OPTIMAL CONTROL AND POWER ELECTRONICS
IN AGRICULTURAL APPLICATIONS

BY

P K MULLEN B.Sc.

A dissertation submitted in partial fulfilment of the regulations governing the award by the Council of National Academic Awards of the degree of

MASTER OF SCIENCE IN CONTROL ENGINEERING

December 1984

DEPARTMENT OF ELECTRICAL, ELECTRONIC AND CONTROL ENGINEERING
SUNDERLAND POLYTECHNIC
ENGLAND
The aim of this project was to design two control systems arising from needs in the agricultural/automotive industry. The first is a baler control system for improved material handling, and the second is a spray control system for even crop coverage.

The baler control system uses a conventional control strategy and was successfully produced to meet very tight deadlines. The electrical control system that was designed involved state-of-the-art power electronics, including newly developed power MOSFET transistors. The baler control system is now on sale worldwide and is currently regarded by the agricultural industry generally as a market leader.

A design study was carried out using optimal control theory for the spray control system, and the work was validated using simulation studies.
ABSTRACT

Comparison of an existing electronic control with a product in development yields insights into the increasing sophistication of agricultural equipment. The existing Round Baler Control combined analog and digital circuits with power electronics, to drive an electric actuator wrapping twine around a bale. Actuator speed was measured in a novel way using motor back e.m.f. and controlled by a proportional plus integral network driving power MOSFET transistors.

The new Fertiliser Sprayer Control required that flow rate be made proportional to speed, for even coverage. A performance index was defined to minimise the error in the flow rate and also the actuation effort (to prolong actuator life). Realisation of the performance index by the Euler-Lagrange equation was not satisfactory, but the Hamilton-Jacobi equation yielded a closed-loop control law. This law was implemented using a low order observer to generate an inaccessible signal. Analog computer tests showed the viability of the control law in the presence of noise and system nonlinearity.
DEDICATION

This book is dedicated with love to my parents.
CONTENTS

ABSTRACT

ACKNOWLEDGEMENTS

LIST OF SYMBOLS

LIST OF FIGURES

1 INTRODUCTION

PART A: AN ELECTRONIC PRODUCT FOR FARM APPLICATIONS:

THE WARNER ELECTRIC ROUND BALER CONTROL

2 TYING TWINE AROUND BALES

2.1 Statement of Problem

2.2 Specification of Control

3 DIGITAL CIRCUIT OF ROUND BALER CONTROL

3.1 Introduction

3.2 Selection of Logic Family

3.3 Sequencer Circuit

  3.3.1 Overview
  3.3.2 Design Method for Sequencer Control
  3.3.3 Sequence Table
  3.3.4 Circuit of Round Baler Sequencer
  3.3.5 Clock Signal to Drive Sequencer
  3.3.6 "In Sequence" Indicator Lamp
  3.3.7 Time Delays in Sequence
  3.3.8 Operation of Relays
  3.3.9 Generation of Initiating Signals
  3.3.10 Suppression of Electrical Noise
  3.3.11 External Start Signal
  3.3.12 Detection of Stall Current

3.4 Power Supply and Suppression of Electrical Interference

  3.4.1 Overview
  3.4.2 Power Source
  3.4.3 Common Earthing Point
  3.4.4 Printed Circuit Board Tracks and Layout
  3.4.5 Transient Suppression
  3.4.6 Radio Frequency Interference
  3.4.7 Protection Against Battery Reversal and Negative Transients
  3.4.8 Voltage Regulator
  3.4.9 Decoupling of Supply Rails
  3.4.10 Voltage Doubler Circuit
  3.4.11 Code of Good Design Practice

3.5 Conclusion

4 ANALOG CIRCUIT

4.1 Introduction

4.2 Speed Measurement

  4.2.1 Overview
  4.2.2 Speed Measurement Principle
  4.2.3 Implications of Speed Measurement Principle
  4.2.4 Duration of Inductive Transient
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.2.5</td>
<td>Window Circuit to Measure Back E.M.F.</td>
<td>31</td>
</tr>
<tr>
<td>4.2.6</td>
<td>Interface Circuit for Motor Voltage</td>
<td>32</td>
</tr>
<tr>
<td>4.2.7</td>
<td>Sample and Hold Circuit</td>
<td>33</td>
</tr>
<tr>
<td>4.3</td>
<td>Speed Control</td>
<td></td>
</tr>
<tr>
<td>4.3.1</td>
<td>Overview</td>
<td>35</td>
</tr>
<tr>
<td>4.3.2</td>
<td>Proportional-plus-Integral Controller</td>
<td>35</td>
</tr>
<tr>
<td>4.3.3</td>
<td>Problems with Operational Amplifiers</td>
<td>37</td>
</tr>
<tr>
<td>4.3.4</td>
<td>Pulse Width Modulation</td>
<td>38</td>
</tr>
<tr>
<td>4.3.5</td>
<td>Generation of Modulating Waveform</td>
<td>39</td>
</tr>
<tr>
<td>4.3.6</td>
<td>Operation of Waveform Generator</td>
<td>39</td>
</tr>
<tr>
<td>4.3.7</td>
<td>Maximum On-Time for Duty Cycle</td>
<td>41</td>
</tr>
<tr>
<td>4.3.8</td>
<td>Minimum Off-Time for Duty Cycle</td>
<td>42</td>
</tr>
<tr>
<td>4.3.9</td>
<td>Improvement of Actuator Performance</td>
<td>42</td>
</tr>
<tr>
<td>4.3.10</td>
<td>Comparator</td>
<td>43</td>
</tr>
<tr>
<td>4.3.11</td>
<td>Full Speed Override</td>
<td>43</td>
</tr>
<tr>
<td>4.4</td>
<td>Power Circuit</td>
<td></td>
</tr>
<tr>
<td>4.4.1</td>
<td>Overview</td>
<td>44</td>
</tr>
<tr>
<td>4.4.2</td>
<td>Comparison of Bipolar and Power MOSFET Transistors</td>
<td>44</td>
</tr>
<tr>
<td>4.4.3</td>
<td>Thermal Coefficient of Resistance, and Secondary Breakdown</td>
<td>46</td>
</tr>
<tr>
<td>4.4.4</td>
<td>Circuit Configuration</td>
<td>47</td>
</tr>
<tr>
<td>4.4.5</td>
<td>Drive Circuit for Power MOSFET Transistors</td>
<td>50</td>
</tr>
<tr>
<td>4.4.6</td>
<td>Clamping of Inductive Transient</td>
<td>51</td>
</tr>
<tr>
<td>4.5</td>
<td>Summary</td>
<td></td>
</tr>
</tbody>
</table>

5 CONCLUSIONS FOR ROUND BALER CONTROL

PART B: A PRODUCT IN DEVELOPMENT:

AN ELECTRONIC CONTROL FOR FERTILISER SPRAYING

6 A BETTER METHOD OF FERTILISER SPRAYING

6.1 Statement of Problem

6.2 Specification of Desired Product

7 SYSTEM MODEL AND PERFORMANCE CRITERION

7.1 Introduction

7.2 System Model

7.3 Specification

7.4 Performance Indices

7.5 Summary

8 IMPLEMENTATION OF PERFORMANCE INDEX BY EULER-LAGRANGE EQUATION

8.1 Introduction

8.2 Optimal Tracking Controller

8.3 Performance Index

8.4 Functional Minimisation

8.5 Calculation of Control Law using Euler-Lagrange Equation

8.6 Simulation of Euler-Lagrange Controller

8.7 Summary

9 IMPLEMENTATION OF PERFORMANCE INDEX USING HAMILTON-JACOBI EQUATION

9.1 Introduction

9.2 Derivation of Control Law using Hamilton-Jacobi Equation

9.3 Numerical Determination of Optimal Control
ACKNOWLEDGEMENTS

I express my thanks to the following:

The Governors of Sunderland Polytechnic for the use of the facilities placed at my disposal;

Warner Electric Limited for the use of facilities and for the opportunities offered to me;

John Hargrave of Sunderland Polytechnic for supervision of this dissertation and for boundless enthusiasm;

Chris Cox and George Brown of Sunderland Polytechnic and Fred Wright and Keith Abbot for casting their pearls before the swine;

Audrey Mullen for typing the script.
## LIST OF SYMBOLS

Most symbols are defined in the text as they occur, but these symbols are listed here for reference.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a)</td>
<td>control signal weighting factor</td>
</tr>
<tr>
<td>(A)</td>
<td>matrix of system state coefficients</td>
</tr>
<tr>
<td>(b)</td>
<td>vector of system input coefficients</td>
</tr>
<tr>
<td>(c)</td>
<td>vector of output coefficients</td>
</tr>
<tr>
<td>(D)</td>
<td>observer state coefficient</td>
</tr>
<tr>
<td>(e)</td>
<td>error signal</td>
</tr>
<tr>
<td>(E)</td>
<td>observer input coefficient</td>
</tr>
<tr>
<td>(F)</td>
<td>vector of system state coefficients</td>
</tr>
<tr>
<td>(g)</td>
<td>vector of constraints</td>
</tr>
<tr>
<td>(H)</td>
<td>error measure</td>
</tr>
<tr>
<td>(H_c)</td>
<td>constrained error measure</td>
</tr>
<tr>
<td>(H)</td>
<td>Hamiltonian function</td>
</tr>
<tr>
<td>(i)</td>
<td>current</td>
</tr>
<tr>
<td>(I)</td>
<td>current in d.c. motor</td>
</tr>
<tr>
<td>(J)</td>
<td>performance index</td>
</tr>
<tr>
<td>(J^*)</td>
<td>optimal cost function</td>
</tr>
<tr>
<td>(k_F)</td>
<td>forward path coefficient</td>
</tr>
<tr>
<td>(k_G)</td>
<td>gearing coefficient</td>
</tr>
<tr>
<td>(k_M)</td>
<td>motor coefficient</td>
</tr>
<tr>
<td>(k_{MT})</td>
<td>feedback coefficient of measured state</td>
</tr>
<tr>
<td>(k_O)</td>
<td>feedback coefficient of output state</td>
</tr>
<tr>
<td>(k_T)</td>
<td>tachogenerator coefficient</td>
</tr>
<tr>
<td>(k_V)</td>
<td>valve coefficient</td>
</tr>
<tr>
<td>(L_M)</td>
<td>Inductance of coil of d.c. motor</td>
</tr>
<tr>
<td>(p)</td>
<td>observer/system speed ratio</td>
</tr>
<tr>
<td>(p_{xx})</td>
<td>elements of (P) matrix</td>
</tr>
<tr>
<td>(P)</td>
<td>matrix of minimum cost coefficients</td>
</tr>
<tr>
<td>(Q)</td>
<td>matrix of state weighting coefficients</td>
</tr>
<tr>
<td>(r)</td>
<td>control signal weighting factor</td>
</tr>
<tr>
<td>(R_M)</td>
<td>resistance of coil of d.c. motor</td>
</tr>
<tr>
<td>(s)</td>
<td>Laplace operator</td>
</tr>
<tr>
<td>(T)</td>
<td>vector of system state coefficients</td>
</tr>
</tbody>
</table>
\( u \)  control signal
\( V \)  voltage
\( V_M \)  voltage across a d.c. motor
\( x \)  vector of system states
\( x_D \)  desired output flowrate
\( x_M \)  measured output flowrate
\( x_0 \)  actual output flowrate
\( \hat{x}_0 \)  system output state estimated by observer
\( y \)  position of linear actuator extension tube
\( \lambda \) (lambda)  vector of Lagrange multipliers
\( \lambda_1, \lambda_2 \)  system eigenvalues
\( \lambda_0 \)  observer eigenvalue
\( \xi (\xi) \)  observer state
\( \tau (\tau) \)  time constant
\( \omega (\omega) \)  angular velocity of d.c. motor
### LIST OF FIGURES

<table>
<thead>
<tr>
<th>Fig.</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Round Baler in Action</td>
</tr>
<tr>
<td>2</td>
<td>Twining Mechanism</td>
</tr>
<tr>
<td>3</td>
<td>Operation of Twiner</td>
</tr>
<tr>
<td>4</td>
<td>Circuit Diagram of Round Baler Control</td>
</tr>
<tr>
<td>5</td>
<td>Temperature Range of Logic Families</td>
</tr>
<tr>
<td>6</td>
<td>Schematic of Versatile Logic Sequencer</td>
</tr>
<tr>
<td>7</td>
<td>Timing Diagram for Versatile Logic Sequencer</td>
</tr>
<tr>
<td>8</td>
<td>Sequence Table for Versatile Logic Sequencer</td>
</tr>
<tr>
<td>9</td>
<td>Sequence Table for Round Baler Control</td>
</tr>
<tr>
<td>10</td>
<td>Astable Multivibrator</td>
</tr>
<tr>
<td>11</td>
<td>Waveforms of Astable Multivibrator</td>
</tr>
<tr>
<td>12</td>
<td>Current Profile during Automatic Operation</td>
</tr>
<tr>
<td>13</td>
<td>Abuse of Actuator during Manual Control</td>
</tr>
<tr>
<td>14</td>
<td>Generation of &quot;Window&quot; Pulse to Sense Back E.M.F.</td>
</tr>
<tr>
<td>15</td>
<td>Proportional plus Integral Controller</td>
</tr>
<tr>
<td>16</td>
<td>Pulse Width Modulation</td>
</tr>
<tr>
<td>17</td>
<td>Limiting Duty Cycle to 90%</td>
</tr>
<tr>
<td>18</td>
<td>Switching Circuit with Bipolar Power Transistor</td>
</tr>
<tr>
<td>19</td>
<td>Switching Circuit with Darlington Power Transistor</td>
</tr>
<tr>
<td>20</td>
<td>Switching Circuit with Power MOSFET Transistor</td>
</tr>
<tr>
<td>21</td>
<td>Configuration of Power Circuit</td>
</tr>
<tr>
<td>22</td>
<td>Clamping of Inductive Transients</td>
</tr>
<tr>
<td>23</td>
<td>Fertiliser Sprayer in Action</td>
</tr>
<tr>
<td>24</td>
<td>Open Loop Control Scheme</td>
</tr>
<tr>
<td>25</td>
<td>Closed Loop Control Scheme</td>
</tr>
<tr>
<td>26</td>
<td>Closed Loop Block Diagram</td>
</tr>
<tr>
<td>27</td>
<td>Cutaway View of Linear Actuator</td>
</tr>
<tr>
<td>28</td>
<td>Simulation of Euler-Lagrange Controller</td>
</tr>
<tr>
<td>29</td>
<td>Simulation of Sprayer System</td>
</tr>
<tr>
<td>30</td>
<td>Analog Computer Diagram of Euler-Lagrange Controller</td>
</tr>
<tr>
<td>31</td>
<td>Analog Computer Diagram of Sprayer System</td>
</tr>
<tr>
<td>32</td>
<td>Newton - Raphson Method of Finding Roots</td>
</tr>
<tr>
<td>33</td>
<td>State Feedback using Estimated State from Luenberger Observer</td>
</tr>
<tr>
<td>34</td>
<td>Analog Computer Diagram of State Feedback Control System</td>
</tr>
<tr>
<td>35</td>
<td>Variation in Performance Index with Feedback Coefficients</td>
</tr>
</tbody>
</table>
Fig. 36  Global Optimum using Output State Feedback
Fig. 37  Local Optimum using Measureable State Feedback
Fig. 38  Analog Computer Diagram of Luenberger Observer
Fig. 39  Performance of Luenberger Observer
Fig. 40  Luenberger Observer with Measurement Noise
Fig. 41  Feedback with Noisy Estimated State
Fig. 42  Non-Linearity of Ball Valve
Fig. 43  Analog Computer Diagram to Simulate Non-Linearity
Fig. 44  Response of System with Non-Linearity
CHAPTER 1

INTRODUCTION

This document will examine and compare two controls for agricultural use. The first control is for wrapping up round bales with twine. The control was field tested in 1983 and production commenced in 1984, so it has an established track record. The project was developed on a very short time scale to meet deadlines, so much of the design was empirical.

The second control is for spraying fertiliser to give constant coverage. At this stage the control is only a design proposal and the work presented here was to assess the feasibility or otherwise of that proposal. The design was undertaken on a much more fundamental basis than the empiricism of the round baler control.

The round baler design is largely presented in a tutorial style in the hope that it will benefit others designing automotive/agricultural controls. Particular emphasis is laid on the sequencer circuit, the power supply, transient suppression, speed measurement, the use of power MOSFET transistors, and a new method of clipping inductive transients.

The desired performance of the fertiliser sprayer was defined mathematically by a performance index. Functional minimisation by various techniques generated a controller. It is important to note that the controller was specified by the mathematics on the basis of the desired performance index. This is in contrast to the round baler speed control, where the form of the controller was specified by the designer on the basis of previous experience.

The two controls will now be discussed separately.
PART A:
AN ELECTRONIC PRODUCT FOR FARM APPLICATIONS:
THE WARNER ELECTRIC ROUND BALER CONTROL.

CHAPTER 2
TYING TWINE AROUND BALE

2.1 Statement of Problem

Round bales or big bales are now a common sight, and they are being used for all sorts of crops. Dairy farmers bale green grass and put it into polythene sacks to mature into silage. They feed their cattle in the winter by unrolling the bales down a ramp in the cow shed. Hay and straw can be baled with a highly compressed outer shell which resists the elements. The bales do not need to be stacked and can be stored in the open. American farmers bale cotton and maize stalks for fuel. The big bales can easily be fork-lifted into furnaces. These examples show that farmers are looking for efficient, labour-saving and economical solutions to their problems.

Round baler machines (fig 1) are drawn by tractor and powered through the power take-off. As the bale is drawn across the cut crop, fingers pick it up and carry it to the baling chamber where belts or rollers rotate the growing bale. When it reaches full size the tractor stops and twine is fed across the rotating bale to produce a long spiral wrap. Extra wraps are used at each end to secure it. The twine is cut and the bale can be ejected.

The twine is normally fed across the bale through a tube (fig 2) which is moved either by a piece of string or by a hydraulic cylinder. Tugging at a string or tweaking a joystick gives poor results because the twine can not be put on evenly. Multiple passes may be needed to ensure that the bale is secure, taking extra time and using more of the expensive twine than necessary.
A solution envisaged by Warner Electric European Marketing was firstly to replace the hydraulic cylinder with a Warner Electric Linear Actuator. This device used a heavy duty 12 volt d.c. motor to rotate a ballscrew through a geartrain. A nut on the ballscrew converted the rotary motion to linear motion, pushing a ram in and out of the cover tube. The actuator was similar in size and load capacity to a hydraulic cylinder but offered the advantage of electrical operation.

Secondly, they suggested that an electronic control should be developed to perform the twining function automatically. This would give consistent twining on every bale. Additional features could readily be incorporated in an electronic control, such as a buzzer to indicate a full bale, or steering lights to warn of a tapered bale. These suggestions were formalised into a specification.

2.2 SPECIFICATION OF CONTROL

The twining mechanism devised by each baler manufacturer is different, but they all tend to incorporate the same features. The sequence of operation is shown in fig 3.

Initially the actuator is retracted and the twine tube is in the parked position. The twining cycle is started either by pushing the start button on the front of the control or by closure of an external switch.

The actuator extends at full speed, sweeping the twine tube across the bale, until it reaches its end stop.

The actuator pauses at its end stop for a period between 6 and 15 seconds, adjustable on the front panel of the control. During this time, the twine catches around the bale and several wraps are put on the end.

Then the actuator starts to retract slowly under speed control. The required speed range is 10% - 100% of the full speed, adjustable on the front panel of the control. This slow retraction puts a long spiral wrap around the bale. The pitch of the wrap can be adjusted to suit the crop, whether it is
precision chopped grass for silage, or maize stalks for burning in furnaces. If the driver sees that the twine has not caught around the bale, then pressing the start button again at this stage will send the actuator out again to the end of stroke, for a further period of 6 to 15 seconds. The actuator brings the twine tube across to the other side of the bale where it strikes a microswitch. This signal stops the actuator for a further period of 6 to 15 seconds, and the second end of the bale receives several wraps during this time. Then the actuator retracts again at full speed. The twine is drawn across a blade and severed. The actuator reaches the end of stroke and the cycle ceases.

This specification was presented to the U K laboratory of Warner Electric Ltd. Some further suggestions were made. For example back e.m.f. sensing was proposed for the measurement of actuator speed. To detect the ends of stroke of the actuator, the recommendation was to monitor the actuator current. A prolonged rise to a high level would indicate that the actuator had stalled at the end of stroke. But apart from this, the responsibility for the design and development of the control was left in the hands of the author.
CHAPTER 3
DIGITAL CIRCUIT OF ROUND BALER CONTROL

3.1 INTRODUCTION
This chapter deals with the predominantly digital aspects of the Round Baler Control. Firstly the choice of logic family is discussed, and it is seen that environmental and economic considerations dictate the type used.

The digital circuit which controls the twining sequence is described and guidelines are given for designing similar circuits. Emphasis is laid on suppressing electrical interference, which could cause erratic operation.

The power supply is also described in this section because it is so important for reliability in the digital circuit. Some of the subtler points used to combat the harsh automotive environment are explained in detail. Finally, a Code of Good Design Practice is given.

3.2 SELECTION OF LOGIC FAMILY
The specification called for a digital circuit to implement the sequence of extension, slow retraction and full speed retraction of the actuator. Various techniques were available for the sequence, including microprocessors, EPROM-based logic, and the TTL or CMOS circuits.

