

This item was submitted to Loughborough University as a PhD thesis by the author and is made available in the Institutional Repository (<u>https://dspace.lboro.ac.uk/</u>) under the following Creative Commons Licence conditions.

COMMONS DEED
Attribution-NonCommercial-NoDerivs 2.5
You are free:
<ul> <li>to copy, distribute, display, and perform the work</li> </ul>
Under the following conditions:
<b>Attribution</b> . You must attribute the work in the manner specified by the author or licensor.
Noncommercial. You may not use this work for commercial purposes.
No Derivative Works. You may not alter, transform, or build upon this work.
<ul> <li>For any reuse or distribution, you must make clear to others the license terms of this work</li> </ul>
<ul> <li>Any of these conditions can be waived if you get permission from the copyright holder.</li> </ul>
Your fair use and other rights are in no way affected by the above.
This is a human-readable summary of the Legal Code (the full license).
<u>Disclaimer</u> 曰

For the full text of this licence, please go to: <u>http://creativecommons.org/licenses/by-nc-nd/2.5/</u>



# Novel Displacement Sensing. Towards Robotic Tunnelling

by

# William Leonard Dudeney

A Doctoral thesis submitted in partial fulfillment of the requirements for the award of Doctor of Philosophy of Loughborough University

May 2001

By William Leonard Dudeney 2001

and a second of the

د. همه کو چې د

Lenghbonzigh
Date Arc or
Class
Acc. 040250947

.

#### Acknowledgement

There are several people who have made my academic achievements possible.

I am indebted first and foremost to my parents who gave so much to provide my education and foundation for postgraduate study. I appreciate and give thanks for the support and patience of my wife, for as long as she's known me I have been working for this degree.

I would like to thank my supervisor, Dr. Mike Jackson for his able management of this project, and for his securing the necessary funding. In particular I would like to thank colleagues: Neil Brown for his advice, and his opinions, Rolph Thomas for changing the subject, Sergio D'Almeida for his enthusiasm about everything, and Richard Carroni because it never seemed so bad when we talked about his research. A general thanks is offered to other members of the Mechatronics Research Group and the wider structure of the Wolfson School, particularly the technical staff for their advice and manufacturing skills.

This PhD is for the provision of my family in the years ahead. I give thanks to The Risen Lord.

#### Abstract

This thesis is presented in two parts.

It was noted that in the second half of the 1990's, a period during which there were long and dry summers in the UK, the water utility companies were openly criticized for their inefficiency and poor distribution-network performance as supplies ran low. It was noted that some of these companies were suffering up to 50% loss of water through leaks in the distribution network. Increasing pressure mounted for the owners of buried infrastructure to more effectively identify and service both small and large leaks alike. Conventional methods are shown to be not effective or efficient in the long-term, and that it would be most useful to adopt a more appropriate strategy in the access to and repair of damaged water mains.

A conceptual solution is developed and proposed about a 'minimal-intrusion' system which would allow efficient access to buried infrastructure for the purposes of repair, whilst reducing effects on the local environment when compared to conventional excavation. The concept is based around biological mimicry as the highly adaptable access vehicle copies both the form of and the locomotion strategy of a common earthworm. The associated technological challenges are identified and it is realized that to achieve a technical content within the project it is necessary to focus on one particular area of interest. The area selected is locomotion of a worm-like vehicle. Suitable actuators are discussed in some detail, as is a method for deriving feedback on the surface form of the vehicle, which may vary greatly according to conditions. A justification is made to pursue the development of a surface mounted or embedded sensor network which is able to profile the form of a complex surface.

The sensor technology is selected after some consideration, and is based on inductive principles. A completely novel sensor in terms of form and electronic/signal-processing design is developed from first principles. A prototype is constructed from suitable materials and subjected to a considered testing strategy. The testing is based around a custom built rig, which is intended to allow extended duration procedures and automated data acquisition. The prototype is subjected to several iterations of refinement before a main test series is conducted. The performance of the prototype is compared with a conventional sensing device where applicable.

Finally, some thought is given to the application of the new sensor in the context of this project, and in other fields.

