the motion

$$\ddot{\theta}(t) + 2\dot{\theta}(t) = 2.$$

The auxiliary equation is $\lambda^2 + 2\lambda = 0$, giving the complementary function $\theta_{\rm c}(t) = \mathsf{A} + \mathrm{B}\exp(-2t)$ and a particular solution is $\theta_{\rm p}(t) = t$. Thus the general solution is

$$\theta(t) = t + \mathsf{A} + \mathrm{B}\exp(-2t).$$

Since $\theta(0) = \dot{\theta}(0) = 0$, $\mathsf{A} + \mathrm{B} = 0$ and $1 - 2\mathrm{B} = 0$. Thus

$$\theta(t) = t - \tfrac{1}{2}[1 - \exp(-2t)].$$

For the second part of the motion

$$\ddot{\theta}(t) + 2\dot{\theta}(t) = -2.$$

This simply changes the sign of the particular solution. So

$$\theta(t) = -t + \mathsf{A} + \mathsf{B}\exp(-2t).$$

Since $\theta(t_{\mathrm{F}}) = \pi$ and $\dot{\theta}(t_{\mathrm{F}}) = 0$, $-t_{\mathrm{F}} + \mathsf{A} + \mathsf{B}\exp(-2t_{\mathrm{F}}) = \pi$ and $-1 - 2\mathsf{B}\exp(-2t_{\mathrm{F}}) = 0$. Thus

$$\theta(t) = t_{\mathrm{F}} - t + \pi + \tfrac{1}{2}[1 - \exp\{2(t_{\mathrm{F}} - t)\}].$$

At the switching time $t = t_{\mathrm{S}}$ both $\theta(t)$ and $\dot{\theta}(t)$ are continuous.

$$t_{\mathrm{S}} - \tfrac{1}{2}[1 - \exp(-2t_{\mathrm{S}})] = t_{\mathrm{F}} - t_{\mathrm{S}} + \pi + \tfrac{1}{2}[1 - \exp[2(t_{\mathrm{F}} - t_{\mathrm{S}})],$$
$$1 - \exp(-2t_{\mathrm{S}}) = -1 + \exp[2(t_{\mathrm{F}} - t_{\mathrm{S}})].$$

Adding the first equation to one half the second gives

$$t_{\mathrm{S}} = \tfrac{1}{2}[\pi + t_{\mathrm{F}}]$$

and substituting back into the second gives

$$1 - \exp(-\pi - t_{\mathrm{F}}) = -1 + \exp(t_{\mathrm{F}} - \pi).$$

This simplifies to $\exp(\pi) = \cosh(t_{\mathrm{F}})$, and thus

$$t_{\mathrm{F}} = \operatorname{arccosh}[\exp(\pi)].$$

Problems 6

1) Find the extremum of

$$\mathcal{I}[x] = \int_0^1 \left\{\tfrac{1}{2}[\dot{x}(\tau)]^2 + x(\tau)\dot{x}(\tau)\right\} \mathrm{d}\tau$$

with $x(0) = 0$ and $x(1) = 5$, subject to the constraint

$$\mathcal{J}[x] = \int_0^1 x(\tau)\mathrm{d}\tau = 2.$$

2) Find the extremum of

$$\mathcal{I}[x] = \int_0^2 \left\{[\dot{x}(\tau)]^2 + x(\tau)\right\} \mathrm{d}\tau$$

with $x(0) = 1$ and $x(2) = 48$, subject to the constraint

$$\mathcal{J}[x] = \int_0^2 x(\tau)\tau\mathrm{d}\tau = 43.$$

3) The equation of a control system is

$$\dot{x}(t) = u(t).$$

Find the input $u(t)$ which minimizes

$$\mathcal{I}[u,x] = \int_0^{t_{\mathrm{F}}} \left\{[x(t)]^2 + [u(t)]^2\right\} \mathrm{d}t,$$

given that $x(0) = x(t_{\mathrm{F}}) = \alpha$. Find the minimum value of $\mathcal{I}[x]$.

4) The equation of a control system is

$$\ddot{x}(t) = u(t).$$

Find the input $u(t)$ which minimizes

$$\mathcal{I}[u] = \int_0^1 [u(t)]^2 \mathrm{d}t,$$

given that $x(0) = x(1) = 0$ and $\dot{x}(0) = \dot{x}(1) = 1$.

5) A system is governed by the equation

$$\dot{x}(t) = u(t) - x(t).$$

With $x(0) = x(t_{\mathrm{F}}) = x_0$, where $t_{\mathrm{F}} > 0$, find the input $u(t)$ which minimizes

$$\mathcal{I}[u] = \frac{1}{2}\int_0^{t_{\mathrm{F}}} [u(t)]^2 \mathrm{d}t.$$

Find the minimized value of $\mathcal{I}[u]$ and show that it is less than the value of $\mathcal{I}$ obtained by putting $x(t) = x_0$ over the whole range of t. Given that the condition $x(t_{\mathrm{F}}) = x_0$ is dropped and $x(t_{\mathrm{F}})$ is unrestricted show that the minimum value of $\mathcal{I}$ is zero.

6) A system is governed by the equation

$$\dot{x}(t) = u(t) - 1.$$

With $x(0) = 0$ find the input $u(t)$ which minimizes

$$\mathcal{I}[u] = \frac{1}{2}\int_0^{t_{\mathrm{F}}} \{[u(t)]^2 + [x(t)]^2\} \mathrm{d}t.$$

when

(a) $x(t_{\mathrm{F}}) = 1$.

(b) $x(t_{\mathrm{F}})$ is unrestricted.

Show that for case (a)

$$u(t) = 1 + \frac{\cosh(t)}{\sinh(t_{\mathrm{F}})}, \qquad x(t) = \frac{\sinh(t)}{\sinh(t_{\mathrm{F}})},$$

and for case (b)

$$u(t) = 1 - \frac{\cosh(t)}{\cosh(t_{\mathrm{F}})}, \qquad x(t) = -\frac{\sinh(t)}{\cosh(t_{\mathrm{F}})}.$$

Without evaluating the minimized $\mathcal{I}$ show from these results that it is smaller in case (b) than in case (a). *Think about why this is what you should expect.*

7) The equation of motion of a flywheel with friction is

$$\ddot{\theta}(t) + \mu\dot{\theta}(t) = u(t),$$

where μ is a positive constant and the input variable u is restricted by $|u(t)| \leq u_{\mathrm{B}}$. It is required to bring the flywheel from $\theta(0) = \theta_{\mathrm{I}}$, $\dot{\theta}(0) = 0$ to $\theta = \theta_{\mathrm{F}}$ with $\dot{\theta} = 0$ in minimum time. Use the *Hamilton-Pontriagin method* with $x_1 = \theta$ and $x_2 = \dot{\theta}$ to show that during the motion either $u = u_{\mathrm{B}}$ or $u = -u_{\mathrm{B}}$, and that the variables $p_1(t)$ and $p_2(t)$ conjugate to $x_1(t)$ an $x_2(t)$ are given by

$$p_1(t) = \mathsf{C}, \qquad p_2(t) = \mu^{-1}\mathsf{C} + \mathsf{B}\exp(\mu t),$$

where C and B are constants. Deduce that there is at most one switch between u_{B} and $-u_{\mathrm{B}}$.

Show from the transversality condition that there is no switch if $\dot{\theta}$ is unrestricted at $t = t_{\mathrm{F}}$ and that in this case t_{F} is give by the implicit equation

$$\mu^2(\theta_{\mathrm{F}} - \theta_{\mathrm{I}}) = u_{\mathrm{B}}[\mu\, t_{\mathrm{F}} - 1 + \exp(-\mu\, t_{\mathrm{F}})].$$

8) A rocket is ascending vertically above the earth. Its equations of motion are

$$\ddot{x}(t) = \frac{\mathsf{C}\, u(t)}{m(t)} - g, \qquad \dot{m}(t) = -u(t).$$

The propellant mass flow can be controlled subject to $0 \leq u(t) \leq u_{\mathrm{U}}$. The mass, height and velocity at time $t = 0$ are all known and it is required to *maximize* the height subsequently reached (when the time is taken to be t_{F}). Show that the optimum control has the form

$$u^*(t) = \begin{cases} u_{\mathrm{U}}, & \mathcal{S}(t) < 0, \\ 0, & \mathcal{S}(t) > 0, \end{cases}$$

where the switching function $\mathcal{S}(t)$ satisfies the equation $\dot{\mathcal{S}}(t) = \mathsf{D}/m(t)$, for some *positive* constant D. Given that the switch occurs at t_{S}, show that

$$\mathcal{S}(t) = \begin{cases} -\dfrac{\mathsf{D}}{u_{\mathrm{U}}} \ln\left\{\dfrac{m(t)}{m(t_{\mathrm{S}})}\right\}, & 0 \le t \le t_{\mathrm{S}}, \\[2ex] -\dfrac{\mathsf{D}(t_{\mathrm{S}} - t)}{m(t_{\mathrm{S}})}, & t_{\mathrm{S}} \le t \le t_{\mathrm{F}}. \end{cases}$$

9) A vehicle moves along a straight road, its distance from the starting point at time t being denoted by $x(t)$. The motion of the vehicle is governed by

$$\ddot{x}(t) = u(t) - k, \qquad k > 0,$$

where $u(t)$, the thrust per unit mass, is the control variable. At time $t = 0$, $x = \dot{x} = 0$ and the vehicle is required to reach $x = L > 0$, with $\dot{x} = 0$, in minimum time, subject to the condition that $|u(t)| < u_{\mathrm{B}}$, where $u_{\mathrm{B}} > k$. Using the state-space variables $x_1 = x$ and $x_2 = \dot{x}$, construct the Hamiltonian for the Hamiltonian-Pontryagin method and show that during the motion either $u(t) = u_{\mathrm{B}}$ or $u(t) = -u_{\mathrm{B}}$.

Show that $u(t)$ cannot switch values more than once and that a switch occurs when

$$x = \frac{L(u_{\mathrm{B}} + k)}{2u_{\mathrm{B}}}.$$

Find the time taken for the journey.

Chapter 7

Complex Variable Methods

7.1 Introduction

At some stage in many problems in linear control theory we need to analyze a relationship of the form

$$\bar{y}(s) = G(s)\bar{u}(s), \tag{7.1}$$

where $\bar{u}(s)$ and $\bar{y}(s)$ are respectively the Laplace transforms of the input $u(t)$ and output $y(t)$. The transfer function $G(s)$ is a rational function of s. That is

$$G(s) = \frac{\psi(s)}{\phi(s)}, \tag{7.2}$$

where $\psi(s)$ and $\phi(s)$ are polynomials in s. Our interest has been, not only in obtaining $y(t)$ for a given form for $u(t)$ for which we have usually used partial fraction methods, but in determining the stability of the system, for which we have the Routh-Hurwitz criterion described in Sect. 5.2.1. In this chapter we shall describe other methods which rely on using the properties of (7.1) in the complex s-plane.

7.2 Results in Complex Variable Theory

7.2.1 Cauchy's Residue Theorem

We consider a meromorphic function $F(s)$ of the complex variable s. Suppose that s_0 is a pole of order m of $F(s)$. Then

$$F(s) = f(s) + \sum_{j=1}^{m} \frac{\beta_j}{(s-s_0)^j}, \tag{7.3}$$

where $f(z)$ is analytic in a neighbourhood of s_0 and $\beta_1, \ldots, \beta_m$ are constants. Then β_1 is called the *residue of* $F(s)$ *at* s_0 and is denoted by $\mathrm{Res}(F; s_0)$.

Example 7.2.1 Suppose

$$F(s) = \frac{\exp(s\,t)}{(s-5)^3}. \tag{7.4}$$

This has a third-order pole at $s = 5$. Expand $\exp(s\,t)$ about $s = 5$

$$\begin{aligned} \exp(s\,t) \quad &= \quad \exp(5\,t) + t(s-5)\exp(5\,t) + \tfrac{1}{2!}t^2(s-5)^2\exp(5\,t) \\ &\quad + \tfrac{1}{3!}t^3(s-5)^3\exp(5\,t) + \mathrm{O}[(s-5)^4]. \end{aligned} \tag{7.5}$$

So dividing by $(s-5)^3$ we see that

$$\mathrm{Res}(F;5) = \tfrac{1}{2!}t^2\exp(5\,t). \tag{7.6}$$

A result which is very useful in this context is:[1]

Theorem 7.2.1 *If $F(s)$ has a pole of order m at s_0*

$$\mathrm{Res}(F;s_0) = \frac{1}{(j-1)!}\lim_{s\to s_0}\frac{\mathrm{d}^{j-1}}{\mathrm{d}s^{j-1}}\left[(s-s_0)^j F(s)\right], \tag{7.7}$$

for any $j \geq m$.

With $j = 3$ you will see that this gives you a quick way to obtain (7.6).[2] A *closed contour* is a closed curve in the complex plane with a direction. Given any closed contour γ and some point s_0 not lying on γ the *winding number or index of γ with respect to* s_0, denoted by $\mathrm{Ind}(\gamma;s_0)$ is the number of times the contour passes around s_0 in the anticlockwise direction minus the number it passes around s_0 in the clockwise direction.

Theorem 7.2.2 *Let γ be a closed contour and s_0 a point not lying on γ then*

$$\mathrm{Ind}(\gamma;s_0) = \frac{1}{2\mathrm{i}\pi}\int_\gamma \frac{\mathrm{d}s}{s-s_0}. \tag{7.8}$$

We can now state the *Cauchy residue theorem.*

Theorem 7.2.3 *Let $F(s)$ be a meromorphic function with poles $s_1, s_2, \ldots, s_n$ and let γ be a contour not passing through any of these points. Then*

$$\frac{1}{2\pi\mathrm{i}}\int_\gamma F(s)\mathrm{d}s = \sum_{j=1}^{n}\mathrm{Res}(f;s_j)\mathrm{Ind}(\gamma;s_j). \tag{7.9}$$

[1] Proofs for this theorem and the Cauchy residue theorem are given in any book on Complex Variable Theory and in particular in the notes for course CM322C.

[2] It is also easy to check that it give the correct result with $j = 4$

7.2.2 The Argument Principle

A useful consequence of the residue theorem is the *argument principle*.[3]

Theorem 7.2.4 *Let $F(s)$ be a meromorphic function and let γ be a simple closed curve transcribed in the anticlockwise direction*[4] *not passing through any of the poles or zeros of $F(s)$. Then*

$$\frac{1}{2\mathrm{i}\pi}\int_\gamma \frac{F'(s)}{F(s)}\mathrm{d}s = \{\textit{Number of zeros in } \gamma\} - \{\textit{Number of poles in } \gamma\}, \tag{7.10}$$

where in each case multiplicity is included in the counting.

This result can be re-expressed in a slightly different way by writing

$$\frac{F'(s)}{F(s)}\mathrm{d}s = \frac{\mathrm{d}F(s)}{F(s)} = \mathrm{d}\left\{\ln[F(s)]\right\}. \tag{7.11}$$

Now let

$$F(s) = |F(s)|\exp(\mathrm{i}\Theta). \tag{7.12}$$

So, since $\ln|F(s)|$ is single-valued,

$$\begin{aligned}
\frac{1}{2\mathrm{i}\pi}\int_\gamma \frac{F'(s)}{F(s)}\mathrm{d}s &= \frac{1}{2\mathrm{i}\pi}\int_\gamma \mathrm{d}\left\{\ln[F(s)]\right\} \\
&= \frac{1}{2\mathrm{i}\pi}\int_\gamma \mathrm{d}\left\{\ln|F(s)|\right\} + \frac{1}{2\pi}\int_\gamma \mathrm{d}\Theta \\
&= \frac{1}{2\pi}\int_\gamma \mathrm{d}\Theta.
\end{aligned} \tag{7.13}$$

This final term measures the change in argument (*hence the name 'argument principle'*) of $F(s)$ along γ in units of 2π. Suppose now we consider the mapping

$$s \Longrightarrow F(s). \tag{7.14}$$

As s describes the curve γ, $F(s)$ will describe some other closed curve Γ_{F} and the last term in (7.13) is just the number of times that Γ_{F} passes around the origin, or simply the winding number $\mathrm{Ind}(\Gamma_{\mathrm{F}};0)$. Thus

$$\mathrm{Ind}(\Gamma_{\mathrm{F}};0) = \{\text{Number of zeros in } \gamma\} - \{\text{Number of poles in } \gamma\}. \tag{7.15}$$

[3] Again proved in the notes for CM322C.

[4] For a simple closed curve described in the anticlockwise direction the winding number of every point in the complex plane is either zero (corresponding to outside) or one (corresponding to inside).

7.3 The Inverse Laplace Transform

According to (2.23) the Laplace transform in (7.1) can be inverted by

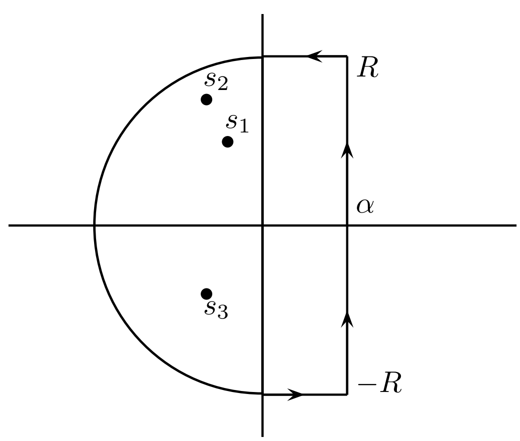
$$y(t) = \frac{1}{2\pi \mathrm{i}} \int_{\alpha - \mathrm{i}\infty}^{\alpha + \mathrm{i}\infty} G(s)\bar{u}(s)\exp(st)\mathrm{d}s, \tag{7.16}$$

where $\alpha > \eta$ and the integral is along the vertical line $\Re\{s\} = \alpha$ in the complex s–plane. According to Sect. 2.3 the parameter η is such that the integral defining the Laplace transform converges when $\Re\{s\} > \eta$. This might seem to be a problem for solving the integral (7.16). However, we do have some information. We know that $G(s)$ is a *meromorphic function*, that is its only singularities are poles. Suppose that these are located at the points $s_1, s_2, \ldots, s_r$ in the complex s–plane. The pole s_j will contribute a factor $\exp(s_j t)$ to the Laplace transform. Thus for convergence we must have $\Re\{s\} > \Re\{s_j\}$ and so $\alpha > \eta > \Re\{s_j\}$. This applies to all the poles of $G(s)$. If we also assume that $\bar{u}(s)$ is also a meromorphic function these must also be included and we final have the conclusion that *the vertical line of integration in (7.16) must be to the right of all the poles of $G(s)\bar{u}(s)$.*

The problem now is to evaluate the integral (7.16). We have assumed the $\bar{u}(s)$ is meromorphic, so the integrand is meromophic with poles denoted by $s_1, s_2, \ldots, s_n$.[5] Now define

$$y_R(t) = \frac{1}{2\pi \mathrm{i}} \int_{\gamma_{\mathrm{R}}} G(s)\bar{u}(s)\exp(st)\mathrm{d}s, \tag{7.17}$$

where γ_{R} is



[5]These are all the poles of $G(s)$ and all the poles of $\bar{u}(s)$ unless one of these functions has a zero which annihilates a pole of the other.

We take R to be sufficiently large so that all the poles of the integrand are within the contour. Then since the winding number of γ_{R} with respect to each of the poles is one

$$\begin{aligned} y_{\mathrm{R}}(t) &= \frac{1}{2\pi\mathrm{i}}\int_{\gamma_{\mathrm{R}}} G(s)\bar{u}(s)\exp(st)\mathrm{d}s \\ &= \sum_{j=1}^{n} \mathrm{Res}(G(s)\bar{u}(s)\exp(st); s_j). \end{aligned} \tag{7.18}$$

The final part of this argument, on which we shall not spend any time, is to show that, subject to certain boundedness conditions on the integrand, in the limit $R \to \infty$ the contributions to the contour integral from the horizontal sections and from the semicircle become negligible. Thus

$$y(t) = \frac{1}{2\pi\mathrm{i}}\int_{\alpha-\mathrm{i}\infty}^{\alpha+\mathrm{i}\infty} G(s)\bar{u}(s)\exp(st)\mathrm{d}s = \sum_{j=1}^{n} \mathrm{Res}(G(s)\bar{u}(s)\exp(st); s_j). \tag{7.19}$$

Example 7.3.1 Consider the final part of Example 3.4.2 where

$$y(t) = \frac{u_0}{2\pi\mathrm{i}}\int_{\alpha-\mathrm{i}\infty}^{\alpha+\mathrm{i}\infty} \left\{\frac{1}{s} - \frac{s + a\omega_0^2}{\left(s+\frac{1}{2}a\omega_0^2\right)^2 + \omega^2}\right\}\exp(st)\mathrm{d}t \tag{7.20}$$

and $\omega^2 = \omega_0^2 - \frac{1}{4}a^2\omega_0^4$.