Microprocessors seemed over-powerful for this application but a simple EPROM-based sequential circuit could easily do the job, while offering a degree of flexibility should modifications or enhancements be necessary at a later date. However both of these technologies and also TTL were ruled out on environmental/economic grounds. Standard commercial grade CMOS circuits can operate over a range -40°C to +85°C, while the upper temperature limit of the other technologies is only 70°C (see fig 5). The specified maximum operating temperature for the Round Baler Control was 70°C but tests with the
Analog circuits had shown that dissipation from the power transistors would raise internal temperature by 10-15°C above this level. While automotive and military grades of circuits cover temperatures up to 85°C and 125°C respectively, these might have been difficult to obtain, and they would have been relatively expensive. Standard CMOS circuits seemed ideal for the job for these reasons.

An additional advantage of using CMOS parts was the low supply current. This gave negligible power dissipation in the logic circuits and also in the voltage regulator, so no further temperature increase occurred inside the control box.

Furthermore, CMOS circuits can operate over a wide voltage range, from 3 to 15 volts, which offered the designer a great deal of flexibility. The best supply voltage for the analog circuit was 8 volts, so the CMOS logic circuit was operated from this rail also.

The total supply current to the analog and CMOS circuits was about 20 milliamps, so the circuit dissipated a negligible 160 milliwatts. The voltage regulator dissipated 80 milliwatts so a heatsink for this was unnecessary.

In fact the reasons for choosing CMOS logic were quite overwhelming, mainly due to the stringent temperature specification and the high dissipation of the power transistors. CMOS logic is not the universal choice for agricultural controls. Every case should be assessed on its merits.

3.3 SEQUENCER CIRCUIT

3.3.1 Overview

A new design method is described for sequencer circuits. A simple example is used to show how a Sequence Table can be directly implemented using logic elements. Then the Sequence Table for the Round Baler Control is given and its circuit is described. Time delays and the operation of relays and lamps
are included. Initiating signals are fed into the circuit, but protection against electrical interference is vital. The stall current of the Linear Actuator gives a small signal which must be amplified in the presence of much electrical noise.

3.3.2 Design Method for Sequencer Circuit

Numerous design methods have been published by academics where the criterion of excellence is the minimisation of the number of logic gates of flip-flops required. These methods are valuable in that a solution can always be derived for a given problem. However, current engineering practice draws heavily on the large body of higher level circuits referred to as MSI (Medium Scale Integration). These devices include such types as counters and shift registers and offer greater complexity at a fraction of the cost, number of integrated circuits, board area and design effort. Designs employing MSI are frequently easier to comprehend than multi-gate equivalents, an important consideration for faultfinding and servicing.

For the problem in question, the versatile clocked digital sequencer shown in fig 6 was a good solution. The Johnson Counter has the property that only one of its outputs can be on at any one time. Clock pulses cause the active output to step onwards from one stage to the next. After the final stage, the first stage is energised again, (hence the alternative name of ring counter). The Count Inhibit input blocks the clock pulses. The Reset input causes the counter to return to its first stage.

A timing diagram (fig 7) is useful to explain the operation of the sequencer circuit. The Johnson Counter is clocked on a regular basis but a NOR gate activates the Count Inhibit input. One stage of the counter (say stage 0) is energised. This active stage can operate on output W. When input A is turned on, the blocking signal on the Count Inhibit input is removed and the next rising edge of the Clock signal causes the counter to increment to
stage 1. Output W is de-energised and output X is energised in its place. Input A is now ignored and stage 1 is maintained until input B is turned on. A complete sequence can easily be "programmed" using this approach. Unused stages (such as 4, 5, 6, and 7) remove Count Inhibit directly, so that the counter steps through them in 4 clock pulses, returning to stage 0.

When power is first applied to the circuit, a short pulse is applied to the Reset input, so that the sequence always starts from stage 0. The sequence 'can also be reset to stage 0 from mid-sequence, should a fault condition or a manual over-ride occur. This is shown in stage 2, where an input R can cause an immediate asynchronous reset to stage 0.

3.3.3 Sequence Table

The operation of the sequencer in fig 6 can be written down in the form of a Sequence Table (fig 8). It is even more apparent from this how closely the circuit represents the function to be performed. Technicians and even apprentices have been able to modify such circuits without difficulty to accomplish similar or completely different functions.

The detailed Sequence Table for the Round Baler Control is shown in fig 9. Relay 1 is used to extend the actuator and Relay 2 is used to retract the actuator under speed control. The preset retraction speed is over-ridden by the "High Speed" output, so that the actuator retracts at full speed. (All these outputs are determined by the analog and power circuits which are described elsewhere).

To initiate time delays, the Timer Enable output removes the reset command from a clocked binary counter circuit. At the start of the timed period, the counter is empty but after $2^N$ clock pulses, an output is given to signal the end of the timed period.
The manual start button can be used either to start the timing sequence, or to re-initiate the sequence if twine is not feeding out around the bale. However, the External Start input can only be used to start the sequence and it has no effect while in cycle. To differentiate between the two, the sequencer generates another output, the External Start Disable, which blocks out the External Start input while in cycle.

A variety of signals are used to advance from one stage to another. The Start Switch, the External Start signal and the End of Timed Period signal have already been discussed. A signal is also obtained from the Stall Current of the Motor, which indicates the end of stroke of the linear actuator. A signal is also obtained from a Microswitch attached to the mechanism. This is used to halt the actuator temporarily at a certain part of the cycle.

During stage 3, a fault condition may occur. The actuator is retracting slowly under speed control and an obstruction such as a blockage in the intake area of the baler may stop the twine tube and cause the actuator motor to stall. The stall current generates a signal which resets the sequencer.

Also during stage 3, the operator may observe that the twine has not caught and is not feeding out around the bale in a long spiral wrap. Pressing the Start button again at this time causes the sequencer to reset immediately. The sequence does not terminate here because at the next clock pulse (within about 1 millisecond) the still-closed Start button causes the cycle to start again. The actuator extends again to the end of stroke so the twine has another opportunity to catch around the bale.

Both of these signals are combined by a separate logic circuit so that if either the actuator stalls or the Start button is pressed, the sequence is reset to stage 0.
The Sequence Table defines the operation of the sequencer in every detail. Referring to the circuit diagram (fig 4) is almost an anti-climax because it reveals little more than the Sequence Table has already told us. The ease of transferring the specifications from the Sequence Table to the circuit diagram is obvious. However, some details of the circuit diagram may be of interest.

3.3.4 Circuit of Round Baler Sequencer

IC5 is a CD4022B chip, an 8 stage Johnson Counter. (Motorola or RCA are specified as suppliers of this device, because the Clock input incorporates a Schmitt trigger. Other manufacturers such as Signetic/Mullard do not include the Schmitt input which leads to problems).

When power is first applied to the circuit, the sequencer is initialised to stage 0 by a pulse applied to the Reset line. A pulse of about 1 second duration is generated by the charging of capacitor C14 by resistor R17. The pulse is "squared-up" by a Schmitt trigger before application to the Reset line.

To guard against momentary losses of power, R17 is paralleled by diode D6. Loss of voltage on the 8 volt logic supply immediately discharges capacitor C14 via diode D6, so that when power is resumed, a reset pulse of 1 second duration is given again.

IC9 is a 4078 chip, an 8-input NOR gate which controls the inhibit input of the Johnson Counter. The incrementing signals are gated by IC7 and IC8 which are quad 2-input AND gates with part number 4081.

3.3.5 Clock Signal to Drive Sequencer

The Clock Signal to the Johnson Counter is provided by an astable multivibrator based on an analog operational amplifier (fig 10). The mode of
operation is shown in Fig. 11. If we assume that the output 2/14 of the amplifier is high, then the non-inverting input 2/12 is pulled up to the positive threshold by positive feedback via resistor R15. The capacitor C13 charges up exponentially via resistor R16 and the potentiometer P1. When the voltage it feeds to the inverting input 2/13 of the amplifier reaches the positive threshold level, the output abruptly switches to low level, aided by positive feedback via R15. The non-inverting input is now at the negative threshold, and capacitor C13 discharges exponentially via R16 and P1. When the voltage it feeds to the inverting input of the amplifier reaches the negative threshold level, the output switches back to the high level.

For an ideal operational amplifier, the on/off times and the frequency can be readily calculated. For a real operational amplifier, the task is complicated by the indeterminate output voltage levels. The low output voltage level is within 200mV of the negative rail but the high output level can be as much as 1.5V below the positive rail. These indeterminate voltage levels vary the thresholds and the rates of charge and discharge of the timing capacitor.

But for a given operational amplifier, the frequency is inversely proportional to both the timing resistance $R_T$ and the timing capacitance $C_T$, so

$$f \propto \frac{1}{R_T C_T}.$$ 

In the circuit, the timing resistor is a combination of a fixed resistor and a variable resistor, to allow the frequency to be adjusted. The reason for this will be explained later.

3.3.6 "In Sequence" Indicator Lamp

Stage 0 of the Johnson Counter represents the inactive stage of the cycle. The cycle awaits a start command either from the Start button or the External
Start signal. The stage 0 signal is inverted so that the red lamp stays off and the External Start input is enabled. After a start command, the red lamp comes on for the duration of the sequence, the External Start input is disabled, and the sequencer steps on to stage 1. At stage 1, relay 1 is activated to extend the actuator at full speed until the end of stroke is reached and the stall current causes the sequencer to advance to stage 2.

3.3.7 Time Delays in Sequence

Stage 2 represents a time delay between 6 and 15 seconds, while a number of wraps are placed around the end of the bale. The time delay is produced by removing the Reset signal from IC10, a CD4020B chip, which is a 14 stage binary counter. After 16,384 clock pulses, the output of the counter goes high and causes the sequencer to step on to stage 3. The clock signal used is exactly the same as that which clocks the sequencer chip IC5. Varying the clock period from 0.37 ms to 0.92 ms has no effect on the operation of the sequencer but after division by the counter chip, it produces the variable delay of 6 to 15 seconds.

3.3.8 Operation of Relays

At stage 3, relay 2 is operated to retract the actuator under speed control. A Darlington Driver chip, IC6, is used to operate the relay. Its output stage is a common emitter transistor, so that output 6/15 can be wire-ORed with output 6/14. This means that both stage 3 and stage 6 can operate relay 2, as shown in the sequence table.

Stage 3 is normally terminated by the operation of the microswitch. However, overcurrent in the actuator or a manual override on the Start button are required to cause a reset. These two signals are combined with an OR gate and fed to the sequencer Reset line during stage 3.
Stages 4 and 7 are unused, and the sequencer steps on to the next stage at the next clock pulse. (Stage 4 was originally used for another function but this was dropped after the project had been costed).

Stage 5 is another time delay, identical to stage 2.

Stage 6 operates relay 2 via the Darlington Driver chip IC6 connected in a wired-OR configuration. Stage 6 also feeds a signal into the analog circuit which overrides the speed preset on the control knob and instead demands maximum speed. The actuator retracts at full speed until it reaches the end of stroke where the stall current causes the sequencer to step onto stage 7. The arrival of the next clock pulse causes the sequencer to step on to stage 0, where the red light goes off, signalling the end of the twining cycle. The sequencer waits at this stage until another start command is given.

Having examined the operation of the sequencer, we will now look at the generation of the actuating signals.

3.3.9 Generation of Initiating Signals
Consider firstly the input from the microswitch. When the microswitch operates, a logic high is to be produced on pin 6 of integrated circuit 8 to increment the sequencer. There are some very important practical considerations here.

The microswitch is mounted on the baler about 3 metres away from the control box. To save wire and to minimise the number of ways in any connectors, it is desirable that only a single wire should go out to the microswitch. When the microswitch operates it shorts this wire down to chassis.

3.3.10 Suppression of Electrical Noise
This long wire going out to the microswitch acts as an extremely effective
aerial. Sources of electrical noise such as alternators, spark plugs, solenoids and wiper motors create interference at frequencies between 500 Hz and 500 MHz. Furthermore, modern tractor cabs are frequently adorned with C.B. radio equipment which further contributes to the unwanted signals being picked up by this wire. It is vital that the electrical noise present on this input should be filtered out to prevent spurious operation.

The circuit diagram shows how these requirements are satisfied. One contact of the microswitch is connected to the chassis and the other contact goes to the control. A pull-up resistor R6 is used to ensure that clearly defined voltages of 0 volts or 8 volts are presented to the control; any noise pulse must then induce an amplitude of at least 4 volts to cause a transition. If the pull-up resistor R6 was not used and the input was allowed to float at an undefined level, noise pulses with a lower voltage amplitude could cause signal transitions. The pull-up resistor also reduces the input resistance from typically 100 Mohm down to 10 Kohm. Magnetic induction can create electrical current in the input circuit. For example, if a current of 10 microAmp is induced in the circuit, a negligible voltage of 100 millivolt is produced in the circuit with the pull-up resistor. Without the pull-up resistor, a large voltage would result which would cause a signal transition.

The input is protected against high frequency noise by an L - C input network. The inductor and capacitor are selected for good performance at high frequencies. The ferrite bead inductor L2 rejects all high frequency signals, and any signal which gets through it is decoupled to ground by polyester capacitor C4. Note that C4 is connected directly to a point referred to as the "Common Earthing Point". The importance of this would soon be realised if it was connected elsewhere, such as to the pin 7 of integrated circuit 12. An induced current pulse would then pass down the zero volt supply line to all the analog and digital circuits. While the resistance of this line is
very small, its impedance is not, and a substantial voltage spike would be produced, "scrambling" the logic and producing glitches in the operational amplifiers. So besides selecting suitable components, the circuit designer must also carefully consider the circuit layout to ensure adequate noise rejection.

The L - C network is followed by two resistor-capacitor networks. Resistor R5 and capacitor C8 attenuate medium frequencies while resistor R7 and C6 attenuate lower frequencies. The net effect is that any signal with a duration shorter than 100 milliseconds will not be detected by the circuit. This heavily filtered signal has a very slow rise-time and to obtain a more acceptable logic signal, it is fed through a Schmitt trigger which "squares up" the waveform.

The Microswitch signal is slightly delayed in its passage through the filter circuit, but this has no significant effect on the operation of the system. Spurious operation by noise signals is completely prevented.

3.3.11 External Start Signal

The External Start signal is interfaced in the same way as the Microswitch signal. A long single wire connector is used so an inductor-capacitor network is used to filter out radio frequency signals. Two cascaded resistor capacitor networks are used to attenuate medium and low frequencies, so that only a valid External Start signal can initiate the twining sequence.

A diode gate consisting of D1 and D2 is used to block out the External Start signal after the sequence has started.

The sequence can also be initiated by a push-button on the front of the control. This input by-passes the inductor-capacitor network and is fed
directly into the resistor-capacitor networks. This is because the Start button and its associated wiring are contained within the metal enclosure of the control in a shielded environment which is relatively free of electromagnetic radiation.

3.3.12 Detection of Stall Current

Another initiating signal is provided by the stall current of the linear actuator. A trip level of 21 Amps was chosen because the full range of linear actuators have stall currents in excess of this. The motor current is sensed by a shunt resistor R42. The value of the shunt resistor was selected to be 0.01 ohms, a very small value indeed. This gave the advantage that the power dissipation was very small, allowing use of a single 7 watt resistor. However it introduced complications which required some care to circumvent. Firstly, low ohmic resistors such as these have a resistance which is comparable with that of their own leads. The manufacturers state that the resistance is only within tolerance when measured at a specific distance along the leads. To guarantee that the correct lead length is used, the girls assembling the boards fit the resistors with insulating fishtail beads, so the resistor is spaced off the board at exactly the right lead length.

The shunt resistor is also comparable in resistance to the tinned copper track on the printed circuit board. The track resistance varies with temperature and also because of self-heating. It is important therefore that a four wire connection is used. This means that two thick copper tracks carry the 30 amps of actuator current to and from the shunt resistor, while two thin tracks carrying about 1 microamp sense the voltage across it. Careful layout of the printed circuit board ensures that each thin track and corresponding thick track meet at only one point, the lead of the shunt resistor. This layout is shown in schematic form in the circuit diagram.
One end of the shunt resistor is connected to the battery, and the other end is connected to the linear actuator. Both leads are quite long and may carry large amounts of electrical noise. To keep high frequency noise out of the control, an inductor-capacitor-inductor filter made up of L6, C22 and L7 is used. Some filtering of medium frequencies is provided by the resistor-capacitor network R45 and C23.

The signal level provided by the shunt resistor is quite small: 30 amps only yields 0.3 volts. An amplifier has to be used to boost the signal to a usable level. A non-inverting amplifier output 2/7 is used, with its gain set to a factor of 23 by the resistors R46 and R47. A capacitor C24 is used to limit the amplification of higher frequencies to unity. After amplification, a current level of 21 Amps produces a signal level of 4.8 volts. The amplified signal is fed into a comparator input. Positive feedback through resistor R50 is used to promote clean switching. The positive threshold is 4.8 volts, but when the input signal exceeds this level, the output goes low and the threshold drops to 3.2 volts.

A further stage of filtering by a resistor-capacitor network R51 and C25 is used, and the signal is finally squared up by a Schmitt trigger with output 4/6. These successive stages of filtering may seem excessive, but they are essential to overcome two big problems; firstly the high frequency noise present on the battery and actuator lines, and secondly the starting transient of the actuator. When the actuator is first switched on, a high current flows which is equal in magnitude to the stall current. After about 100 milliseconds, the motor accelerates up to speed and the high starting current falls to the lower running current. If the actuator is heavily loaded, the starting current may linger for up to 200 milliseconds before the motor reaches its normal running speed. Clearly the starting transient can not be distinguished from the stall current purely by magnitude. It is necessary
also to assess the duration. The circuit uses two first order lags with time constants of 200 milliseconds. Starting transients are too short to propagate through the circuit, and only a genuine sustained stall current will give an output.

3.4 POWER SUPPLY AND SUPPRESSION OF ELECTRICAL NOISE

3.4.1 Overview

The noisy and unstable supply from the tractor battery must be filtered and regulated to feed the logic and linear circuits. The importance of correct earthing is stressed and details are given of the separate low current and high current earth tracks in the control. Protection against transients, supply reversal and radio frequency interference is discussed. Regulation and decoupling of the supply rail is explained. An auxiliary voltage is generated using a voltage doubler circuit. Finally a Code of Good Design Practice is given.

3.4.2 Power Source

The twining system is energised by the tractor battery which is charged by an alternator. The voltage is nominally 12V but this may vary between 10.6V and 14.4V depending on the age and state of charge of the battery. A large amount of electrical noise is present due to the alternator and the other electrical accessories on the vehicle.

From this unstable and noisy source, it is essential to generate a steady supply for the analog and digital circuits which are used in the control. Some cunning design is necessary to achieve this.

It must first be recognised that the negative supply rail is not a zero-resistance link between all parts of the circuit. Standing currents or current impulses acting on the distributed impedance of the negative rail may
give significant voltage differences between parts of the circuit which should be separated into sensitive low current and more robust high current areas. By routing the currents in these areas independently, the effect of high currents on sensitive circuits can be minimised. The low current and high current negative rails must be joined at only one point to prevent interaction.

The circuit diagram shows clearly how the current paths are segregated. The line called BAT- is the negative line from the battery and carries a high current. It is routed through a current sensing resistor R42 to the linear actuator, and on the printed circuit board is made as short as possible.

The negative connections of the operational amplifier (IC1 and IC2) and the digital circuits (IC3 - 5 and IC7 - 12) are all joined to a common zero volt rail. The current in this rail is small at about 20 milliamps and relatively steady, so there is no interaction between the integrated circuits.

3.4.3 Common Earthing Point

The BAT- line and the zero volt rail are joined at a point in the circuit denoted by "Common Earthing Point". Several other devices in the circuit are also connected directly to this point. The negative connection of integrated circuit IC6 is routed directly to it, instead of a straightforward link to the zero volt rail. This is because IC6 is an array of seven Darlington transistors which operate the two relays and the indicator lamp. Each relay takes 150 milliamps and the incandescent lamp takes 180 milliamps with an inrush current of perhaps 400 milliamps at switch-on. Injecting a steady current of this magnitude into the zero volt rail would certainly upset the sensitive operational amplifiers. But the supply to the lamp and the relays comes from the noisy positive battery rail, and the current passing through IC6 is also noisy as a result. Injecting this noisy current into the zero volt rail could interfere with the digital integrated circuits also.
The earth connection of the voltage regulator VR1 is also returned directly to the common earthing point, as are the two decoupling capacitors C28 and C29 which are associated with it. This ensures that the 8 volt supply to the integrated circuits is stable and noise-free.

The inputs from the microswitch and the External Start are both filtered by inductor-capacitor networks. Due to the high noise currents induced in these inputs, the capacitors are returned directly to the common earthing point. The noise current therefore has no effect on the integrated circuits.

Such detailed discussion of a bit of wire and a few copper tracks on a printed circuit board may seem tedious. However, it is attention to detail on this scale which can spell the difference between success and failure of an electronic control in the field.

3.4.4 Printed Circuit Board Tracks and Layout

On the circuit diagram, heavy current tracks which supply the linear actuator are shown in bold. P.C.B. tracks are sized according to the r.m.s. current and maximum acceptable temperature rise in the track. In this circuit the current does not maintain a steady value, and a typical current profile during automatic operation is shown in fig 12. Some consideration must also be given to abuse, either accidental or deliberate, during manual operation (fig 13). If the actuator is ran to its end stop and held there under power, a high stall current flows. After a period of 15 - 30 seconds, an automatically resetting thermal trip built into the motor of the linear actuator will operate, shutting off the current. After cooling down for a few seconds, the thermal trip will reset and pass current again. This heating/cooling limit cycle will be maintained as long as the manual switch is operated.

A conservative design would size the track on the basis of the full stall
current of 30 amps with a track temperature rise of 20°C. This gives an unrealistic track width of 1 inch using 2 oz. per square foot copper foil. (This thickness of copper foil is unsuitable for fine tracks because the etching solution tends to undercut the tracks giving poor track adhesion to the fibreglass substrate). After much thought, an estimated current rating of 18 Amps was coupled with a track temperature rise of 45°C to give a track width of 0.25 inch using standard 1oz. per square foot copper foil. The estimated current rating was based on the full stall current with some consideration given to the estimated duty cycle.