Keywords Tunnelling Displacement sensor Embedded control Robot Transducer

Microcontroller Measurement Data acquisition Mechatronics

# Table of contents

1	A Problem of Access to Buried Infrastructure	6
	1.1 Introduction	
	Accessing buried infrastructure, a typical scenario	
	Types of failure requiring access to buried infrastructure	
	1.3.1 Serviceable failure	
	1.3.1.1 Serviceable failure at a specific point	
	1.3.1.2 Serviceable failure over a region	
	1.3.2 Failure due to 3 <sup>rd</sup> party intrusion	
	1.3.3 The obligation of buried infrastructure owners to maintain their assets	
	1.4 Materials and techniques used in the maintenance of buried infrastructure	
	1.4.1 Maintaining water distribution networks	
	1.4.1.1 Diagnosing a water distribution problem	
	1.4.1.2 Locating leaks in the distribution network	
	1.4.1.3 Pipe vibration:	
	1.4.1.3.1 Soil Impact:	
	1.4.1.3.2 Fluid Circulation:	
	1.4.1.4 Actions to repair/replace faulty water distribution infrastructure	
	1.4.1.4.1 Operations on a specific network point.	
	1.4.1.4.2 Operations on a complete section,	
	14.1.4.2.1 Supine	
	1.4.1.4.2.2 Die draw,	
	$1.4.1.4.2.5  \text{Pipt Dursting}, \dots, 14$	
	1.4.1.4.2.4 Centent Mortal Lining. 15 1 $4$ 1 $4$ 2 5 Enovy Resin Lining. 17	
	14.14.26 Narrow Trenching 17	
	141427 Other area refurbishment processes	
	1415 The bias of water distribution on this research project 17	
	1 4 2 Maintaining gas distribution networks	
	1 4 3 Maintaining cable networks	
	1.4.3.1 Power distribution	
	1.4.3.2 Communications cable networks	
	5 Kou phases of the generic process to access and renair huried infrastructure 20	
	Key phases of the generic process to access and repair buried infrastracture	
	1.6 Conclusion	
	1.6.1 Summary of chapter 1	
	1.6.2 Research opportunity arising	
2	A novel concept on access and repair of buried infrastructure	22
	2.1 Introduction	
2	2.2 An existing system to exploit	
	2.2.1 Earthworm Locomotion Mechanics	
	2.2.2 Passage through a solid	
	27 Technical challenges of a man-made earthworm	
	2.3.1 Locomotion	
	2.3.2 Location and positioning	
	2.3.3 Ground cutting	
	2.3.4 Spoil processing	
	2.3.5 Underground vision/obstacle identification	
	2.3.5.1 Global vision:	
	2.3.5.2 Local vision	
	2.3.6 Power supply	
	2.3.7 Control	
	2.3.7.1 Locomotion control	