The first term has a simple pole at the origin and the second has two simple poles at

$$s = -\tfrac{1}{2}a\omega_0^2 \pm \mathrm{i}\omega. \tag{7.21}$$

According to sign of a we now need to choose α so that the line of integration is to the right of these poles. Then the contribution to the pole at the origin, which can be treated separately is just u_0. The contributions from the other poles are

$$\begin{aligned} & -u_0\left\{(s+\tfrac{1}{2}a\omega_0^2 \mp \mathrm{i}\omega)\frac{(s+a\omega_0^2)\exp(st)}{\left(s+\frac{1}{2}a\omega_0^2\right)^2+\omega^2}\right\}_{s=-\frac{1}{2}a\omega_0^2\pm\mathrm{i}\omega} \\ = \quad & -u_0\left\{\frac{(s+a\omega_0^2)\exp(st)}{s+\frac{1}{2}a\omega_0^2\pm\mathrm{i}\omega}\right\}_{s=-\frac{1}{2}a\omega_0^2\pm\mathrm{i}\omega} \\ = \quad & -\frac{1}{2}u_0\left[\left(1\mp\frac{\mathrm{i}a\omega_0^2}{2\omega}\right)\exp\left(\pm\mathrm{i}\omega t - \tfrac{1}{2}a\omega_0^2 t\right)\right]. \end{aligned} \tag{7.22}$$

Adding these contributions together gives (3.47).

7.4 The Stability of a System with Unit Feedback

7.4.1 The Frequency Response Function

In Example 7.3.1 we considered a system with $u(t) = u_0$, a constant input turned on at $t = 0$. The effect on the output was a constant contribution given by the pole at the origin. If, for the general case (7.1), we choose $u(t) = \exp(\mathrm{i}\omega t)$, then $\bar{u}(s) = (s - \mathrm{i}\omega)^{-1}$ and

$$y(t) = \frac{1}{2\pi\mathrm{i}} \int_{\alpha-\mathrm{i}\infty}^{\alpha+\mathrm{i}\infty} \frac{G(s)}{s - \mathrm{i}\omega} \exp(st)\mathrm{d}s. \tag{7.23}$$

Unless $G(s)$ itself has a pole at $s = \mathrm{i}\omega$, which would produce a resonance effect, the contribution to $y(t)$ from this pole is $G(\mathrm{i}\omega)\exp(\mathrm{i}\omega t)$. The amplitude factor $G(\mathrm{i}\omega)$ is called the *frequency response function.*

7.4.2 The Nyquist Criterion

Now suppose that $G(s)$ is the open-loop transfer function and we introduce a unit feedback to give

$$\bar{y}(s) = G_{\mathrm{CL}}(s)\bar{u}(s), \tag{7.24}$$

where

$$G_{\mathrm{CL}}(s) = \frac{G(s)}{1 + G(s)}. \tag{7.25}$$

We consider the mapping

$$s \Longrightarrow G(s). \tag{7.26}$$

The imaginary axis in the s plane is $s = \mathrm{i}\omega$, $-\infty \le \omega \le +\infty$. This will be mapped into a curve Γ_{G} in the $Z = X + \mathrm{i}Y$ plane given by

$$X(\omega) + \mathrm{i}Y(\omega) = G(\mathrm{i}\omega), \qquad -\infty \le \omega \le +\infty. \tag{7.27}$$

Assuming that $\psi(s)$ and $\phi(s)$ in (7.2) are both polynomials with real coefficients with $\psi(s)$ a lower degree that $\phi(s)$ then Γ_{G} is a closed curve[6] in the Z–plane with $G(\mathrm{i}\infty) = G(-\mathrm{i}\infty) = 0$ and symmetry about the X–axis. This curve is called the *Nyquist locus*[7] of $G(s)$. We now prove the following theorem:

Theorem 7.4.1 *If the number of poles of $G(s)$ with $\Re\{s\} > 0$ is equal to* $\mathrm{Ind}(\Gamma_G; -1)$, *the index of the Nyquist locus of $G(s)$ with respect to $Z = -1$, then the closed-loop transfer function $G_{CL}(s)$ is asymptotically stable.*

[6] Although, in Example 7.4.2, we see a case where this 'closed' curve has a discontinuous jump from $+\mathrm{i}\infty$ to $-\mathrm{i}\infty$ as ω passes zero.

[7] Or sometimes 'plot' or 'diagram'.

Proof: Let

$$F(s) = 1 + G(s) = \frac{1}{1 - G_{\mathrm{CL}}(s)}. \tag{7.28}$$

The poles of $G_{\mathrm{CL}}(s)$ will be the zeros of $F(s)$.

Let γ_{R} be the the closed contour (traversed in the clockwise direction) consisting of the imaginary axis in the s-plane from $s = -\mathrm{i}R$ to $s = \mathrm{i}R$, together with a semicircle, centre the origin, of radius R to the right of the imaginary axis. We assume for simplicity that $G(s)$ has no poles on the imaginary axis.[8] Then for sufficiently large R all the poles and zeros of $G(s)$ and $F(s)$ with $\Re\{s\} > 0$ will be enclosed within γ_{R}. This closed curve in the s–plane is called the *Nyquist contour*. Now plot $F(\mathrm{i}\omega) = U(\omega) + \mathrm{i}V(\omega)$ in the $W = U + \mathrm{i}V$ plane. The contour Γ_{F}, produced by this is, apart from an infinitesimal arc near the origin the image of γ_{R} in the s–plane. (With very large R the arc of radius R is contracted into a very small arc around the origin in the Z–plane.) Thus, from (7.15),[9]

$$\begin{aligned} \mathrm{Ind}(\Gamma_{\mathrm{F}};0) \quad = \quad & \{\text{Number of poles of } F(s) \text{ with } \Re\{s\} > 0\} \\ & - \{\text{Number of zeros of } F(s) \text{ with } \Re\{s\} > 0\}. \end{aligned} \tag{7.29}$$

When $F(\mathrm{i}\omega) = 0$, $G(\mathrm{i}\omega) = -1$; so

$$\mathrm{Ind}(\Gamma_{\mathrm{F}};0) = \mathrm{Ind}(\Gamma_{\mathrm{G}};-1). \tag{7.30}$$

Since it is also the case that the poles of $F(s)$ and $G(s)$ coincide, if

$$\mathrm{Ind}(\Gamma_{\mathrm{G}};-1) = \{\text{Number of poles of } G(s) \text{ with } \Re\{s\} > 0\} \tag{7.31}$$

then

$$\{\text{Number of zeros of } F(s) \text{ with } \Re\{s\} > 0\} = 0 \tag{7.32}$$

and thus

$$\{\text{Number of poles of } G_{\mathrm{CL}}(s) \text{ with } \Re\{s\} > 0\} = 0, \tag{7.33}$$

which means that $G_{\mathrm{CL}}(s)$ is asymptotically stable.

An immediate consequence of this is the *Nyquist criterion*[10] that: *If $G(s)$ is itself asymptotically stable, and thus has no poles with $\Re\{s\} > 0$, then the closed-loop transfer function is asymptotically stable if Γ_G does not encircle the point -1.*

Example 7.4.1 Let

$$G(s) = \frac{\mathsf{K}}{s - \mathsf{Q}}, \qquad \mathsf{K} > 0. \tag{7.34}$$

[8] If it does the contour can be diverted around them.

[9] The change of sign, as compared to (7.15), is because the Nyquist contour is transcribed in the clockwise direction.

[10] In some texts the more general result Thm. 7.4.1 is called the Nyquist criterion.

with the closed-loop transfer function

$$G_{\mathrm{CL}}(s) = \frac{G(s)}{1+G(s)} = \frac{\mathsf{K}}{s-\mathsf{Q}+\mathsf{K}}. \tag{7.35}$$

With $X(\omega)+\mathrm{i}Y(\omega)=G(\mathrm{i}\omega)$

$$X(\omega) = -\frac{\mathsf{KQ}}{\mathsf{Q}^2+\omega^2}, \qquad Y(\omega) = -\frac{\mathsf{K}\omega}{\mathsf{Q}^2+\omega^2}. \tag{7.36}$$

This gives

$$[X+(\mathsf{K}/2\mathsf{Q})]^2+Y^2 = (\mathsf{K}/2\mathsf{Q})^2. \tag{7.37}$$

Thus Γ_{G} is a circle centre $X=-(\mathsf{K}/2\mathsf{Q})$, $Y=0$ and radius $|\mathsf{K}/2\mathsf{Q}|$. Now $G(s)$ has a pole at $s=\mathsf{Q}$. So, for $\mathsf{Q}>0$, the open-loop transfer function is unstable and we need to use Thm. 7.4.1 rather than the special case which is the Nyquist criterion. For the closed-loop transfer function to be stable Γ_{G} must pass around -1 once, which will be the case when $\mathsf{K}>\mathsf{Q}$. The reasoning for $\mathsf{Q}<0$ follows in the same way with the open-loop transfer function now stable, leading to an application of the Nyquist criterion. In all cases the results agree with the straightforward deductions made by our usual methods.

Example 7.4.2 In Problem Sheet 7, Example 2, we consider the system with open-loop transfer function

$$G(s) = \frac{\mathsf{K}(\alpha+\beta s)}{s(1+2s)^2}. \tag{7.38}$$

and unit feedback. We used the Routh-Hurwitz method to show that the closed-loop transfer function was asymptotically stable if $\beta<\alpha$ and $0<\mathsf{K}<(\alpha-\beta)^{-1}$, or $\alpha<\beta$ and $\mathsf{K}>0$. We now investigate this result using MAPLE and the Nyquist method.

In this case $G(s)$ has no poles so $G_{\mathrm{CL}}(s)$ will be asymptotically stable if the Nyquist plot does not pass around $Z=-1$. We first use MAPLE to find the real and imaginary parts of $G(\mathrm{i}\omega)$.

```
> G:=(s,K,a,b)->K*(a+b*s)/(s*(1+2*s)^2):
> X:=(w,K,a,b)->simplify(evalc(Re(G(I*w,K,a,b)))):
> X(w,K,a,b);
```

$$-\frac{K\,(-b+4\,b\,w^2+4\,a)}{1+8\,w^2+16\,w^4}$$

```
> Y:=(w,K,a,b)->simplify(evalc(Im(G(I*w,K,a,b)))):
> Y(w,K,a,b);
```

$$\frac{K\,(-a+4\,a\,w^2-4\,b\,w^2)}{w\,(1+8\,w^2+16\,w^4)}$$

We see that $Y(-\omega)=-Y(\omega)$. The Nyquist plot is symmetric about the X–axis with both ends of the curve at the origin ($X(\pm\infty)=Y(\pm\infty)=0$). Unless $\alpha=0$, $Y(\omega)\to\mp 0\times(\mathsf{K}\alpha)$, as $\omega\to\pm 0$, giving an infinite discontinuity in the plot as ω passes through zero. If $\alpha=0$ then $Y(\omega)\to 0$ as, $\omega\to 0$, with $X(\omega)\to\mathsf{K}\beta$.

In the case $\alpha\neq 0$ the plot cuts the X–axis when $Y(\omega)=0$ giving

$$\omega=\pm\sqrt{\frac{\alpha}{4(\alpha-\beta)}}. \tag{7.39}$$

If $\alpha>0$ and $\alpha>\beta$ or $\alpha<0$ and $\beta>\alpha$ the two branches of the plot cross at this value of ω. We calculate the point on the X–axis where this occurs for $\beta=\alpha/2>0$. In this case (7.39) gives $\omega=1/\sqrt{2}$.

```
> X1:=(w,K,a)->simplify(X(w,K,a,a/2)):
> X1(w,K,a);
```

$$-\frac{1}{2}\,\frac{K\,a\,(7+4\,w^2)}{1+8\,w^2+16\,w^4}$$

```
> Y1:=(w,K,a)->simplify(Y(w,K,a,a/2)):
> Y1(w,K,a);
```

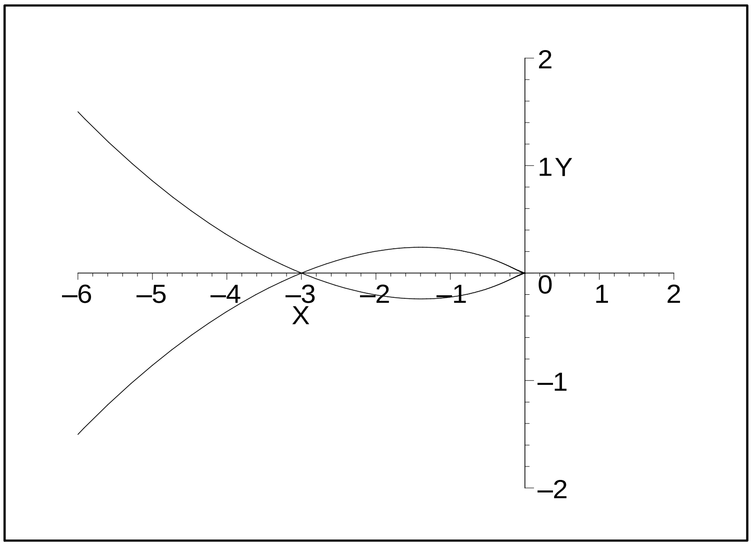
$$\frac{K\,a\,(-1+2\,w^2)}{w\,(1+8\,w^2+16\,w^4)}$$

```
> simplify(X1(1/sqrt(2),K,a));
```

$$-\frac{1}{2}\,K\,a$$

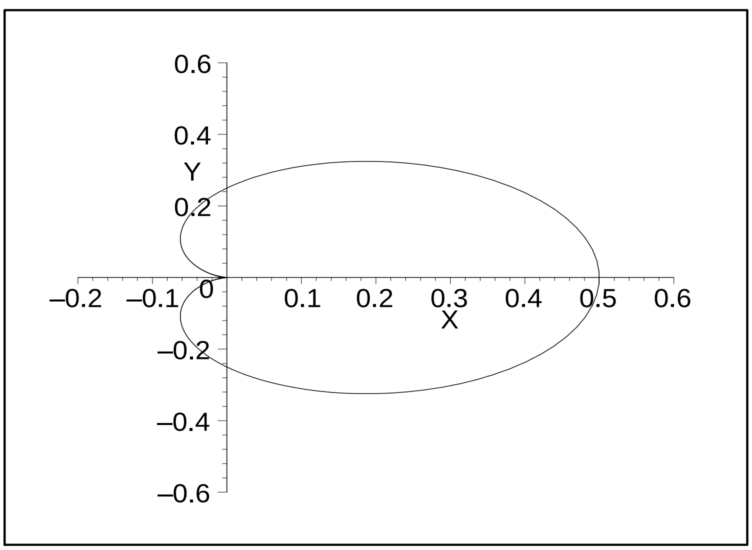
So, for $\beta=\alpha/2>0$, the closed-loop transfer function is stable if $-\frac{1}{2}\mathsf{K}\alpha>-1$. That is if $\mathsf{K}<2/\alpha$, which is the result obtained by the Routh-Hurwitz method. We compute the Nyquist plot for an unstable case when $\mathsf{K}=1$, $\alpha=6$, $\beta=3$.

```
> with(plots):
> plot([X(w,1,6,3),Y(w,1,6,3),
> w=-infinity..infinity],X=-6..2,Y=-2..2,numpoints=1000);
```



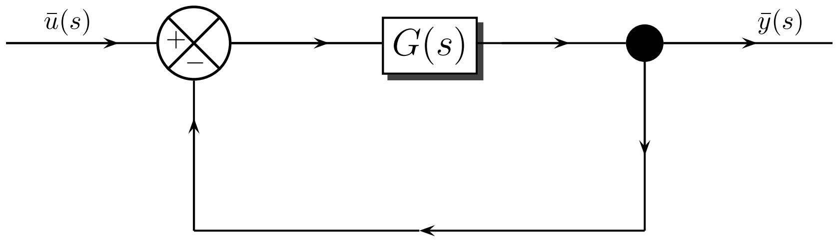
When $\alpha = 0$, the system will be stable if the point where the plot cuts the X–axis is to the right of the origin. That is $\mathsf{K}\beta > 0$. We plot the case $\mathsf{K} = \frac{1}{2}$, $\beta = 1$.

```
> plot([X(w,0.5,0,1),Y(w,0.5,0,1),

> w=-infinity..infinity],X=-0.2..0.6,Y=-0.6..0.6,numpoints=1000);
```



Problems 7

1) For the system with block diagram:



$$G(s) = \frac{\mathsf{K}}{(1+s)^3}.$$

Determine the closed loop transfer function $G_{\text{CL}}(s)$ and use the Routh-Hurwitz stability criteria to show that the system is stable for $-1 < \mathsf{K} < 8$.

Find the functions $X(\omega)$ and $Y(\omega)$ so that $G(\mathrm{i}\omega) = X(\omega) + \mathrm{i}Y(\omega)$. Define the Nyquist plot and state the *Nyquist criterion* which relates the form of this curve to the stability of $G_{\text{CL}}(s)$. Show that, for the given example, the result obtained from the Nyquist criterion confirms the result obtained by the Routh-Hurwitz procedure.

Chapter 8

Non-Linear Systems

8.1 Introduction

In Sects. 1.5 and 1.6 we discussed systems of differential equations most of which were non-linear. As we have seen, it is no restriction to concentrate on a first-order system since higher-order equations governing a system can be expressed as a system of first-order equations by introducing additional variables. For simplicity we shall again consider single input/single output systems and we shall also suppose the system is autonomous. A realization will then be of the form

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{X}(u(t), \boldsymbol{x}(t)), \tag{8.1}$$

$$y(t) = Y(\boldsymbol{x}(t)). \tag{8.2}$$

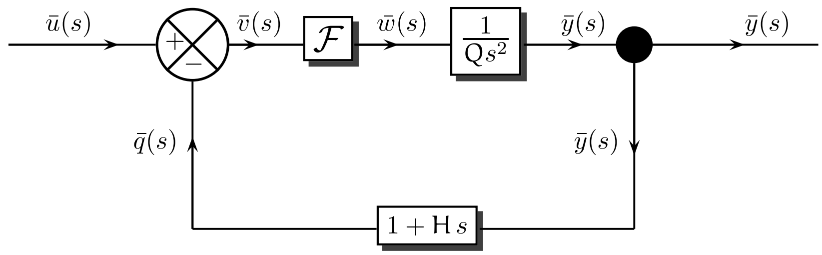
where $\boldsymbol{x}(t)$ is the n-dimensional state-space vector, just as in the linear version (4.44)–(4.45) of these equations.