The figures given refer to copper foil only. In fact copper tracks are electro-tinned during manufacture and are coated with solder during the flow solder process. The track is therefore much thicker and temperature rises are lower than anticipated.

3.4.5 Transient Suppression

The noise present on the battery lines must be filtered out to ensure reliable operation of the circuit. This filtering is performed in a number of stages. The first line of protection against gross overvoltage is a voltage dependent resistor type V22ZA3. This device clamps low and medium energy transients to an acceptable voltage level, and gives excellent protection under most conditions.

However two situations can occur in automotive systems where very high energy transients are produced. Field Decay transients can occur when the starter motor is cranking over the tractor engine. Should one of the battery leads become disconnected, then the inductive energy in the starter motor produces a large negative spike. Papers published by the Society of Automotive Engineers (in America) say that voltages of -120V with a source impedance of only 1 ohm can occur, decaying with a time constant of 200 milliseconds.
Load Dump transients can occur when the tractor engine is running and the alternator is charging the battery. Should one of the battery leads become disconnected, then the inductive energy in the alternator produces a large positive spike. The S.A.E. describe transients of 150V with a source impedance of 1 ohm, decaying with a time constant of up to 4 seconds. (ref. 1)

Both of these transients will cause a current surge through the voltage dependent resistor which may blow the 25 Amp fuse. The v.d.r. will clamp Load Dump transients to a level of 60 volts; this may damage the voltage regulator which is only rated to handle 40 volts. The high instantaneous power dissipation may rupture the package of the voltage dependent resistor and the unclamped spike will rise to its full level of 150 volts. This would certainly break down the voltage regulator and destroy all the integrated circuits on the board.

Clearly the control is not adequately protected against Load Dump transients. Fortunately these transients do not occur very often, and the voltage dependent resistor can guard against the more common minor transients. Hopefully, better voltage dependent resistors will become available in the future and give an answer to this vexing problem.

3.4.6 Radio Frequency Interference

Within the band of voltage passed by the voltage dependent resistor, radio frequency interference may be present. An inductor-capacitor-inductor filter made up of L4, C26 and L5 is used to prevent this interference from passing into the integrated circuits. Inductor L5 may seem superfluous but in fact it is quite important. It effectively blocks radio frequency noise from the zero volt line to the integrated circuits. Without it, strong sources of electromagnetic noise such as 27 MHz C.B. radios could virtually establish a standing wave on the zero volt line. The consequences of this are difficult to predict but they are unlikely to be beneficial.
3.4.7 Protection against Battery Reversal and Negative Transients

Reversal of supply voltage can occur due to mis-connection or because of negative transients. Negative transients also cause momentary power loss during which the control should continue to function. A diode can be used to protect against reverse polarity, and a capacitor can give momentary power storage. In the Round Baler control, the current drawn from the 12 volt rail by the lamp, relays and integrated circuits was 350 milliamps. To maintain power to the circuit during an interruption of 25 milliseconds would have required a storage capacitor of 10,000 microfarads. This component would have measured 8 cm. long by 3 cm. diameter and would have been too large to fit in the case.

The problem was solved by connecting the storage capacitor to only the integrated circuits, where the current drain was only 20 milliamps. This allowed use of a 470 microfarad capacitor measuring only 2 cm. long by 1 cm. diameter which could readily be fitted on the printed circuit board. This capacitor maintains the supply to the integrated circuits during a momentary power loss of up to 25 milliseconds. The lamp and the relays are unpowered during this period, but resume their functions after power is restored.

Diode D14 ensures that the integrated circuits and the relay and lamp circuits cannot be reverse biased. Capacitor C27 provides back-up for the integrated circuits during momentary power drop-outs. Diode D15 blocks the relay and lamp circuits from the back-up capacitor, so that the stored energy is only supplied to the integrated circuits. The positive battery line BAT+ is not protected by a diode, and is connected directly to the power MOSFET transistors TR2, 3 and 4. A reversed connection causes no damage because the relays cannot then be energised, so the transistors are not reverse biased. If a negative transient occurs while the control is in cycle, then the transistors may be reverse biased for up to 30 milliseconds before the relay contact opens, due to the relay's mechanical inertia. The parallel reverse
diode which is an integral part of the power MOSFET transistor can safely handle any reverse current which may occur at such a time; in fact the current rating of the diode is identical to that of the transistor itself.

3.4.8 Voltage Regulator

The voltage supply to the integrated circuits is regulated at a level of 8 volts by a proprietary voltage regulator type LM7808CP. This voltage level is sufficiently high to allow the operational amplifiers an adequate swing in output voltage, allowing for saturation at the upper and lower limit. The input voltage to the regulator must remain above a level of 9.2 volts in order to maintain regulation at a level of 8 volts. This means that the input voltage to the control must stay above 10.6 volts, because 0.6 volts are dropped across each of the diodes D14 and D15. The input voltage may vary over the range of 10.6 volts to 13.4 volts, so even at low input voltages, regulation is still maintained.

These figures may seem to disagree with the manufacturer's data sheet which specifies an "overhead" voltage of 2 volts at a current rating of 1 Amp. But the actual current drawn from the regulator is only 20 milliamps, and the overhead voltage at this current level is in the region of only 1.2 volts. (This figure is not specified by the manufacturers, but has been confirmed by extrapolating their graphs and testing various samples).

Use of the voltage regulator brings various advantages. The output is short circuit protected, so no fuse is necessary for the printed circuit board. To test the board, it is simply connected to a 1 Amp power supply, in the certain knowledge that no damage will result due to solder-bridged tracks. Another benefit is the rejection of input noise, which exceeds 40 dB over the range 10 Hz to 100 kHz.
Capacitors C28 and C29 are connected across the input and output of the regulator to prevent parasitic oscillation and improve transient response.

3.4.9 Decoupling of Supply Rails

The voltage regulator provides a stable noise-free voltage source to supply the integrated circuits. More decoupling is used to ensure that the voltage rails remain noise-free. Capacitor C30 gives bulk decoupling of the rails at low frequencies, although this function is also performed by the regulating ability of the voltage regulator VR1. To guard against transients at higher frequencies, capacitors C31 – 35 decouple the sensitive integrated circuits such as the operational amplifiers IC1 and IC2 and the Johnson Counter chip IC5, and also "unguarded" areas such as the vicinity of IC4 and IC9 on the printed circuit board.

The integrated circuits are not just receivers of electrical noise; they also generate electrical noise when their outputs switch from one level to another. Operational amplifier and CMOS circuits are much better in this respect than TTL circuits which tend to generate "current-gulps" on the negative rail and require even more careful decoupling than this design.

3.4.10 Voltage Doubler Circuit

One final item in the power supply remains to be discussed. The common drain configuration of the power MOSFET transistors Tr 2, 3 and 4 requires a gate voltage of 10 volts above the positive rail to turn them on efficiently. From an input voltage of nominally 12V, an auxiliary rail of 22V must be generated. The current required from this rail is only a few milliamps, so a simple diode pump circuit is adequate, and the expense of an inductor in a switching d.c./d.c. converter is avoided.

An astable multivibrator output 4/8 drives a Darlington amplifier output 6/16 which charges a capacitor C11 up to about 11 volts via diode D5. When the
Darlington amplifier is off, capacitor C11 discharges into capacitor C12 by the path R12, C11, D4, C12, and both capacitors attain a voltage of 5.5 volts. The Darlington amplifier then charges C11 up to a level of 11 volts again. When C11 discharges again, sharing its charge with C12, the voltage on both capacitors becomes 8.2 volts. With each successive charge/discharge cycle, the voltage on storage capacitor C12 is pumped up until it reaches an equilibrium level of 11 volts, and a voltage of 22 volts is available on the auxiliary rail. (In fact, voltage drops across the diodes reduce this level slightly).

Selection of a suitable value for the resistor R12 in the discharge path was difficult. Ideally, this resistor should be low in value, to give a fast efficient discharge. But a low resistance would pass a high current when the Darlington amplifier was on, resulting in excessive power dissipation. This problem was solved by designing an asymmetrical astable multivibrator which drove the Darlington amplifier with a 10% duty cycle. The low resistance path of diode D5 and the Darlington amplifier charged up capacitor C11 fully in only 50 microseconds, while the high resistance path of diode D4 and resistor R12 allowed a relatively sedate discharge in 450 microseconds. The current in resistor R12 was minimised by this scheme and the resulting power dissipation was acceptably low.

3.4.11 Code of Good Design Practice

To summarise, the following points require careful attention for reliable noise-immune operation:

1) Protection against overvoltage and transients.
2) Filtering out radio frequency noise.
3) Protection against polarity reversal.
4) Protection against momentary drop-out of the supply.
5) Regulation and active filtering of low/medium frequency transients.
6) Decoupling of supply rails to integrated circuits.
7) Careful track layout by segregation of high and low current paths with only one common point.

3.5 CONCLUSION

The harsh environmental conditions encountered by Round Balers ruled out the use of several logic families (including microprocessors) but CMOS circuits were suitable. The consequent low power dissipation simplified the power supply.

A simple design technique for sequencer circuits was presented, which allowed the written specification in the Sequence Table to be implemented directly. The sequencer circuit was used to drive lamps and relays, and time delays were included. Input signals were heavily filtered to remove electrical noise.

The power supply was also examined. Emphasis was placed on the importance of correct earthing, and the provision of separate power earths and signal earths. Suppression of electrical noise and protection against other hazards in the automotive environment were also explained. These considerations were summarised in a Code of Good Design Practice.

Now that the digital circuit has been analysed, the analog circuit can be considered.
4.1 INTRODUCTION

Our discussion of the digital circuit has already touched on the functions to be performed by the analog circuit, namely speed control and control of the direction of the actuator. These topics are enlarged upon in this chapter.

Speed measurement is first described. The principle of measuring the speed of a d.c. motor by sampling its back e.m.f. is explained and the implications are discussed. The practical problems encountered and their solutions are given in detail.

Having measured the speed, a speed control using a proportional-plus-integral amplifier is devised, to give zero steady state error in spite of load variations. Power amplification is achieved by Pulse Width Modulation and the practical details of this are discussed.

The power circuit consists of 2 relays and power MOSFET transistors which are needed to control the direction and speed of the linear actuator. Design details are given to explain the circuit configuration. The relative merits of power MOSFET transistors and bipolar transistors are discussed. The paralleling of power MOSFET transistors to increase the power handling is also described. Finally, a new approach to clamping inductive transients is revealed.

4.2 SPEED MEASUREMENT

4.2.1 Overview

The need for speed control of the linear actuator is first discussed. Sensing the speed is difficult because tachogenerators are rejected on various grounds. A new scheme is devised to sample back e.m.f. during momentary
switch-off of the actuator. The scheme is complicated by the inductive transient occurring at the moment of switch-off. A window circuit is designed which gives a gating pulse after the inductive transient has decayed. This pulse operates a sample and hold circuit which samples motor voltage through an interface circuit and stores the back e.m.f. signal. The implications of the speed sensing method are quite profound, and they shape the rest of the analog circuit.

4.2.2 Speed Measurement Principle

In the Round Baler application, speed control of the linear actuator is very important, to give even wrapping of twine around the bale. Furthermore the speed must be adjustable by the farmer to suit the crop being baled. For example, straw can be coarsely wrapped with up to 6 inches between strands of twine. This uses little twine and can be done cheaply and rapidly. Finer materials such as precision-chopped silage must be more closely wrapped to prevent the bale falling apart, and this costs money for the twine and for the extra time taken.

The d.c. motor used in the linear actuator has the usual characteristics of falling speed with increasing load. This means that any open loop scheme for speed control will fail because the varying load on the actuator during the twining cycle will result in a varying speed. A closed loop speed control is essential to give consistent wrapping. This implies that the speed of the d.c. motor must be measured in some way.

Use of a tachogenerator was undesirable because a new non-standard linear actuator would have to be designed. Inventory levels would increase. The re-fit market, where controls would be added to old balers using existing actuators, would be inhibited.

It was realised that the motor of the linear actuator could itself function
as a tachogenerator. The voltage across the motor is

\[ V_M = K_T \omega + L_M \frac{dI}{dt} + IR_M \]

where

- \( V_M \) = motor voltage
- \( K_T \) = tachogenerator constant
- \( \omega \) = motor speed
- \( L_M \) = motor inductance
- \( R_M \) = motor resistance
- \( I \) = motor current.

By reducing the motor current to a steady zero level,

\[ V_M = K_T \omega, \]

i.e. the motor generates a back e.m.f. which is directly proportional to speed.

Clearly under these conditions, the motor will come to a halt. To keep it turning, the supply voltage must be pulsed on and off repeatedly. During the on periods, the motor is energised momentarily. During the off periods, after the current has decayed to zero, the back e.m.f. can be measured.

### 4.2.3 Implications of Speed Measurement Principle

This simple scheme has a large number of implications:

a) Obviously, a switching supply must be used.

b) At some time, the supply to the motor must be switched off to allow measurement of speed. Therefore the torque available from the motor will be reduced. To obtain the highest torque, the sampling time should be as short as possible, and samples should be taken as seldom as possible.

c) After switching off the supply to the motor, the inductive energy stored in the motor field must be dissipated before a valid measurement of the motor back e.m.f. can be made. To agree with implication (b), this should be done as quickly as possible. The motor field has the characteristic

\[ \frac{dI}{dt} = \frac{1}{L_M} V_M. \]
So to reduce the inductive energy and motor current to zero as quickly as possible, the clamping voltage placed across the motor during the off period should be large and negative.

d) Dissipation of the inductive energy will call for a special semiconductor device which can handle the high instantaneous power.

e) The torque available from the motor will be less than that implied by the duty cycle. This is because the current is reduced to zero every cycle, so the average current is less than the figure which might be expected.

4.2.4 Duration of Inductive Transient

Before designing the speed measurement circuit, a power amplifier stage was built to switch the linear actuator on and off, to discover the behaviour of this device. These tests showed that after the linear actuator was switched off, the motor inductance produced a voltage spike, and after this spike had decayed the back e.m.f. could be measured. The duration of the voltage spike depended on the type of linear actuator and the loading upon it.

To discover the maximum duration of the voltage spike, worst case conditions were imposed. A heavy duty actuator with a large field inductance and a high operating current was obtained, and it was ran to the end of its stroke and stalled there. The resulting transient at switch off had a duration of 0.58 milliseconds.

Following the transient, the back e.m.f. could be measured briefly before switching the actuator on again. These measurements therefore defined a "window" when a valid measurement of back e.m.f. could be made. A circuit was designed to provide a pulse during this window. (fig 14).

4.2.5 Window Circuit to Measure Back E.M.F.

The output of IC1 pin 14 is a square wave whose rising edge is coincidental with actuator turn-off under worst conditions. (Under lightly loaded
conditions, actuator turn-off will occur before this rising edge). The rising edge triggers a monostable comprising C20, R37, D12 and Schmitt trigger input 3/9. The output of the Schmitt trigger 3/8, which is normally high, goes low for a period of 0.68 milliseconds, which is determined by C20 and R37. After this period, the output 3/8 returns to the high state, triggering off another monostable comprising C21, R38, D13 and Schmitt trigger input 3/11. The output of the Schmitt trigger 3/10 goes low for a period of 0.15 milliseconds. A final inversion by a gate with output 3/12 provides a positive pulse 0.15 milliseconds long occurring 0.68 milliseconds after the turn-off of the linear actuator.

The time delay before the gating pulse occurs is deliberately made longer than the time measured in the experiments with the linear actuator. This is because variations may occur in the time delay due to tolerances in the passive components or differing threshold characteristics in the Schmitt triggers.

4.2.6 Interface Circuit for Motor Voltage

The gating pulse is used to operate a sample and hold circuit. The waveform being sampled, i.e. the voltage across the motor, cannot be directly connected to integrated circuits. The on periods of +12 volts and the negative transients of -40 volts would damage the input stages, so an interface is necessary to limit the voltages to a narrow band acceptable to the integrated circuits. Furthermore, the back e.m.f. can vary over the range 0 to 12 volts, but the logic circuits and operational amplifiers are working from a voltage rail of only 0 to 8 volts, so it is preferable that the back e.m.f. should be scaled down to the 0 to 8 volt range.

In the speed control mode, relay RL2 is operated and the voltage across the motor is controlled by the power transistors TR2, 3 and 4. The voltage is attenuated by resistors R19, 20 and 21 so that the scaled down back e.m.f. is within the 0 to 8 volt range.
Positive and negative transients are suppressed by the 10 volt zener diode Z1 which clips any voltages outside the range -0.6 to +10 volts.

Attenuation by R20 and R21 ensure that the voltage presented to the integrated circuit is less than 8 volts, the positive supply rail, to prevent damage to it.

To prevent any negative excursion whatsoever, an extra circuit consisting of R22, D7 and D8 is used. Resistor R22 and diode D8 form a low impedance voltage source of +0.6 volts. If the input to the integrated circuit attempts to go negative, then diode D7 turns on and holds the input at zero.

A single germanium diode connected across resistor R21 was considered for this job, as it held the input voltage to an acceptable -0.1 volts. However, at 85°C, the maximum operating temperature of the control, the leakage current of the germanium diode had risen to such a level that it was affecting the accuracy of the speed control. For this reason, only silicon semiconductors were used in this application.

4.2.7 Sample and Hold Circuit

The scaled and clipped voltage is connected to IC11 pin 1. This device is a 4066 CMOS analog switch. Logic voltages applied to pin 13, the gate, can either open or close the semiconductor switch between pins 1 and 2. When open, the switch has a high resistance exceeding 100 Megohm. When closed, the switch resistance is less than 600 ohms.

The analog switch is controlled by the gating pulse which defines the back e.m.f. "window". When the analog switch closes, the back e.m.f. charges up the capacitor C15. When the analog switch opens again, the capacitor has no path through which it can discharge, so it holds the voltage imposed on it. If a low resistance circuit was connected to the capacitor, it would rapidly
discharge the capacitor, degrading the accuracy of the back e.m.f. sample. Therefore a high input impedance unity gain amplifier, input IC1/3, is connected to the capacitor to buffer it.

When selecting component values, the first priority was to ensure that the capacitor was large enough to keep its charge during the hold period without drooping. Data sheets indicated that the unity gain buffer had an input bias current of 0.1 microamps. The hold period was 10 milliseconds and during this time a droop of 0.1 volts was acceptable. The capacitor value was calculated by

$$C = \frac{i \Delta t}{\Delta V} = 0.1 \text{ microamps} \times \frac{10 \text{ milliseconds}}{0.1 \text{ volts}}.$$  

$$C = 10 \times 10^{-9} = 10 \text{ nanofarads}.$$  

A standard value of 100 nanofarads was used, this being a standard value used elsewhere in the circuit.

Having selected the capacitor, the resistors R19, 20 and 21 had to be chosen. Their ratio was already known but their absolute values were not yet decided. Ideally they should be low in value, to allow a fast charge up of the capacitor, but this would have dissipated too much power. To charge up the capacitor, a time constant of 5 milliseconds was required, to give adequately fast tracking of the back e.m.f.

The resistor value was not calculated by

$$\tau = RC.$$  

Instead, because the waveform was being sampled, the equation was used:

$$\tau = \frac{\text{Sampling Period}}{\text{Sample Duration}} RC.$$  

The resistance required was

$$R = \frac{\text{Sample Duration}}{\text{Sampling Period}} \frac{\tau}{C} = \frac{0.15 \text{ ms}}{10 \text{ ms}} \frac{5 \text{ ms}}{0.1 \text{ microfarad}} = 750 \text{ ohms}.$$
That is to say, the equivalent charging resistance calculated according to Thevenin's Theorem should not exceed 750 ohms.

For the resistor values selected, the equivalent charging resistance is

$$R_{\text{charge}} = \frac{(R_{19} + R_{20}) \times R_{21}}{R_{19} + R_{20} + R_{21}} = 659 \text{ ohms}.$$ 

Fortunately, the power dissipation with the selected resistor values was acceptable, while giving adequately fast tracking of the back e.m.f. Resistors with power ratings of \(\frac{1}{2}\) watt were used, similar to those in the rest of the circuit. Efforts were made, in this way, to use as many common component types as possible. The dividend of this can be seen during assembly of the printed circuit boards. 67 out of 69 resistors and diodes fit in the board with a pitch of 0.5 inch, and these bandoliered components can be fed through a crop and form machine without intermediate adjustment.

By the means described, a voltage proportional to the actuator speed was obtained. The next step was to use this measured signal to control the actuator speed.

### 4.3 SPEED CONTROL

#### 4.3.1 Overview

A proportional-plus-integral controller is selected to give zero steady state error despite load variations. The theoretical performance of the proportional-plus-integral circuit is analysed, and then some hardware problems are discussed. Pulse Width Modulation is used to drive the output stage, and the modulator circuit is described in detail. The modulating frequency and the maximum duty cycle depend upon the system, and the procedure for selecting them is given. An override circuit for maximum speed is also incorporated.

#### 4.3.2 Proportional-plus-Integral Controller

Having devised a sensing system to measure actuator speed, a speed control
system was required. Actuator speed had to be maintained constant, whether the actuator was unloaded or heavily loaded. A proportional-only controller could only achieve this if a high gain was used, which would cause stability problems. Instead it was realised that a proportional plus integral controller would have to be used. The circuit of the P + I controller is shown in fig 15.