2.3.7.2 2.3.7.3	Ground cutting and spoil handling	
2.4 The n	nechatronic device – some applications	
2.5 The s	election of a suitable research problem	39
2.6 An ex	xpansion on locomotion	40
2.6.1	Spoil displacement	
2.6.2	Spoil extraction	41
2.6.3	Compliance:	
2.6.3.1	Closed loop control of actuators	
2.6.4	Methods of achieving worm-like actuation	
2.6.4.1	Natural worm actuator topography	
2.6.4.2	Custom-designed worm actuator topography	
2.6.5	Examples of appropriate actuation technology	
2.6.5.1	Polymer gel.	50
2.6.5.2	Conventional Chemistry actuation:	
2.6.5.3	Shape memory alloy.	
2.6.5.4	Coupled mechanisms	
2.6.6	Positional feedback for a mechatronic earthworm segment	
2.6.6.1	Worm shape by direct actuator state feedback	59
2.6.6.2	Indirect feedback	60
267	The Sensor Mesh	60
2.0.7	High strain transducers	
2.0.7.1	1 1 Fluid filled tubes	
2.0.7	1.2 Conductive rubber	
2.0.7	1.2 Adaptation of line displacement sensor	
2.0.7.		
2.7 Estim	nation of some operational requirements	62
2.8 Conc	lusion	65
3 Sensor Te	echnology and Development of a Novel Surface Measurement System	66
3 Sensor Te	echnology and Development of a Novel Surface Measurement System	
3 Sensor Te 3.1 Introd	echnology and Development of a Novel Surface Measurement System	65 66
3 Sensor Te 3.1 Introd 3.2 Meas	echnology and Development of a Novel Surface Measurement System duction	65 66 66
3 Sensor Te 3.1 Introd 3.2 Meas 3.2.1 S	echnology and Development of a Novel Surface Measurement System duction turement and Instrumentation fundamentals	65 66 66 66
3 Sensor Te 3.1 Introd 3.2 Meas 3.2.1 S 3.2.2 D	echnology and Development of a Novel Surface Measurement System duction curement and Instrumentation fundamentals Sensor definitions Definition of measurement terms as used in this text	
3 Sensor Te 3.1 Introd 3.2 Meas 3.2.1 S 3.2.2 I 3.2.2.1	echnology and Development of a Novel Surface Measurement System duction curement and Instrumentation fundamentals Sensor definitions Definition of measurement terms as used in this text Transducer	
3 Sensor Te 3.1 Introd 3.2 Meas 3.2.1 S 3.2.2 J 3.2.2.1 3.2.2.1 3.2.2.2	echnology and Development of a Novel Surface Measurement System duction curement and Instrumentation fundamentals Sensor definitions Definition of measurement terms as used in this text Transducer Primary Transducer	
3 Sensor Te 3.1 Introd 3.2 Meas 3.2.1 S 3.2.2 I 3.2.2.1 3.2.2.2 3.2.2.3	echnology and Development of a Novel Surface Measurement System duction curement and Instrumentation fundamentals Sensor definitions Definition of measurement terms as used in this text Transducer Primary Transducer Active Transducer	
3 Sensor Te 3.1 Introd 3.2 Meas 3.2.1 S 3.2.2 I 3.2.2.1 3.2.2.2 3.2.2.3 3.2.2.4	echnology and Development of a Novel Surface Measurement System duction nurement and Instrumentation fundamentals Sensor definitions Definition of measurement terms as used in this text Transducer Primary Transducer Active Transducer Sensor	
3 Sensor Te 3.1 Introd 3.2 Meas 3.2.1 S 3.2.2 I 3.2.2.1 3.2.2.2 3.2.2.3 3.2.2.4 3.2.2.5	echnology and Development of a Novel Surface Measurement System duction nurement and Instrumentation fundamentals Sensor definitions Definition of measurement terms as used in this text Transducer Primary Transducer Active Transducer Sensor Instrument	
3 Sensor Te 3.1 Introd 3.2 Meas 3.2.1 S 3.2.2 I 3.2.2.1 3.2.2.2 3.2.2.3 3.2.2.4 3.2.2.5 3.2.2.6	echnology and Development of a Novel Surface Measurement System duction curement and Instrumentation fundamentals	
3 Sensor Te 3.1 Introd 3.2 Meas 3.2.1 S 3.2.2 S 3.2.2.1 3.2.2.2 3.2.2.3 3.2.2.4 3.2.2.5 3.2.2.6 3.2.3	echnology and Development of a Novel Surface Measurement System duction murement and Instrumentation fundamentals	
3 Sensor Te 3.1 Introd 3.2 Meas 3.2.1 S 3.2.2 J 3.2.2.1 3.2.2.2 3.2.2.3 3.2.2.4 3.2.2.5 3.2.2.6 3.2.3 ( 3.2.3.1)	echnology and Development of a Novel Surface Measurement System duction murement and Instrumentation fundamentals	
3 Sensor Te 3.1 Introd 3.2 Meas 3.2.1 S 3.2.2 S 3.2.2 S 3.2.2.1 3.2.2.2 3.2.2.3 3.2.2.4 3.2.2.5 3.2.2.6 3.2.3.1 3.2.3.1 3.2.3.2	echnology and Development of a Novel Surface Measurement System duction mement and Instrumentation fundamentals	
3 Sensor Te 3.1 Introd 3.2 Meas 3.2.1 3 3.2.2 1 3.2.2.1 3.2.2.2 3.2.2.3 3.2.2.4 3.2.2.5 3.2.2.6 3.2.3 3.2.3.1 3.2.3.2 3.2.3.1 3.2.3.2 3.2.3.3	echnology and Development of a Novel Surface Measurement System duction murement and Instrumentation fundamentals	
3 Sensor Te 3.1 Introd 3.2 Meas 3.2.1 S 3.2.2 I 3.2.2.1 3.2.2.2 3.2.2.3 3.2.2.4 3.2.2.5 3.2.2.6 3.2.3 3.2.3.1 3.2.3.2 3.2.3.3 3.2.3.4	echnology and Development of a Novel Surface Measurement System duction murement and Instrumentation fundamentals	
3 Sensor Te 3.1 Introd 3.2 Meas 3.2.1 S 3.2.2 S 3.2.3 S 3.2.	echnology and Development of a Novel Surface Measurement System duction mement and Instrumentation fundamentals	
3 Sensor Te 3.1 Introd 3.2 Meas 3.2.1 3 3.2.2 1 3.2.2 3 3.2.2 4 3.2.3 1 3.2.3 2 3.2.3 1 3.2.3 1 3.2.3 2 3.2.3 1 3.2.3 1 3.2.	echnology and Development of a Novel Surface Measurement System duction mement and Instrumentation fundamentals	