8.2 Laplace Transforms and Transfer Functions

Because the Laplace transform is a *linear transform* its use in non-linear problems is very limited. To illustrate this point consider the system with equations

$$\begin{aligned}
\bar{y}(s) &= \frac{\bar{w}(s)}{\mathsf{Q}s^2}, \\
\bar{w}(s) &= \mathcal{F}\{\bar{v}(s)\}, \\
\bar{v}(s) &= \bar{u}(s) - \bar{q}(s), \\
\bar{q}(s) &= (1 + \mathsf{H}\, s)\bar{y}(s).
\end{aligned} \tag{8.3}$$

and block diagram



where $\mathcal{F}$ is some non-linear operator. Assuming all the variables and necessary derivatives are zero for $t < 0$ we eliminate all the intermediate variables to give

$$\mathsf{Q}\,\ddot{y}(t) = \mathcal{L}^{-1}\left\{\mathcal{F}\left\{\mathcal{L}\{u(t) - \mathsf{H}\,\dot{y}(t) - y(t)\}\right\}\right\}. \tag{8.4}$$

For a linear system the operator $\mathcal{F}$ would simply apply a multiplicative rational function of s to the Laplace transform of $u(t) - \mathsf{H}\,\dot{y}(t) - y(t)$ and the final effect of the sequence of operators $\mathcal{L}^{-1}\left\{\mathcal{F}\left\{\mathcal{L}\{\cdot\}\right\}\right\}$ would be to produce some linear combination of $u(t)$ and $y(t)$ and their derivatives. For a non-linear system we define the non-linear function

$$f(\cdot) = \frac{1}{\mathsf{Q}}\mathcal{L}^{-1}\left\{\mathcal{F}\left\{\mathcal{L}\{\cdot\}\right\}\right\}. \tag{8.5}$$

Now introduce the two state-space variables

$$\begin{aligned} x_1(t) &= y(t) + \mathsf{H}\,\dot{y}(t), \\ x_2(t) &= -\dot{y}(t) \end{aligned} \tag{8.6}$$

and we have the realization

$$\begin{aligned} \dot{x}_1(t) &= \mathsf{H}\,f(u(t) - x_1(t)) - x_2(t), \\ \dot{x}_2(t) &= -f(u(t) - x_1(t)), \\ y(t) &= x_1(t) + \mathsf{H}\,x_2(t), \end{aligned} \tag{8.7}$$

which is of the form (8.1)–(8.2) with $n = 2$.

8.3 Constant Control-Variable Systems

In the cases where the realization is one-dimensional there is not much scope or advantage for $y(t)$ to be anything other than $x(t)$, so we assume that (8.2) is simply $y(t) = x(t)$ and (8.1) is

$$\dot{x}(t) = X(u(t), x(t)). \tag{8.8}$$

In Sect. 1.6.1 we considered a number of examples where, although we didn't speak of it in these terms, we had a constant input $u(t) = u_0$, for which we usually used the variable a. We examined the equilibrium points of the system to determine their stability and showed that this can alter as u_0 changes. This leads to bifurcations at particular values of u_0, where the stability of an equilibrium point changes and/or new equilibrium points appear. The simplest case we considered was Example 1.6.1, which in our present notation is

$$\dot{x}(t) = u_0 - x^2(t). \tag{8.9}$$

We saw that, for $u_0 > 0$, there were two equilibrium points $x = \sqrt{u_0}$ being stable and $x = -\sqrt{u_0}$ being unstable. These merged at a turning-point bifurcation when $u_0 = 0$ and there were no equilibrium points for $u_0 < 0$. The modification of (8.9) with u_0 replaced by a variable $u(t)$ leads in even the simplest cases, $u(t) = u_0 t$ say, to very complicated problems. So we shall concentrate on systems with constant control. Examples 1.6.1–1.6.3 can all now be interpreted as examples of this situation where we examined the structure of the equilibrium points for different ranges of value of the control. In most cases a detailed solution of the problem, to give an explicit form of $x(u_0, t)$, would have been difficult. The only case for which we gave a full solution was (8.9) with $u_0 = 0$. In Examples 1.6.6 and 1.6.7 we consider two-dimensional systems and linearized about equilibrium points to determine their stability from the eigenvalues of the stability matrix according to Thm. 1.6.1. In Example 1.6.7 we also found a *limit cycle* or periodic orbit, which is another possible equilibrium solution to a system of non-linear equations. We now concentrate on an investigation of the equilibrium solutions of the system

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{X}(u_0, \boldsymbol{x}(t)). \tag{8.10}$$

Such an equilibrium solution $\boldsymbol{x} = \tilde{\boldsymbol{x}}(u_0, t)$ will be a solution of

$$\boldsymbol{X}(u_0, \boldsymbol{x}(t)) = 0. \tag{8.11}$$

8.3.1 The Stability of Trajectories

In this section we consider the general stability properties of a solution $\boldsymbol{x}(t)$ of (8.10). With $\boldsymbol{x}(t_{\mathrm{I}}) = \boldsymbol{x}_{\mathrm{I}}$ specifying the solution at time t_{I}, $\boldsymbol{x}(t)$ defines a *trajectory*[1] in the space $\mathit{\Gamma}_n$ of the n variables $x_1, x_2, \ldots, x_n$.

The map ϕ_t: $\mathit{\Gamma}_n \to \mathit{\Gamma}_n$ for all $t \geq 0$ is defined by

$$\phi_t[\boldsymbol{x}(t_{\mathrm{I}})] = \boldsymbol{x}(t_{\mathrm{I}} + t) \tag{8.12}$$

and the set of maps $\{\phi_t : t \geq 0\}$ is called a *flow*. Since

$$\phi_{t_1}[\phi_{t_2}[\boldsymbol{x}(t_{\mathrm{I}})]] = \boldsymbol{x}(t_{\mathrm{I}} + t_1 + t_2) \qquad t_1, t_2 \geq 0 \tag{8.13}$$

the flow satisfies the conditions

$$\phi_{t_1}\phi_{t_2} = \phi_{t_1+t_2} = \phi_{t_2}\phi_{t_1}. \tag{8.14}$$

[1] Also called the *path* or *orbit*.

It thus has all the properties of an Abelian (commutative) group apart from the possible non-existence of an inverse; it is therefore an *Abelian semigroup*.

The important question concerning a solution $\boldsymbol{x}(t)$ of (8.10) is whether it is *stable*. There are many different definitions of stability in the literature. As we did for equilibrium points in Sect. 1.6 we shall use the one due to Lyapunov:

The solution $\boldsymbol{x}(t)$ to (8.10), with $\boldsymbol{x}(t_I) = \boldsymbol{x}_I$, is said to be ***uniformly stable*** *or* ***stable in the sense of Lyapunov*** *if there exists, for every $\varepsilon > 0$, a $\delta(\varepsilon) > 0$, such that any other solution $\tilde{\boldsymbol{x}}(t)$, for which $\tilde{\boldsymbol{x}}(t_I) = \tilde{\boldsymbol{x}}_I$ and*

$$|\boldsymbol{x}_I - \tilde{\boldsymbol{x}}_I| < \delta(\varepsilon), \tag{8.15}$$

satisfies

$$|\boldsymbol{x}(t) - \tilde{\boldsymbol{x}}(t)| < \varepsilon, \tag{8.16}$$

for all $t \geq t_I$. If no such $\delta(\varepsilon)$ exists then $\boldsymbol{x}(t)$ is said to be ***unstable in the sense of Lyapunov****. If $\boldsymbol{x}(t)$ is uniformly stable and*

$$\lim_{t\to\infty} |\boldsymbol{x}(t) - \tilde{\boldsymbol{x}}(t)| = 0. \tag{8.17}$$

it is said to be ***asymptotically stable in the sense of Lyapunov****.*

Lyapunov stability could be characterized by saying that, for stability, the two solutions are forced to lie in a 'tube' of thickness ε, for $t > t_I$, by the initial condition (8.15). The following definitions are also useful:

The solution $\boldsymbol{x}(t)$ to (8.10), with $\boldsymbol{x}(t_I) = \boldsymbol{x}_I$, is a ***periodic solution*** *of period T if, $\boldsymbol{x}(t+T) = \boldsymbol{x}(t)$, for all $t > t_I$, and there does not exist a $T' < T$ with $\boldsymbol{x}(t+T') = \boldsymbol{x}(t)$, for all $t > t_I$.*

A ***cluster (or limit) point*** *$\boldsymbol{x}_\infty$ of the solution $\boldsymbol{x}(t)$ to (8.10), with $\boldsymbol{x}(t_I) = \boldsymbol{x}_I$, is such that, for all $\tau > 0$ and $\varepsilon > 0$, there exists a $t_1(\varepsilon) > \tau$ with*

$$|\boldsymbol{x}_\infty - \boldsymbol{x}(t_1)| < \varepsilon. \tag{8.18}$$

The set of cluster points is called the ***$\boldsymbol{\omega}$-limit set*** *of the trajectory.*

Given that the solution $\boldsymbol{x}(t)$ to (8.10) is defined for all (positive and negative) t and $\boldsymbol{x}(0) = \boldsymbol{x}_I$ the ***reverse trajectory*** *$\boldsymbol{x}^R(t)$ is defined by $\boldsymbol{x}^R(t) = \boldsymbol{x}(-t)$. The set of cluster points of the reverse trajectory is called the* ***$\boldsymbol{\alpha}$-limit set*** *of the trajectory $\boldsymbol{x}(t)$.*

It is clear that the existence of a cluster point $\boldsymbol{x}_\infty$ implies the existence of a sequence $t_1 < t_2 < \cdots < t_m \to \infty$ such that, for the specified trajectory,

$$\boldsymbol{x}(t_m) \to \boldsymbol{x}_\infty, \quad \text{as } m \to \infty. \tag{8.19}$$

Let $\mathfrak{A}$ be the ω-limit set of a particular solution $\boldsymbol{x}(t)$ to (8.10). If there exists a region $\mathcal{D}(\mathfrak{A})$, in Γ_n, which contains $\mathfrak{A}$ and for which the trajectories with $\boldsymbol{x}(0) = \boldsymbol{x}_I$, for all $\boldsymbol{x}_I$ in $\mathcal{D}(\mathfrak{A})$, have $\mathfrak{A}$ as their ω-limit set, then $\mathfrak{A}$ is called an ***attractor with basin (or domain)*** *$\mathcal{D}(\mathfrak{A})$. An α-limit with the same property for reverse trajectories is called a* ***repellor****.*

8.3.2 The Lyapunov Direct Method

An interesting method for establishing the stability of an equilibrium point is given by *Lyapunov's first theorem for stability*:

Theorem 8.3.1 *Let $\boldsymbol{x}^*(u_0)$ be an equilibrium point of (8.10). Suppose that there exists a continuous differentiable function $\mathcal{L}(\boldsymbol{x})$ such that*

$$\mathcal{L}(\boldsymbol{x}^*) = 0 \tag{8.20}$$

and for some $\mu > 0$

$$\mathcal{L}(\boldsymbol{x}) > 0, \quad \textit{when } 0 < |\boldsymbol{x}^* - \boldsymbol{x}| < \mu. \tag{8.21}$$

Then $\boldsymbol{x}^*$ is

(i) stable if

$$\boldsymbol{X}(u_0, \boldsymbol{x}).\boldsymbol{\nabla}\mathcal{L}(\boldsymbol{x}) \le 0, \quad \textit{when } |\boldsymbol{x}^* - \boldsymbol{x}| < \mu, \tag{8.22}$$

(ii) asymptotically stable if

$$\boldsymbol{X}(u_0, \boldsymbol{x}).\boldsymbol{\nabla}\mathcal{L}(\boldsymbol{x}) < 0, \quad \textit{when } |\boldsymbol{x}^* - \boldsymbol{x}| < \mu, \tag{8.23}$$

(iii) unstable if

$$\boldsymbol{X}(u_0, \boldsymbol{x}).\boldsymbol{\nabla}\mathcal{L}(\boldsymbol{x}) > 0, \quad \textit{when } |\boldsymbol{x}^* - \boldsymbol{x}| < \mu. \tag{8.24}$$

Proof: From (8.10) along a trajectory

$$\frac{\mathrm{d}\mathcal{L}(\boldsymbol{x})}{\mathrm{d}t} = \boldsymbol{\nabla}\mathcal{L}(\boldsymbol{x}).\dot{\boldsymbol{x}}(t) = \boldsymbol{X}(u_0, \boldsymbol{x}).\boldsymbol{\nabla}\mathcal{L}(\boldsymbol{x}). \tag{8.25}$$

From (8.20) and (8.21), $\boldsymbol{x}^*$ is a local minimum of $\mathcal{L}(\boldsymbol{x})$. So we can find an $R > 0$, with $\mu \ge R$, such that, for all $R > |\boldsymbol{x}^* - \boldsymbol{x}_1| > |\boldsymbol{x}^* - \boldsymbol{x}_2|$, $\mathcal{L}(\boldsymbol{x}_1) > \mathcal{L}(\boldsymbol{x}_2)$. Then if (8.22) applies, it follows from (8.25) that a trajectory cannot move further from $\boldsymbol{x}^*$ and, given any $\varepsilon > 0$, (1.114) can be satisfied by choosing $\delta(\varepsilon)$ in (1.113) to be the smaller of ε and R. If the strict inequality (8.23) applies it follows from (8.25) that the trajectory must converge to $\boldsymbol{x}^*$. The condition for $\boldsymbol{x}^*$ to be unstable is established in a similar way.

A function $\mathcal{L}(\boldsymbol{x})$ which satisfies (8.22) is called a *Lyapunov function* and which satisfies (8.23) a *strict Lyapunov function*. The method of establishing stability of an equilibrium point by finding a Lyapunov function is called the *Lyapunov direct method.* If

$$\boldsymbol{X}(u_0,\boldsymbol{x}) = -\boldsymbol{\nabla}U(u_0,\boldsymbol{x}), \tag{8.26}$$

(cf. (1.108)) and $U(u_0,\boldsymbol{x})$ has a local minimum at $\boldsymbol{x}^*$, for some fixed u_0^*. Then the choice

$$\mathcal{L}(\boldsymbol{x}) = U(u_0^*,\boldsymbol{x}) - U(u_0^*,\boldsymbol{x}^*), \tag{8.27}$$

satisfies (8.20) and (8.21), with

$$\boldsymbol{X}(u_0^*,\boldsymbol{x}).\boldsymbol{\nabla}\mathcal{L}(\boldsymbol{x}) = -|\boldsymbol{\nabla}U(u_0^*,\boldsymbol{x})|^2 < 0. \tag{8.28}$$

So a local minimum of $U(u_0,\boldsymbol{x})$ is, as we might expect, an asymptotically stable equilibrium point. To establish that a local maximum is an unstable equilibrium point simply make the choice

$$\mathcal{L}(\boldsymbol{x}) = U(u_0^*,\boldsymbol{x}^*) - U(u_0^*,\boldsymbol{x}). \tag{8.29}$$

Example 8.3.1 *Show that* $(0,0)$ *is a stable equilibrium point of*

$$\dot{x}(t) = -2x(t) - y^2(t), \qquad \dot{y}(t) = -x^2(t) - y(t). \tag{8.30}$$

Try

$$\mathcal{L}(x,y) = \alpha x^2 + \beta y^2. \tag{8.31}$$

For α and β positive (8.20) and (8.21) are satisfied and

$$\begin{aligned}\boldsymbol{X}(x,y).\boldsymbol{\nabla}\mathcal{L}(x,y) &= -\{2\alpha x(2x+y^2) + 2\beta y(y+x^2)\}\\ &= -2x^2(2\alpha+\beta y) - 2y^2(\beta+2\alpha x).\end{aligned} \tag{8.32}$$

So in the neighbourhood $|x| < \beta/\alpha$, $|y| < 2\alpha/\beta$ of the origin (8.22) is satisfied and the equilibrium point is stable.

The problem in this method is to find a suitable Lyapunov function. This in general can be quite difficult. There are, however, two cases where the choice is straightforward:

A conservative system given by

$$\ddot{\boldsymbol{x}}(t) = -\boldsymbol{\nabla}V(u_0,\boldsymbol{x}), \tag{8.33}$$

(cf. (1.102) and (1.104)), which in terms of the $2n$ variables $x_1,\ldots,x_n,v_1,\ldots,v_n$ can be expressed in the form

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{v}, \qquad \dot{\boldsymbol{v}}(t) = -\boldsymbol{\nabla}V. \tag{8.34}$$

An equilibrium point with $u_0 = u_0^*$ is given by $\boldsymbol{v} = \boldsymbol{0}$ and a value $\boldsymbol{x} = \boldsymbol{x}^*$ which satisfies $\boldsymbol{\nabla} V = \boldsymbol{0}$. Now try

$$\mathcal{L}(\boldsymbol{x}, \boldsymbol{v}) = \frac{1}{2}\boldsymbol{v}.\boldsymbol{v} + V(u_0^*, \boldsymbol{x}) - V(u_0^*, \boldsymbol{x}^*). \tag{8.35}$$

With

$$\boldsymbol{\nabla}\mathcal{L}(\boldsymbol{x}) = \begin{pmatrix} \boldsymbol{\nabla} V \\ \boldsymbol{v} \end{pmatrix} \tag{8.36}$$

$$\boldsymbol{X}(u_0^*, \boldsymbol{x}).\boldsymbol{\nabla}\mathcal{L}(\boldsymbol{x}) = 0. \tag{8.37}$$

Since, from (8.35) $\mathcal{L}(\boldsymbol{x}^*, 0) = 0$ it follows from (8.37) that the equilibrium point is stable (but not asymptotically stable) if (8.21) holds. From (8.35) this will certainly be the case if $\boldsymbol{x}^*$ is a local minimum of $V(u_0^*, \boldsymbol{x})$. It can be shown that such a minimum of the potential is a centre, which is stable in the sense of Lyapunov.

A Hamiltonian system given by (1.100), in terms of the $2n$ variables $x_1, \ldots, x_m, p_1, \ldots, p_m$. If the system is autonomous and we have an equilibrium point $(\boldsymbol{x}^*, \boldsymbol{p}^*)$ then, with

$$\mathcal{L}(\boldsymbol{x}, \boldsymbol{p}) = H(\boldsymbol{x}, \boldsymbol{p}) - H(\boldsymbol{x}^*, \boldsymbol{p}^*) \tag{8.38}$$

we have, from (1.101)

$$\frac{\mathrm{d}\mathcal{L}}{\mathrm{d}t} = \frac{\mathrm{d}H}{\mathrm{d}t} = \boldsymbol{X}(\boldsymbol{x}, \boldsymbol{p}).\boldsymbol{\nabla}\mathcal{L}(\boldsymbol{x}, \boldsymbol{p}) = 0. \tag{8.39}$$

The equilibrium point is stable if it is a local minimum of the Hamiltonian. An example where this is true is the equilibrium point at the origin for the simple harmonic oscillator with Hamiltonian

$$H(x, p) = \frac{1}{2m}p^2 + \frac{1}{2}\omega^2 x^2. \tag{8.40}$$

Even when the equilibrium point is not a local minimum of the Hamiltonian, its form can often be a guide to finding an appropriate Lyapunov function.

Example 8.3.2 *Consider the stability of the equilibrium point at the origin for the system with Hamiltonian*

$$H(u_0, x_1, x_2, p_1, p_2) = \frac{1}{2}\{x_1^2 + x_2^2 + p_1^2 + p_2^2\} + u_0\{p_1 x_2 - p_2 x_1\}. \tag{8.41}$$

From (1.100) the equations of motion for this system are

$$\begin{aligned} \dot{x}_1(t) &= \frac{\partial H}{\partial p_1} = p_1 + u_0 x_2, & \dot{p}_1(t) &= -\frac{\partial H}{\partial x_1} = -x_1 + u_0 p_2, \\ \dot{x}_2(t) &= \frac{\partial H}{\partial p_2} = p_2 - u_0 x_1, & \dot{p}_2(t) &= -\frac{\partial H}{\partial x_2} = -x_2 - u_0 p_1. \end{aligned} \tag{8.42}$$

The origin is clearly an equilibrium point. However, in the plane $x_2 = p_1 = 0$,

$$\begin{vmatrix} \dfrac{\partial^2 H}{\partial x_1^2} & \dfrac{\partial^2 H}{\partial x_1 \partial p_2} \\ \dfrac{\partial^2 H}{\partial p_2 \partial x_1} & \dfrac{\partial^2 H}{\partial p_2^2} \end{vmatrix} = 1 - {u_0}^2. \tag{8.43}$$

So the origin is a saddle point in this plane when $|u_0| > 1$. However, the function

$$\mathcal{L}(x_1, x_2, p_1, p_2) = H(0, x_1, x_2, p_1, p_2) \tag{8.44}$$

has a minimum at the origin with

$$\boldsymbol{X}(u_0, x_1, x_2, p_1, p_2).\boldsymbol{\nabla}\mathcal{L}(x_1, x_2, p_1, p_2) = 0. \tag{8.45}$$

So we have found a Lyapunov function which establishes the stability of the equilibrium point.