The circuit can be analysed by considering the current flowing into and out of Node A. The inverting input has a high impedance and takes negligible current from the node, so

$$i = \frac{V_{IN} - V_I}{R} + C_I \frac{d(V_{IN} - V_I)}{dt} + C_F \frac{d(V_{OUT} - V_I)}{dt} = 0.$$  

Taking Laplace transforms,

$$\frac{V_{IN} - V_I}{R} + C_I s V_{IN} - C_I s V_I + C_F s V_{OUT} - C_F s V_I = 0.$$  

A useful property of high gain operational amplifiers is that negative feedback causes the voltage on the inverting input to become equal to that on the non-inverting input, so

$$V_I = V_{REF}.$$  

Substituting this into the equation above,

$$C_F s V_{OUT} = \frac{V_{REF} - V_{IN}}{R} + C_I s (V_{REF} - V_{IN}) + C_F s V_{REF}.$$  

$$V_{OUT} = \frac{V_{REF} - V_{IN}}{sRC_F} + \frac{C_I (V_{REF} - V_{IN})}{C_F} + V_{REF},$$

i.e. $V_{OUT} = \text{integral part} + \text{proportional part} + \text{constant term}.$

It can therefore be seen that the circuit is a true proportional-plus-integral controller.

The reference voltage for the speed control is obtained from potentiometer P2. Resistors $R_{24}$ and $R_{25}$ are used at the ends of the potentiometer so that when the wiper sweeps from end to end a swing from 8% to 83% of the 8 volt
rail is obtained. The reference voltage is fed into a high impedance non-inverting input of the operational amplifier which gives the proportional plus integral function. The components in the P + I controller were selected on an empirical basis. The capacitors chosen were simply those used elsewhere in the circuit, and the resistor was "selected on test". The circuit was connected up and different values of resistor were tried until satisfactory results were obtained for the speed control loop. No formal design technique was used for this circuit; it was lashed up in a few minutes and quickly tested to meet a deadline. However this empirical approach was backed up by 10 years of engineering design experience.

4.3.3 Problems with Operational Amplifiers

Under certain circumstances, the output of the operational amplifier IC1 pin 7 can go negative, outside the voltage rails. This occurs when IC1 pin 1 falls from +8 volts to zero volts, pushing C16 and C17 negative. The phenomenon is therefore not caused by amplifier 1/7 creating a negative voltage, but is simply a capacitor pumping effect. This negative voltage must not be passed on to amplifier IC2 pin 10. This amplifier type LM324 exhibits the undesirable characteristic of making its output go high when a slightly negative voltage is presented to its non-inverting input. Even a voltage of as little as -0.3 volts below the negative supply rail can produce this effect, which is exactly the opposite of what a well-behaved operational amplifier should do. Silicon diodes clamp at a voltage of 0.6 volts which is too high. Germanium diodes clamp at only 0.1 volts and are ideal at room temperature, but at elevated temperatures their leakage current increases, rendering them useless. The solution is to use a silicon diode, which is not connected to ground, but instead to a potential of +0.6 volts. Should the amplifier output of IC1 pin 7 go negative, then diode D9 turns on, and clamps the voltage coming out of resistor R27 at a firm zero level.

Capacitor C18 is necessary to absorb any high speed transients occurring at
this point. It was suspected that noise was entering the circuit from the actuator, and was being capacitively coupled through diodes D7 and D9.

4.3.4 Pulse Width Modulation

The voltage at test pin 12 can be described as the "actuating signal". It is, loosely speaking, an analog measure of the amount of effort required from the actuator. Turning this low power analog signal into an on-off signal to switch the power transistors is performed by a Pulse Width Modulation circuit.

Pulse width modulation is achieved by comparing the analog input signal against a modulating waveform. In Fig 16 a triangular modulating waveform is used. As the magnitude of the input signal varies, the mark-to-space ratio or duty cycle of the modulated waveform varies also. The slopes of the triangular waveform need not be identical, and assymetrical triangular waves or even sawtooths may be used. The only stipulation is that a linear slope should be used, to give a linear modulator.

As can be seen from Fig 16, if the input voltage is higher than the maximum voltage of the modulating triangular wave, then the mark-to-space ratio is 100%. This is not suitable for our purposes because the actuator must be switched off for part of the time to allow the back e.m.f. to be measured. Clamping the input voltage at a level corresponding to say 90% duty cycle is one solution. But the input voltage would have to be clamped very accurately; 5% tolerance zener diodes would not be good enough. Also the triangular modulating waveform would have to be generated very accurately, to ensure the amplitude was correct. A production batch of 5% tolerance zener diodes could vary the duty cycle between 83% and 97%. To look at this from a more relevant viewpoint, the off-period would vary between 3% and 17% of the cycle time. The off-period should be the shortest time possible to reduce the motor current to zero and measure the back e.m.f. with the sample and hold circuit. An off-period of only 3% would not allow sufficient time for the measurement;
conversely, an off-period of 17% would waste time and reduce maximum power available from the actuator. Clearly a more accurate way of defining the maximum duty cycle is needed.

4.3.5 Generation of Modulating Waveform

The approach which was followed was to generate an asymmetrical triangular wave so that the ramp-up time occupied 90% and the ramp-down time occupied 10% of the total cycle time. The circuit used to produce this waveform also had a second output, a rectangular waveform. By adding the rectangular waveform to the triangular wave, a composite waveform (fig 17) was produced which solved the problem in a most satisfactory manner.

The waveform generator is based on two operational amplifiers, outputs 1/8 and 1/14. Amplifier 1/8 is connected as an integrator and amplifier 1/14 is connected as a comparator with hysteresis. The circuit is self-oscillating and produces a triangular wave at output 1/8 and a rectangular wave at output 1/14.

A "pseudo-earth" is established at 3.2 volts by resistors R30 and R31. This voltage is used in preference to the mid-rail voltage of 4 volts in order to centre the triangular wave around this level. The output of operational amplifier 1/14 can switch between low and high levels. The low level is well defined, being less than 200 millivolts above the negative rail. The high level can be up to 1.8 volts below the positive rail. This spread in the high level can cause some variation in the shape of the triangular waveform, but fortunately this does not affect the overall action of the control.

4.3.6 Operation of Waveform Generator

The waveform generator works in the following way. Assume that the output of amplifier 1/14 is high. Then current flows through resistors R32 and R33 and diode D10 to the virtual earth 1/9, and then through capacitor C19, charging
it in a linear manner. The voltage at the integrator output 1/8 decreases linearly. Initially the voltage on the output 1/8 is at a high level, and this produces a high level at the non-inverting input 1/12 of the comparator, which maintains the high output 1/14 which we originally assumed. But as the integrator output 1/8 reduces, the comparator input 1/12 reduces also, until it reaches the voltage of 3.2 volts present on the inverting input 1/13. Then the comparator output 1/14 abruptly switches low, aided by positive feedback via resistor R35. The voltage on the non-inverting input 1/12 of the comparator has now reduced from 3.2 volts to about 0.5 volts, and the comparator output 1/14 has gone well and truly low. This low output draws current through resistor R33. Diode D10 is reverse biased and blocks any current flow through R32. This smaller current charges capacitor C19 in the reverse direction, and the integrator output 1/8 slowly increases also until a level of 3.2 volts is reached, equal to that on the inverting input 1/13. Thereupon the comparator output switches abruptly high, aided by positive feedback through R35, which pulls the voltage on the non-inverting input 1/12 up to a high level of about 6 volts. The cycle then repeats again.

The upper and lower threshold voltages for the comparator are established by the resistors R30, 31, 34 and 35, and also to a lesser extent by the high and low level saturation voltages of the comparator output 1/14. The upper and lower threshold voltages are equal to the high and low limits of the triangular waveform.

The triangular waveform is made asymmetrical mainly by the diode D10, which passes or blocks the current which charges the capacitor C19. The saturation levels of the amplifier output 1/14 have a slight effect but mainly the waveform is established by the magnitude of the resistors R32 and R33. These items are relatively stable and in a production batch using 5% tolerance resistors, the off-period in the cycle would only vary between 9% and 11%.
The stability of the off-period is thus very much better using this circuit than with the method previously discussed.

The triangular wave varies between 0.6 volts and 5.2 volts, and the rectangular wave switches between 0.1 volts and 6.6 volts, (the upper limit being somewhat dependent on the particular integrated circuit being used). The two waveforms are combined by resistor R36 and diode D11 to give a composite waveform between 0.6 volts and 6.0 volts. The analog voltage must still be clamped to prevent it rising above 6.0 volts. The analog voltage must occur at any voltage between 5.2 volts and 5.9 volts, and the actual clamping voltage does not affect the duration of the off-period in any way. Clamping is performed by a 5.6 volt zener diode Z2, with 5% tolerance.

4.3.7 Maximum On-Time for Duty Cycle

So far, the maximum duty cycle has been referred to as 90%, but how was this figure determined? To obtain maximum performance from the linear actuator, the highest possible duty cycle should be provided by the control. As described previously, an off-time of 0.58 milliseconds is necessary to allow the actuator speed to be measured from the back e.m.f.. For maximum duty cycle, the maximum on-time should therefore be used. As the on-time is increased, the duty cycle increases but the period between speed measurements increases also. In effect, the sampling frequency is being reduced and ultimately a situation is reached where the speed is not being measured sufficiently often, and the control system becomes unstable.

An empirical method was used to discover the maximum on-time. The speed control circuit was connected up to an actuator and components in the triangular wave generator were adjusted to vary the sampling frequency. Sampling periods greater than 80 milliseconds were guaranteed to cause instability in the control loop; the speed oscillated up and down, and the
actuator moved in a peculiar jerky fashion. Sampling periods less than 20 milliseconds always gave smooth stable speed control. A selection of actuators were used under a variety of load conditions, and finally a sampling period of 10 milliseconds was selected. This value gave a large safety margin and ensured that no control loop stability problems ever occurred in production. Round baler controls are all tested first with a toy motor which has a stall current of only 1 Amp, and are then tested with a linear actuator with a stall current in the region of 30 Amps. Stable operation is obtained with both types of motor.

4.3.8 Minimum Off-Time for Duty Cycle

The duty cycle can not be made larger by increasing the on-time, because the increasing sampling period causes instability. The only way to increase the duty cycle is to reduce the off-time. This could be achieved by dumping the inductive energy in the motor armature more quickly. This would occur at a higher voltage, so 100 volt rated power MOSFET transistors would be used instead of 60 volt devices. However, only a small increase in duty cycle would occur, from say 90% to 94%, so the increase in performance would be negligible. The main factor which reduces the actuator performance is the back e.m.f. measurement method, which reduces the current to zero during the off-time. The current then takes several milliseconds to build up again when current is restored, so the average current is very much less than would be expected from consideration of the duty cycle alone.

4.3.9 Improvement of Actuator Performance

Actuator performance could be improved in another way, by increasing the supply voltage to the control. While the battery can only deliver a nominal 12 volts, a switching step-up regulator could be used to boost the voltage to a level of say 18 volts. At this voltage level, the actuator performance would be back to its original level before the speed control was interposed.
Investigation of such a regulator is currently being undertaken for two important areas: for high power applications where all the force available from the actuator is needed, and also for re-fits, where an actuator on an existing baler is upgraded by the addition of the automatic twine-tying control. On many existing balers, small actuators were fitted to save costs and these small actuators may only just have sufficient power for the application. A switching step-up regulator will maintain or even improve the performance of the actuator when the speed control is added.

4.3.10 Comparator
The analog "actuating signal" must now be compared against the composite modulating waveform. This is achieved by operational amplifier output 2/8 which is configured as a comparator with hysteresis. The hysteresis or positive feedback is very small; the ratio of R27 plus R28 to R29, which determines the amount of feedback, is only 3%. This amount of hysteresis has no significant effect on the voltage at which the comparator switches, but it does ensure that switching occurs cleanly and decisively. Without it, multiple transitions could occur at the output if either of the input signals was contaminated by say 1 millivolt of noise. Keeping out noise of this magnitude is not feasible, and the use of positive feedback to give a single clean transition at the output is a much easier solution.

4.3.11 Full Speed Override
Another analog switch or transmission gate in IC11 is connected to the non-inverting input 2/10 of the comparator. When this analog switch is closed, it connects the non-inverting input to the positive rail, above the level of the modulating waveform, and the duty cycle then becomes 100%. By this means, the logic sequencer can override the speed control and call for full actuator speed by simply presenting a high logic level to the control input 11/6 of the transmission gate.
4.4 POWER CIRCUIT

4.4.1 Overview

The previous section described the generation of the pulse-width modulated signal which is to drive the power circuit.

The selection procedure for the semiconductors used in the power stage is described. The relative merits of bipolar transistors and power MOSFET transistors are discussed for this application. The importance of a positive temperature coefficient of resistance is stressed, to obtain satisfactory operation when the transistors are paralleled. The power circuit configuration is explained, and the interface circuit between the pulse-width modulated signal and the power transistors is described.

Dissipation of the inductive transient at motor switch off must be done quickly to allow sensing of back e.m.f.. Conventional clamping circuits are assessed and a new clamping technique is used to collapse the transient rapidly.

4.4.2 Comparison of Bipolar and Power MOSFET Transistors

The linear actuator was a difficult load to drive with semiconductors. The current required was under 5 Amps when running freely, without load, but rising to perhaps 15 Amps under load. Under starting or stall conditions, currents between 25 Amps and 40 Amps would be drawn. The voltage supplied to the actuator was 12 volts, but the inductive nature of the load would produce a transient at switch-off with an amplitude of 40 volts. The semiconductor would have to withstand 52 volts.

Bipolar power transistors were originally evaluated for this application. 40 Amp devices were available at a reasonable price, and their collector-emitter saturation voltage was sufficiently low to ensure an acceptable power
dissipation. But to turn on the power transistor would have required a base current of between 5 and 8 Amps. Providing such a current was quite impossible. If the current came from the opposite voltage rail (fig 18) then the current limiting resistor would dissipate between 60 and 100 watts - far too much. Using a Darlington configuration (fig 19) reduced the base current to an acceptable 200 milliamps, but the collector-emitter saturation voltage rose to a much higher level, and the dissipation of the power transistor rose to a level above 60 watts. With bipolar transistors there was no simple way to solve this problem.

The advent of power MOSFET transistors was the breakthrough in this application. MOSFETs give saturation voltages comparable with the best bipolar transistors but require negligible gate current to drive them (fig 20). In the on state, gate currents are in the region of microamps. Additional current is required to charge and discharge the gate capacitance, but this is only needed when the MOSFET is turned on or off. At the low switching frequency used in this application, a current of 5 milliamps and a pull-down resistor of 2.2 kilohms gave acceptable switching.

Another great advantage of the power MOSFET transistor was its ability to be paralleled to share the load current. Single power MOSFET transistors were available which could handle the load requirement of 60 volts and 40 Amps, but these devices were fairly expensive and were packaged in TO-3 metal cans which were relatively expensive to mount on printed circuit boards. Instead, three transistors rated at 60 volts and 14 Amps could be connected in parallel to do the same job. These three plastic TO-220 devices were cheaper than the single metal can transistor and were fairly easy to mount. The assembly method used was very straightforward. The three power transistors have their leads cropped and formed. A heatsink bracket is placed on the edge of the board and the three power transistors are bolted down, forming a board/bracket/transistor sandwich. Three silastomer gaskets and three plastic
bushes are used in the assembly to provide thermal transfer with electrical isolation. The silastomer gaskets are a much tidier alternative to the old-fashioned mica washer plus heatsink compound. During the assembly, the transistor leads are located through holes in the printed circuit board. The remainder of the components on the printed circuit board are inserted and then the complete board is flow soldered in one operation.

Having three transistor mounting holes gives a high degree of flexibility in selecting power transistors. More efficient transistors with higher ratings are coming onto the market at lower cost. When these are more readily available, the three transistors will be replaced by two higher grade transistors.

4.4.3 Thermal Coefficient of Resistance, and Secondary Breakdown

Bipolar transistors cannot be paralleled to share a load because of their tendency to current-hog instead of current-share, due to a negative temperature coefficient of resistance. If bipolar transistors are connected in parallel, then one of the transistors may pass a little more current than the others. This dissipates more power, so the transistor becomes hotter than the others, and because of the negative temperature coefficient of resistance, it passes more current still. This produces a thermal runaway, with one red hot transistor hogging all the current and the others passing no current at all.

Power MOSFET transistors exhibit the opposite behaviour due to their positive temperature coefficient of resistance. If one of the power MOSFET transistors takes more current than the others, it becomes hotter and so tends to share the current with the others. In a very high performance application, it would be necessary to match the power MOSFET transistors for transconductance. This application does not push the transistors anywhere near their limits and I have even witnessed three unmatched transistors from different
manufacturers working together quite happily on the same heatsink.

The temperature coefficient of resistance has another implication besides paralleling separate transistors. Imagine that a single power transistor is made up of 100 mini-transistors in individual cells or in the bulk of the semiconductor. The behaviour of one of these mini-transistors depends upon the temperature coefficient of conductance as much as the complete power transistor does. This means that in bipolar power transistors, hot-spotting may occur. Individual cells may become very hot and burn out due to thermal runaway. This places a greater load on the remaining cells which in turn may burn out, causing a snowball effect leading to catastrophic failure of the entire device.

This effect is referred to as secondary breakdown of bipolar transistors, and curtails the useful operating range of the device, or Safe Operating Area. For example, a bipolar transistor rated at 60 volts and 40 Amps cannot handle these two parameters simultaneously for even the smallest fraction of a second, due to secondary breakdown.

On the other hand, if a power MOSFET transistor is rated at 60 volts and 40 Amps, then it can handle these two parameters simultaneously, providing due consideration is given to power dissipation and the transient thermal impedance of the device. Hot-spotting does not occur and the load current is evenly divided across the chip. Generally speaking, the power MOSFET transistor is far more robust, and handles its load current in a more intelligent manner than its bipolar counterpart.

4.4.4 Circuit Configuration

To return to the question of circuit configuration, it was desired that the analog circuits should be kept simple. This decided the configuration of the current sensing resistor, the two relays and the power transistors (fig 21).
First of all, the current sensing resistor which detects stall current in the motor had to be connected to the negative rail. This was because the voltage across the current sensing resistor would be very small and an amplifier would be needed to boost the signal up to a sensible level. Referring one end of the current sensing resistor to the negative rail would allow the use of a simple non-inverting operational amplifier circuit. If the current sensing resistor had not been connected to the negative rail then a more complicated instrumentation-type amplifier would have been necessary to strip off the common mode voltage and refer the signal back to the ground rail.

Measurement of back e.m.f. posed a similar problem. It was desirable to measure the voltage across the motor with one end connected to the negative rail, so that a voltage relative to ground would be obtained. Otherwise, an instrumentation amplifier would have been needed to remove the common mode voltage and produce a signal relative to ground. These two requirements seemed incompatible; both the motor and the current sensing resistor had to be connected to ground to give useful signals, yet the two had to be connected in series so that the current sensing resistor could detect the motor current.

Fortunately, it was remembered that during the measurement of back e.m.f., the motor current had been reduced to zero. This meant that the voltage across the current sensing resistor would be zero at the instant of back e.m.f. measurement, and therefore the motor and the current sensing resistor could be connected in series as shown in fig 21 without one measurement affecting the other.

The direction of actuator travel is controlled by two relays which switch the supply voltage to the motor either positive, negative, or off. Relays were chosen in preference to semiconductors because they were technically far superior for this function. The two relays were easy to drive and gave a much lower voltage drop than four semiconductors in an H-bridge configuration.
The relays selected for the job were printed circuit mounted skeleton relays, manufactured by Bosch for automotive applications such as intermittent windscreen wiper controls. Bosch produce these relays in such volume that the price is remarkably low (80 pence).

When relay RL1 is operated, the motor is connected directly across the supply and the actuator extends at full speed. When relay RL1 is de-energised, the contact first removes the supply from the motor and then places a short circuit directly across the motor via relay RL2. This produces a high degree of dynamic braking, and the actuator stops dead instead of coasting to a halt.

When relay RL2 is operated, the supply to the motor is reversed and the actuator retracts. Instead of switching the motor directly to the positive rail, the power transistors TR1, 2 and 3 are interposed, and the pulse width modulation of these devices governs the motor speed.

Both P-channel and N-channel MOSFETs were assessed for this application. The P-channel transistors were easier to drive but enquiries soon revealed that P-channel transistors were only available with low current ratings. Perhaps in a few years the situation will be different but at the present time only N-channel transistors can handle the necessary currents.

N-channel transistors presented a difficulty. When they were turned on, the voltage at their drains would be about 11 volts (assuming a 12 volt supply). To turn them on, the voltage at the gate relative to the drain would have to be at least 10 volts. Therefore it would be necessary to apply at least 21 volts to the gate relative to ground to turn on the transistors. In a system operating from a 12 volt battery supply, the problem was obvious.

The 22 volts needed to drive the gates of the power MOSFET transistors was
obtained from a voltage doubler circuit. The current requirement was very small (less than 10 mA) and regulation or smoothing was not very important. This indicated that a simple diode pump could be used instead of a switching step-up regulator which would have needed a relatively expensive inductor.

For most efficient operation, an active pull up and pull down should be used to drive the diode pump. For a small power requirement such as this, efficiency may be less important and a passive pull up may be satisfactory. The circuit selected used a Darlington transistor in the IC6 array as an active pull down, and R12, a 220 ohm resistor as a passive pull up.

The operation of the circuit is fully described in section 3.4.10.

4.4.5 Drive Circuit for Power MOSFET Transistors

Turning the power MOSFET transistors on and off was achieved by switching the 22 volt supply to the gates. The output of the pulse width modulating comparator IC2 pin 8 swung between 0 and 8 volts and this was used to switch a Darlington transistor in the IC6 array. The common emitter Darlington transistor was used in turn to switch between 22 and 0 volts, to turn transistor TR1 off and on, through resistor R40. When the Darlington transistor in the IC6 array is off a small leakage current may flow through it, and may turn on transistor TR1. To prevent this possibility, an additional resistor R39 is used to divert this leakage current away from the transistor.