3 Sensor Te 3.1 Introd 3.2 Meas 3.2.1 3 3.2.2 1 3.2.2 3 3.2.2 4 3.2.3 1 3.2.3 2 3.2.3 1 3.2.3 1 3.2.3 1 3.2.3 2 3.2.3 1 3.2.3 2 3.2.3 1 3.2.3 2 3.2.3 1 3.2.3 2 3.2.3 1 3.2.3 2 3.2.3 2 3.2.	echnology and Development of a Novel Surface Measurement System duction murement and Instrumentation fundamentals	66         66         66         66         68         69         69         69         69         70         70         70         71         73         74
3 Sensor Te 3.1 Introd 3.2 Meas 3.2.1 3 3.2.2 1 3.2.2 1 3.2.3 2 3.2.3 1 3.2.3 2 3.2.3 1 3.2.3 2 3.2.3 1 3.2.3 2 3.2.3 1 3.2.3 2 3.2.3 2 3.2.	echnology and Development of a Novel Surface Measurement System duction murement and Instrumentation fundamentals	66         66         66         66         68         69         69         69         69         70         70         70         70         71         73         74
3 Sensor Te 3.1 Introd 3.2 Meas 3.2.1 3 3.2.2 1 3.2.2 1 3.2.3 2 3.2.3 1 3.2.3 1 3.2.3 2 3.2.3 1 3.2.3 2 3.2.3 1 3.2.3 2 3.2.3 1 3.2.3 2 3.2.3 2 3.2.	echnology and Development of a Novel Surface Measurement System duction murement and Instrumentation fundamentals	66           66           66           66           66           68           69           69           69           70           70           70           70           70           70           70           70           70           71           71           71           71           71           71           71           71           71           71           71           71           71           71           71           71           71           71           73           74           74           75
3 Sensor Te 3.1 Introd 3.2 Meas 3.2.1 3 3.2.2 1 3.2.2 1 3.2.3 2 3.2.3 1 3.2.3 2 3.2.3 1 3.2.3 2 3.2.3 1 3.2.3 2 3.2.3 1 3.2.3 2 3.2.3 1 3.2.3 2 3.2.3 1 3.2.3 1 3.2.	echnology and Development of a Novel Surface Measurement System duction marement and Instrumentation fundamentals	66         66         66         66         66         68         69         69         69         69         70         70         71         73         74         75         75          75
3 Sensor Te 3.1 Introd 3.2 Meas 3.2.1 3 3.2.2 1 3.2.2 1 3.2.3 1 3.2.3 1 3.2.3 1 3.2.3 1 3.2.3 1 3.2.3 1 3.2.3 1 3.2.3 1 3.2.3 2 3.2.3 1 3.2.3 1 3.2.3 2 3.2.3 1 3.2.3 1 3.2.	echnology and Development of a Novel Surface Measurement System duction curement and Instrumentation fundamentals	66         66         66         66         66         67         69         69         69         69         70         70         71         72         73         74         75         75         75         75
3 Sensor Te 3.1 Introd 3.2 Meas 3.2.1 3 3.2.2 1 3.2.2 1 3.2.3 1 3.2.3 1 3.2.3 1 3.2.3 1 3.2.3 1 3.2.3 1 3.2.3 1 3.2.3 1 3.2.3 2 3.2.3 1 3.2.3 1 3.2.3 2 3.2.3 1 3.2.3 1 3.2.3 2 3.2.3 1 3.2.3 1 3.2.	echnology and Development of a Novel Surface Measurement System duction	
3 Sensor Te 3.1 Introd 3.2 Meas 3.2.1 3 3.2.2 1 3.2.2 1 3.2.3 2 3.2.3 1 3.2.3 1 3.2.3 2 3.2.3 1 3.2.3 2 3.2.3 1 3.2.3 1 3.2.	echnology and Development of a Novel Surface Measurement System duction	66         66         66         66         66         67         68         69         69         69         70         70         70         70         70         71         72         73         74         75         75         75         76         77
3 Sensor Te 3.1 Introd 3.2 Meas 3.2.1 S 3.2.2 S 3.2.2 S 3.2.2 S 3.2.2 S 3.2.2 S 3.2.2 S 3.2.2 S 3.2.2 S 3.2.2 S 3.2.3 S 3.2.	echnology and Development of a Novel Surface Measurement System duction	66         66         66         66         66         67         68         69         69         69         70         70         70         70         70         71         72         73         74         75         75         75         76         77         78
3 Sensor Te 3.1 Introd 3.2 Meas 3.2.1 S 3.2.2 S 3.2.3 S 3.2.	echnology and Development of a Novel Surface Measurement System duction	66         66         66         66         66         67         68         69         69         69         70         70         70         70         70         71         72         73         74         75         75         75         76         77         78          70          70          70          71          72          73
3 Sensor Te 3.1 Introd 3.2 Meas 3.2.1 3 3.2.2 1 3.2.2 1 3.2.3 2 3.2.3 1 3.2.3 2 3.2.3 1 3.2.3 1 3.2.3 2 3.2.3 1 3.2.3 1 3.2.3 2 3.2.3 1 3.2.3 2 3.2.3 1 3.2.3 2 3.2.3 2 3.2.3 1 3.2.3 1 3.2.3 2 3.2.3 1 3.2.3 1 3.2.3 2 3.2.3 1 3.2.3 1 3.2.	echnology and Development of a Novel Surface Measurement System duction	65         66         66         66         66         67         68         69         69         69         70         70         70         70         70         71         72         73         74         75         75         75         76         77         78         90           90          90          91          92          93