8.4 The Stability of Periodic Solutions

In Example 1.6.7 we investigated the Hopf bifurcation at which a stable limit cycle emerged from a stable equilibrium point. It is clear that a limit cycle is a type of periodic orbit but we have yet to give a more formal definition. This can be done using the definitions of stability of trajectories given in Sect. 8.3.1.

The periodic solution $\boldsymbol{x}(t)$ to (8.10) is a ***stable limit cycle*** *if it is asymptotically stable and an* ***unstable limit cycle*** *if it is unstable.*

We now consider the case of periodic solutions for a two-dimensional system

$$\dot{x}_1(t) = X_1(u_0, x_1, x_2), \qquad \dot{x}_2(t) = X_2(u_0, x_1, x_2). \tag{8.46}$$

which we suppose to have a unique solution at all points in $\{x_1, x_2\}$ which are not equilibrium points $[X_1(u_0, x_1, x_2) = X_2(x_1, x_2) = 0]$. We state two important results for such systems. The second of these, which is the *Poincaré-Bendixson theorem* will be shown to be a consequence of the first result, which is stated without proof.

Theorem 8.4.1 *If a trajectory of (8.46) has a bounded ω-set, then that set is either an equilibrium point or a periodic trajectory.*

Theorem 8.4.2 *Let $\mathcal{C}$ be a closed, bounded (i.e. compact) subset of the x_1–x_2 plane. If there exists a solution $\gamma = (x_1(t), x_2(t))$ of (8.46), which is contained in $\mathcal{C}$ for all $t \geq 0$, then it tends either to an equilibrium point or to a periodic solution as $t \to \infty$.*

Proof: Consider the infinite sequence $(x_1(t_0+n\varepsilon), x_2(t_0+n\varepsilon))$ of points of γ, with $t_0>0$, $\varepsilon>0$, $n=0,1,2,\ldots$. All these points lie in the compact set $\mathcal{C}$ so it follows from the *Bolzano-Weierstrass theorem* that the sequence has at least one limit point. This point must belong to the ω-limit set of γ, which is thus non-empty. From Thm. 8.4.1 this ω-limit set is an equilibrium point or a periodic solution to which γ tends.

It follows from the Poincaré-Bendixson theorem that the existence of a trajectory γ of the type described in the theorem guarantees the existence of either a periodic trajectory or an equilibrium point in $\mathcal{C}$. It is clear that a periodic solution which is the ω-set of γ cannot be an unstable limit cycle, but it also need not be a stable limit cycle.

Example 8.4.1

$$\begin{aligned}\dot{x}_1(t) &= x_1 - x_2 - x_1(x_1^2+2x_2^2),\\ \dot{x}_2(t) &= x_1 + x_2 - x_2(x_1^2+x_2^2).\end{aligned} \tag{8.47}$$

In polar coordinates (8.47) take the form

$$\dot{r}(t) = r - r^3\left\{1+\frac{1}{4}\sin^2(2\theta)\right\}, \tag{8.48}$$

$$\dot{\theta}(t) = 1 + r^2\sin^2(\theta)\cos(\theta). \tag{8.49}$$

From (8.48)

$$r - \frac{5}{4}r^3 \le \frac{\mathrm{d}r}{\mathrm{d}t} \le r - r^3, \quad \text{for all } \theta, \tag{8.50}$$

and thus

$$\begin{aligned}&\dot{r}(t) < 0, \qquad \text{for all } \theta, \text{ if } r > r_1 = 1,\\ &\dot{r}(t) > 0, \qquad \text{for all } \theta, \text{ if } r < r_2 = 2/\sqrt{5} = 0.8944.\end{aligned} \tag{8.51}$$

So any trajectory with $(x_1(0), x_2(0))$ in the annulus $\mathcal{C} = \{(x_1,x_2) : r_2 \le \sqrt{x_1^2+x_2^2} \le r_1\}$ remains in this region for all $t>0$. The minimum value of $1+r^2\sin^2(\theta)\cos(\theta)$ as θ varies at constant r is $1-2r^2/(3\sqrt{3})$ and thus

$$\dot{\theta}(t) > 1 - \frac{2r_2^2}{3\sqrt{3}} = 1 - \frac{8}{15\sqrt{3}} \simeq 0.69208. \tag{8.52}$$

So $\dot{\theta}(t)$ is never zero and there are no equilibrium points in $\mathcal{C}$. Thus, from the Poincaré-Bendixson theorem there is at least one periodic orbit.

Problems 8

1) Systems are given by

(i) $\dot{x}(t) = -x - 2y^2, \qquad \dot{y}(t) = xy - y^3,$

(ii) $\dot{x}(t) = y - x^3, \qquad \dot{y}(t) = -x^3.$

Using a trial form of $\mathcal{L}(x, y) = x^n + \alpha y^m$ for the Lyapunov function show that, in each case the equilibrium point $x = y = 0$ is asymptotically stable.

2) A system is given by

$$\dot{x}(t) = x^2y - xy^2 + x^3, \qquad \dot{y}(t) = y^3 - x^3$$

Show that $x = y = 0$ is the only equilibrium point and, using a trial form of $\mathcal{L}(x, y) = x^2 + \alpha xy + \beta y^2$ for the Lyapunov function, show that it is unstable.

3) Consider the system

$$\dot{x}(t) = F(x, y), \qquad \dot{y}(t) = G(x, y).$$

Let $\mathcal{C}$ be a closed bounded subset of the $\{x, y\}$ plane. Show that if there exists a solution $\gamma = (x(t), y(t))$ to these equations which is contained in $\mathcal{C}$ for all $t \le 0$ then $\mathcal{C}$ contains either an equilibrium point or a periodic solution of the system. For the particular case

$$F(x, y) = -x - y + x(x^2 + 2y^2), \qquad G(x, y) = x - y + y(x^2 + 2y^2)$$

show that the origin is the only equilibrium point and determine its type. Express the equations in polar form and, by considering the directions in which trajectories cross suitable closed curves, show that the system has at least one periodic solution. *As an optional extra solve the equations and determine the equation of the periodic solution. Try plotting it in* MAPLE .

Chapter 9

Solutions

9.1 Problems 1

1) (a) This equation is separable and can be rearranged to give

$$\int \frac{\mathrm{d}x}{2x} = \int \frac{\mathrm{d}t}{t} + \text{constant}.$$

This gives

$$\frac{1}{2} \ln |x| = \ln |t| + \text{constant},$$

and hence

$$x = \mathsf{A} t^2,$$

for any constant A.

(b) This equation is homogeneous so write $y = x/t$ with

$$\frac{\mathrm{d}x}{\mathrm{d}t} = y + t\frac{\mathrm{d}y}{\mathrm{d}t}.$$

This is now separable and can be rearranged to give

$$\int \frac{\cos(y)}{\sin(y)} \mathrm{d}y = -\int \frac{\mathrm{d}t}{t} + \text{constant}.$$

Integrating

$$\ln |\sin(y)| = -\ln |t| + \text{constant}.$$

Solving for y and substituting $x = yt$ gives

$$x = t \arcsin(\mathsf{A}/t),$$

for any constant A.

(c) Again let $x = yt$ to give

$$2y\left\{y + t\frac{\mathrm{d}y}{\mathrm{d}t}\right\} = y^2 + 1$$

This can be rearranged to give

$$\int \frac{2y\mathrm{d}y}{1-y^2} = \int \frac{\mathrm{d}t}{t} + \text{constant.}$$

Integrating

$$\ln|1-y^2| = -\ln|t| + \text{constant.}$$

Thus

$$y = \pm\sqrt{1 - \frac{\mathsf{A}}{t}},$$

giving

$$x = \pm\sqrt{t^2 - t\mathsf{A}}.$$

(d) Rearrange the equation in the form

$$\frac{\mathrm{d}x}{\mathrm{d}t} - x\left(\frac{1+t}{t}\right) = \frac{\exp(t)}{t}.$$

The integrating factor is

$$\exp\left\{-\int(1+1/t)\mathrm{d}t\right\} = \frac{\exp(-t)}{t}.$$

So

$$\frac{\mathrm{d}}{\mathrm{d}t}\left\{\frac{x\exp(-t)}{t}\right\} = \frac{1}{t^2},$$

giving

$$x = (\mathsf{A}t - 1)\exp(t).$$

(e) Rearrange the equation in the form

$$\frac{\mathrm{d}x}{\mathrm{d}t} + \frac{xt}{t^2-1} = -\frac{t}{t^2-1}.$$

The integrating factor is

$$\exp\left\{\int\frac{t\mathrm{d}t}{t^2-1}\right\} = \exp\left\{\frac{1}{2}\ln|t^2-1|\right\} = \sqrt{|t^2-1|}.$$

Take first the case $|t| > 1$. Then

$$\frac{\mathrm{d}}{\mathrm{d}t}\left\{x\sqrt{t^2-1}\right\} = -\frac{t}{\sqrt{t^2-1}},$$

giving

$$x\sqrt{t^2-1} = \mathsf{A} - \sqrt{t^2-1}$$

and thus

$$x = \frac{\mathsf{A}}{\sqrt{t^2 - 1}} - 1.$$

If you now repeat the calculation for $|t| < 1$ you will get the solution

$$x = \frac{\mathsf{A}}{\sqrt{1 - t^2}} - 1.$$

So the solution for all $t \neq \pm 1$ is

$$x = \frac{\mathsf{A}}{\sqrt{|t^2 - 1|}} - 1.$$

In fact with some care you could retain the modulus signs throughout and do both cases together.

2) (a) The auxiliary equation for

$$\frac{\mathrm{d}^2 x}{\mathrm{d}t^2} - 5\frac{\mathrm{d}x}{\mathrm{d}t} + 6x = 2\exp(t) + 6t - 5$$

is

$$\lambda^2 - 5\lambda + 6 = 0$$

with roots $\lambda = 2, 3$. So the complementary function is

$$x_{\mathrm{c}}(t) = \mathsf{A}\exp(2t) + \mathsf{B}\exp(3t).$$

To find a particular integral we construct a trial function $\mathsf{T}(t)$ from $f(t) = 2\exp(t) + 6t - 5$. Since $\lambda = 1$ is not a root of the auxiliary equation the trial function for the exponential term $2\exp(t)$ is $\mathsf{C}\exp(t)$. Since $\lambda = 0$ is not a root of the auxiliary equation the trial function for $6t$ is $\mathsf{E}t + \mathsf{G}$ and the constant just adds a constant. So the total trial function is

$$\mathsf{T}(t) = \mathsf{C}\exp(t) + \mathsf{E}t + \mathsf{G}.$$

Now substitute into the equation

$$\{\mathsf{C}\exp(t)\} - 5\{\mathsf{C}\exp(t) + \mathsf{E}\} + 6\{\mathsf{C}\exp(t) + \mathsf{E}t + \mathsf{G}\} = 2\exp(t) + 6t - 5.$$

Equating coefficients on the left and right

$$2\mathsf{C} = 2, \qquad 6\mathsf{E} = 6, \qquad \mathsf{G} = 0.$$

So the particular integral is

$$x_{\mathrm{p}}(t) = \exp(t) + t$$

and the general solution is

$$x(t) = \mathsf{A}\exp(2t) + \mathsf{B}\exp(3t) + \exp(t) + t.$$

(b) The auxiliary equation for

$$\frac{\mathrm{d}^2x}{\mathrm{d}t^2} + x = 2\sin(t)$$

is

$$\lambda^2 + 1 = 0$$

with roots $\lambda = \pm\mathrm{i}$. So the complementary function is

$$x_{\mathrm{c}}(t) = \mathsf{A}\cos(t) + \mathsf{B}\sin(t).$$

Now $\lambda^2 + 1$ is a factor of the auxiliary equation of multiplicity one; (in fact it is the whole function). So the trial function is

$$\mathsf{T}(t) = t[\mathsf{C}\cos(t) + \mathsf{E}\sin(t)].$$

Substituting into the equation

$$-t[\mathsf{C}\cos(t) + \mathsf{E}\sin(t)] - 2[\mathsf{C}\sin(t) - \mathsf{E}\cos(t)] + t[\mathsf{C}\cos(t) + \mathsf{E}\sin(t)] = 2\sin(t).$$

Equating coefficients $\mathsf{C} = -1$ and $\mathsf{E} = 0$. So the general solution is

$$x(t) = -t\cos(t) + \mathsf{A}\cos(t) + \mathsf{B}\sin(t).$$

(c) For

$$\frac{\mathrm{d}^3x}{\mathrm{d}t^3} + 2\frac{\mathrm{d}^2x}{\mathrm{d}t^2} + 6\frac{\mathrm{d}x}{\mathrm{d}t} = 1 + 2\exp(-t),$$

the auxiliary equation is

$$\lambda^3 + 2\lambda^2 + 6\lambda = 0.$$

This equation has the real root $\lambda = 0$ and the complex pair

$$\lambda = -1 \pm \mathrm{i}\sqrt{5}.$$

So the complementary function is

$$x_{\mathrm{c}}(t) = \mathsf{A} + \exp(-t)[\mathsf{B}\cos(\sqrt{5}t) + \mathsf{C}\sin(\sqrt{5}t)].$$

Since $\lambda = -1$ is not a root of the auxiliary equation, the trial function for $2\exp(-t)$ is $\mathsf{E}\exp(-t)$. However, $\lambda = 0$ is a root of multiplicity one, so the trial function for 1 is $\mathsf{G}t$. So

$$\mathsf{T}(t) = \mathsf{G}t + \mathsf{E}\exp(-t).$$

Substituting in the equation

$$-\mathsf{E}\exp(-t) + 2\mathsf{E}\exp(-t) - 6\mathsf{E}\exp(-t) + 6\mathsf{G} = 1 + 2\exp(-t).$$

So $\mathsf{G} = 1/6$ and $\mathsf{E} = -2/5$ and the general solution is

$$x(t) = \frac{1}{6}t - \frac{2}{5}\exp(-t) + \mathsf{A} + \exp(-t)[\mathsf{B}\cos(\sqrt{5}t) + \mathsf{C}\sin(\sqrt{5}t)].$$

3) The auxiliary equation of

$$\frac{\mathrm{d}^2x}{\mathrm{d}t^2} - 3\frac{\mathrm{d}x}{\mathrm{d}t} + 4x = 0$$

is

$$\lambda^2 - 3\lambda + 4 = 0$$

with roots $\lambda = \frac{3}{2} \pm \mathrm{i}\frac{\sqrt{7}}{2}$. So

$$x_{\mathrm{c}}(t) = \exp\left(\frac{3}{2}t\right)\left[\mathsf{A}\cos\left(\frac{\sqrt{7}}{2}t\right) + \mathsf{B}\sin\left(\frac{\sqrt{7}}{2}t\right)\right].$$

This is the complementary function for

$$\frac{\mathrm{d}^2x}{\mathrm{d}t^2} - 3\frac{\mathrm{d}x}{\mathrm{d}t} + 4x = t^2\exp(t).$$

Since $\lambda = 1$ is not a root of the auxiliary equation the trial function is

$$\mathsf{T}(t) = \exp(t)[\mathsf{C}t^2 + \mathsf{G}t + \mathsf{H}].$$

Now

$$\frac{\mathrm{d}\mathsf{T}(t)}{\mathrm{d}t} = \exp(t)[\mathsf{C}t^2 + (2\mathsf{C} + \mathsf{G})t + (\mathsf{G} + \mathsf{H})],$$
$$\frac{\mathrm{d}^2\mathsf{T}(t)}{\mathrm{d}t^2} = \exp(t)[\mathsf{C}t^2 + (4\mathsf{C} + \mathsf{G})t + (2\mathsf{C} + 2\mathsf{G} + \mathsf{H})].$$

and substituting into the equation and equating coefficients $2\mathsf{C} = 1$, $-2\mathsf{C} + 2\mathsf{G} = 0$, $2\mathsf{C} - \mathsf{G} + 2\mathsf{H} = 0$ giving the general solution

$$x(t) = \frac{1}{4}\exp(t)[2t^2 + 2t - 1] + \exp\left(\frac{3}{2}t\right)\left[\mathsf{A}\cos\left(\frac{\sqrt{7}}{2}t\right) + \mathsf{B}\sin\left(\frac{\sqrt{7}}{2}t\right).\right]$$

Now applying the conditions $x = 0$ and $\mathrm{d}x/\mathrm{d}t = 1$ at $t = 0$

$$0 = -\tfrac{1}{4} + \mathsf{A},$$
$$1 = \tfrac{1}{4} + \tfrac{3}{2}\mathsf{A} + \tfrac{\sqrt{7}}{2}\mathsf{B}.$$

This gives $\mathsf{A} = \frac{1}{4}$, $\mathsf{B} = 3\sqrt{7}/28$ and

$$x(t) = \frac{1}{4}\exp(t)[2t^2 + 2t - 1] + \frac{1}{4}\exp\left(\frac{3}{2}t\right)\left[\cos\left(\frac{\sqrt{7}}{2}t\right) + \frac{3}{\sqrt{7}}\sin\left(\frac{\sqrt{7}}{2}t\right).\right]$$

4) (i) The equilibrium points are given by $x = 0$ and $x = x^* = (a-c)/ab$. Linearizing about $x = 0$

$$\frac{\mathrm{d}\triangle x}{\mathrm{d}t} = (a-c)\triangle x,$$

with solution

$$\triangle c = \mathsf{C}\exp[(a-c)t].$$

So this solution is stable if $a < c$ and unstable if $a > c$. Linearize about $x = x^*$

$$\frac{\mathrm{d}\triangle x}{\mathrm{d}t} = (c-a)\triangle x,$$

with solution

$$\triangle c = \mathsf{C}\exp[(c-a)t].$$

So this solution is stable if $a > c$ and unstable if $a < c$. There are five different cases:

When $c = 0$, $x^* = 1/b$ and the lines of equilibrium points are parallel to the a-axis. There is no bifurcation but the stability changes at $a = 0$. (Fig. 9.1.)

When $b > 0$ and $c > 0$, there is a *transcritical* bifurcation at $x = 0$, $a = c$ on one branch of $x = x^*(a)$. The second branch is unstable. (Fig. 9.2.) The case $b < 0$, $c > 0$ is the mirror image of this in the vertical axis.

When $b < 0$ and $c < 0$, there is a *transcritical* bifurcation at $x = 0$, $a = c$ on one branch of $x = x^*(a)$. The second branch is stable. (Fig. 9.3.) The case $b > 0$, $c < 0$ is the mirror image of this in the vertical axis. The equation is separable so

$$\int \frac{\mathrm{d}x}{x(a-c-abx)} = t + \text{constant}.$$

Using partial fractions it is easy to do the integration and the final solution is

$$x(t) = \frac{\mathsf{C}(a-c)\exp[(a-c)t]}{1 + ab\mathsf{C}\exp[(a-c)t]},$$

for some constant C. If $a < c$, $x \to 0$ as $t \to \infty$ and, if $a > c$, $x \to (a-c)/ab$ as $t \to \infty$.

(ii) The equilibrium solutions are $x = 0$ and

$$x = x^* = \begin{cases} a/b, & \text{if } c = 0, \\[2ex] \dfrac{b \pm \sqrt{b^2 - 4ac}}{2c}, & \text{if } c \neq 0, \end{cases}$$

$c = 0$

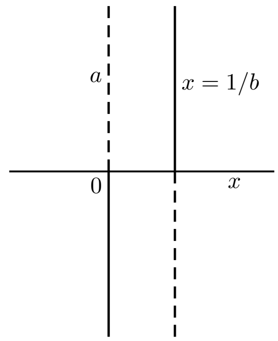


Figure 9.1: Bifurcation diagram for question 4(i) with $c = 0$.

$c > 0, \quad b > 0$

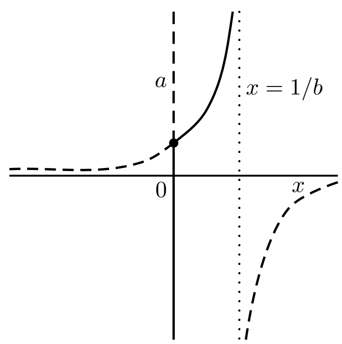


Figure 9.2: Bifurcation diagram for question 4(i) with $c > 0$, $b > 0$.