The signal from transistor TR1 is suitable for driving the gates of the power MOSFET transistors TR2, 3 and 4. But if the gates are simply connected in parallel, parasitic oscillation may occur, with the transistors triggering each other into instability. Stability can be ensured by fitting a small resistor or a ferrite bead in series with each gate. Both methods work; a piece of test equipment for the Round Baler Control contains six paralleled
power MOSFET transistors, with the gates decoupled by ferrite beads. But in this control, 220 ohm resistors (R8, 9 and 10) were used.

Transistor TR1 quickly charges up the gate capacitance to turn on the power MOSFET transistors. To discharge the gate capacitance quickly when TR1 turns off, a pull-down resistor R41 is used.

The three power MOSFET transistors are connected in close proximity on the double sided printed circuit board. Very thick tracks on both sides of the board connect the terminals of the transistors and transfer current from the battery to the transistors and then through relay RL2 to the actuator.

4.4.6 Clamping of Inductive Transient

When relay RL2 is closed and the transistors are turned on, 12 volts are placed across the actuator, and the actuator current climbs exponentially to a level between 30 and 50 Amps. Inductive energy is stored in the motor field according to

\[ E = \frac{1}{2}LI^2. \]

When the transistors turn off, a large negative voltage spike is induced by the motor field. If left unchecked, it could easily plunge to a level of -200 volts (fig 22a). This would overstress the power transistors which are only rated to withstand 60 volts, and they would be destroyed. Numerous methods exist for diverting the inductive spike into an auxiliary device, and dissipating the inductive energy there. Some of these methods were assessed for this application.

Fitting a 40 Amp diode to limit the negative spike to about -2 volts was the first method tried (fig 22b). This undoubtedly prevented damage to the power transistors, but dissipating the inductive energy took a long time at such a low voltage, and this interfered with the measurement of the motor speed from
the back e.m.f.. Some calculations showed that clamping the inductive spike
to about 40 volts would dissipate the energy sufficiently quickly.

Placing a 39 volt zener diode in series with the 30 Amp diode would have
performed this function (fig 22c), but the power rating of the zener diode
would have to be very large, and its cost would be prohibitive. Coupling a
low power zener diode with a bipolar power transistor seemed to provide a
fairly cheap solution to the problem (fig 22d). In fact a prototype was
built and demonstrated at the SIMA exhibition in Paris, before it was realised
that the safe operating area of the power transistor was being exceeded. The
simultaneous application of 40 volts and 30 Amps to the bipolar transistor
would guarantee failure from secondary breakdown, sooner or later. When the
transistor was re-assessed and rated for this duty, it was realised that the
cost would again be prohibitive.

Replacing the bipolar transistor with a power MOSFET transistor finally gave
a reliable but more expensive voltage clamp.

Then after all this work, a breakthrough was made which confirms once again
that "An engineer can do for a shilling what any fool can do for a pound".
It was realised that a separate voltage clamping device was not needed: the
primary switching device could perform the clamping function also. This was
achieved by partly turning on the switching device during the inductive
spike, so that it operated momentarily in the linear mode with high
dissipation.

The method used was to connect a low power 30 volt zener diode $Z_3$ to the
gates of the power MOSFET transistors (fig 22e). During the negative
inductive spike, the zener diode turns on when the voltage across it exceeds
30 volts, so the transistor gates are clamped at $-30$ volts. The inductive
spike pulls the transistor sources even lower, to $-38$ or $-40$ volts. At this
stage the gate to source voltage is sufficiently high to turn on the power transistors, and the inductive spike is clamped at $-40$ volts. The motor current decays linearly to zero, whereupon the inductive spike collapses, all its energy having been dissipated in the power transistors.

An additional diode D16 is needed to block off the zener diode during the on-period. But the overall additional cost is under 10 pence. External clamping devices such as power zener diodes or power MOSFET circuits could cost between £3 and £6. This method is therefore a great advance in clamping inductive transients. The method is only applicable to power MOSFET transistors because the simultaneous voltage and current stress in a bipolar transistor would result in secondary breakdown.

4.5 SUMMARY

Measurement of the speed of the d.c. motor in the linear actuator was achieved by sampling the back e.m.f. shortly after the voltage supply was switched off. This delay was needed to allow the inductive transient to decay. The duration of the transient was known so a timer circuit triggered by voltage switch-off generated a pulse just after the transient to operate a sample-and-hold circuit. An interface circuit was needed to protect the sample-and-hold circuit from the damaging voltages and transients across the motor.

The motor speed was controlled by a proportional-plus-integral amplifier to give zero steady state error despite load variations. The control circuit was analysed to confirm its operation but its parameters were selected empirically. Some practical problems were encountered with the operational amplifiers used in this circuit but these were overcome.

The control signal was amplified by pulse width modulation. A special modulator was devised which could only output a signal between 0% and 90%
duty cycle. The remaining period of at least 10% was needed to allow back e.m.f. sampling.

The power of the actuator was reduced considerably because of this duty cycle limitation and also because the motor current was reduced to zero during every cycle to allow back e.m.f. measurement. The maximum duty cycle and the switching frequency were determined by the length of the inductive transient and by the stability of the speed control.

For switching power to the linear actuator both bipolar transistors and power MOSFET transistors were assessed. The power MOSFET transistors were far superior: they were easier to drive, they could be paralleled to share the current, and they were free from secondary breakdown. Much of their advantage stemmed from their positive temperature coefficient of resistance.

The circuit configuration was described from the designer's viewpoint of "which bit goes where". A number of conflicting requirements existed, but eventually everything fitted into its place.

It was important that the inductive transient of the motor should be dissipated quickly at switch-off to allow back e.m.f. sampling. Various auxiliary devices were placed across the motor in the time-honoured manner to clamp the voltage spike, but then a new method was discovered. The power MOSFET transistors which performed the switching could be triggered into the linear mode during the inductive transient. This rapidly dissipated the inductive energy with negligible additional cost.
The volume of sales of round baler controls is increasing rapidly because the basic concept is acceptable to most of the baler manufacturers. They generally request cosmetic changes to the control but the basic printed circuit board has remained substantially the same. Some balers have unconventional cycles, but the sequencer circuit has allowed minor changes to be made very easily.

The importance of a good power supply and good transient suppression has been emphasised. The control aspects of this project, such as the proportional plus integral controller, and the sampling frequency, were selected on an empirical basis. However tests were performed to ensure these selections were satisfactory.

A great deal of experience has been gained with power MOSFET transistors. These devices undoubtedly have a very important future in the switching of heavy inductive loads.

The measurement of motor speed from the back e.m.f. has some useful implications. Integrating the speed will give information about the position of the actuator. This may be useful for the next generation of round baler controls. Some existing balers give uneven wrapping with closer wraps at one end than at the other, due to the nonlinear mechanism. If the actuator position is known, the speed can be varied across the stroke in such a way that even wrapping is produced.

Having examined the round baler control, the design proposals for the fertiliser sprayer control will be discussed.
6.1 STATEMENT OF PROBLEM

Liquid fertilisers and insecticides are commonly distributed in Europe by tractor. A tank is slung behind the tractor and spray booms allow coverage of up to 40 ft. at one pass (fig 23).

As the tractor runs up and down the field at 40 ft. intervals, its wheels destroy the crop below creating "tramlines". If more fertiliser is applied later these tramlines are used again to avoid further damage.

More and more farmers are using tramlines to their advantage by creating them when the crop is sown. A typical seed drill may be 8 ft. wide with coulters at intervals of 1 ft. Normally all the coulters are open so that the field is fully seeded but on every fifth bout, two of the coulters (directly behind the tractor wheels) are closed. The tramlines become visible as soon as the crop starts to grow.

Some farmers still operate their coulters with string on every fifth bout but the increasing tendency is to use purpose made counters such as the Warner Electric Tramline Controllers. These devices contain an electronic counter which is incremented automatically at the end of each bout. On the fifth bout a signal is given to operate Warner Electric Solenoids or Linear Actuators to shut the tramline coulters.
Accurate tramlines have made possible very precise spraying. But in spite of them, results can be disappointing or very costly. This is because the spray rate is set by a manual control valve which adjusts the flow from a pump driven by the Power Take Off (PTO) on the tractor. Clearly for constant "gallons per acre", the tractor must be driven at constant speed, and this is difficult to achieve in practice. Errors occur during acceleration and deceleration and also because of the difficulty in maintaining a constant speed by referring to a coarsely graduated speedometer while steering through the tramlines. Errors can also occur on uneven fields where the oscillation induced in the long spray booms may cause the driver to slow the tractor to a very slow speed until the oscillation ceases. Finally, some farm labourers or contractors may not have the high degree of skill and motivation required to achieve good results.

Underspraying of pesticides and fungicides can reduce the crop immunity, and overspraying of fertilisers or herbicides can cause localised "burning" of the crop. A conceptual solution to these problems is as follows:

Make the flow rate proportional to the tractor speed.

A method suggested by Warner Electric European Marketing was to monitor the tractor speed with a transducer and use an electronic control to adjust a Linear Actuator connected to a flow control valve. This method could be implemented quite easily using analog circuits.

However, it was felt that a digital microprocessor-based circuit could offer both technical and marketing advantages. For example, a digital display could show any of these parameters:

- speed (miles/hour or kilometres/hour)
- gallons/acre setpoint (litres/hectare)
- gallons/acre delivery (litres/hectare)
- gallons/minute delivery (litres/minute)
- total gallons delivered or gallons remaining. (litres)

Key system parameters such as boom width and tank capacity could be entered and held in non-volatile memory to allow the display parameters to be computed.

In addition to this multi-function controller, a stripped-down version could be offered for bottom of the range spraying systems. However all models would have a control for setting the gallon/acre setpoint and also an On/Off control. The On/Off control would be essential while turning around at the headland at the end of the furrow and also for driving around obstructions such as telegraph poles in the middle of the fields.

The commercial potential of any agricultural control device can be quite staggering, and the concept can be extended to other spraying systems also, such as salting and gritting of roads in Scandinavia.

6.2 SPECIFICATION OF DESIRED PRODUCT

The brief given at the start of the project was very broadly based. The intention was to test the viability of the concept, rather than to produce a specific piece of hardware. However all electronic controls especially those for agricultural applications must conform to certain environmental conditions and details of these will be given also.

1) A robust transducer is to provide an electrical signal which is a function of tractor speed. An optical shaft encoder mounted on one of the tractor wheel bearings would give a suitable digital signal. The wheel chosen should preferably be undriven so that wheel slip does not affect accuracy. However the increasing trend towards four wheel drive tractors could undermine this scheme. Big field sprayers are now trailer-mounted and drawn behind the tractor and the undriven wheels on these are an ideal location for the transducer.
2) An electronic control is to adjust the position of a Warner Electric Linear Actuator to vary the flow rate through a ball valve. The flow rate is to be maintained exactly proportional to the tractor speed so that coverage is kept constant.

Ultimately an "all-singing all-dancing" control is required which will first interrogate the tractor driver for key system parameters such as:

- Diameter of wheel fitted with speed transducer,
- Number of pulses per revolution provided by speed transducer,
- Spray boom width,
- Spray tank capacity,
- Gallons/acre setpoint.

On the basis of these parameters the controller can calculate exactly the flow rate required at any particular tractor speed.

3) Accuracy of flow rate can be ensured either in a closed-loop or an open-loop fashion. Fitting a flow transducer is the simplest solution but could be relatively expensive. If flow rate could be inferred from the position of the Linear Actuator then this would be much cheaper. Any non-linear relationship between the actuator position and the flow rate could easily be corrected by the microcomputer.

An investigation of this area clearly forms an important part of the project.

4) The speed transducer, the electronic control, the Linear Actuator and the flow control valve should conform to the usual agricultural requirements:
   - Rainproof and dustproof to International Protection rating IP66.
   - Operating temperature range 0°C to 70°C (from freezing conditions to temperatures achieved by "black boxes" in blazing Continental sunshine).
   - Resistant to humidity and condensation.
   - Control knobs and switches to be rugged.
Displays and lamps to be visible in full sunlight.

The control should be proof against short circuits on the output side: if a fuse is used then it should be readily accessible for replacement.

The control should be proof against reverse polarity on its supply input.

The control should resist a 1 metre drop test in any plane onto concrete.
7.1 INTRODUCTION

The need for an accurate spraying system has already been discussed. To turn these requirements into reality, a mathematical representation must be formulated. In this chapter, the system is modelled and the specification is written in a mathematical form.

For the Round Baler Control, knowledge of the system had suggested that a proportional-plus-integral controller would give good performance. The approach taken with the Fertiliser Sprayer is exactly the opposite. The performance to be obtained is specified in advance in the form of a Performance Index. Later chapters will use various techniques to turn the Performance Index into a controller.

7.2 SYSTEM MODEL

A decision had to be made at an early stage between an open loop (fig 24) and a closed loop approach (fig 25). The open loop approach held some attractions because no flow transducer would be necessary, the flow rate being inferred from the position of the Linear Actuator. Furthermore, techniques are being developed for inferring the actuator position directly from the voltage and current fed to it. Thus, for a clearly defined mechanical system, no feedback signals would be necessary. However if the mechanical system was changed or if the characteristics of the hydraulic system altered (perhaps due to fouling or blockage), then the accuracy of the spraying system would be jeopardised with no warning to the driver.

On these grounds the closed loop approach was selected because it offered much higher system integrity.
The diagram of the closed loop (fig 26) shows all the functional blocks in the system. It was necessary to formulate a mathematical model of the system before a controller could be designed. Four blocks had to be modelled.

The first block represented the linear actuator. This was a 12 volt d.c. permanent magnet motor. The equation describing the motor was

\[ V = iR + L \frac{di}{dt} + k_τ \omega. \]  \[1\]

The motor output was geared down to turn an acme screw. A nut on the acme screw converted the rotary motion into linear motion (fig 27). The gears and the pitch of the acme screw gave a low output speed but a high load capability, so high that the load needed to turn the rotary control valve was negligible. Because of the light loading, a good approximation was

\[ V = k_τ \omega. \]  \[2\]

Because of the fixed gearing, the motor speed was proportional to the linear speed of the output shaft of the linear actuator, so

\[ \omega = k_τ \dot{y}. \]  \[3\]

Also the motor voltage \( V \) could be regarded as the control signal \( u \). Therefore

\[ u = k_τ k_G \dot{y}, \]  \[4\]

or

\[ \dot{y} = \frac{1}{k_τ k_G} \int u \, dt. \]  \[5\]

This equation described the behaviour of the linear actuator, and filled the first block of the block diagram.

The next block represented the linkage of the linear actuator to the rotary flow control valve. This block did not contain any significant dynamics but was of course slightly nonlinear. The next block was the rotary flow control valve itself. This element had a double ended saturation characteristic with a slightly nonlinear relationship between the zones of saturation. Again, the dynamics of this block were negligible but the nonlinearity posed severe analytical difficulties.
To simplify the design of the controller, it was necessary to linearise these two blocks. Linearisation was a valid approach because when the control was in use it was operating on a substantially linear part of the characteristic. The highly nonlinear saturation characteristic was only traversed during start-up and close-down, and this could be handled by special routines in the program of the digital controller. Linearising the two blocks made the output flowrate proportional to the actuator position, giving

\[ x_0 = k_V y \]

so \[ x_0 = \frac{k_V}{k_T k_D} \int u \, dt \]  

or \[ x_0 = k_F \int u \, dt . \]  

Taking Laplace transforms,

\[ x_0(s) = \frac{k_F}{s} u(s) . \]  

This defined the linearised transfer function of the forward path.

The signal from the flowrate transducer had to be filtered, and the dynamics of this had to be included in the control loop. A Pelton wheel transducer would provide a pulse train whose frequency was proportional to the flowrate. To turn the pulse train into a useable analog signal, the pulse train would have to be fed to a frequency-to-voltage converter. To remove the pulsating component from the output of the frequency-to-voltage converter would require a filter with a relatively long time constant. Accordingly, the transfer function of the feedback path was defined as

\[ x_M(s) = \frac{1}{1 + s \tau} x_0(s) . \]  

7.3 SPECIFICATION

The system specification had naively requested that flow rate should be directly proportional to the tractor speed. Such a requirement could not be satisfied exactly. Suppose the tractor speed increased slightly, then the flow rate would be in error; the control would energise the linear actuator
and after a short interval the flow rate would again be proportional to the tractor speed. To call for zero error at all times was unrealistic. The best that could be hoped for was that the error should be as small as possible during any particular time interval.

This new specification was more realistic but might have led to problems. The tractor speed would not be constant as it drove across a field. The flow rate must therefore be continuously adjusted and the linear actuator must be in constant backward and forward motion. Continuous energisation of the actuator would cause overheating and might reduce its working life. Actuator life could be extended by ignoring small flow rate errors and taking no action. The actuator would only be energised if the error had grown to an unacceptably large value.

The specification now called for minimum flow rate error over any time interval, subject also to minimum energisation of the linear actuator. These requirements were of course mutually incompatible, and a trade-off had to be made as to how much one requirement and how little of the other requirement were to be satisfied.

7.4 PERFORMANCE INDICES

A word statement such as that given above was useful but lacked precision. It was therefore necessary to define a Performance Index, a quantitative measure of quality or goodness.

Important criteria in the choice of a Performance Index are

a) the ability to express the Performance Index in a mathematical form so that computational or analytic procedures may be devised to assess its magnitude,

b) the simplicity of that form, to avoid intractable functions during subsequent operations,
c) the physical significance of the Performance Index in relation to the system and in relation to measurable parameters of the system, and
d) the ability to apply the Performance Index as a control on the system.

The parameters most frequently used in Performance Indices are functions of error \( e \), functions of control effort \( u \), and time. Typical Performance Indices are:

- **Error Functions**
  \[ |e(t)|, e^2, \int_0^T |e| \, dt \]
- **Control functions**
  \[ |u|, u^2, \int_0^T |u| \, dt \]
- **Combined functions**
  \[ \int_0^T e^2 \, dt \text{ subject to the constraint } |u| \leq U_{\text{max}} \]
  \[ \int_0^T (e^2 + au^2) \, dt \text{, where } a \text{ is a weighting function} \]
- **Time functions**
  \[ T \]

The modules or the square of the control and error functions are used to ensure that any control effort or error, of either polarity, cause the Performance Index to be increased. The modulus is a little easier to use for computational solution or "number-crunching", but the square is easier to use for analytic solution.

In the extremum form, three of these Performance Indices have particular significance and are designated

- "The minimum fuel criterion": \( \min \int_0^T |u| \, dt \)
- "The minimum time criterion": \( \min \int_0^T dt \)
- "The quadratic criterion": \( \min \int_0^T (e^2 + au^2) \, dt \).
The Quadratic Criterion is most amenable to mathematical manipulation and has attained great popularity in academic circles. For the case in question, the Quadratic Cost Function was ideal. Squaring the error and the control function ensured that small values were ignored while large values were heavily penalised. The weighting function determined the relative importance of the system error to the control effort.

7.5 SUMMARY
An open loop system was briefly examined but the advantages of a closed loop control system were overwhelming. The spraying system was nonlinear with saturation characteristics, but normally operated on a substantially linear part of its curve. Linearised equations were derived to represent the system and the flowrate feedback transducer.

A Quadratic Criterion was chosen as Performance Index so that flowrate errors could be made small without excessive operation of the actuator.
CHAPTER 8
IMPLEMENTATION OF PERFORMANCE INDEX
BY EULER-LAGRANGE EQUATION

8.1 INTRODUCTION
Having established a Performance Index for the Fertiliser Sprayer, an optimal control was required to achieve the criterion.

The nature of the optimal control is first described, followed by further discussion of the Performance Index. The constraints imposed by the system state equations cause difficulties in analysis so the state equations are incorporated in the error measure using Lagrange multipliers. Functional minimisation of this error measure by the Calculus of Variations gives three sets of equations:

a) the control equations,
b) the state equations,
and c) the costate or Euler-Lagrange equations.

The control equations are expressed in terms of the costates but some re-arrangement yields a dynamical equation for the control in terms of the system states and the control itself. The situation is complicated by a Two Point Boundary Value Problem, but Transversality offers some assistance.

Analog computer simulation of the system and the Euler-Lagrange controller show that fears about the stability of the controller are well founded.

8.2 OPTIMAL TRACKING CONTROLLER
The Fertiliser Spreader could be described as a tracking control system or linear servomechanism, in which the desired output rather than being constant (as with a regulator) is continually changing with time. The output tracks or follows a desired output state. The control input \( u \) is determined by a controller with compares actual output \( x_0 \) with the desired output \( x_D \).
The error measure was the difference,
\[ H(t) = x_D(t) - x_Q(t) \] \[20\]
and a conventional feedback controller would generate a control which was a function of the error measure,
\[ u(t) = u\{H\} = u\{x_D(t) - x_Q(t)\} \] \[21\]
It is important to note that the instantaneous control \( u(t) \) was a function of the instantaneous output \( x_0(t) \) and the instantaneous desired output \( x_D(t) \). As such, the conventional controller took no explicit account of possible future errors between actual and desired output.

In contrast with this situation, an optimal controller would take account of the errors between the predicted actual output and the future desired output, and predict a future control function \( u(\tau), (t \leq \tau \leq t+T) \) for the whole of the future control interval \( (t, t+T) \). To be satisfied that such a predicted control was suitable, a scalar measure of the predicted errors was required. This scalar measure was referred to as a Performance Index.

So the essential difference between a conventional controller which endeavoured to minimise the error measure and the optimal controller attempting to minimise the performance index was that the optimal controller generated not only the present control input but also predicted the future control input.

8.3 PERFORMANCE INDEX

Regulator control systems require that the actual output of the system should be maintained as close as possible to a constant desired output. A quadratic error measure of form
\[ H = [x_D - x_0(t)]^2 \] \[22\]
would satisfy the essential requirements of a regulator.