\_

3	3.2.4.1.5 Non-contact surface measurement devices	80
3.3 S	Surface mounted/embedded systems for surface measurement	82
3.3.1	Principle of employment	82
3.3.2	Application to specific technology	84
3.3.	2.1 Resistive displacement devices.	85
3.3.	2.2 Inductive displacement devices.	85
3.3.	2.3 Reluctive based displacement devices	87
3.3.	.2.4 Capacitance based displacement devices.	88
3.3.3	The suitability of contact sensing devices to flexible surface measurements	88
3.4 T	The development of a flexible inductive displacement sensor	89
3.4.1	Considerations of coil geometry	89
3.4.2	Consideration of core design	91
3.4.	2.1 A flexible ferrite core	93
3.4.3	Design and manufacture of inductor	94
3.4.	3.1 The Coil	94
3.4.	.3.2 The Tapered core	95
3.4.	.3.3 Modulator housing	95
3.4.	3.4 Circuit Design	97
3.4.4	Anticipated performance of prototype sensor	97
3.4.	4.1 Usefulness of the output signal	101
3.5 C	Conclusion	101
4 Exper	rimental design	102
4 T T		103
4.1 1	niroauciion	102
4.2 B	Background to experimental design	102
4.2.1	Predecessors to the prototype	103
13 L	dentification of test criteria	105
4.3 Ia 431	dentification of test criteria	105
4.3 Io 4.3.1 4.3.2	dentification of test criteria What is the purpose of testing the prototype? Relevant performance categories	<i>105</i> 105 106
4.3 Ia 4.3.1 4.3.2	dentification of test criteria What is the purpose of testing the prototype? Relevant performance categories	<i>105</i> 105 106
4.3 IA 4.3.1 4.3.2 4.4 E	dentification of test criteria What is the purpose of testing the prototype? Relevant performance categories Experimental design – Apparatus	105 105 106 106
4.3 Ia 4.3.1 4.3.2 4.4 Ea 4.4.1	dentification of test criteria What is the purpose of testing the prototype? Relevant performance categories Experimental design – Apparatus Measurand drive	105 105 106 106 108
4.3 IA 4.3.1 4.3.2 4.4 E 4.4.1 4.4.1	dentification of test criteria	105 105 106 106 108 108
4.3 IA 4.3.1 4.3.2 4.4 E 4.4.1 4.4. 4.4.	dentification of test criteria	105 105 106 106 108 108 108
4.3 In 4.3.1 4.3.2 4.4 E 4.4.1 4.4. 4.4. 4.4. 4.4.2	dentification of test criteria	105 105 106 106 108 108 108 108 110
4.3 IA 4.3.1 4.3.2 4.4 E 4.4.1 4.4. 4.4. 4.4. 4.4. 4.4.	dentification of test criteria	105 105 106 106 108 108 108 110 111
4.3 In 4.3.1 4.3.2 4.4 E 4.4.1 4.4. 4.4. 4.4. 4.4. 4.4. 4.4. 4.	dentification of test criteria	105 105 106 106 108 108 108 110 111 112
4.3 IA 4.3.1 4.3.2 4.4 E 4.4.1 4.4. 4.4. 4.4. 4.4. 4.4. 4.4. 4.	dentification of test criteria	105 105 106 106 108 108 108 110 111 112 113
4.3 In 4.3.1 4.3.2 4.4 E 4.4.1 4.4. 4.4. 4.4. 4.4. 4.4. 4.4. 4.	dentification of test criteria	105 105 106 106 108 108 108 110 111 111 112 113 114
4.3 In 4.3.1 4.3.2 4.4 E 4.4.1 4.4. 4.4. 4.4. 4.4. 4.4. 4.4. 4.	dentification of test criteria	105 105 106 106 108 108 108 110 111 111 112 113 114 115
4.3 In 4.3.1 4.3.2 4.4 E 4.4.1 4.4. 4.4. 4.4. 4.4. 4.4. 4.4. 4.	dentification of test criteria	105 105 106 106 108 108 108 110 111 112 113 114 115 115
4.3 In 4.3.1 4.3.2 4.4 E 4.4.1 4.4. 4.4. 4.4. 4.4. 4.4. 4.4. 4.	dentification of test criteria	105 105 106 106 108 108 108 110 111 112 113 114 115 115 123
4.3 IA 4.3.1 4.3.2 4.4 E 4.4.1 4.4.	dentification of test criteria	105 105 106 106 108 108 108 108 110 111 112 113 114 115 115 123 123
4.3 IA 4.3.1 4.3.2 4.4 E 4.4.1 4.4.	dentification of test criteria	105 105 106 106 108 108 108 110 111 112 113 114 115 115 123 124 124
4.3 IA 4.3.1 4.3.2 4.4 E 4.4.1 4.4.	dentification of test criteria	105 105 106 106 108 108 108 108 110 111 112 113 114 115 115 123 124 124 124