Linearizing about $x = 0$

$$\frac{\mathrm{d}\triangle x}{\mathrm{d}t} = a\triangle x.$$

So this equilibrium point is stable if $a < 0$ and unstable if $a > 0$. Linearizing about $x = x^*$

$$\frac{\mathrm{d}\triangle x}{\mathrm{d}t} = x^*(2x^*c - b)\triangle x.$$

So x^* is stable if $x^*(2x^*c - b) < 0$ and unstable if $x^*(2x^*c - b) > 0$. When $c = 0$ these conditions reduce to $a > 0$ and $a < 0$ respectively.

When $c = 0$ and $b > 0$, there is a transcritical bifurcation at the origin. (Fig. 9.4.) For $c = 0$ and $b < 0$ the bifurcation diagram is obtained from this by reflection in the vertical axis.

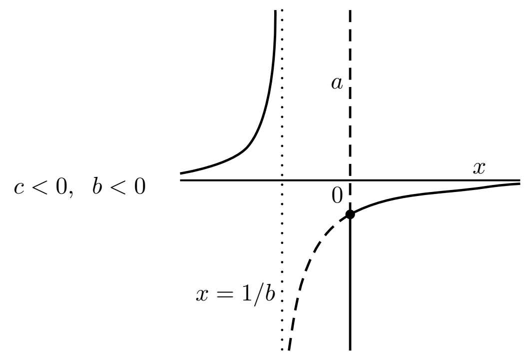


Figure 9.3: Bifurcation diagram for question 4(i) with $c < 0$, $b < 0$.

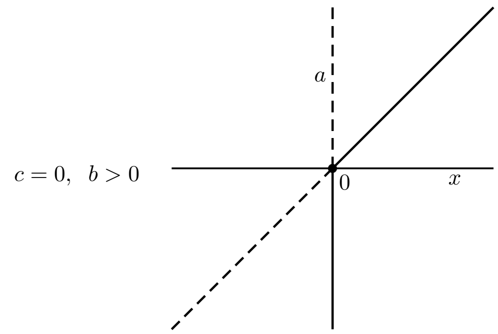


Figure 9.4: Bifurcation diagram for question 4(ii) with $c = 0$, $b > 0$.

When $c > 0$ and $b > 0$, there is a transcritical bifurcation at the origin and a turning-point bifurcation at $x = b/2c$, $a = b^2/4c$. (Fig. 9.5.) The case $c > 0$, $b < 0$ is obtained from this by reflection in the vertical axis.

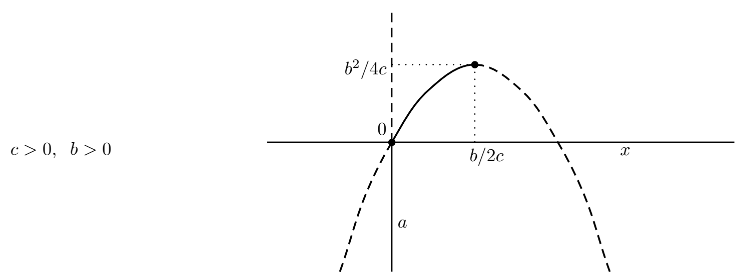


Figure 9.5: Bifurcation diagram for question 4(ii) with $c > 0$, $b > 0$.

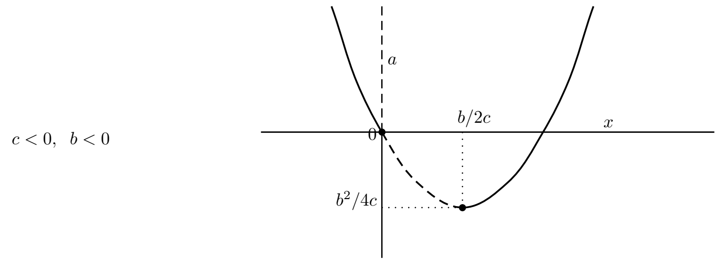


Figure 9.6: Bifurcation diagram for question 4(ii) with $c < 0$, $b < 0$.

When $c < 0$ and $b < 0$, there are again a transcritical and a turning-point bifurcation at the same locations. (Fig. 9.6.) The case $c < 0$ and $b > 0$ is obtained from this by reflection in the vertical axis.

Each of these $c \neq 0$ systems of bifurcations goes into a pitchfork bifurcation when $b \to 0$. Denoting the two branches of x^* by $x^{(\pm)}$, the equation can separated into

$$\int \frac{\mathrm{d}x}{x[x - x^{(+)}][x - x^{(-)}]} = \text{constant} + at$$

Decomposing into partial fractions and integrating gives

$$x^{\alpha}[x - x^{(+)}]^{\gamma^{(+)}}[x - x^{(-)}]^{\gamma^{(-)}} = \mathrm{C}\exp(at).$$

where $\alpha = x^{(+)}x^{(-)}$, $\gamma^{(\pm)} = x^{(\pm)}[x^{(\pm)} - x^{(\mp)}]$. The limiting behaviour as $t \to \infty$ can be obtained by considering the various signs of the parameters.

5) In both parts of this problem the only equilibrium point is $x = y = 0$ and in a neighbourhood of the origin

$$\frac{\mathrm{d}\triangle\boldsymbol{x}}{\mathrm{d}t} = \boldsymbol{A}\boldsymbol{x}, \qquad (\star) \qquad \text{where} \qquad \boldsymbol{x} = \begin{pmatrix} x \\ y \end{pmatrix}$$

and

(i)

$$\boldsymbol{A} = \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix}.$$

This matrix has eigenvalues $\lambda = -2, 4$. The equilibrium point is unstable because it has a positive (real) eigenvalue, but since it also has one negative (real) eigenvalue it is a *saddle-point*.

(ii)

$$\boldsymbol{A} = \begin{pmatrix} 3 & 2 \\ 1 & 2 \end{pmatrix}.$$

This matrix has eigenvalues $\lambda = 1, 4$. The equilibrium point is unstable because it has two positive (real) eigenvalues. In this case since both eigenvalues are positive it is called an *unstable node*.

6) The right-hand sides of these two equations are both zero when $x = y = 0$. Now the Taylor expansions of $\sin(x)$ and $\cos(x)$ give

$$\sin(\triangle x) = \triangle x + \mathrm{O}(\triangle x^3), \qquad \cos(\triangle x) = 1 + \mathrm{O}(\triangle x^2).$$

So when linearized to the same form as $(\star)$ we have

$$\boldsymbol{A} = \begin{pmatrix} 1 & 1 \\ 0 & -2 \end{pmatrix}.$$

This matrix has eigenvalues $\lambda = -2, 1$. The equilibrium point is a *saddle-point*.

7) All the equilibrium points are given by the simultaneous solutions of

$$x^2 = y, \qquad 8x = y^2.$$

This gives $x^4 = 8x$, which has the solutions

$$x = 0, \quad \text{implying} \quad y = 0, \qquad (1)$$
$$x = 2, \quad \text{implying} \quad y = 4. \qquad (2)$$

For (1)

$$\boldsymbol{A} = \begin{pmatrix} 0 & 1 \\ 8 & 0 \end{pmatrix}.$$

This matrix has eigenvalues $\lambda = \pm\sqrt{8}$ giving a *saddle-point*.

For (2)

$$\boldsymbol{A} = \begin{pmatrix} -4 & 1 \\ 8 & -8 \end{pmatrix}.$$

This matrix has eigenvalues $\lambda = -6 \pm 2\sqrt{3}$. Both these eigenvalues are negative so the equilibrium point is a *stable node*.

9.2 Problems 2

1) (i) From lines 7 and 11 of the table

$$\mathcal{L}\{t\sin(\omega t)\} = -\frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{\omega}{s^2+\omega^2}\right) = \frac{2\omega s}{(\omega^2+s^2)^2},$$

(ii) From lines 6, 7 and 11 of the table

$$\mathcal{L}\{\sin(\omega t) - \omega t\cos(\omega t)\} = \frac{\omega}{s^2+\omega^2} + \omega\frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{s}{s^2+\omega^2}\right) = \frac{2\omega^3}{(\omega^2+s^2)^2}.$$

Taking the Laplace transform of the differential equation gives

$$\bar{x}(s)[s^2+\omega^2] = \frac{\omega}{\omega^2+s^2},$$

giving

$$\bar{x}(s) = \frac{\omega}{(\omega^2+s^2)^2}.$$

So

$$x(t) = \frac{1}{2\omega^2}[\sin(\omega t) - \omega t\cos(\omega t)].$$

2) Taking the Laplace transform of the differential equation gives

$$\bar{x}(s) = \frac{1}{s(1+s^3)} = \frac{1}{s(s+1)(s^2-s+1)} = \frac{1}{s(s+1)[(s-\frac{1}{2})^2+\frac{3}{4}]}.$$

Resolving into partial fractions

$$\frac{1}{s(s+1)[(s-\frac{1}{2})^2+\frac{3}{4}]} = \frac{1}{s} - \frac{1}{3(s+1)} - \frac{2s-1}{3(s^2-s+1)}$$

and

$$\begin{aligned}
\mathcal{L}^{-1}\left\{\frac{1}{s}\right\} &= 1, \\
\mathcal{L}^{-1}\left\{\frac{1}{3(s+1)}\right\} &= \frac{1}{3}\exp(-t), \\
\mathcal{L}^{-1}\left\{\frac{2s-1}{3(s^2-s+1)}\right\} &= \mathcal{L}^{-1}\left\{\frac{2\left(s-\frac{1}{2}\right)}{3\left[\left(s-\frac{1}{2}\right)^2+\frac{3}{4}\right]}\right\} \\
&= \frac{2}{3}\exp\left(\tfrac{1}{2}t\right)\cos\left(\tfrac{\sqrt{3}}{2}t\right).
\end{aligned}$$

So

$$x(t) = 1 - \frac{1}{3}\exp(-t) - \frac{2}{3}\exp\left(\tfrac{1}{2}t\right)\cos\left(\tfrac{\sqrt{3}}{2}t\right).$$

3) Using

$$\mathcal{L}^{-1}\left\{\frac{1}{(s-1)^2}\right\} = t\exp(t)$$

and the convolution integral formula

$$y(t) = -\int_0^t u(t-u)\exp(t-u)\mathrm{d}u.$$

Now just use integration by parts and you will find that

$$y(t) = -\{2 + t + \exp(t)[t-2]\}.$$

4) From the second line of the table and the value of $\Gamma(\frac{1}{2})$ given below equation (2.15)

$$\mathcal{L}\{t^{-\frac{1}{2}}\} = \frac{\Gamma\left(\frac{1}{2}\right)}{s^{\frac{1}{2}}} = \frac{\sqrt{\pi}}{s^{\frac{1}{2}}}.$$

For part two of this question there are (at least) two methods:

Method 1:

$$\begin{aligned}
\mathcal{L}\left\{\mathrm{Erf}(t^{\frac{1}{2}})\right\} &= \frac{2}{\sqrt{\pi}}\int_0^\infty \mathrm{d}t\ \exp(-st)\int_0^{t^{\frac{1}{2}}} \mathrm{d}u\ \exp(-u^2) \\
&\quad \text{Let } v = u^2. \\
&= \frac{1}{\sqrt{\pi}}\int_0^\infty \mathrm{d}t\ \exp(-st)\int_0^t \mathrm{d}v\, v^{-\frac{1}{2}}\exp(-v) \\
&\quad \text{Now go through the procedure for changing} \\
&\quad \text{the order of integration as in pages 44 and 45 of the notes.} \\
&= \lim_{\lambda\to\infty}\frac{1}{\sqrt{\pi}}\int_0^\lambda \mathrm{d}t\ \exp(-st)\int_0^t \mathrm{d}v\, v^{-\frac{1}{2}}\exp(-v) \\
&= \lim_{\lambda\to\infty}\frac{1}{\sqrt{\pi}}\int_0^\lambda \mathrm{d}v v^{-\frac{1}{2}}\exp(-v)\int_v^\lambda \mathrm{d}t\exp(-st) \\
&\quad \text{Now let } t = w + v.
\end{aligned}$$

$$= \lim_{\lambda\to\infty} \frac{1}{\sqrt{\pi}} \int_0^\lambda \mathrm{d}v v^{-\frac{1}{2}} \exp(-v) \int_0^{\lambda-v} \mathrm{d}w \exp[-s(w+v)]$$

Now take the λ limit.

The integral factors into two parts.

$$= \frac{1}{\sqrt{\pi}} \left[\int_0^\infty \mathrm{d}v v^{-\frac{1}{2}} \exp[-v(s+1)]\right] \left[\int_0^\infty \mathrm{d}w \exp(-sw)\right]$$

The second integral contributes $1/s$

and the first is the 'shifted' Laplace transform of $t^{-\frac{1}{2}}$.

$$= \frac{1}{s(s+1)^{\frac{1}{2}}}.$$

Method 2:

Using the same change of variable $v = u^2$

$$\mathcal{L}\left\{\mathrm{Erf}(t^{\frac{1}{2}})\right\} = \frac{1}{\sqrt{\pi}}\mathcal{L}\left\{\int_0^t \frac{1}{\sqrt{v}} \exp(-v)\mathrm{d}v\right\}.$$

Using the shift theorem

$$\mathcal{L}\left\{\frac{1}{\sqrt{t}} \exp(-t)\right\} = \sqrt{\frac{\pi}{1+s}}$$

and the result follows from the formula for the Laplace transform of a convolution with $y(t-u) = 1$.

5) (i)

$$\begin{aligned} \mathcal{Z}^{-1}\left\{\frac{z}{(z-1)(z-2)}\right\} &= \mathcal{Z}^{-1}\left\{z\left[\frac{1}{z-2} - \frac{1}{z-1}\right]\right\} \\ &= \mathcal{Z}^{-1}\left\{\frac{z}{z-2}\right\} - \mathcal{Z}^{-1}\left\{\frac{z}{z-1}\right\} \\ &= 2^k - 1. \end{aligned}$$

(ii)

$$\mathcal{Z}^{-1}\left\{\frac{z}{(z^2+a^2}\right\} = \mathcal{Z}^{-1}\left\{\frac{z}{2\mathrm{i}a}\left[\frac{1}{z-\mathrm{i}a} - \frac{1}{z+\mathrm{i}a}\right]\right\}$$

$$\begin{aligned}
&= \frac{1}{2\mathrm{i}a}\left[\mathcal{Z}^{-1}\left\{\frac{z}{z-\mathrm{i}a}\right\} - \mathcal{Z}^{-1}\left\{\frac{z}{z+\mathrm{i}a}\right\}\right] \\
&= \frac{1}{2\mathrm{i}a}[(\mathrm{i}a)^k - (-\mathrm{i}a)^k] \\
&= a^{k-1}\frac{1}{2\mathrm{i}}\left[\exp\left(\frac{\mathrm{i}k\pi}{2}\right) - \exp\left(-\frac{\mathrm{i}k\pi}{2}\right)\right] \\
&= a^{k-1}\sin\left(\frac{k\pi}{2}\right)
\end{aligned}$$

(iii) $$\frac{z^3+2z^2+1}{z^3} = 1 + 2z^{-1} + z^{-3}.$$

So $x(0)=1$, $x(1)=2$, $x(2)=0$, $x(3)=1$ and $x(k)=0$ for $k>3$.

6) (i) Taking the $\mathcal{Z}$ transform

$$8z^2\tilde{x}(z) - 8z^2 - 12z - 6z\tilde{x}(z) + 6z + \tilde{x}(z) = \frac{9z}{z-1}.$$

So

$$\begin{aligned}
\tilde{x}(z) &= \frac{z}{8(z-\frac{1}{2})(z-\frac{1}{4})}\left[\frac{9}{z-1} + 8z + 6\right] \\
&= \frac{2z}{z-\frac{1}{4}} - \frac{4z}{z-\frac{1}{2}} + \frac{3z}{z-1}.
\end{aligned}$$

Giving

$$x(k) = 2\left(\tfrac{1}{4}\right)^k - 4\left(\tfrac{1}{2}\right)^k + 3.$$

(ii) Taking the $\mathcal{Z}$ transform

$$z^2\tilde{x}(z) - z^2 - z\sqrt{2} + 2\tilde{x}(z) = 0.$$

So

$$\begin{aligned}
\tilde{x}(z) &= \frac{z(z+\sqrt{2})}{z^2+2} \\
&= \frac{z(1+\mathrm{i})}{2(z+\sqrt{2}\mathrm{i})} + \frac{z(1-\mathrm{i})}{2(z-\sqrt{2}\mathrm{i})}
\end{aligned}$$

Giving

$$x(k) = 2^{\frac{k}{2}}\left[\cos\left(\frac{k\pi}{2}\right) + \sin\left(\frac{k\pi}{2}\right)\right].$$

7) Taking the $\mathcal{Z}$ transform

$$z\tilde{x}(z) - x_0 z = 1.5\tilde{x}(z) - \tilde{y}(z),$$
$$z\tilde{y}(z) - y_0 z = 0.21\tilde{x}(z) + 0.5\tilde{y}(z).$$

Eliminating $\tilde{y}(z)$

$$\begin{aligned}\tilde{x}(z) &= \frac{zx_0(z-0.5) - y_0 z}{(z-1.5)(z-0.5)+0.21} \\ &= z\left[\frac{x_0(z-0.5)-y_0}{z^2-2z+0.96}\right] \\ &= \frac{z(1.75x_0-2.5y_0)}{z-1.2} - \frac{z(0.75x_0-2.5y_0)}{z-0.8}.\end{aligned}$$

Giving

$$x(k) = (1.75x_0 - 2.5y_0)(1.2)^k - (0.75x_0 - 2.5y_0)(0.8)^k.$$

In the limit of large k

$$\frac{x(k+1)}{x(k)} = 1.2,$$

which is a 20% increase or decrease according to the sign of $(1.75x_0 - 2.5y_0)$.