Tracking control systems or servomechanisms have outputs which track or follow
a desired output state. Instead of being constant, they are continually changing with time. A quadratic error measure for a tracking system would be

\[ H = [x_d(t) - x_0(t)]^2. \]  

[23]

So far only the plant output error has been considered, on the assumption that it would be possible to find control inputs which would minimise the corresponding quadratic error index. It quickly becomes apparent, however, that to minimise such an error index regardless of other considerations might call for infinitely large control signals, and thus give rise to impractical situations. A convenient way round this difficulty was to incorporate a quadratic control term in the error measure, such as

\[ H = [x_d(t) - x_0(t)]^2 + a u(t)^2. \]  

[24]

By giving sufficient weight to the control term in this expression, the amplitude of control which would minimise the corresponding error index could be kept within practical bounds, although at the expense of increased output errors.

8.4 FUNCTIONAL MINIMISATION

In mathematical terms the optimal control problem was one of functional minimisation. It was similar to the ordinary minimisation problem in calculus, where the necessary condition to be satisfied by the independent variable of a scalar function was established by equating to zero the derivative of the function. In functional minimisation, rather than obtain an algebraic equation which was satisfied by the independent variable at the minimum of the scalar function, we should obtain instead a set of differential equations satisfied over the control interval by the control and state variable functions of time at the extremum of the error index function.

The functional minimisation was further complicated by the need to satisfy the constraints existing among the control and state variables due to the
process dynamics, as specified by the state equations. State equation constraints could be introduced into the functional minimisation problem by means of the mathematical technique of Lagrange multipliers. (ref. 3,4,5)

8.5 CALCULATION OF CONTROL LAW USING EULER-LAGRANGE EQUATION

The system whose performance was to be optimised was described by the following equations.

The output was
\[ x_0(s) = \frac{k_p}{s} u(s). \quad [25] \]

The output was not directly available because of the filtering effect of the flowrate transducer. The measured output was
\[ x_m(s) = \frac{1}{1 + s\tau} x_0(s). \quad [26] \]

The system was a tracking control which attempted to follow a desired output state \( x_D \).

The error measure was the difference between desired and actual output states,
\[ e(t) = x_D(t) - x_0(t). \quad [27] \]

A quadratic error measure for the tracking control system was
\[ H = [x_D(t) - x_0(t)]^2. \quad [28] \]

As discussed previously, it was necessary to incorporate a quadratic control term to avoid infinitely large control signals, so
\[ H = [x_D(t) - x_0(t)]^2 + a u(t)^2. \quad [29] \]

The performance index was
\[ J = \int_0^T \left\{ [x_D(t) - x_0(t)]^2 + a u(t)^2 \right\} dt. \quad [30] \]

The problem was to find the control signal \( u(t) \) which minimised the performance index over the control interval \( 0 \leq t \leq T \), subject to satisfying
the state equations
\[ \dot{x}_0(t) = k_F u(t) \]  \[31\]
and
\[ \dot{x}_M(t) = \frac{1}{k} x_0(t) - \frac{1}{k} x_M(t) . \]  \[32\]

The constraints imposed by the state equations were
\[ g_1 = x_0(t) - k_F u(t) = 0 \]  \[33\]
and
\[ g_2 = \tau \dot{x}_M(t) + x_M(t) - x_0(t) = 0 . \]  \[34\]

Constraints cause difficulty in analysis. The problem was changed from a constrained functional minimisation to an unconstrained one by introducing a constrained error measure
\[ H_c(x, \lambda, u, t) = H + \lambda^T g \]  \[35\]
where \( \lambda \) was a vector of Lagrange multipliers.
\[ H_c = (x_0 - x_0)^2 + au^2 + \lambda_1 (\dot{x}_0 - k_F u) + \lambda_2 (\tau \dot{x}_M + x_M - x_0) . \]  \[36\]

The Calculus of Variations showed that the following conditions were satisfied when the corresponding constrained performance index was minimised:

a) the control equations
\[ \frac{\delta H_c}{\delta u} = 0 \]  \[37\]
b) the state equations
\[ \frac{\delta H_c}{\delta \lambda} = 0 \]  \[38\]
c) the costate or Euler-Lagrange equations
\[ \frac{\delta H_c}{\delta x} - \frac{d}{dt} \left( \frac{\delta H_c}{\delta \dot{x}} \right) = 0 . \]  \[39\]

For the tracking control system, the control equation was
\[ \frac{\delta H_c}{\delta u} = 2au - k_F \lambda_1 = 0 \]  \[40\]
or
\[ u = \frac{k_F}{2a} \lambda_1 . \]  \[41\]

The state equations were
\[ \frac{\delta H_c}{\delta \lambda_1} = \dot{x}_0 - k_F u = 0 \]  \[42\]
and
\[ \frac{\delta H_c}{\delta \lambda_2} = \dot{x}_M + x_M - x_0 = 0 . \]  \[43\]
The costate equations were
\[
\frac{\partial H_c}{\partial x_0} - \frac{d}{dt} \left( \frac{\partial H_c}{\partial \dot{x}_0} \right) = -2x_D + 2x_0 - \lambda_2 - \dot{\lambda}_1 = 0 \quad [44]
\]
and
\[
\frac{\partial H_c}{\partial x_M} - \frac{d}{dt} \left( \frac{\partial H_c}{\partial \dot{x}_M} \right) = \lambda_2 - \tau \dot{\lambda}_2 = 0 \quad [45]
\]

From these equations we wished to find an expression for the control effort \( u \) in terms of \( x_D \) and \( x_M \) which were known or measurable quantities rather than in terms of the costate \( \lambda_1 \).

From equation 44,
\[
\lambda_2 = 2x_0 - 2x_D - \dot{\lambda}_1. \quad [46]
\]

From equation 43,
\[
x_0 = x_M + \tau \dot{x}_M \quad [47]
\]
so
\[
\lambda_2 = 2(x_M - x_D) + 2\tau \dot{x}_M - \dot{\lambda}_1. \quad [48]
\]

From equation 41,
\[
\lambda_1 = \frac{2a}{k_F} \dot{u} \quad [49]
\]
so
\[
\lambda_2 = 2(x_M - x_D) + 2\tau \dot{x}_M - \frac{2a}{k_F} \dot{u}. \quad [50]
\]

Integrating equation 45,
\[
\int \lambda_2 \, dt - \tau \lambda_2 = 0. \quad [51]
\]

Substituting the expression for \( \lambda_2 \) into this gave
\[
2 \int (x_M - x_D) \, dt + 2\tau x_M - \frac{2a}{k_F} \dot{u} - 2\tau (x_m - x_D) - 2\tau^2 \dot{x}_M + \frac{2\tau a}{k_F} \dot{u} = 0. \quad [52]
\]
Simplifying,
\[
k_F \int (x_M - x_D) \, dt + k_F \tau x_D - k_F \tau^2 \dot{x}_M - au + \tau \dot{au} = 0. \quad [52]
\]

Integrating and re-arranging,
\[
\tau au = \int u \, dt + k_F \int (x_D - x_M) \, dt + k_F \int \tau^2 x_M \, dt + k_F \int \tau x_M \, dt. \quad [53]
\]

The control equation was
\[
u = \frac{1}{\tau} \int u \, dt + \frac{k_F}{\tau a} \int (x_D - x_M) \, dt + \frac{k_F}{a} \int x_D \, dt + \frac{k_F}{a} \tau x_M. \quad [54]
\]

The control signal \( u \) was a function of \( x_D \), \( x_M \) and \( u \). The variables \( x_D \) and \( x_M \) were obtained from the system without difficulty. But the positive integral
feedback of the control signal $u$ was potentially a recipe for disaster. There was a danger that the control signal would zoom off exponentially.

Furthermore the initial value of the control signal $u(0)$ was not revealed by the Calculus of Variations. Transversality gave some limited information about conditions of the states and costates at the boundaries of the control interval:

$$
\begin{align*}
  x_D(t_o) &= \text{known} \\
  x_D(t_f) &= \text{unknown at } t_o \\
  x_M(t_o) &= \text{known} \\
  x_M(t_f) &= \text{unknown at } t_o \\
  \lambda_1(t_o) &= \text{free} \\
  \lambda_1(t_f) &= 0 \\
  \lambda_2(t_o) &= \text{free} \\
  \lambda_2(t_f) &= 0.
\end{align*}
$$

Attempts to integrate the state and costate equations simultaneously were thwarted by this Two point Boundary Value Problem, in which some of the boundary conditions were known at the initial time and the remainder at the final time. The solution (by simulation) could only be obtained by iteration.

One approach would be to guess initial values for the costates and run the simulation. The values of the costates would be examined after the run to find if they were zero. If not the initial values would be adjusted and another simulation run made. The initial values would be progressively adjusted until the terminal conditions were satisfied. The iterative adjustment of the initial values could result in its own instability problems, and a clever algorithm might be needed to achieve convergence to the desired terminal conditions.

8.6 SIMULATION OF EULER-LAGRANGE CONTROLLER

From the previous discussion, it was realised that certain problems could result from the use of a controller designed by the Euler-Lagrange equations in conjunction with Lagrange multipliers. Nevertheless, it was decided to evaluate the controller by simulation on an analog computer because "hands-on"
experience of the controller might have revealed solutions to the problems.

Certainly the initial values were not regarded as a serious problem. Suppose a simulation run was started with all states of the system at zero values. Then obviously the initial value of control effort would be zero also. Therefore a simulation run could be initiated and the response of the controller and the system to a step input or other driving function in the desired output state could be assessed.

After a few false starts, the control law (equation 54) was implemented by the scheme in fig 28. Previous attempts had included integrators which generated terms such as $\int x_d \, dt$; in response to a step input, these integrators had saturated after a few seconds, making the simulation useless.

The system itself, represented by equations 31 and 32, was implemented as shown in fig 29.

Numerical values were required for the parameters used in the simulation. For the flowrate transducer, the time constant $\tau$ was estimated to be 0.5 seconds.

The weighting factor $a$ could be varied over a wide range, say 0.1 up to 10, at the discretion of the designer. A nominal value of unity was used to give equal weighting to flowrate error and control effort. After seeing the results of the simulation, the weighting factor could be adjusted if necessary.

The linearised system gain was also needed. From equation 31,

$$x_o' = k_p \ u$$

or

$$\frac{\Delta x_o}{\Delta t} = k_p \ u.$$
The linear actuator selected for test could traverse its 4 inch stroke in 2 seconds. For part of the stroke, the ball valve was fully closed or fully open, and it was estimated that the ball valve would go from fully closed to fully open with only one inch travel on the actuator. This would occur in 0.5 seconds, so $\Delta t = 0.5$.

The flow rate changed from zero to maximum, so the normalised flow rate changed from zero to unity. The change in normalised flow rate was $\Delta \xi = 1$.

From equation 56,

$$k_p = \frac{\Delta \xi}{u \Delta t} = \frac{1}{1 \times 0.5} = 2.$$

This linearised value of system gain represented a band of operation where movement of the linear actuator resulted in opening or closing of the ball valve. At the ends of stroke of the linear actuator, where the ball valve was fully opened or closed, a highly nonlinear double ended saturation characteristic resulted.

Using these numerical values, the analog computer diagrams in fig 30 and fig 31 were produced for the controller and the system. The diagrams were patched up on the Sunderland Polytechnic EAI 180 Analog Computer, as a unified closed loop control system.

Initial values of all integrators were set to zero. The desired output $x_d$ was also set to zero. This solved the problem of selecting the initial values. The intention was to inject a step waveform into the input $x_d$, and observe how well the actual output $x_o$ tracked or followed the desired output. In fact this could not be done because after a short period of time, the control signal $u$ zoomed off into either positive or negative saturation. Steady operation could be obtained for up to 5 or 8 seconds. Doubtless if the amplifiers and integrators had been balanced, a longer steady period
could have been obtained. But the message was quite obvious. The Euler-Lagrange controller was unusable for an infinite time control period. The positive integral feedback used to generate the control signal made the controller itself unstable.

The Euler-Lagrange approach was abandoned at this stage.

8.7 SUMMARY

An optimal tracking controller was required, rather than a conventional feedback controller. A quadratic control term was incorporated in the error measure to moderate the control signal. The resulting performance index was to be minimised, subject to satisfying the state equations which acted as constraints on the minimisation.

The problem was changed from a constrained functional minimisation to an unconstrained one by introducing the state equations into the error measure using Lagrange multipliers. The Calculus of Variations yielded three sets of differential equations, the control equations, the state equations, and the costate or Euler-Lagrange equations. From these equations, an expression for the control signal was derived as a function of the system states and of the control signal itself, a potentially unstable situation. The initial value of the control signal was not revealed by the Calculus of Variations, but analog computer simulation could give an answer by an iterative procedure.

The problem of finding an initial value for the control signal was avoided when the system and controller were tested on the analog computer, by starting with all states and signals at zero. However it was found that the control signal ran off exponentially into saturation after a short period of time. This was due to the positive integral feedback used in the generation of the control signal. The Euler-Lagrange equation can undoubtedly yield optimal controls over a finite time interval, but its unstable controllers are not
suitable for the infinite time period. Another technique was needed to give an optimal tracking controller.
CHAPTER 9
IMPLEMENTATION OF PERFORMANCE INDEX
USING HAMILTON-JACOBI EQUATION

9.1 INTRODUCTION

The Euler-Lagrange equation had failed to yield an infinite-time stable controller. This chapter describes an alternative strategy, the Hamilton-Jacobi equation. The Hamiltonian is formed and substituted into the Hamilton-Jacobi equation. For the infinite time period, the controller uses state feedback with time-invariant coefficients. The coefficients are determined numerically by the Newton-Raphson method.

Analysis of the coefficients show that an inaccessible system state is needed. A low order Luenberger observer is used to reconstruct it.

9.2 DERIVATION OF CONTROL LAW USING HAMILTON-JACOBI EQUATION

An alternative to variational procedures for determining an optimal control was to use the method of dynamic programming. In this a minimum error function would be defined as an explicit function of the state variables and time. Its value at each point in state space and time would be equal to the least value of the error index that could be achieved during the remainder of the control interval. From a knowledge of the minimum error function, the optimal control could be obtained.

The first step in this method was to determine the functional form and coefficient values of the minimum error function. This was done by setting up a partial differential equation, the Hamilton-Jacobi equation, whose solution was the minimum error function. ( ref. 4,5,6 )

This method was limited in its application to linear systems with quadratic error indices. But for the system under consideration, it gave a simple
solution to the linear regulator problem, and by extension, the linear servomechanism problem.

The linearised state equations of the system were

\[ \dot{x}_0 = k_F u \]  \[57\]

and \[ \dot{x}_m = \frac{1}{c} x_d - \frac{1}{c} x_m \].  \[58\]

These state equations were expressed more elegantly in vector notation

\[ \dot{x} = Ax + bu \] \[59\]

where \( A \) was an \( n \times n \) matrix and \( b \) was an \( n \) vector. Writing the vector equation in full,

\[
\begin{bmatrix}
\dot{x}_0 \\
\dot{x}_m
\end{bmatrix} =
\begin{bmatrix}
0 & 0 \\
\frac{1}{c} & -\frac{1}{c}
\end{bmatrix}
\begin{bmatrix}
x_0 \\
x_m
\end{bmatrix} +
\begin{bmatrix}
k_F \\
0
\end{bmatrix} u .
\] \[60\]

The cost function to be minimised was

\[ J = \frac{1}{2} \int_0^\infty \begin{bmatrix} x^T Q x + ru^2 \end{bmatrix} dt , \] \[61\]

where \( Q = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} . \)

The optimal cost function was denoted by the symbol \( J^* \). Partial derivatives of the optimal cost function were represented as shown:

\[ J^*_t = \frac{\partial J^*}{\partial t} \]

and \[ J^*_x = \frac{\partial J^*}{\partial x} . \]

To use the Hamilton-Jacobi equation, the Hamiltonian had to be formed

\[ H(x,u,J^*_x,t) = \frac{1}{2} x^T Q x + \frac{1}{2} ru^2 + J^*_x [Ax + bu] . \] \[62\]

The necessary condition for \( u(t) \) to minimise \( H \) was

\[ \frac{\partial H}{\partial u} = 0 . \] \[63\]

\[ \frac{\partial H}{\partial u} = ru + J^*_x b = 0 \] \[64\]
so the optimal control was

\[ u = -r^{-1} J^* x \quad \text{and} \quad b = -r^{-1} b^T J^* x \]  

[65]

since \( J^* x b = b^T J^* x \) is a scalar quantity.

This procedure could also have yielded a maximum instead of a minimum, so the nature of this extremum had to be established. The second partial derivative was

\[ \frac{\partial^2 \mathcal{H}}{\partial u^2} = r. \]

The control weighting factor \( r \) was a positive quantity and \( \mathcal{H} \) was a quadratic form in \( u \), so the control \( u = -r^{-1} b^T J^* x \) did minimise \( \mathcal{H} \) globally.

Substituting the optimal control into the Hamiltonian,

\[ \mathcal{H}(x, J^*, t) = \frac{1}{2} x^T Q x - \frac{1}{2} r^{-1} J^* x b b^T J^* x + J^* x A x. \]  

[66]

The Hamilton-Jacobi equation was

\[ 0 = J^* _t + \frac{1}{2} x^T Q x - \frac{1}{2} r^{-1} J^* x b b^T J^* x + J^* x A x. \]  

[67]

Since the system and the \( Q \) and \( r \) terms were time invariant, and since the optimisation was for a period of infinite duration, it followed that \( J^* \) would depend only on the initial state. This implied that

\[ J^*_t = \frac{\partial J^*}{\partial t} = 0, \]

so the Hamilton-Jacobi equation assumed the simpler form

\[ 0 = \frac{1}{2} x^T Q x - \frac{1}{2} r^{-1} J^* x b b^T J^* x + J^* x A x. \]  

[68]

The procedure then taken was to assume a solution. It was known that the minimum cost for linear regulators was a quadratic function of the states, so it seemed reasonable to guess as a solution the form

\[ J^* = \frac{1}{2} x^T P x \]  

[69]

where \( P \) was a real symmetric positive-definite matrix, whose value had to be determined.
The partial derivative was
\[ \frac{\partial J}{\partial x} = Px \, . \]

The Hamilton-Jacobi equation was
\[ 0 = \frac{1}{2}x^TQx - \frac{3}{2}r^{-1}(Px)^Tb\, b^T(Px) + (Px)^TAX \, . \]  

[70]

Remembering that the P matrix was symmetric,
\[ P = P^T \, , \]

and the Hamilton-Jacobi equation became
\[ x^T \left[ \frac{1}{2}Q + \frac{1}{2}A^TP + \frac{3}{2}r^{-1}Pbb^TP \right] x = 0 \, . \]  

[71]

This said that for any non-zero \( x(t) \), the matrix P had to satisfy the \( n(n+1)/2 \) algebraic equations
\[ Q + PA + A^TP - r^{-1}Pbb^TP = 0 \]  

[72]

These equations had to be solved for the elements of P so that the control could be computed from
\[ u = -\frac{b^T}{x^P x^{-1}} = -b^TPr^{-1} \, . \]  

[73]

9.3 NUMERICAL DETERMINATION OF OPTIMAL CONTROL

Substituting numerical values into equation 72 gave
\[
\begin{bmatrix}
1 & 0 \\
0 & 0
\end{bmatrix} + \begin{bmatrix}
p_{11} & p_{12} \\
p_{21} & p_{22}
\end{bmatrix} \begin{bmatrix}
0 & 0 \\
1 & -1
\end{bmatrix} \begin{bmatrix}
p_{11} & p_{12} \\
p_{21} & p_{22}
\end{bmatrix} - r^{-1}P_{kk} \, [k, k_0] \, P = 0. \]  

[74]

\[ \tau \begin{bmatrix}
1 & 0 \\
0 & 0
\end{bmatrix} + \begin{bmatrix}
p_{12} & -p_{12} \\
p_{22} & -p_{22}
\end{bmatrix} + \begin{bmatrix}
p_{21} & p_{22} \\
-p_{21} & -p_{22}
\end{bmatrix} - \begin{bmatrix}
p_{11} & p_{12} \\
p_{21} & p_{22}
\end{bmatrix} \, k_F \, [k_0] \, Pr^{-1} \tau = 0. \]  

[75]

\[ \tau \begin{bmatrix}
1 & 0 \\
0 & 0
\end{bmatrix} + \begin{bmatrix}
p_{12} & -p_{12} \\
p_{22} & -p_{22}
\end{bmatrix} + \begin{bmatrix}
p_{21} & p_{22} \\
-p_{21} & -p_{22}
\end{bmatrix} - \begin{bmatrix}
p_{11} & 0 \\
p_{21} & 0
\end{bmatrix} \, k_F^2 \, \tau \, r^{-1} \, \tau = 0. \]  

[76]

\[ \tau \begin{bmatrix}
1 & 0 \\
0 & 0
\end{bmatrix} + \begin{bmatrix}
p_{12} & -p_{12} \\
p_{22} & -p_{22}
\end{bmatrix} + \begin{bmatrix}
p_{21} & p_{22} \\
-p_{21} & -p_{22}
\end{bmatrix} - \begin{bmatrix}
p_{11} & 0 \\
p_{21} & 0
\end{bmatrix} \, k_F^2 \, \tau \, r^{-1} \, \tau = 0. \]  

[77]

For simplicity, put \( g = k_F^2 \, \tau \, r^{-1} \).