4.3 In 4.3.1 4.3.2 4.4 E 4.4.1 4.4.	dentification of test criteria	105 105 106 106 108 108 108 108 110 111 112 113 114 115 115 123 123 124 124 124 124
4.3 In 4.3.1 4.3.2 4.4 E 4.4.1 4.4.	dentification of test criteria.         What is the purpose of testing the prototype?         Relevant performance categories         Experimental design – Apparatus         Measurand drive         1.1         Fabrication material.         1.2       Topographic design         Data logging         2.1       Prototype sensor         2.2       Reference LVDT sensor         2.3       Reference temperature sensors         2.4       The real-time controller for data logging         2.5       Interfacing sensors to the real-time controller         .4.2.5.1       Prototype sensor interface         .4.2.5.3       Temperature sensor interface         .4.2.5.3       Temperature sensor interface         .4.3.1.1       Configuration of test parameters         .4.3.1.2       Observation/processing of 'real-time' data         .4.3.1.3       Observation/processing of stored data	105 105 106 106 108 108 108 108 110 111 112 113 114 115 123 123 124 124 124 124
4.3 In 4.3.1 4.3.2 4.4 E 4.4.1 4.4.	dentification of test criteria.         What is the purpose of testing the prototype?         Relevant performance categories         Experimental design – Apparatus.         Measurand drive.         1.1         Fabrication material.         1.2       Topographic design.         Data logging.         2.1       Prototype sensor.         2.2       Reference LVDT sensor.         2.3       Reference temperature sensors.         2.4       The real-time controller for data logging .         2.5       Interfacing sensors to the real-time controller.         .4.2.5.1       Prototype sensor interface.         .4.2.5.2       LVDT sensor interface         .4.2.5.3       Temperature sensor interface.         .4.2.5.4       Thereface         .4.3.1       Configuration of test parameters         .4.3.1.2       Observation/processing of 'real-time' data.         .4.3.1.3       Observation/processing of stored data.         .4.3.1.4       Ultimate system control and automation override.	105 105 106 106 108 108 108 108 110 111 112 113 114 115 123 123 124 124 124 124 124 124 125 126
4.3 In 4.3.1 4.3.2 4.4 E 4.4.1 4.4.	dentification of test criteria	105 105 106 106 108 108 108 108 110 111 112 113 114 115 115 123 124 124 124 124 124 124 125 126
4.3 IA 4.3.1 4.3.2 4.4 E 4.4.1 4.4.	dentification of test criteria	105 105 106 106 108 108 108 108 110 111 112 113 114 115 115 123 124 124 124 124 124 124 125 126 126
4.3 IA 4.3.1 4.3.2 4.4 E 4.4.1 4.4.	dentification of test criteria	105 105 106 106 108 108 108 108 110 111 112 113 114 115 123 123 124 124 124 124 124 126 126 126 129
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	dentification of test criteria	105 105 106 106 108 108 108 108 110 111 112 113 114 115 123 124 124 124 124 124 124 126 126 129 129
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	dentification of test criteria	105 105 106 106 108 108 108 108 110 111 112 113 114 115 115 123 124 124 124 124 124 125 126 126 129 129 129
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	dentification of test criteria	105 105 106 106 108 108 108 108 110 111 112 113 114 115 123 123 124 124 124 124 124 124 125 126 126 129 129 129 129
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	dentification of test criteria	105 105 106 106 108 108 108 108 110 111 112 113 114 115 123 124 124 124 124 124 124 124 125 126 126 129 129 129 129 129 129