9.3 Problems 3

1) Since there is a unit feedback the block diagram represents the situation

$$\bar{y}(s) = G(s)[\bar{u}(s) - \bar{y}(s)],$$

where

$$G(s) = \frac{\mathsf{K}}{s(s+\mathsf{Q})}.$$

Thus

$$\bar{y}(s) = \frac{G(s)\bar{u}(s)}{G(s)+1} = \frac{\mathsf{K}\bar{u}(s)}{s(s+\mathsf{Q})+\mathsf{K}}.$$

If $u(t) = u_0$, $\bar{u}(s) = u_0/s$ and

$$\bar{y}(s) = \frac{\mathsf{K}u_0}{s(s-s^{(+)})(s-s^{(-)})},$$

where

$$s^{(\pm)} = \tfrac{1}{2}\left[-\mathsf{Q} \pm \sqrt{\mathsf{Q}^2 - 4\mathsf{K}}\right].$$

(i) when $\mathsf{K} - \frac{1}{4}\mathsf{Q}^2 = \omega^2 > 0$,

$$s^{(\pm)} = -\frac{1}{2}\mathsf{Q} \pm \mathrm{i}\omega$$

and

$$\begin{aligned}\bar{y}(s) &= \frac{\mathsf{K}u_0}{s\left[\left(s+\frac{1}{2}\mathsf{Q}\right)^2 + \omega^2\right]}\\ &= \frac{\mathsf{A}}{s} + \frac{\mathsf{B}\left(s+\frac{1}{2}\mathsf{Q}\right)+\mathsf{C}}{\left(s+\frac{1}{2}\mathsf{Q}\right)^2+\omega^2}.\end{aligned}$$

Then recombining the partial fractions $\mathsf{A} = u_0$, $\mathsf{B} = -u_0$ and $\mathsf{C} = -\frac{1}{2}\mathsf{Q}u_0$, giving

$$y(t) = u_0\left[1 - \exp\left(-\tfrac{1}{2}\mathsf{Q}t\right)\left\{\cos(\omega t) + \frac{\mathsf{Q}}{2\omega}\sin(\omega t)\right\}\right],$$

(ii) when $\frac{1}{4}\mathsf{Q}^2 - \mathsf{K} = \zeta^2 > 0$,

$$s^{(\pm)} = -\frac{1}{2}\mathsf{Q} \pm \zeta$$

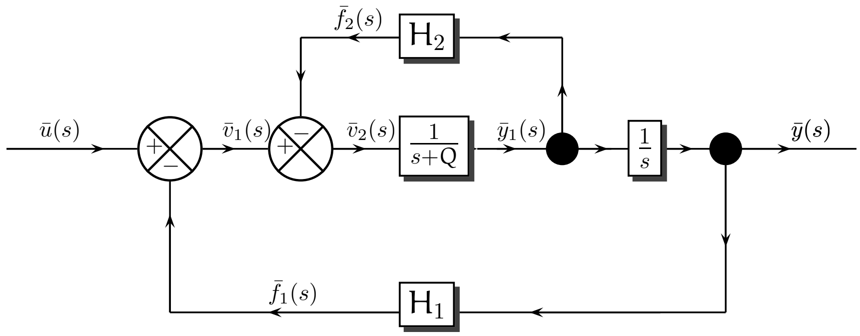
and

$$\begin{aligned}\bar{y}(s) &= \frac{\mathsf{K}u_0}{s\left[\left(s+\frac{1}{2}\mathsf{Q}\right)+\zeta\right]\left[\left(s+\frac{1}{2}\mathsf{Q}\right)-\zeta\right]}\\ &= \frac{\mathsf{A}}{s} + \frac{\mathsf{B}}{\left(s+\frac{1}{2}\mathsf{Q}\right)+\zeta} + \frac{\mathsf{C}}{\left(s+\frac{1}{2}\mathsf{Q}\right)-\zeta}.\end{aligned}$$

Then recombining the partial fractions $\mathsf{A} = u_0$, $\mathsf{B} = u_0(\frac{1}{2}\mathsf{Q}-\zeta)/(2\zeta)$ and $\mathsf{C} = -u_0(\frac{1}{2}\mathsf{Q}+\zeta)/(2\zeta)$, giving

$$y(t) = u_0\left[1 - \frac{1}{2\zeta}\exp\left(-\tfrac{1}{2}\mathsf{Q}t\right)\left\{\left[\tfrac{1}{2}\mathsf{Q}+\zeta\right]\exp(\zeta t) - \left[\tfrac{1}{2}\mathsf{Q}-\zeta\right]\exp(-\zeta t)\right\}\right].$$

2) Put the intermediate variables on the diagram as follows:



Then

$$\bar{v}_1(s) = \bar{u}(s) - \bar{f}_1(s),$$
$$\bar{v}_2(s) = \bar{v}_1(s) - \bar{f}_2(s),$$
$$\bar{y}_1(s) = \frac{1}{s+\mathsf{Q}}\bar{v}_2,$$
$$\bar{f}_2(s) = \mathsf{H}_2\bar{y}_1(s),$$
$$\bar{y}(s) = \frac{1}{s}\bar{y}_1(s),$$
$$\bar{f}_1(s) = \mathsf{H}_1\bar{y}(s).$$

Thus

$$\begin{aligned}\bar{u}(s) &= \bar{v}_1(s) + \bar{f}_1(s)\\ &= \bar{f}_2 + \bar{v}_2(s) + \bar{f}_1(s)\\ &= \mathsf{H}_2\bar{y}_1(s) + (s+\mathsf{Q})\bar{y}_1(s) + \mathsf{H}_1\bar{y}(s)\\ &= s\mathsf{H}_2\bar{y}(s) + s(s+\mathsf{Q})\bar{y}(s) + \mathsf{H}_1\bar{y}(s).\end{aligned}$$

Giving the required result.

3) (i) Taking the $\mathcal{Z}$ transform of $y(k) - 2y(k-1) = u(k-1)$ gives

$$\tilde{y}(z) - 2z^{-1}\tilde{y}(z) = z^{-1}\tilde{u}(z).$$

Thus

$$\tilde{y}(z) = \frac{\tilde{u}(z)}{z-2}.$$

If $u(k) = 1$, for all k, $\tilde{u}(z) = z/(z-1)$ So

$$\begin{aligned}\tilde{y}(z) &= \frac{z}{(z-2)(z-1)}\\ &= \frac{z}{z-2} - \frac{z}{z-1}\end{aligned}$$

So

$$y(k) = 2^k - 1.$$

(ii) Taking the $\mathcal{Z}$ transform of $y(k) + 5y(k-1) + 6y(k-2) = u(k-1) + u(k-2)$,

$$\tilde{y}(z) + 5z^{-1}\tilde{y}(z) + 6z^{-2}\tilde{y}(z) = z^{-1}\tilde{u}(z) + z^{-2}\tilde{u}(z),$$

giving

$$\tilde{y}(z) = \frac{\tilde{u}(z)(z+1)}{z^2 + 5z + 6}.$$

With the same $\tilde{u}(z)$ as in (i)

$$\begin{aligned}\tilde{y}(z) &= \frac{z(z+1)}{(z-1)(z^2+5z+6)}\\ &= z\left[\frac{\mathsf{A}}{z-1} + \frac{\mathsf{B}}{z+3} + \frac{\mathsf{C}}{z+2}\right].\end{aligned}$$

Recombining the partial fractions and equating gives $\mathsf{A} = \frac{1}{6}$, $\mathsf{B} = -\frac{1}{2}$ and $\mathsf{C} = \frac{1}{3}$. So inverting the $\mathcal{Z}$ transform

$$y(k) = \frac{1}{6} - \frac{1}{2}(-3)^k + \frac{1}{3}(-2)^k.$$

9.4 Problems 4

1) If the polynomial

$$\exp(zt) = \mathsf{B}(t) + \mathsf{C}(t)z$$

is satisfied by each of the the eigenvalues of $\boldsymbol{A}$ it is also satisfied by $\boldsymbol{A}$ itself. Thus

$$\begin{aligned}\exp(\lambda t) &= \mathsf{B}(t) + \mathsf{C}(t)\lambda,\\ \exp(\mu t) &= \mathsf{B}(t) + \mathsf{C}(t)\mu.\end{aligned}$$

So

$$\mathsf{C}(t) = \frac{\exp(\lambda t) - \exp(\mu t)}{\lambda - \mu}, \qquad \mathsf{B}(t) = \frac{\lambda\exp(\mu t) - \mu\exp(\lambda t)}{\lambda - \mu},$$

giving

$$\exp(\boldsymbol{A}t) = \frac{[\lambda \exp(\mu t) - \mu \exp(\lambda t)]\boldsymbol{I} + [\exp(\lambda t) - \exp(\mu t)]\boldsymbol{A}}{\lambda - \mu}.$$

When λ is a double root it also satisfies the derivative of the equation

$$t \exp(\lambda t) = \mathsf{C}(t).$$

Thus

$$\mathsf{C}(t) = t \exp(\lambda t), \qquad \mathsf{B}(t) = \exp(\lambda t)[1 - \lambda t],$$

giving

$$\exp(\boldsymbol{A}t) = \exp(\lambda t)[1 - \lambda t]\boldsymbol{I} + t \exp(\lambda t)\boldsymbol{A}.$$

2) The eigenvalues of the matrix

$$\boldsymbol{A} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2 & 1 & -2 \end{pmatrix}$$

are $1, -1, -2$. So if

$$\exp(zt) = \mathsf{B}(t) + \mathsf{C}(t)z + \mathsf{D}(t)z^2$$

is satisfied by each of the the eigenvalues of $\boldsymbol{A}$ it is also satisfied by $\boldsymbol{A}$ itself. Thus

$$\begin{aligned} \exp(t) &= \mathsf{B}(t) + \mathsf{C}(t) + \mathsf{D}(t), \\ \exp(-t) &= \mathsf{B}(t) - \mathsf{C}(t) + \mathsf{D}(t), \\ \exp(-2t) &= \mathsf{B}(t) - 2\mathsf{C}(t) + 4\mathsf{D}(t), \end{aligned}$$

giving

$$\begin{aligned} \mathsf{B}(t) &= \frac{1}{3}[\exp(t) + 3\exp(-t) - \exp(-2t)], \\ \mathsf{C}(t) &= \frac{1}{2}[\exp(t) - \exp(-t)], \\ \mathsf{D}(t) &= \frac{1}{6}[\exp(t) - 3\exp(-t) + 2\exp(-2t)]. \end{aligned}$$

We need to substitute

$$\exp(\boldsymbol{A}t) = \mathsf{B}(t)\boldsymbol{I} + \mathsf{C}(t)\boldsymbol{A} + \mathsf{D}(t)\boldsymbol{A}^2$$

into

$$y(t) = \int_0^t \boldsymbol{c}^{\mathrm{T}} \exp[\boldsymbol{A}(t-\tau)]\boldsymbol{b}u(\tau)\mathrm{d}\tau,$$

[see (4.134)]. Now with $u(t) = \mathsf{K}t$ and the given forms for $\boldsymbol{b}$ and $\boldsymbol{c}^{\mathrm{T}}$

$$\begin{aligned} \boldsymbol{c}^{\mathrm{T}} \exp[\boldsymbol{A}(t-\tau)]\boldsymbol{b}u(\tau) &= \mathsf{K}\tau[2\mathsf{C}(t-\tau) - 3\mathsf{D}(t-\tau)] \\ &= \frac{1}{2}\mathsf{K}\tau[\exp(t-\tau) + \exp(\tau - t) - 2\exp(2\tau - 2t)]. \end{aligned}$$

Substituting into the integral

$$y(t) = -\frac{1}{4}\mathsf{K}[3 + 2t - 2\exp(t) - 2\exp(-t) + \exp(-2t)].$$

3) $$(s\boldsymbol{I} - \boldsymbol{A}) = \begin{pmatrix} s + \frac{3}{4} & \frac{1}{4} \\ \frac{1}{2} & s + \frac{1}{2} \end{pmatrix}.$$

$$(s\boldsymbol{I} - \boldsymbol{A})^{-1} = \begin{pmatrix} 2\,\dfrac{2\,s+1}{4\,s^2+5\,s+1} & -\dfrac{1}{4\,s^2+5\,s+1} \\ -2\,\dfrac{1}{4\,s^2+5\,s+1} & \dfrac{4\,s+3}{4\,s^2+5\,s+1} \end{pmatrix}.$$

$$G(s) = \boldsymbol{c}^{\mathrm{T}}(s\boldsymbol{I} - \boldsymbol{A})^{-1}\boldsymbol{b} = \frac{6}{s+1}.$$

$$\boldsymbol{U} = \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}, \qquad \boldsymbol{V} = \begin{pmatrix} 4 & 2 \\ -4 & -2 \end{pmatrix}.$$

(Since both these matrices are singular the realization is neither controllable nor observable. In fact we knew that it couldn't be both since, from the form of $G(s)$, $n_{\min} = 1$.)

4) $$\boldsymbol{U} = \begin{pmatrix} 1 & -4 \\ 3 & -10 \end{pmatrix}, \qquad \mathrm{Det}\{\boldsymbol{U}\} = 2, \qquad \boldsymbol{U}^{-1} = \frac{1}{2}\begin{pmatrix} -10 & 4 \\ -3 & 1 \end{pmatrix}.$$

So the system is controllable.

$$\boldsymbol{T} = \begin{pmatrix} -\frac{3}{2} & \frac{1}{2} \\ \frac{5}{2} & -\frac{1}{2} \end{pmatrix}, \qquad \boldsymbol{T}^{-1} = \begin{pmatrix} 1 & 1 \\ 5 & 3 \end{pmatrix}$$

and

$$\boldsymbol{T}\boldsymbol{A}\boldsymbol{T}^{-1} = \begin{pmatrix} 0 & 1 \\ -6 & -5 \end{pmatrix}.$$

5) (i) Decomposing into partial fractions gives

$$\bar{y}(s) = \frac{\bar{u}(s)}{1+s} - \frac{\bar{u}(s)}{(1+s)^2} + \frac{\bar{u}(s)}{(1+s)^3}.$$

Then let

$$\begin{aligned} \bar{x}_1(s) &= \frac{\bar{u}(s)}{1+s}, \\ \bar{x}_2(s) &= \frac{\bar{x}_1(s)}{1+s} = \frac{\bar{u}(s)}{(1+s)^2}, \\ \bar{x}_3(s) &= \frac{\bar{x}_2(s)}{1+s} = \frac{\bar{u}(s)}{(1+s)^3}. \end{aligned}$$

Then

$$\begin{aligned} &\bar{y}(s) = \bar{x}_1(s) - \bar{x}_2(s) + \bar{x}_3(s), \\ &y(t) = x_1(t) - x_2(t) + x_3(t), \\ &\boldsymbol{c}^{\mathrm{T}} = \begin{pmatrix} 1 & -1 & 1 \end{pmatrix}. \end{aligned}$$

$$\begin{aligned} s\bar{x}_1(s) &= -\bar{x}_1(s) + \bar{u}(s), \\ s\bar{x}_2(s) &= -\bar{x}_2(s) + \bar{x}_1(s), \\ s\bar{x}_3(s) &= -\bar{x}_3(s) + \bar{x}_2(s), \end{aligned}$$

$$\begin{aligned} \dot{x}_1(t) &= -x_1(t) + u(t), \\ \dot{x}_2(t) &= -x_2(t) + x_1(t), \\ \dot{x}_3(t) &= -x_3(t) + x_2(t), \end{aligned}$$

So

$$\boldsymbol{A} = \begin{pmatrix} -1 & 0 & 0 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix}, \qquad \boldsymbol{b} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

(ii) Decomposing into partial fractions gives

$$\bar{y}(s) = -\frac{\bar{u}(s)}{1+s} + \frac{2\bar{u}(s)}{(1+s)^2} + \frac{\bar{u}(s)}{(3+s)}.$$

Then let

$$\bar{x}_1(s) = \frac{\bar{u}(s)}{1+s},$$
$$\bar{x}_2(s) = \frac{\bar{x}_1(s)}{1+s} = \frac{\bar{u}(s)}{(1+s)^2},$$
$$\bar{x}_3(s) = \frac{\bar{u}(s)}{3+s}.$$

Then

$$\bar{y}(s) = -\bar{x}_1(s) + 2\bar{x}_2(s) + \bar{x}_3(s),$$
$$y(t) = -x_1(t) + 2x_2(t) + x_3(t),$$
$$\boldsymbol{c}^{\mathrm{T}} = \begin{pmatrix} -1 & 2 & 1 \end{pmatrix}.$$
$$s\bar{x}_1(s) = -\bar{x}_1(s) + \bar{u}(s),$$
$$s\bar{x}_2(s) = -\bar{x}_2(s) + \bar{x}_1(s),$$
$$s\bar{x}_3(s) = -3\bar{x}_3(s) + \bar{u}(s),$$
$$\dot{x}_1(t) = -x_1(t) + u(t),$$
$$\dot{x}_2(t) = -x_2(t) + x_1(t),$$
$$\dot{x}_3(t) = -3x_3(t) + u(t),$$

So

$$\boldsymbol{A} = \begin{pmatrix} -1 & 0 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & -3 \end{pmatrix}, \qquad \boldsymbol{b} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$$

9.5 Problems 5

1) In matrix forms the equations become

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{b}u(t),$$
$$y(t) = \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}(t),$$

where

$$\boldsymbol{A} = \begin{pmatrix} -2 & 4 \\ 4 & -4 \end{pmatrix}, \qquad \boldsymbol{b} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \qquad \boldsymbol{c}^{\mathrm{T}} = \begin{pmatrix} 1 & 0 \end{pmatrix}.$$

The characteristic equation of $\boldsymbol{A}$ is

$$\Delta(\lambda) = \lambda^2 + 6\lambda - 8.$$

The roots are $\lambda^{(\pm)} = -3 \pm \sqrt{17}$. Since $\lambda^{(+)} > 0$ the system is unstable. Now

$$(s\boldsymbol{I} - \boldsymbol{A})^{-1} = \frac{1}{s^2 + 6s - 8} \begin{pmatrix} s+4 & 4 \\ 4 & s+2 \end{pmatrix};$$

giving

$$G(s) = \boldsymbol{c}^{\mathrm{T}}(s\boldsymbol{I} - \boldsymbol{A})^{-1}\boldsymbol{b} = \frac{4}{s^2 + 6s - 8}.$$

Change the input to $u(t) - \gamma x_1(t)$.

$$\begin{aligned} \dot{\boldsymbol{x}}(t) &= \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{b}\left[u(t) - \gamma \begin{pmatrix} 1 & 0 \end{pmatrix} \boldsymbol{x}(t)\right], \\ &= \left[\boldsymbol{A} - \gamma \boldsymbol{b} \begin{pmatrix} 1 & 0 \end{pmatrix}\right] \boldsymbol{x}(t) + \boldsymbol{b}u(t), \\ &= \boldsymbol{A}'\boldsymbol{x}(t) + \boldsymbol{b}u(t), \end{aligned}$$

where

$$\boldsymbol{A}' = \begin{pmatrix} -2 & 4 \\ 4-\gamma & -4 \end{pmatrix}.$$

$$(s\boldsymbol{I} - \boldsymbol{A}')^{-1} = \frac{1}{s^2 + 6s - 8 + 4\gamma} \begin{pmatrix} s+4 & 4 \\ 4-\gamma & s+2 \end{pmatrix};$$

giving

$$G'(s) = \boldsymbol{c}^{\mathrm{T}}(s\boldsymbol{I} - \boldsymbol{A}')^{-1}\boldsymbol{b} = \frac{4}{s^2 + 6s - 8 + 4\gamma} = \frac{G(s)}{1 + G(s)\triangle H(s)}.$$

Thus $\triangle H(s) = \gamma$. For the system to be asymptotically stable the real parts of the roots of

$$\phi(s) = s^2 + 6s - 8 + 4\gamma$$

must be negative. The roots are $s^{(\pm)} = -3 \pm \sqrt{17 - 4\gamma}$. The larger root is real and negative if $2 < \gamma \leq 17/4$ and both roots are complex with a negative real part if $\gamma > 17/4$. So the system is asymptotically stable if $\gamma > 2$. When $\gamma = 5$, $u(t) = u_0$ and $x_1(0) = x_2(0) = 0$

$$\bar{y}(s) = \frac{4u_0}{s(s^2 + 6s + 12)} = \frac{u_0}{3s} - \frac{u_0[(s+3)+3]}{3[(s+3)^2+3]}$$

giving

$$y(t) = \tfrac{1}{3}u_0 \left\{1 - \exp(-3t)\left[\cos(\sqrt{3}t) + \sqrt{3}\sin(\sqrt{3}t)\right]\right\}.$$

2) With unit feedback

$$G_{\mathrm{CL}}(s) = \frac{G_{\mathrm{OL}}(s)}{1+G_{\mathrm{OL}}(s)} = \frac{\psi(s)}{\phi(s)},$$

where

$$\psi(s) = \mathsf{K}(\alpha+\beta s),$$
$$\phi(s) = \mathsf{K}(\alpha+\beta s) + s(1+2s)^2 = 4s^3+4s^2+(1+\mathsf{K}\beta)s+\mathsf{K}\alpha.$$

(If you really prefer the leading coefficient in $\phi(s)$ to be one, divide these expressions by four.) Now

$$\Phi_3 = \begin{vmatrix} 4 & 4 & 0 \\ \mathsf{K}\alpha & (1+\mathsf{K}\beta) & 4 \\ 0 & 0 & \mathsf{K}\alpha \end{vmatrix}.$$

For asymptotic stability we must have $\mathsf{K}\alpha > 0$ and since $\alpha > 0$ this implies $\mathsf{K} > 0$. The only remaining condition is $\Phi_3^{(1)} > 0$, which is

$$1 - \mathsf{K}(\alpha-\beta) = (\alpha-\beta)[(\alpha-\beta)^{-1} - \mathsf{K}] > 0.$$

So if $\alpha > \beta$ the system is stable if $(\alpha-\beta)^{-1} > \mathsf{K}$. If $\alpha < \beta$ the system is stable for all positive K. When $\alpha = 1$, $\beta = 2$, $\mathsf{K} = -6$

$$\bar{y}(s) = \frac{-6\bar{u}(s)}{(2s-3)(s+2)} = \frac{6\bar{u}(s)}{7(s+2)} - \frac{12\bar{u}(s)}{7(2s-3)}.$$