P was symmetrical so \( p_{12} = p_{21} \).
The problem therefore reduced to three algebraic equations
\[ Z + 2p_{12} - gp_{11}^2 = 0 . \]  \[ 77a \]
\[ -p_{12} + p_{22} - gp_{11}p_{12} = 0 . \]  \[ 77b \]
\[ 2p_{22} + gp_{12}^2 = 0 . \]  \[ 77c \]

These equations were manipulated to obtain a quartic equation in one of the elements. From equation 77c,
\[ p_{22} = -\frac{gp_{12}^2}{2} . \]  \[ 78 \]
Substituting this in equation 77b,
\[ p_{12} + \frac{gp_{12}^2}{2} + gp_{11}p_{12} = 0 . \]  \[ 79 \]
From equation 77a,
\[ p_{12} = \frac{gp_{11}^2 - Z}{2} . \]  \[ 80 \]
Substituting into equation 79 yielded an equation in which \( p_{11} \) was the only variable,
\[ \frac{gp_{11}^2 - Z}{2} + \frac{g}{2} \left( \frac{gp_{11}^2 - Z}{4} \right)^2 + gp_{11} \frac{gp_{11}^2 - Z}{2} = 0 . \]  \[ 81 \]
Manipulating,
\[ 4gp_{11}^2 - 4Z + g^3p_{11}^4 - 2rg^2p_{11}^2 + r^2g + 4g^2p_{11}^3 - 4grp_{11} = 0 . \]  \[ 82 \]
So the quartic equation in \( p_{11} \) was
\[ g^3p_{11}^4 + 4g^2p_{11}^3 + (4g - 2rg^2)p_{11}^2 - 4grp_{11} + (r^2g - 4Z) = 0 . \]  \[ 83 \]

Determining roots of quartic equations can be quite tedious. Accordingly, a numerical solution using the Newton-Raphson method was devised. A computer program in the BASIC language was written to yield the root. But once the root was available, the remaining elements of the P matrix could be determined. And once the P matrix was known, the feedback coefficients could be computed from equation 37. So the program was expanded to give optimal feedback coefficients for a given set of system parameters. The final extension to the program was to select a wide combination of system parameters and determine optimal feedback coefficients for all of these. For a map of system parameters, the optimal control could then be determined at a glance.
9.4 FINDING ROOTS BY NEWTON-RAPHSON METHOD

Before using the Newton-Raphson method, the approximate location of the root is needed. The gradient of the function at that location is then used to "home-in" and provide a closer approximation. The gradient at this closer approximation is again used to "home-in" on the true location of the root. This iterative procedure can be pursued until the root is known to any degree of accuracy. (ref. 7)

Suppose that \( x_0 \) is an approximation to the root of the equation \( f(x) = 0 \), then the Taylor series expansion of \( f(x) \) in the neighbourhood of \( x_0 \) is given by

\[
f(x) = f(x_0) + (x - x_0) f'(x_0) + (x - x_0)^2 \frac{f''(x_0)}{2} + \ldots. \tag{84}\]

If \( x \) is a value of \( x \) nearer to the exact root, then \( (x - x_0) \) will be small, \( f(x) \) is approximately zero, and providing that \( f''(x_0) \) is not too large the Taylor series may be replaced by

\[
0 = f(x_0) + (x - x_0)f'(x_0) \tag{85}
\]

from which

\[
x = x_0 - \frac{f(x_0)}{f'(x_0)}. \tag{86}\]

This is the Newton-Raphson rule, and repeating for successive approximations \( x, x, x, \ldots \) leads to the iterative relationship

\[
x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \tag{87}
\]

where \( x_n \) is the \( n \)th iterate.

A geometrical interpretation is shown in Fig 32. The ordinate at \( A_0 \) is \( f(x_0) \) and since \( \tan \theta_0 = f'(x_0) \) the tangent at \( A_0 \) intersects the \( x \)-axis at the point \( x \). The process illustrated is seen to be convergent.
9.5 ALGORITHM FOR NEWTON-RAPHSON METHOD

The roots of the polynomial
\[ f(p_{11}) = p_{11}^4 + \frac{4}{g} p_{11}^3 + \frac{4 - 2\tau g}{g^2} p_{11}^2 - \frac{4\tau}{g^2} p_{11} + \frac{\tau}{g^3} (g - 4) = 0 \] were required

To simplify the computation the polynomial was expressed as
\[ f(p_{11}) = (((p_{11} + \frac{4}{g}) p_{11} + \frac{4 - 2\tau g}{g^2}) p_{11} - \frac{4\tau}{g^2}) p_{11} + \frac{\tau}{g^3} (\tau g - 4) = 0. \]

This was written in BASIC as
\[ FP11 = (((P11 + A) P11 + B) P11 + C) P11 + D \]
where
\[ A = 4/G \]
\[ B = (-2*TAU*G + 4)/G/G \]
\[ C = -4*TAU/G/G \]
\[ D = (TAU*G -4)*TAU/G/G/G \]
and
\[ G = KF*KF*TAU/R. \]

The gradient was
\[ f'(p_{11}) = 4p_{11}^3 + \frac{12}{g} p_{11}^2 + \frac{8 - 4\tau g}{g^2} p_{11} - \frac{4\tau}{g^2}. \]

This polynomial was expressed as
\[ f'(p_{11}) = (((4p_{11} + 12/g) p_{11} + 8 - 4\tau g)/g^2) p_{11} - \frac{4\tau}{g^2}. \]

The expression was written in BASIC as
\[ GRAD = (((4*P11 + E) P11 + F) P11 + C \]
where
\[ E = 3*A \]
\[ F = 2*B. \]

The computer program first calculated parameters A,B,C,D,E and F from the data input. Functions FP11 and GRAD were then calculated. The program is listed in Appendix 1. It was ran on a GenRad microprocessor development system used to support Z80-based tension control systems at the Warner Electric UK plant. The language used was a sub-set of BASIC called "8085 S-BASIC".
The Newton-Raphson rule (equation 87) was written in BASIC as
\[ \text{NEWP}11 = -\text{FP}11/\text{GRAD} + \text{P}11. \]
This equation was applied repeatedly until the difference between consecutive estimates was sufficiently small. The criterion chosen was that the modulus of the difference between consecutive estimates should be less than 0.01. To guard against divergent or oscillatory iteration, a limit of 100 iterations was also incorporated after which the computation would cease. For each computation, the number of iterations was printed out so that an invalid root could be identified. In fact the number of iterations to yield a root of the desired accuracy varied between one and nine.

9.6 CALCULATION OF FEEDBACK COEFFICIENTS

Having computed \( \text{p}11 \), the remaining elements of the \( \text{P} \) matrix could be calculated without difficulty. The optimal control was given by
\[
\text{u} = -b^{T}p_{x}r^{-1}
\]
or
\[
\text{u} = -\begin{bmatrix} k_{p} & 0 \end{bmatrix} \begin{bmatrix} p_{11} & p_{12} \end{bmatrix} \begin{bmatrix} x_{0} \\ x_{M} \end{bmatrix} r^{-1}.
\]

\[
\text{u} = -\frac{k_{p}}{r} \begin{bmatrix} p_{11} & p_{12} \end{bmatrix} \begin{bmatrix} x_{0} \\ x_{M} \end{bmatrix}
\]

\[
\text{u} = -\frac{k_{p}}{r} (p_{11} x_{0} + p_{12} x_{M})
\]

Clearly, only elements \( p_{11} \) and \( p_{12} \) were of importance to the optimal control and the other elements could be ignored.

The element \( p_{12} \) was calculated from equation 80
\[ p_{12} = \frac{g_{p}11^{2} - c}{2}. \]

This was written in BASIC as
\[ \text{P}12 = (G*\text{P}11*\text{P}11 - \text{T}AU)/2. \]

The two feedback coefficients were written in BASIC as
\[
k_{0}, \text{ coefficient of } x_{0} \quad K_{0} = K_{F} * \text{P}11 / R
\]
\[
k_{M}, \text{ coefficient of } x_{M} \quad K_{M} = K_{F} * \text{P}12 / R.
\]
The feedback of both state variables is negative in sign.
9.7 ANALYSIS OF OPTIMAL CONTROL

A selection of system parameters were put into the computer program:

- gain $k_p : 0.5, 1, 1.5, 2,$
- time constant $\tau : 0.1, 0.2, 0.5, 1,$
- control weighting factor $r : 0.1, 0.2, 0.5, 1, 2, 5, 10.$

These 112 sets of system parameters were evaluated by the computer program and optimal feedback coefficients were printed out for each set. The print-out is given in Appendix 2. This "shot-gun" approach might be considered clumsy but afforded a comprehensive overview of the optimal control. One conclusion drawn from this analysis was quite significant.

All coefficients of $x_M$ were virtually zero, whereas coefficients of $x_0$ were "reasonable" values. The physical interpretation of this result was quite simple; the optimal control was striving to minimise the actual value $x_0$, but had no interest whatsoever in the measured value $x_M$. This result was quite understandable but caused difficulty in realisation. The actual value $x_0$ was not directly measurable and could only be accessed via the measured value $x_M$. Apart from this problem, the results obtained from the Hamilton-Jacobi equation seemed satisfactory so it was decided that an observer should be used to reconstruct the missing state.

Other conclusions drawn from the analysis were also of interest. Neither the system gain $k_p$ nor the time constant $\tau$ had any influence on the optimal feedback coefficients. Only the control weighting factor $r$ had any affect, according to the law

$$\text{Optimal Feedback Coefficient of } x_0 = \frac{1}{\sqrt{r}}.$$  

9.8 RECONSTRUCTION OF INACCESSIBLE STATE

For the estimation of system state variables from noisy observation of output variables, the Kalman filter is undoubtedly the best technique. Where no measurement noise is present, a limited form of the Kalman filter could of
course be used. However it is possible to develop a state estimator for the measurement noise-free case with considerably less complexity than the Kalman filter.

Both Kalman and Luenberger have suggested methods for state variable reconstruction. Kalman's method yields a full observer which reconstructs an estimate for the entire state vector. Some of these state variables may be already available by direct measurement so parts of the full observer may be redundant. To deal with such cases, Luenberger suggested a low order observer. (ref. 5)

The system under investigation was

\[ x = Ax + bu \]  \hspace{1cm} [96]

or, in full,

\[
\begin{bmatrix}
    x_0 \\
    x_M
\end{bmatrix}
= \begin{bmatrix}
    0 & 0 \\
    \frac{1}{T} & -\frac{1}{T}
\end{bmatrix}
\begin{bmatrix}
    x_0 \\
    x_M
\end{bmatrix}
+ \begin{bmatrix}
    k_F \\
    0
\end{bmatrix}
\begin{bmatrix}
    u
\end{bmatrix} .
\]  \hspace{1cm} [97]

The accessible state variables could be described by

\[ y = c^T x \]  \hspace{1cm} [98]

or, in full,

\[ y = \begin{bmatrix}
    0 & 1
\end{bmatrix}
\begin{bmatrix}
    x_0 \\
    x_M
\end{bmatrix} , \]  \hspace{1cm} [99]

because the second order system had only one measurable output. A theorem presented by Luenberger stated that a state observer of order \( n - m \) could be constructed having arbitrary eigenvalues for any \( n \)th order completely controllable linear time invariant system having \( m \) linearly independent outputs. Accordingly, a first order observer was synthesised to create an estimate of the missing state \( x_0 \). The dynamics of the linear observer were described by

\[ \dot{\hat{x}} = D\hat{x} + F^T \hat{x} + Bu , \]  \hspace{1cm} [100]

that is to say, the observer was a dynamic system driven by the accessible system states and the system input.
The characteristic polynomial of the system was

$$\det(A - \lambda I) = \begin{vmatrix} -\lambda & 0 \\ \frac{1}{\tau} & -\lambda \end{vmatrix} = \lambda(\lambda + \frac{1}{\tau}) = 0$$

so the eigenvalues were $\lambda_1 = 0$, $\lambda_2 = \frac{1}{\tau}$.

The observer eigenvalue had to be made larger than the system eigenvalues to ensure that any errors between the observer and the system would rapidly decay. However it was known that if the observer eigenvalue was too large, excessively large signals would result and saturation would occur. This would reduce rather than increase the tracking accuracy of the observer.

Accordingly, the observer eigenvalue was set to

$$\lambda_0 = p\lambda_2 = \frac{p}{\tau}.$$  \[102\]

A numerical value of $3 < p < 10$ was expected to give reasonable observer performance.

Some numerical values could now be assigned to the observer equation 100.

$$D = \lambda_0 = -\frac{p}{\tau}.$$  \[103\]

Only $x_M$ could be measured, and $x_0$ was inaccessible, so

$$F^T = [0 \ 1].$$

The remaining parameter $E$ of the observer equation was not obtained so readily. It was necessary to adjust the observer so that

$$\mathbf{g} = \mathbf{T}^T \mathbf{x}.$$  \[104\]

Clearly,

$$\dot{\mathbf{g}} = \mathbf{T}^T \mathbf{Ax} = \mathbf{T}^T \mathbf{T}^T \mathbf{x} + \mathbf{T}^T \mathbf{bu}.$$  \[105\]

The observer equation was

$$\dot{\mathbf{g}} = D\mathbf{g} + F^T \mathbf{x} + \mathbf{E} \mathbf{u}$$

so

$$\dot{\mathbf{g}} = \mathbf{T}^T \mathbf{x} + \mathbf{F}^T \mathbf{x} + \mathbf{E} \mathbf{u}.$$  \[107\]

From equations 106 and 108,

$$\mathbf{T}^T \mathbf{Ax} + \mathbf{T}^T \mathbf{bu} = (\mathbf{D}^T + \mathbf{F}^T) \mathbf{x} + \mathbf{E} \mathbf{u}.$$  \[109\]
For nontrivial \( x(t) \) and \( u(t) \), it was necessary that

\[
T^T A = DT^T + F^T
\]  
[110a]

and \( E = T^T b \). \[110b\]

If the \( T \) vector could be solved, the \( E \) parameter would be known and the observer could be built.

Numerical values were inserted in equation 110a:

\[
\begin{bmatrix} t_1 & t_2 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ \frac{1}{\tau} & \frac{1}{\tau} \end{bmatrix} = -\frac{\tau}{\chi} \begin{bmatrix} t_1 & t_2 \end{bmatrix} + [0 \ 1] .
\]
[111]

This gave two simultaneous equations,

\[
\frac{t_2}{\chi} = -\frac{\chi}{\tau} t_1
\]  
[112a]

and \( -\frac{t_2}{\chi} = -\frac{\chi}{\tau} t_2 + 1 \)
[112b]

From the second equation,

\[
pt_2 - \tau = t_2 .
\]

\[
t_2(p - 1) = \tau .
\]

\[
t_2 = \frac{\tau}{p-1} .
\]
[113]

From the first equation,

\[
t_1 = -\frac{t_2}{\chi}.
\]

\[
t_1 = -\frac{\tau}{p(p - 1)} .
\]
[114]

The \( T \) vector was

\[
T^T = \begin{bmatrix} t_1 & t_2 \end{bmatrix} = \left[ -\frac{\tau}{p(p - 1)} \frac{\tau}{(p - 1)} \right] .
\]
[115]

The desired \( E \) parameter was

\[
E = T^T b = \begin{bmatrix} t_1 & t_2 \end{bmatrix} \begin{bmatrix} k_F \\ 0 \end{bmatrix} = t_1 k_F .
\]
[116]

\[
E = -\frac{k_F \tau}{p(p - 1)} .
\]
[117]
The complete observer equation was obtained by substituting the $D$ and $E$ parameters and the $F$ vector into equation 100

$$\dot{\xi} = - \frac{p}{\tau} \xi + [0 \ 1] \begin{bmatrix} x_0 \\ x_M \end{bmatrix} - \frac{k_p \tau}{p(p-1)} u,$$  \[118\]

or

$$\dot{\xi} = - \frac{p}{\tau} \xi + x_M - \frac{k_p \tau}{p(p-1)} u.$$  \[119\]

Having synthesised the observer, the estimate of the inaccessible state could be created. The observer state variable had been defined by equation 104 to be a linear combination of the system state variables,

$$\xi = \frac{1}{t} x.$$  \[120\]

$$\xi = \begin{bmatrix} t_1 & t_2 \end{bmatrix} \begin{bmatrix} x_0 \\ x_M \end{bmatrix}.$$  \[121\]

$x_M$ was the measurable system state variable but the inaccessible state $x_0$ could be replaced by $\hat{x}_0$ which was the estimate of the state $x_0$.

$$\hat{x} = \begin{bmatrix} t_1 & t_2 \end{bmatrix} \begin{bmatrix} \hat{x}_0 \\ x_M \end{bmatrix}.$$  \[122\]

Re-arranging,

$$\hat{x}_0 = \begin{bmatrix} \frac{1}{t_1} \xi - \frac{t_2}{t_1} x_M \end{bmatrix}.$$  \[123\]

Therefore the estimate of the inaccessible state was

$$\hat{x}_0 = - \frac{p(p-1)}{\tau} \xi + px_M.$$  \[124\]

The low order Luenberger observer was completely described by the observer equation 119 and the reconstruction equation 124. The observer is shown in block diagram form in figure 33, connected to the system and the Hamilton-Jacobi controller.

9.9 SUMMARY

The state equations and the performance index were formulated using vector notation. The Hamiltonian was formed, and minimising with respect to the
control signal gave an intractable expression for the optimal control in terms of a partial derivative of the optimal cost function. This expression allowed use of the Hamilton-Jacobi equation. The Hamilton-Jacobi equation was simplified because the optimisation was for the infinite time interval, implying that the optimal cost function was time-invariant. An assumption was made that the optimal cost was a quadratic function of the states, and the Hamilton-Jacobi equation decomposed to a set of algebraic equations in the elements of the optimal cost. If these algebraic equations were solved, the optimal cost would be known, and the originally intractable expression would yield the optimal control.

The three algebraic equations for this system were manipulated to give a fourth order polynomial in one of the elements. The relevant root was obtained numerically by the Newton-Raphson method. An algorithm was written in BASIC to compute the root. Having obtained the first element, the remaining elements of the optimal cost were also solved by the computer, enabling the optimal control to be calculated.

The optimal control took the form of state variable feedback and the computer program worked out the feedback coefficients for a wide variety of system parameters. Analysis of the results showed that the measured state was ignored, and the optimal control only used the output state which was inaccessible for direct measurement. This was understandable because the performance index had called for the output state to be minimised. The results were otherwise promising so a decision was taken to reconstruct the missing state with an observer.

A low order Luenberger observer was designed with attention to the placement of the observer eigenvalue. The dynamic observer equation and the algebraic reconstruction equation yielded an estimate of the missing output state.
An independent method was now needed to confirm the operation of the Hamilton-Jacobi controller and the Luenberger observer.
CHAPTER 10
TESTING OPTIMAL CONTROLLER ON ANALOG COMPUTER

10.1 INTRODUCTION
The Hamilton-Jacobi controller and the necessary observer which were designed in the previous chapter must now be tested. The analog computer is used to model the system and check that the Performance Index is indeed minimised by the controller. All of the system states are available on the computer, including the "inaccessible" state $x_0$. This allows the performance of the observer to be assessed very easily. The reaction of the observer and the system to measurement noise is discovered. The saturation characteristic of the control valve is simulated to find if any problems will occur at system start-up.

10.2 MEASUREMENT OF PERFORMANCE INDEX
Optimal state feedback coefficients had been derived from the Hamilton-Jacobi equation. As a check on these results, an independent simulation was performed on an analog computer.

The system was

$$\dot{x}_0 = k_F u = 2u$$

and

$$\dot{x}_M = \frac{1}{c} x_0 - \frac{1}{c} x_M = 2x_0 - 2x_M.$$

The system was modelled by amplifiers A00, A01 and A02 in the analog computer diagram (fig 34), which was patched up on the Sunderland Polytechnic EAI 580 Analog Hybrid Computer.

The Hamilton-Jacobi equations had shown that feedback of the single state variable $x_0$ was sufficient to minimise the performance index. But as a check, a state feedback controller was devised on the analog computer which used both state variables:

$$u = -k_0 x_0 - k_M x_M.$$
The system was configured initially as a regulator, with
\[ x_D = 0. \]
The controller was represented by amplifier A08.
Coefficients \( k_0 \) and \( k_M \) were represented by potentiometers P06 and P07. These were hand-alterable potentiometers (rather than servo-alterable) for ease of adjustment.

Because the system was set up as a regulator, an initial disturbance was needed to which the system could respond.
Initial conditions were established by potentiometers P20 and P21 as
\[ x_0(0) = x_M(0) = 0.5. \]

The chosen Performance Index was
\[ J = \int_0^\infty [(x_D - x_0)^2 + ru^2]. \, dt. \]
In this instance \( x_D = 0 \). The period of integration could be terminated after the dynamics had ceased. A control weighting factor \( r \) of unity was arbitrarily selected, to give
\[ J = \int_0^T [x_0^2 + u^2]. \, dt. \]
This performance index was patched up on the analog computer, using two multipliers M42 and M62, and an integrator A10.

All the necessary features were now present to assess the performance index for various feedback coefficients. When the simulation was run, amplifier A10 built up a voltage which became steady after a few seconds (in real time), and the performance index could be measured. The feedback coefficient potentiometers P06 and P07 could then be adjusted in the hope that on the next simulation run, the performance index would be smaller.

Chart recordings of the various signals were obtained using a Hewlett Packard 7004B X - Y Recorder. To drive the X channel, a time base was generated by amplifier A11 on the analog computer.
To accelerate the determination of the optimal feedback coefficients, some of the advanced features of the EIA 580 computer were used. Firstly the time scale was increased by a factor of 1000, so a simulation run would be completed in about 5 milliseconds. Then the computer was placed in the Repetitive Operation mode, so that simulation runs were performed every 5 milliseconds. The signals representing \( u, x_0, x_M \) and \( J \) were fed to the EAI Repetitive Operation Display Unit model 34-035. This meant that if any changes were made to the feedback coefficient potentiometers P06 and P07, the effect on the system and the performance index could immediately be visualised.