Then define

$$\bar{x}_1(s) = \frac{\bar{u}(s)}{(s+2)}, \qquad \bar{x}_2(s) = \frac{\bar{u}(s)}{2s-3},$$

giving

$$\bar{y}(s) = \frac{6}{7}\bar{x}_1(s) - \frac{12}{7}\bar{x}_2(s).$$

Inverting the Laplace transforms

$$\dot{x}_1(t) = -2x_1(t) + u(t), \qquad \dot{x}_2(t) = \frac{3}{2}x_2(t) + \frac{1}{2}u(t),$$
$$y(t) = \frac{6}{7}x_1(t) - \frac{12}{7}x_2(t).$$

Thus the matrix form has

$$\boldsymbol{A} = \begin{pmatrix} -2 & 0 \\ 0 & \frac{3}{2} \end{pmatrix}, \qquad \boldsymbol{b} = \begin{pmatrix} 1 \\ \frac{1}{2} \end{pmatrix}, \qquad \boldsymbol{c}^{\mathrm{T}} = \begin{pmatrix} \frac{6}{7} & -\frac{12}{7} \end{pmatrix}.$$

3) With $H(s) = \gamma$

$$G_{\text{CL}}(s) = \frac{G_{\text{OL}}(s)}{1+\gamma G_{\text{OL}}(s)} = \frac{\psi(s)}{\phi(s)},$$

where

$$\psi(s) = 1,$$
$$\phi(s) = s^3 + s^2 + s + (\gamma + 1)$$

and

$$\Phi_3 = \begin{vmatrix} 1 & 1 & 0 \\ (1+\gamma) & 1 & 1 \\ 0 & 0 & (1+\gamma) \end{vmatrix}.$$

From the Routh-Hurwitz criterion the two conditions for asymptotic stability are $1+\gamma > 0$ and $1 > 1+\gamma$. The latter cannot be satisfied with $\gamma > 0$. If output feedback is included

$$\begin{aligned} \phi(s) &= s^3 + s^2(1+\alpha) + s(1+\beta) + 1 + \gamma \\ &= (s+1)(s+2)(s+3) \\ &= s^3 + 6s^2 + 11s + 6. \end{aligned}$$

Thus $\alpha = 5$, $\beta = 10$, $\gamma = 5$.

$$\bar{y}(s) = \frac{u_0}{s(s+1)(s+2)(s+3)} = \frac{u_0}{6s} - \frac{u_0}{2(s+1)} + \frac{u_0}{2(s+2)} - \frac{u_0}{6(s+3)}.$$

Inverting the Laplace transform

$$y(t) = \tfrac{1}{6}u_0 \left\{1 - 3\exp(-t) + 3\exp(-2t)) - \exp(-3t)\right\}.$$

9.6 Problems 6

1) $$\mathcal{I}[x] + p\mathcal{J}[x] = \int_0^1 \left\{ \tfrac{1}{2}[\dot{x}(\tau)]^2 + x(\tau)\dot{x}(\tau) + px(\tau) \right\} \mathrm{d}\tau$$

The Euler-Lagrange equation is

$$\frac{\mathrm{d}[\dot{x}(\tau) + x(\tau)]}{\mathrm{d}\tau} - \dot{x}(\tau) - p = 0,$$

giving

$$\ddot{x}(\tau) = p.$$

Thus

$$x(\tau) = \tfrac{1}{2}p\tau^2 + \mathsf{A}\tau + \mathsf{B}.$$

From the initial and final conditions $\mathsf{B} = 0$, $\mathsf{A} = 5 - \frac{1}{2}p$. Substituting into the constraint

$$\int_0^1 \left\{\tfrac{1}{2}p\tau^2 + [5 - \tfrac{1}{2}p]\tau\right\} \mathrm{d}\tau = 2$$

which gives $p = 6$ and hence

$$x(\tau) = 3\tau^2 + 2\tau.$$

2)
$$\mathcal{I}[x] + p\mathcal{J}[x] = \int_0^2 \left\{[\dot{x}(\tau)]^2 + x(\tau)[1 + p\tau]\right\} \mathrm{d}\tau.$$

The Euler-Lagrange equation gives

$$\ddot{x}(\tau) = \tfrac{1}{2}[1 + p\tau].$$

Thus

$$x(\tau) = \tfrac{1}{12}p\tau^3 + \tfrac{1}{4}\tau^2 + \mathsf{A}\tau + \mathsf{B}.$$

From the initial and final conditions $\mathsf{A} = 23 - \frac{1}{3}p$, $\mathsf{B} = 1$. Substituting into the constraint and performing the integral gives $p = 60$ and thus

$$x(\tau) = 5\tau^3 + \tfrac{1}{4}\tau^2 + 3\tau + 1.$$

3)
$$\mathcal{I}_p[u, x] = \int_0^{t_{\mathrm{F}}} \left\{[x(t)]^2 + [u(t)]^2 + p(t)[\dot{x}(t) - u(t)]\right\} \mathrm{d}t.$$

The Euler-Lagrange equations give

$$\begin{aligned}\dot{p}(t) - 2x(t) &= 0,\\ 2u(t) - p(t) &= 0,\end{aligned}$$

giving

$$\dot{u}(t) = x(t).$$

At this point you could use Laplace transforms to solve this equation with that given in the question. It is probably easier just to note that they give $\ddot{x}(t) = x(t)$, which has the general solution

$$x(t) = \mathsf{A}\exp(t) + \mathsf{B}\exp(-t).$$

From the initial and final conditions

$$\mathsf{A} = \frac{\alpha[1-\exp(-t_{\mathrm{F}})]}{\exp(t_{\mathrm{F}})-\exp(-t_{\mathrm{F}})}, \qquad \mathsf{B} = \frac{\alpha[\exp(t_{\mathrm{F}})-1]}{\exp(t_{\mathrm{F}})-\exp(-t_{\mathrm{F}})}.$$

Thus

$$\begin{aligned} u(t) &= \mathsf{A}\exp(t) - \mathsf{B}\exp(-t) \\ &= \frac{\alpha\{[1-\exp(-t_{\mathrm{F}})]\exp(t) - [\exp(t_{\mathrm{F}})-1]\exp(-t)\}}{\exp(t_{\mathrm{F}})-\exp(-t_{\mathrm{F}})}. \end{aligned}$$

$$[x(t)]^2 + [u(t)]^2 = 2\mathsf{A}^2\exp(2t) + 2\mathsf{B}^2\exp(-2t).$$

So

$$\begin{aligned} \mathcal{I}[x^*] &= \mathsf{A}^2[\exp(2t_{\mathrm{F}})-1] - \mathsf{B}^2[\exp(-2t_{\mathrm{F}})-1] \\ &= 2\alpha^2\,\frac{\exp(t_{\mathrm{F}})-1}{\exp(t_{\mathrm{F}})+1}. \end{aligned}$$

4) Let $x_1(t) = x(t)$ and $x_2(t) = \dot{x}(t)$. Then

$$\dot{x}_1(t) = x_2(t),$$

$$\dot{x}_2(t) = u(t).$$

We have two constraints so

$$\mathcal{I}_p[u, x_1, x_2] = \int_0^1 \left\{[u(t)]^2 + p_1(t)[\dot{x}_1(t) - x_2(t)] + p_2(t)[\dot{x}_2(t) - u(t)]\right\} \mathrm{d}t$$

and the Euler-Lagrange equations are

$$\dot{p}_1(t) = 0,$$

$$\dot{p}_2(t) + p_1(t) = 0,$$

$$2u(t) - p_2(t) = 0.$$

This gives $p_1(t) = \mathsf{A}$, $p_2(t) = \mathsf{B} - \mathsf{A}t$ and thus

$$u(t) = \tfrac{1}{2}(\mathsf{B} - \mathsf{A}t),$$

$$\dot{x}(t) = \tfrac{1}{2}\mathsf{B}t - \tfrac{1}{4}\mathsf{A}t^2 + \mathsf{C},$$

$$x(t) = \tfrac{1}{4}\mathsf{B}t^2 - \tfrac{1}{12}\mathsf{A}t^3 + \mathsf{C}t + \mathsf{D}.$$

Applying the initial a final conditions gives $\mathsf{A} = -24$, $\mathsf{B} = -12$, $\mathsf{C} = 1$ and $\mathsf{D} = 0$. Thus

$$u(t) = 12t - 6.$$

5) $$\mathcal{I}_p[u(t),x(t)] = \int_0^{t_{\mathrm{F}}} \left\{ \tfrac{1}{2}[u(t)]^2 + p(t)[\dot{x}(t) + x(t) - u(t)] \right\} \mathrm{d}t.$$

The Euler-Lagrange equations are

$$\dot{p}(t) - p(t) = 0,$$
$$u(t) - p(t) = 0$$

Eliminating $p(t)$ gives

$$\dot{u}(t) = u(t),$$

which has the general solution

$$u^*(t) = 2\mathsf{C}\exp(t).$$

Substituting into

$$\dot{x}(t) = u(t) - x(t),$$

gives

$$\dot{x}(t) + x(t) = 2\mathsf{C}\exp(t).$$

This becomes

$$\frac{\mathrm{d}}{\mathrm{d}t}[x(t)\exp(t)] = 2\mathsf{C}\exp(2t)$$

with the solution

$$x^*(t) = \mathsf{C}\exp(t) + \mathsf{D}\exp(-t).$$

From the conditions at the boundaries

$$\mathsf{C} = \frac{x_0[1-\exp(-t_{\mathrm{F}})]}{\exp(t_{\mathrm{F}}) - \exp(-t_{\mathrm{F}})}, \qquad \mathsf{D} = \frac{x_0[\exp(t_{\mathrm{F}}) - 1]}{\exp(t_{\mathrm{F}}) - \exp(-t_{\mathrm{F}})}.$$

Now

$$\begin{aligned} \mathcal{I}[u^*] = 2\mathsf{C}^2 \int_0^{t_{\mathrm{F}}} \exp(2t)\mathrm{d}t &= \mathsf{C}^2[\exp(2t_{\mathrm{F}}) - 1] \\ &= x_0^2 \frac{\exp(t_{\mathrm{F}}) - 1}{\exp(t_{\mathrm{F}}) + 1} \\ &= x_0^2 \tanh\left(\tfrac{1}{2}t_{\mathrm{F}}\right). \end{aligned}$$

If $x(t) = x_0$ for the whole time then $u(t) = x_0$ for the whole time and

$$\mathcal{I}[u] = \tfrac{1}{2}x_0^2 t_{\mathrm{F}}.$$

Now consider

$$g(t_{\mathrm{F}}) = \frac{2}{x_0^2}\{\mathcal{I}[u] - \mathcal{I}[u^*]\} = t_{\mathrm{F}} - 2\tanh\left(\tfrac{1}{2}t_{\mathrm{F}}\right).$$

Since $g(0) = 0$ and $g'(t_{\mathrm{F}}) = 1 - \operatorname{sech}^2\left(\tfrac{1}{2}t_{\mathrm{F}}\right) > 0$, for $t_{\mathrm{F}} > 0$,

$$\mathcal{I}[u] \geq \mathcal{I}[u^*] \qquad \text{for} \qquad t_{\mathrm{F}} \geq 0.$$

If $x(t_{\mathrm{F}})$ is unrestricted the same Euler-Lagrange equations apply but the condition at $t = t_{\mathrm{F}}$ is replaced by the transversality condition $p(t_{\mathrm{F}}) = u(t_{\mathrm{F}}) = 0$. From the general solution for $u(t)$ this implies that $u(t) = 0$ for all t and hence that $\mathcal{I}[u] = 0$.

6) $$\mathcal{I}_p[u(t), x(t)] = \int_0^{t_{\mathrm{F}}} \left\{\frac{1}{2}[u(t)]^2 + \frac{1}{2}[x(t)]^2 + p(t)[\dot{x}(t) + 1 - u(t)]\right\} \mathrm{d}t.$$

The Euler-Lagrange equations are

$$\dot{p}(t) - x(t) = 0,$$
$$u(t) - p(t) = 0$$

Eliminating $p(t)$ gives

$$\dot{u}(t) = x(t).$$

Using

$$\dot{x}(t) = u(t) - 1$$

gives

$$\ddot{x}(t) = x(t),$$

which has the solution

$$x(t) = \mathsf{A}\sinh(t) + \mathsf{B}\cosh(t).$$

With $x(0) = 0$ and $x(t_{\mathrm{F}}) = 1$, $\mathsf{B} = 0$ and $\mathsf{A} = 1/\sinh(t_{\mathrm{F}})$ which gives the solution quoted in the problem for (a). $u(t)$ is given from $u(t) = 1 + \dot{x}(t)$.

In case (b) we still have $\mathsf{B}=0$ but the transversality condition gives $p(t_{\mathrm{F}}) = u(t_{\mathrm{F}}) = 0$ which gives $\mathsf{A} = -1/\cosh(t_{\mathrm{F}})$. Since $\cosh(y)$ is an increasing function greater the $\sinh(y)$ for $y > 0$

$$\left[1 - \frac{\cosh(t)}{\cosh(t_{\mathrm{F}})}\right]^2 \quad < \quad \left[1 + \frac{\cosh(t)}{\sinh(t_{\mathrm{F}})}\right]^2$$

$$\left[\frac{\sinh(t)}{\cosh(t_{\mathrm{F}})}\right]^2 \quad < \quad \left[\frac{\sinh(t)}{\sinh(t_{\mathrm{F}})}\right]^2.$$

So $\mathcal{I}$ is less in case (b) than case (a). This is to be expected since for (b) the minimization is over a range of values for $x(t_{\mathrm{F}})$ and not just $x(t_{\mathrm{F}}) = 0$.

7) The two constraints are

$$\dot{x}_1(t) = x_2(t),$$
$$\dot{x}_2(t) = u(t) - \mu\, x_2(t)$$

Since we are looking for an extremum of time $f(u(t), x_1(t), x_2(t); t) = 1$ and the Hamiltonian is

$$H(u(t), x_1(t), x_2(t), p_1(t), p_2(t)) = p_1(t)x_2(t) + p_2(t)[u(t) - \mu\, x_2(t)] - 1.$$

Then the Hamiltonian-Pontriagin equations are

$$\dot{p}_1(t) = 0, \qquad \dot{p}_2(t) = \mu\, p_2(t) - p_1(t),$$

with

$$\frac{\partial H}{\partial u} = p_2(t).$$

Thus

$$p_1(t) = \mathsf{C}$$

and

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[p_2(t)\exp(-\mu t)\right] = -\mathsf{C}\exp(-\mu t),$$

giving

$$p_2(t) = \mu^{-1}\mathsf{C} + \mathsf{B}\exp(\mu t).$$

This is the switching function. Since it is a monotonically increasing function it has at most one zero in the range $0 \le t \le t_{\mathrm{F}}$, so there will be at most one switch between $u_{\rm B}$ and $-u_{\rm B}$. In fact there must be one switch in order for the wheel to begin from rest and return to rest.

If $x_2 = \dot{\theta}$ is not restricted at $t = t_{\rm F}$ then the transversality condition $p_2(t_{\rm F}) = 0$ applies. Thus the switching function is

$$p_2(t) = \mathsf{B}[\exp(\mu t) - \exp(\mu t_{\rm F})].$$

The zero is now at the end of the range so no switching occurs. So $u(t) = u_{\rm B}$ for the whole motion. So θ satisfies

$$\ddot{\theta}(t) + \mu\dot{\theta}(t) = u_{\rm B},$$

which has the solution

$$\theta = \mathsf{A} + \mathsf{B}\exp(-\mu t) + t u_{\rm B}/\mu.$$

This must satisfy the conditions

$$\begin{aligned}
\theta_{\rm I} &= \mathsf{A} + \mathsf{B}, \\
0 &= -\mu\mathsf{B} + u_{\rm B}/\mu, \\
\theta_{\rm F} &= \mathsf{A} + \mathsf{B}\exp(-\mu t_{\rm F}) + t_{\rm F} u_{\rm B}/\mu.
\end{aligned}$$

Eliminating A and B gives the required condition.

8) With $x_1(t) = x(t)$ and $x_2(t) = \dot{x}(t)$ the three constraints are

$$\begin{aligned}
\dot{x}_1(t) &= x_2(t), \\
\dot{x}_2(t) &= \frac{\mathsf{C}\,u(t)}{m(t)} - g, \\
\dot{m}(t) &= -u(t)
\end{aligned}$$

If $t_{\rm F}$ is the time to reach top of the flight then the height reached is

$$\mathcal{I}[x_2] = \int_0^{t_{\rm F}} \dot{x}_1(t)\mathrm{d}t = \int_0^{t_{\rm F}} x_2(t)\mathrm{d}t$$

and it is required to maximize this. The Hamiltonian is

$$\begin{aligned}
H(u(t), x_1(t), x_2(t)m(t), p_1(t), p_2(t), p_3(t) \quad &= \quad p_1(t)\,x_2(t) + p_2(t)\left[\frac{\mathsf{C}\,u(t)}{m(t)} - g\right] \\
&\quad -p_3(t)u(t) - x_2(t).
\end{aligned}$$

The Hamiltonian-Pontriagin equations are

$$\dot{p}_1(t) = 0,$$

$$\dot{p}_2(t) = 1 - p_1(t),$$

$$\dot{p}_3(t) = \frac{p_2(t)\mathsf{C}\,u(t)}{[m(t)]^2},$$

and the switching function is

$$\mathcal{S}(t) = \frac{\partial H}{\partial u} = \frac{p_2(t)\,\mathsf{C}}{m(t)} - p_3(t).$$

Since we are looking for a maximum rather than a minimum the inequality (6.113) in the notes is reversed and we require

$$\mathcal{S}(t)\delta u > 0.$$

Assuming for the moment that $\mathcal{S}(t)$ as a function of $u(t)$ does not have a zero in the allowed range of $u(t)$ and that this is a situation of bang-bang control we must have

$$u^*(t) = \begin{cases} u_{\mathrm{U}}, & \mathcal{S}(t) < 0, \\ 0, & \mathcal{S}(t) > 0, \end{cases}$$

Now

$$p_1(t) = \mathsf{A}, \qquad p_2(t) = (1 - \mathsf{A})t + \mathsf{B},$$

for some constants A and B and

$$\dot{\mathcal{S}}(t) = \mathsf{D}/m(t), \qquad \text{where } \mathsf{D} = \mathsf{C}(1 - \mathsf{A}).$$

Thus $\mathcal{S}(t)$ is a monotonic function of t with at most one switch. Since the rocket must be initially propelled upwards we must have $\mathcal{S}(0) < 0$ and if a switch is to occur $\mathsf{D} > 0$. If a switch occurs then prior to its occurring $\dot{m}(t) = -u_{\mathrm{U}}$

$$\dot{\mathcal{S}}(t) = -\frac{\dot{m}(t)\mathsf{D}}{u_{\mathrm{U}}m(t)}.$$

Integrating and using the condition $\mathcal{S}(t_{\mathrm{S}}) = 0$ gives

$$\mathcal{S}(t) = -\frac{\mathsf{D}}{u_{\mathrm{U}}} \ln\left\{\frac{m(t)}{m(t_{\mathrm{S}})}\right\}, \qquad 0 \le t \le t_{\mathrm{S}}.$$

After the switch $\dot{m}(t) = 0$ and

$$\mathcal{S}(t) = -\frac{\mathsf{D}(t_{\mathrm{S}} - t)}{m(t_{\mathrm{S}})}, \qquad t_{\mathrm{S}} \le t \le t_{\mathrm{F}}.$$

9) Let $x_1(t) = x(t)$ and $x_2(t) = \dot{x}(t)$. Then

$$\dot{x}_1(t) = x_2(t),$$
$$\dot{x}_2(t) = u(t) - k.$$

If t_{F} is the time of the drive

$$\mathcal{I}[u, x_1, x_2] = t_{\mathrm{F}} = \int_0^{t_{\mathrm{F}}} \mathrm{d}t.$$

This gives $f(u(t), x_1(t), x_2(t); t) = 1$ and the Hamiltonian is

$$H(u(t), x_1(t), x_2(t), p_1(t), p_2(t)) = p_1(t)x_2(t) + p_2(t)[u(t) - k] - 1.$$

Then the Hamiltonian-Pontriagin equations are

$$\dot{p}_1(t) = 0,$$
$$\dot{p}_2(t) = -p_1(t)$$

giving

$$p_1(t) = \mathsf{A},$$
$$p_2(t) = \mathsf{B} - \mathsf{A}t$$

and

$$\frac{\partial H}{\partial u} = p_2(t) = \mathsf{B} - \mathsf{A}t.$$

Thus H is a monotonic strictly increasing or strictly decreasing function of $u(t)$ for all t except at $t = \mathsf{B}/\mathsf{A}$ if this lies in the interval of time of the journey. There can be at most one switch. Since the vehicle starts from rest at $t = 0$ and comes to rest at $t = t_{\mathrm{F}}$ it must be the case that $\ddot{x}(0) = u(0) > 0$ and $\ddot{x}(t_{\mathrm{F}}) = u(t_{\mathrm{F}}) < 0$. So in the early part of the journey $u(t) = u_{\mathrm{B}}$ and in the later part of the journey $u(t) = -u_{\mathrm{B}}$. The switch over occurs when $p_2(t)$ changes sign. So $p_2(t)$ is the switching function $\mathcal{S}(t)$. For the first part of the journey

$$\ddot{x}(t) = u_{\mathrm{B}} - k,$$
$$\dot{x}(t) = (u_{\mathrm{B}} - k)t,$$
$$x(t) = \tfrac{1}{2}(u_{\mathrm{B}} - k)t^2.$$

For the second part of the journey

$$\ddot{x}(t) = -(u_{\mathrm{B}} + k),$$
$$\dot{x}(t) = (u_{\mathrm{B}} + k)(t_{\mathrm{F}} - t),$$
$$x(t) = L - \tfrac{1}{2}(u_{\mathrm{B}} + k)(t_{\mathrm{F}} - t)^2.$$

Since both $\dot{x}(t)$ and $x(t)$ are continuous over the whole journey the switch occurs at

$$t = t_{\mathrm{S}} = \frac{(u_{\mathrm{B}} + k)t_{\mathrm{F}}}{2u_{\mathrm{B}}},$$

with

$$t_{\mathrm{F}} = 2\sqrt{\frac{u_{\mathrm{B}}L}{u_{\mathrm{B}}^2 - k^2}}.$$

The distance travelled when switching occurs is

$$x = \frac{L(u_{\mathrm{B}} + k)}{2u_{\mathrm{B}}}.$$

9.7 Problems 7

1) The closed-loop transfer function is

$$G_{\mathrm{CL}}(s) = \frac{G(s)}{1 + G(s)} = \frac{\mathsf{K}}{(1+s)^3 + \mathsf{K}} = \frac{\mathsf{K}}{\phi(s)},$$

where

$$\phi(s) = s^3 + 3s^2 + 3s + (1 + \mathsf{K}).$$

From, the Routh-Hurwitz criterion, for stability we must have $(1 + \mathsf{K}) > 0$ and

$$\Phi_3^{(1)} = \begin{vmatrix} 3 & 1 \\ (1+\mathsf{K}) & 3 \end{vmatrix} > 0.$$

(The remaining condition is $a_2 > 0$, which is true.) Thus, for stability $-1 < \mathsf{K} < 8$.