The minimum exhibited by the performance index was fairly shallow and the true minimum could not be determined very easily from the display tube. To assist with this, the length of the simulation run was increased from 5 milliseconds to 1 second. For 99.5% of the time, amplifier A05 held a steady voltage corresponding to the performance index. This voltage was connected to a digital voltmeter to get a high resolution reading. By this ruse, the true minimum could readily be determined.

The results of these tests are shown in fig 35, which charts the value of the performance index \( J \) against various values of the feedback coefficients \( k_0 \) and \( k_M \). The global optimum was

\[
k_0 = 0.97, \quad k_M = 0, \quad J = 0.122.
\]

The location of the global optimum agreed with that predicted by the Hamilton-Jacobi equation and the computer program in Appendix 1. This agreement between two independent methods was most welcome.

Fig 35 gives another very important piece of information which was not revealed by the Hamilton-Jacobi equation. A local optimum is found at

\[
k_0 = 0, \quad k_M = 0.55, \quad J = 0.134.
\]

The global optimum used feedback from the inaccessible state \( x_0 \), implying
that an observer would be needed to reconstruct the state. But the local optimum tells us that feedback from the measurable state $x_M$ will also give quite good results. Comparing the performance indices for the two optima shows that the local optimum is only 10% worse than the global optimum. A sub-optimal controller could be created very easily indeed.

Chart recordings of system behaviour are shown in fig 36 for the global optimum and in fig 37 for the local optimum.

10.3 LUENBERGER OBSERVER

Testing the observer was very simple indeed because the inaccessible state $x_0$ was available in the analog computer simulation, and could be compared against the reconstructed state $\hat{x}_0$.

The observer and reconstruction equations were

$$\dot{x} = - \frac{p}{\tau} x + x_M - \frac{k_p \tau}{p(p-1)} u,$$

and

$$\dot{\hat{x}}_0 = \frac{p(p-1)}{\tau} x + px_M.$$

A factor $p = 3$ was chosen to give good tracking without saturation problems, so the equations became

$$\dot{x} = - 6 x + x_M - \frac{1}{6} u,$$

and

$$\dot{\hat{x}}_0 = 12 x + 3x_M.$$

These equations were patched up on the analog computer as shown in fig 38.

It is worth noting that if a factor $p = 10$ was used to attain higher accuracy, then $\hat{x}_0 = 180 x + 10x_M$. The high gain of 180 would clearly cause saturation in the real system.

Having established that the parameters $k_0 = 1$ and $k_M = 0$ minimised the performance index, these parameters were programmed into the analog computer by adjusting potentiometers P06 to zero and P07 to 0.1.
The controller was reconfigured as a tracking regulator:

\[ u = x_D - x_0. \]

A driving function of a step input was injected into amplifier A08 as shown in fig 38. The magnitude of the step was adjusted by potentiometer P08 so that saturation did not occur in any of the subsequent stages of the analog computer.

Initial system states were set to zero

\[ x_0(0) = x_M(0) = 0, \]

so the initial conditions on amplifiers A00 and A01 (fig 34) were reduced to zero.

When a simulation run was made in response to the step input, it was seen with some delight that the observed estimated state \( \hat{x}_0 \) tracked the real "inaccessible" state \( x_0 \) exactly. On the X - Y chart recording in fig 39, it was impossible to distinguish between the two.

To examine the sensitivity of the observer to variations in the coefficient of \( u \), further simulation runs were made with P30 set to 0 and 0.5 instead of the specified 0.167. The zero coefficient of \( u \) was of particular interest because it represented a "stripped-down" observer which was driven purely by the accessible system state instead of by both the accessible system state and the system input. These other runs are also recorded in fig 39 but neither is a satisfactory estimate. However it must be admitted that the estimate from the "stripped-down" observer with P30 = 0 is closer to the "inaccessible" state \( x_0 \) than is the measured state \( x_M \).

The setting of P30 was adjusted to the specified 0.167, so the estimate \( \hat{x}_0 \) again gave perfect tracking of the state \( x_0 \). The system at this stage still used feedback of the inaccessible state \( x_0 \) to close the loop. The state \( x_0 \) was disconnected from potentiometer P07 and the estimate \( \hat{x}_0 \) was connected in its place, so the loop was closed by an estimated state generated by the
Further simulation runs soon proved that the performances with real state feedback and with estimated state feedback were indistinguishable.

10.4 EFFECT OF MEASUREMENT NOISE

A problem with Luenberger observers is their susceptibility to measurement noise. This is made clear in fig 33. If the measured state $x_M$ is contaminated by noise, the noise is amplified by the factor $p$ and added to the estimated output $\hat{x}_0$. There are no dynamics and no filtering in the signal path whatsoever.

To discover the effect of noise on this observer, multi-level pseudo-random noise from a Hewlett Packard 54410A Analog to Digital Converter was injected into the input of amplifier A34 which added the noise to the signal $x_M$. The noise contaminated signal was available on amplifier A33. The amplitude of the noise was measured by a Hewlett Packard 400EL A.C. Voltmeter. The values had to be estimated because the needle of the voltmeter was flickering, due to the low frequency content of the noise. The noise input to amplifier A34 was 300 millivolts r.m.s. (± 5%). The amplified noise at the output of the observer A38 was 900 millivolts r.m.s. (± 5%). This was expected because the factor $p$ was 3.

Fig 40 shows the effect of measurement noise upon the observer output during a step response. The feedback loop was closed with the real state $x_0$. Firstly the noise alone was charted. The recording of the noise-contaminated state $x_M$ showed that the amplitude of the noise had not been changed at that stage. However the estimated state $\hat{x}_0$ contained a much higher noise level. Despite this, it is clear that the average of the estimate is still very close to the real state $x_0$, which is also shown. (Keen-eyed critics will notice that there is no correlation between the noise peaks on the three noise-contaminated signals. This is because the recordings were made sequentially instead of simultaneously.)
Now the feedback loop could be closed not with the real state $x_0$ but with the estimated state $\hat{x}_0$. The step response using feedback from the observer is shown in fig 41. The measurable state $x_M$ was contaminated by 300 millivolts r.m.s. of noise. The noise was amplified by the observer so the estimated state $\hat{x}_0$ contained 900 millivolts r.m.s. of noise, which was fed back to the controller. The output of the controller was the control signal $u$, which contained 920 millivolts r.m.s. (5%) of noise. This noisy control signal was applied to the system where the first block, an integrator, filtered out most of the noise. The resulting system state $x_0$ was virtually noise free, with only 5.5 millivolts r.m.s. of noise present.

Comparing traces of the system state $x_0$ for both the noisy and the noise-free case, it can be seen that the response is identical. However the control signal $u$ in the noisy case contains a great deal of noise, which may cause overheating or dither in the actuator. Neither of these is desirable; in fact the performance index is specifically trying to minimise the control signal to reduce these effects and prolong actuator life. Fortunately, the measurement noise which propagates through the system is likely to be in a higher band of frequencies than the system can respond to. A simple filter could be used to remove the majority of the noise from the control signal $u$ before it is applied to the system.

10.5 SIMULATION OF SYSTEM NON-LINEARITY

Having established that measurement noise did not present too much of a problem, it only remained to assess the performance of the system and observer with the non-linearity included within the loop. The non-linearity was the double-ended saturation characteristic of the ball valve, shown in fig 42. The system normally operated in the linear region. An end stop or microswitch would be fitted to the actuator to prevent excursions into the upper saturation region. The only occasion which required any consideration was at start-up, when the tractor set off from rest. Then the actuator moved from
the fully retracted position for a short period of time before the ball valve started to open. It was feared that this could cause a transient or upset the observer, so simulation was used to find out what would happen.

To simulate the non-linearity, a logic controlled switch was used to initially disconnect the control signal $u$ from subsequent stages, and then after a short period of time to connect it. This meant that the initial gain of the stage was zero, but after a short time delay, it rose to $k_F$. This would simulate the opening of the ball valve exactly.

This scheme was programmed into the analog computer as shown in fig 43. A time base was generated by amplifier $A_{11}$, and fed into a comparator. When the time base reached a level set by potentiometer $P_{10}$, the comparator output changed state, causing the logic controlled switch to close, and feeding the control signal $u$ through to subsequent stages. Potentiometers $P_{09}$ and $P_{10}$ were adjusted to give a time delay of 2 seconds.

The response of the system to a step input with the non-linearity included is shown in fig 44. For the first 2 seconds, the loop was open and the system output stayed at zero. However the observer still thought that the loop was closed, and an incorrect estimate $\hat{x}_0$ was produced. Because the estimated signal was fed back to the controller, the controller also produced an incorrect value for the control signal to the actuator, $u_H$.

After 2 seconds, the loop was closed and the control signal to the valve $u_V$ became active. The observer could now generate an accurate estimate and it can be seen that the estimate $\hat{x}_0$ rapidly converged to the real state $x_0$. This confirmed that the observer eigenvalue was larger that the system eigenvalues. The system output $x_0$ rose smoothly to its desired level. It was now clear that the non-linearity in the ball valve characteristic would not cause any control problems.
10.6 SUMMARY

The system and the Hamilton-Jacobi controller were simulated on the Analog Computer. An additional circuit was patched in to measure the performance index. The performance indices for various controller feedback coefficients were noted over a number of simulation runs. A global optimum was identified which agreed with that predicted by the Hamilton-Jacobi equation. In addition, a local optimum was found for which the performance index was only 10% higher than the global optimum. This local optimum used feedback of the measured state, so a sub-optimal controller could have been implemented without an observer.

The observer and reconstruction equations were patched up on the analog computer. Simulation runs showed that the reconstructed state matched the "inaccessible" state exactly, provided that the system parameters remained constant. The system feedback loop could be closed with either the output state or the estimated state with no noticeable difference in performance.

To check the effect of measurement noise on the observer, white noise was added to the measured state. This was amplified by the observer and the estimated state contained a much higher level of noise. Despite this, the average of the estimate was still very close to the real state. When the noisy estimate was used to close the feedback loop, the control signal was also noisy, but the dynamics of the system filtered out the noise, leaving the output state virtually noise-free. The system response in both the noisy and the noise-free cases was therefore identical. However it was suggested that the noisy control signal should be filtered to prevent overheating or dither of the actuator due to excessive high frequency operation.

The nonlinear saturation characteristic of the ball valve was modelled and included in the system simulation. At system start-up, the ball valve had to turn a short distance before it started to pass fluid and during this period
the observer generated an incorrect estimate. However, once the ball valve
started to open, the observer estimate rapidly homed in on the output state,
as expected from the eigenvalue placement specified in the observer design
procedure. The simulation showed that no problems would be caused by the
non-linear characteristic.
The most disappointing part of this work was the Euler-Lagrange controller, which was unstable over the infinite time period. However the Hamilton-Jacobi equation yielded a stable state feedback controller. A Luenberger observer was necessary to generate an inaccessible state for the feedback.

The analog computer proved invaluable for testing the controllers. It gave an independent confirmation of the correctness of the Hamilton-Jacobi optimal control. Most important of all, the analog computer provided some "hands-on" experience of the system, and revealed the presence of a sub-optimal control. The sub-optimal control could be implemented without an observer and gave a performance only 10% worse than the true optimal control. The sub-optimal control had of course been ignored by the Hamilton-Jacobi equation.

Numerical solution of the optimal state feedback coefficients was done in BASIC. A "shotgun" method was used to find optimal coefficients for a wide range of system parameters. A similar method had been used with the analog computer, and had revealed the sub-optimum. The method may seem clumsy but it affords an excellent overview of the situation.

Certain difficulties such as noise and non-linearity had been anticipated. Simulation on the analog computer had shown that these difficulties did not really present any problem.

This work has been for a continuous time controller. In fact it is probable that a discrete time controller would be used. A microprocessor could handle the task and perform other functions also (see chapter 6). Microprocessor-based sprayer monitors are appearing on the market, but this sprayer control could also be configured as a sprayer monitor or acreage meter.
Specific observations about the two controls have already been made in their own sections.

I would like to comment on the two approaches to design that have been used in this work, the empirical and the theoretical. The experienced engineer working on a well-understood system can short circuit the standard design methods and lash together low performance systems in next to no time. On the other hand, if performance is to be truly optimised, intuition must take second place to a fundamental theoretical analysis of the situation. But the theoretical work must then be confirmed, preferably by analog computer simulation. If the simulation is undertaken with an open mind, other useful phenomena may also be noticed.

Finally, I hope that the Fertiliser Sprayer Control will achieve the same success as the Round Baler Control.
REFERENCES


FIG 1: ROUND BALER IN ACTION
FIG 2: TWINING MECHANISM

FIG 3: OPERATION OF TWINER
FIG 5: TEMPERATURE RANGE OF LOGIC FAMILIES

FIG 6: SCHEMATIC OF VERSATILE LOGIC SEQUENCER
FIG 7: TIMING DIAGRAM FOR VERSATILE LOGIC SEQUENCER

FIG 8: SEQUENCE TABLE FOR VERSATILE LOGIC SEQUENCER
<table>
<thead>
<tr>
<th>STAGE</th>
<th>RED LAMP</th>
<th>RELAY 1</th>
<th>RELAY 2</th>
<th>HIGH SPEED</th>
<th>TIMER ENABLE</th>
<th>EXTERNAL START DISABLE</th>
<th>INCREMENTING SIGNAL</th>
<th>RESET</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>START SWITCH OR EXTERNAL START</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>STALL CURRENT IN MOTOR</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>END OF TIMED PERIOD</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>CONTACT WITH MICROSWITCH</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>STAGE 4</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>END OF TIMED PERIOD</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>STALL CURRENT IN MOTOR</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>STAGE 7</td>
<td></td>
</tr>
</tbody>
</table>

FIG 9: SEQUENCE TABLE FOR ROUND BALER CONTROL
FIG 10: ASTABLE MULTIVIBRATOR

Clock Signal $V_{CLOCK}$

Threshold Voltage $V_{THRESH}$

Capacitor Voltage $V_{CAP}$

FIG 11: WAVEFORMS OF ASTABLE MULTIVIBRATOR
FIG 12: CURRENT PROFILE DURING AUTOMATIC OPERATION

FIG 13: ABUSE OF ACTUATOR DURING MANUAL CONTROL
PULSE WIDTH MODULATION SIGNAL

Motor Voltage

ON

OFF

+12V

0V

-40V

Valid measurement of back e.m.f.

1/14

3/9

FIRST TIME DELAY 3/8

3/11

SECOND TIME DELAY 3/10

"WINDOW" SIGNAL 3/12

FIG 14: GENERATION OF "WINDOW" PULSE TO SENSE BACK E.M.F.

FIG 15: PROPORTIONAL PLUS INTEGRAL CONTROLLER
INPUT VOLTAGE

TRIANGULAR MODULATING WAVEFORM

PULSE WIDTH MODULATED SIGNAL

FIG 16: PULSE WIDTH MODULATION

TRIANGULAR WAVEFORM

RECTANGULAR WAVEFORM

COMPOSITE MODULATING WAVEFORM

INPUT VOLTAGE

MODULATING WAVEFORM

PULSE WIDTH MODULATED SIGNAL

CLAMP LEVEL

90% MAX

FIG 17: LIMITING DUTY CYCLE TO 90%
**FIG 18: SWITCHING CIRCUIT WITH BIPOLAR POWER TRANSISTOR**

\[ P = 11.4\text{V} \times 5\text{A} + 1\text{V} \times 25\text{A} = 58\text{W} + 25\text{W} = 83\text{W} \]

**FIG 19: SWITCHING CIRCUIT WITH DARLINGTON POWER TRANSISTOR**

\[ P = 10.8\text{V} \times 0.2\text{A} + 2.4\text{V} \times 25\text{A} = 2\text{W} + 60\text{W} = 62\text{W} \]

**FIG 20: SWITCHING CIRCUIT WITH POWER MOSFET TRANSISTOR**

\[ P = (25\text{A})^2 \times 0.06\text{ Ohm} = 36\text{W} \]

**FIG 21: CONFIGURATION OF POWER CIRCUIT**

Signal to MOSFETs

Back e.m.f. signal

Current signal
FIG 22a: NO PROTECTION

FIG 22b: DIODE PROTECTION

FIG 22c: ZENER PROTECTION

FIG 22d: ACTIVE ZENER PROTECTION

FIG 22e: A NEW METHOD OF SUPPRESSION

FIG 22: CLAMPING OF INDUCTIVE TRANSIENTS
FIG 23: FERTILISER SPRAYER IN ACTION
FIG 24: OPEN LOOP CONTROL SCHEME
Fig 25: Closed Loop Control Scheme
FIG. 26 CLOSED LOOP BLOCK DIAGRAM
Fig 27: CUTAWAY VIEW OF LINEAR ACTUATOR
FIG 28: SIMULATION OF EULER-LAGRANGE CONTROLLER

\[-\frac{1}{c} \int \int (x_D - x_M) \cdot dt \cdot dt + \int x_D \cdot dt\]

FIG 29: SIMULATION OF SPRAYER SYSTEM
FIG 30: ANALOG COMPUTER DIAGRAM OF EULER-LAGRANGE CONTROLLER

FIG 31: ANALOG COMPUTER DIAGRAM OF SPRAYER SYSTEM
FIG 32: NEWTON-RAPHSON METHOD OF FINDING ROOTS

\[ y = f(x) \]

**CONTROLLER** | **SYSTEM**
--- | ---
\[ x_0 \] | \[ K_F \frac{\tau}{s} \]
\[ u \] | \[ \frac{1}{1 + s \tau} \]
\[ x_0 \] | \[ x_M \]

**OBSERVER**

\[ \frac{K_F \tau}{p(p-1)} \]

**FIG 33: STATE FEEDBACK USING ESTIMATED STATE FROM LUENBERGER OBSERVER**
FIG 34: ANALOG COMPUTER DIAGRAM OF STATE FEEDBACK CONTROL SYSTEM
$K_M$, coefficients of $x_M$

<table>
<thead>
<tr>
<th>$K_0$, coefficients of $x_0$</th>
<th>0</th>
<th>0.25</th>
<th>0.5</th>
<th>0.75</th>
<th>1.0</th>
<th>1.25</th>
<th>1.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>J=0.207</td>
<td>J=0.136</td>
<td>J=0.147</td>
</tr>
<tr>
<td>0.25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>J=0.138</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.5</td>
<td>$X_5$</td>
<td>J=0.149</td>
<td>J=0.124</td>
<td>J=0.133</td>
<td>J=0.196</td>
<td>J=0.295</td>
<td></td>
</tr>
<tr>
<td>0.75</td>
<td></td>
<td>J=0.126</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.0</td>
<td>$X_5$</td>
<td>J=0.122</td>
<td>J=0.131</td>
<td>J=0.152</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.25</td>
<td></td>
<td>J=0.126</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.5</td>
<td></td>
<td>J=0.134</td>
<td>J=0.152</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Global Optimum
$K_M=0, K_0=0.97, J=0.122$

Local Optimum
$K_M=0.55, K_0=0, J=0.134$

FIG 35: VARIATION OF PERFORMANCE INDEX WITH FEEDBACK COEFFICIENTS
FIG. 36: GLOBAL OPTIMUM USING OUTPUT STATE FEEDBACK

1. $x_M$, measured state
2. $x_g$, output state
3. $J$, performance index
4. $u$, control signal
1 $x_M$, measured state
2 $x_o$, output state
3 $J$, performance index
4 $u$, control signal

FIG 37: LOCAL OPTIMUM USING MEASURABLE STATE FEEDBACK
FIG 38: ANALOG COMPUTER DIAGRAM OF LUNBERGER OBSERVER
1. $\hat{x}_0$, estimated output with coefficient of $u$ set to 0.5.
2. $x_0$, output state.
3. $\hat{x}_0$, estimated output with coefficient of $u$ set to 0.167.
4. $\hat{x}_0$, estimated output with coefficient of $u$ set to zero (i.e. "stripped down" observer).
5. $x_M$, measured output.
6. $u$, control signal.

FIG 39: PERFORMANCE OF LUENBERGER OBSERVER
1. Noise
2. \( x_N \), measured state with noise
3. \( \hat{x}_O \), estimated output with noise
4. \( \hat{x}_O \), estimated output without noise

FIG. 40: LUENBERGER OBSERVER WITH MEASUREMENT NOISE
1. \( u \), control signal
2. \( \hat{x}_g \), estimated output state
3. \( x_M \), measured state
4. \( x_g \), output state

FIG. 41 : FEEDBACK WITH NOISY ESTIMATED STATE
nonlinear mechanism and ball valve

FIG 42: NONLINEARITY OF BALL VALVE

FIG 43: ANALOG COMPUTER DIAGRAM TO SIMULATE NONLINEARITY
1 $u_M$, drive to motor
2 $\hat{x}_0$, estimated output
3 $u_V$, drive to valve
4 $x_0^*$, output state

FIG 44: RESPONSE OF SYSTEM WITH NONLINEARITY
APPENDIX 1: PROGRAM TO CALCULATE OPTIMAL STATE FEEDBACK COEFFICIENTS

A flow chart is given below, followed by the program listing.

Repeat for values of gain of 0.5, 1, 1.5, 2

Repeat for values of time constant of 0.1, 0.2, 0.5, 1

Repeat for values of control term weighting of 0.1, 0.2, 0.5, 1, 2, 5, 10

- Calculate factors A, B, C, D, E, F, G from input parameters
- Reset iteration counter
- Set initial approximation of root to unity
- Use Newton-Raphson rule to calculate new iterate
- Increment iteration counter
- Repeat until difference between iterations is less than 0.01 or until 100 iterations have been made
- Condition satisfied
- Condition not satisfied
- Calculate state feedback coefficients
- Print input parameters, coefficients of P matrix, number of iterations and state feedback coefficients