Now

$$G(\mathrm{i}\omega) = \frac{\mathsf{K}}{(1 + \mathrm{i}\omega)^3} = \frac{\mathsf{K}(1 - \mathrm{i}\omega)^3}{(1 + \omega^2)^3}.$$

So

$$X(\omega) = \frac{\mathsf{K}(1 - 3\omega^2)}{(1 + \omega^2)^3}, \qquad Y(\omega) = \frac{\mathsf{K}\omega(\omega^2 - 3)}{(1 + \omega^2)^3}.$$

The Nyquist locus Γ_{G} is the curve in the $Z = X + \mathrm{i}Y$ plane given by

$$X(\omega) + \mathrm{i}Y(\omega) = G(\mathrm{i}\omega), \qquad -\infty \le \omega \le +\infty.$$

The Nyquist criterion states that: *If $G(s)$ is itself asymptotically stable, and thus has no poles with $\Re\{s\} > 0$, then the closed-loop transfer function is asymptotically stable if Γ_G does not encircle the point -1.*

As is normally the case the ends of the curve, where $\omega = \pm\infty$ are at the origin and $Y(-\omega) = -Y(\omega)$. The curve cuts the X–axis at $\omega = 0$, when $X = \mathsf{K}$ and at $\omega = \pm\sqrt{3}$. These two parameter values coincide with $X = -\mathsf{K}/8$. When $\mathsf{K} < 0$ the single crossing point is at negative values of X. So $\mathrm{Ind}(\Gamma_{\mathrm{G}}; -1) = 1$ leading to instability if $\mathsf{K} < -1$ and $\mathrm{Ind}(\Gamma_{\mathrm{G}}; -1) = 0$ leading to stability if $\mathsf{K} > -1$. If $\mathsf{K} > 0$ the double crossing point is at negative values of X. So $\mathrm{Ind}(\Gamma_{\mathrm{G}}; -1) = 2$ leading to instability if $\mathsf{K} > 8$ and $\mathrm{Ind}(\Gamma_{\mathrm{G}}; -1) = 0$ leading to stability if $\mathsf{K} < 8$. The following MAPLE program calculates $X(\omega)$ and $Y(\omega)$ and display stable and unstable cases for both signs of K.

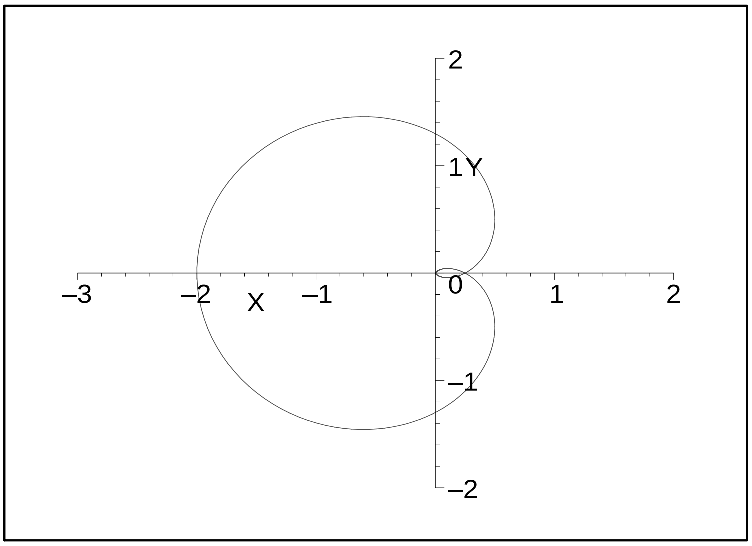
```
> G:=(s,K)->K/((1+s)^3):

> X:=(w,K)->simplify(evalc(Re(G(I*w,K)))):

> X(w,K);
```

$$-\frac{K\,(-1+3\,w^2)}{1+3\,w^2+3\,w^4+w^6}$$

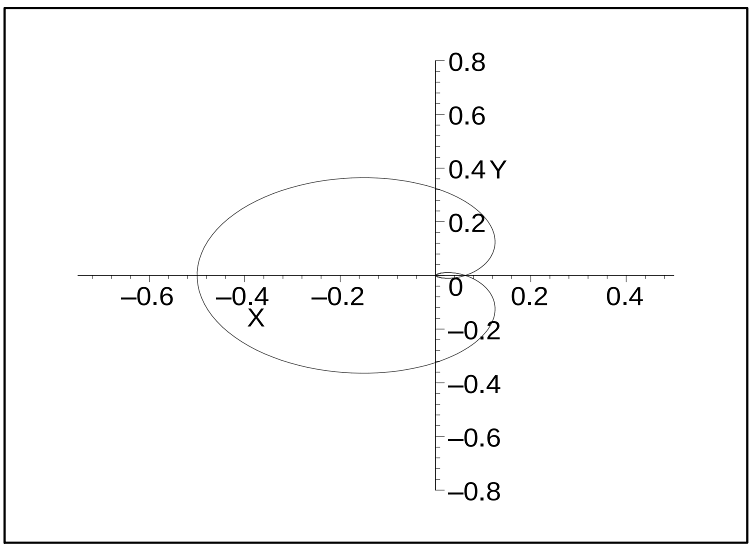
```
> Y:=(w,K)->simplify(evalc(Im(G(I*w,K)))):

> Y(w,K);
```

$$\frac{K\,w\,(-3+w^2)}{1+3\,w^2+3\,w^4+w^6}$$

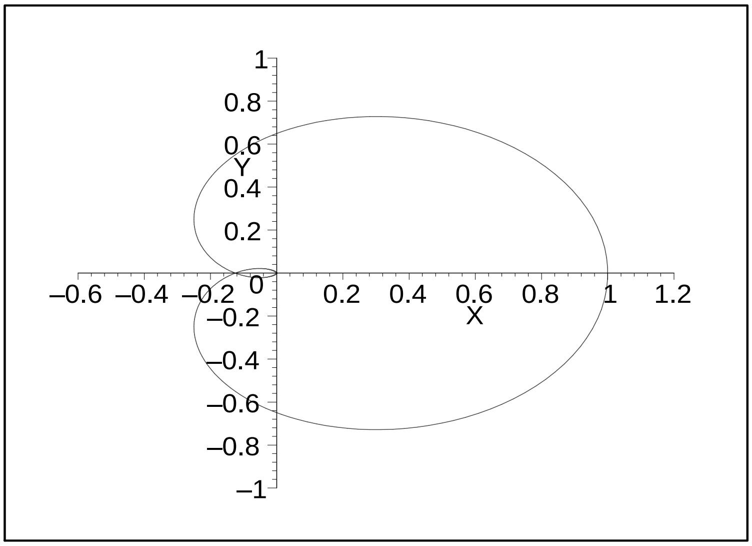
```
> with(plots):

> plot([X(w,-2),Y(w,-2),

> w=-infinity..infinity],X=-3..2,Y=-2..2,numpoints=1000);
```



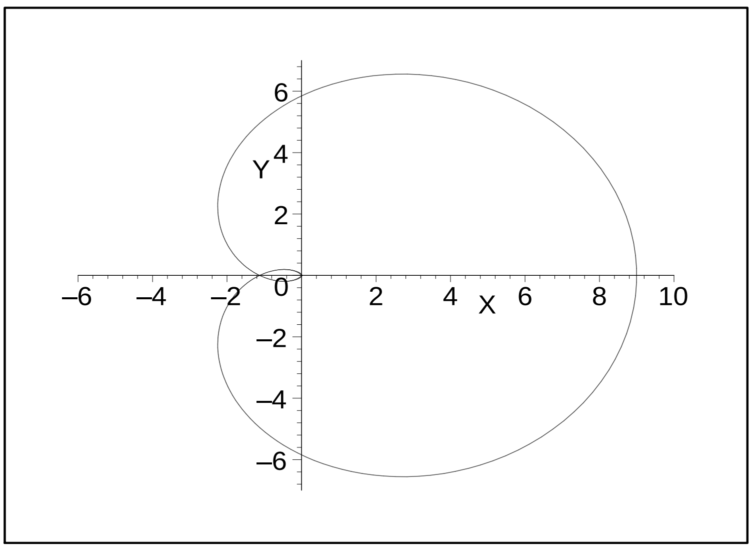

```
> plot([X(w,-0.5),Y(w,-0.5),
> w=-infinity..infinity],X=-0.75..0.5,Y=-0.8..0.8,numpoints=1000);
```




```
> plot([X(w,1),Y(w,1),
> w=-infinity..infinity],X=-0.6..1.2,Y=-1..1,numpoints=1000);
```




```
> plot([X(w,9),Y(w,9),

> w=-infinity..infinity],X=-6..10,Y=-7..7,numpoints=1000);
```



9.8 Problems 8

1) $\nabla\mathcal{L} = (nx^{n-1}, \alpha m y^{m-1})$.

(i)

$$\begin{aligned}\boldsymbol{F}.\boldsymbol{\nabla}\mathcal{L} &= -nx^{n-1}(x+2y^2)+\alpha m y^{m-1}(xy-y^3)\\ &= -nx^n-2nx^{n-1}y^2+\alpha m x y^m-\alpha m y^{m+2}.\end{aligned}$$

If $n=m=2$ and $\alpha=2$ then

$$\boldsymbol{F}.\boldsymbol{\nabla}\mathcal{L}=-2x^2-4y^4<0$$

and $\mathcal{L}(0,0)=0$, with $\mathcal{L}(x,y)$ having a minimum at $(0,0)$. So the system is asymptotically stable.

(ii)

$$\begin{aligned}\boldsymbol{F}.\boldsymbol{\nabla}\mathcal{L} &= nx^{n-1}(y-x^3)-\alpha m y^{m-1}x^3\\ &= nx^{n-1}y-nx^{n+2}-\alpha m y^{m-1}x^3.\end{aligned}$$

If $n=4$, $m=2$ and $\alpha=2$ then

$$\boldsymbol{F}.\boldsymbol{\nabla}\mathcal{L}=-4x^6<0$$

and $\mathcal{L}(0,0)=0$, with $\mathcal{L}(x,y)$ having a minimum at $(0,0)$. So the system is asymptotically stable.

2) From the second equation $y^3=x^3$ so the equilibrium point is on $x=y$. From the first equation $0=x^2y-xy^2+x^3=x^3$. So the only equilibrium point is $x=y$.

$$\boldsymbol{\nabla}\mathcal{L}=(2x+\alpha y,2\beta y+\alpha x).$$

So

$$\begin{aligned}\boldsymbol{F}.\boldsymbol{\nabla}\mathcal{L} &= (2x+\alpha y)(-xy^2+x^2y+x^3)+(2\beta y+\alpha x)(y^3-xz^3)\\ &= x^4(2-\alpha)+2y^4\beta+x^2y^2(\alpha-2)+x^3y(2+\alpha-2\beta).\end{aligned}$$

If $\alpha=\beta=2$ then

$$\boldsymbol{F}.\boldsymbol{\nabla}\mathcal{L}=4y^4>0.$$

Also $x^2+2xy+2y^2=0$ has no real roots. So $\mathcal{L}(x,y)$ has a zero minimum at $(0,0)$ and the system is therefore unstable.

3) Consider the reverse trajectory obtained by replacing t by $-t$. The trajectory $\gamma=(x(-t),y(-t))$ satisfies the conditions of the Poincaré–Bendixson theorem for $(-t)\geq 0$. So the reverse trajectory tends to a periodic solution or equilibrium point in $\mathcal{C}$. Such a period solution or equilibrium point is

also a periodic solution or equilibrium point of the forward trajectory. The equilibrium points of the equation are given by

$$
\begin{aligned}
0 &= -x - y + x(x^2 + 2y^2), \qquad (1) \\
0 &= x - y + y(x^2 + 2y^2). \qquad (2)
\end{aligned}
$$

Multiplying (1) by y and (2) by x and subtracting gives $x^2 + y^2 = 0$. The only solution to this (and thus the only equilibrium point) is $x = y = 0$. Linearizing about $(0, 0)$ gives the stability matrix

$$
\boldsymbol{J}^* = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}.
$$

The eigenvalues of this matrix are $-1 \pm \mathrm{i}$, so the origin is a stable focus. From the given equations

$$
\begin{aligned}
r\frac{\mathrm{d}r}{\mathrm{d}t} &= x\frac{\mathrm{d}x}{\mathrm{d}t} + y\frac{\mathrm{d}y}{\mathrm{d}t} \\
&= -x^2 - y^2 + (x^2 + y^2)(x^2 + 2y^2) \\
&= -r^2 + r^4[1 + \sin^2(\theta)].
\end{aligned}
$$

So

$$
\dot{r}(t) = -r + r^3[1 + \sin^2(\theta)].
$$

Transforming the first equation into polar form

$$
\begin{aligned}
\frac{\mathrm{d}x}{\mathrm{d}t} &= \frac{\mathrm{d}r}{\mathrm{d}t}\cos(\theta) - r\sin(\theta)\frac{\mathrm{d}\theta}{\mathrm{d}t} \\
&= -r\cos(\theta) - r\sin(\theta) + r^3\cos(\theta)[1 + \sin^2(\theta)].
\end{aligned}
$$

Substituting for $\dot{r}(t)$ gives

$$
\dot{\theta}(t) = 1
$$

and this

$$
\theta(t) = \theta_0 + t.
$$

The angular velocity of the system is constant. With $r = 1 + \delta$ and $\delta > 0$

$$
\dot{r}(t) = (1 + \delta)(2\delta + \delta^2) + (1 + \delta)^3\sin^2(\theta) > 0.
$$

Rearranging the formula for $\dot{r}(t)$ gives

$$\dot{r}(t) = r(2r^2 - 1) + r^3[\sin^2(\theta) - 1].$$

With $r = 1/\sqrt{2} - \delta$, and $\delta > 0$,

$$\dot{r}(t) = -\left(\frac{1}{\sqrt{2}} - \delta\right)\left(\frac{4\delta}{\sqrt{2}} - 2\delta^2\right) + \left(\frac{1}{\sqrt{2}} - \delta\right)^2 [\sin^2(\theta) - 1],$$

which is negative for sufficiently small δ. So the region

$$\frac{1}{\sqrt{2}} - \delta \le r \le 1 + \delta$$

satisfies the conditions of the first part of the problem and thus contains a periodic solution. (It can't contain an equilibrium point, since the only equilibrium point $(0, 0)$ is outside the region.)

From the expression for $\dot{r}(t)$

$$\frac{1}{r^3}\frac{\mathrm{d}r}{\mathrm{d}t} + \frac{1}{r^2} = 1 + \sin^2(t + \theta_0).$$

The integrating factor is $\exp(-2t)$ and we obtain

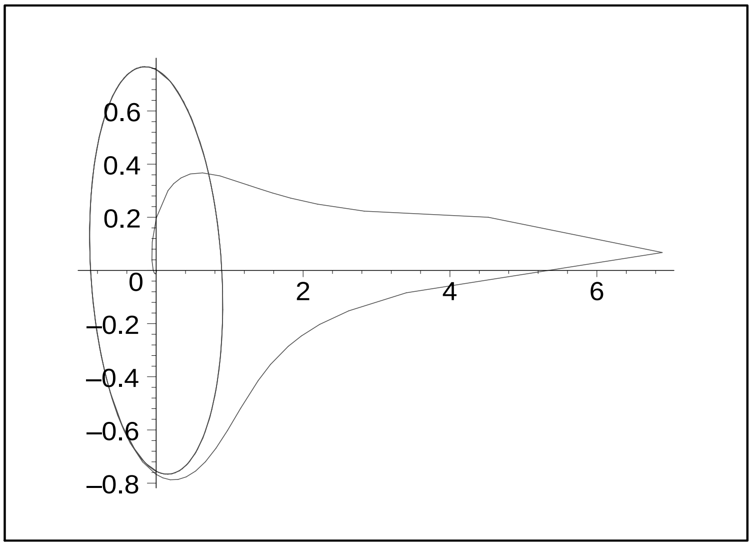
$$\frac{1}{r^2} = 2\mathsf{A}\exp(2t) + \frac{1}{4}\{5 + 2\sin^2(t + \theta_0) + 2\sin(t + \theta_0)\cos(t + \theta_0)\}.$$

A is evaluated by setting $r = t_0$ at $t = 0$. So A is finite and as $t \to \infty$, $r \to 0$. As $r \to -\infty$ the trajectory approaches the curve

$$r(\theta) = 2\{6 + 2\sin(2\theta) - \cos(2\theta)\}^{-1/2}.$$

We can use MAPLE not only to obtain the curve $r(\theta)$ but the whole of the solution, parametrized by t. With $r = r_0 = 5$ and $\theta = \theta_0 = 0$ at $t = 0$ we have we have

```
> A:=(r0,theta0)->

> 1/(2*r0^2)-(1/8)*(5+2*(sin(theta0))^2+2*sin(theta0)*cos(theta0)):
> W:=(t,r0,theta0)->

> 2*A(r0,theta0)*exp(2*t)

> +(1/4)*(5+2*(sin(t+theta0))^2+2*sin(t+theta0)*cos(t+theta0)):
> r:=(t,r0,theta0)->1/sqrt(abs(W(t,r0,theta0))):
> theta:=(t,theta0)->t+theta0:
> with(plots):
> plot([r(t,5,0),theta(t,0),t=25..-25],coords=polar);
```


0.6
0.4
0.2
0
2
4
6
–0.2
–0.4
–0.6
–0.8