4.5.2 Te	sting active performance of sensor system	132
4.6 Conclus	sion	132
5 Results and	discussion	13
5.1 Introdu	ction	133
5.2 Prelimi	nary Results	133
5.2.1 Co	ncerning properties of elastic ferrite core	133
5.2.1.1	Magnetic properties of elastic ferrite core	133
5.2.1.2	Mechanical properties of elastic ferrite core	135
5.3 Main R	esults	135
5.3.1 Re	sults to determine optimal configuration	135
5.3.1.1	Increase the microcontroller clock speed.	138
5.3.1.2	Count more than 16 cycle periods from the sensor output signal	138
5.3.1.3	Increase the sensor output signal frequency range.	138
5.3.2 IVI8	an-stream results for first version of prototype	139 120
J.J.4.1 5371	1 Steady state drift test	139
5.3.2.1.	2 Variable measurand tests	141
5.3.2.1.	3 Linearity	142
5.3.2.1.	4 Hysteresis	143
5.3.2.1.	5 Resolution	143
5.3.2.1.0	6 Repeatability	144
5.3.2.2		144
5.3.2.3	Test results for the third version of the prototype	148
5.3.2.3.	1 Steady state drift test	148
5324	Z Range test	151
5.3.2.4	1 Correction of spurious data points	152
5.3.2.4.	2 Static drift test.	155
5.3.2.4.	3 Range test	157
5.3.2.4.4	4 Linearity	158
5.3.2.4.	5 Hysteresis	158
5.3.2.4.0	6 Repeatability	159
5.3.2.5	Results from the fifth test series	160
5.3.2.3. 5.3.2.5/	<ol> <li>Steady state test result (temperature compensation)</li> <li>Range test</li> </ol>	101
5.3.2.5	3 Application of linearity correction in real-time	165
5.3.2.5.4	4 Application of metric calibration and tolerance setting	172
5.4 Perform	nance Under Bending	175
5.5 Tests or	onducted	177
5.5.1 Te	sting sensor system stability	177
5.5.2 Te	sting active performance of sensor system	177
5.6 Conclus	ions	178
6 Refinements	and Applications	
6.1 Introdu	ction	180
67 Further	modelling of coil inductance	180
6.2.1 Re	fined parameter quantification	180
6.2.2 Ea	uation refinement	181
6.2.3 Re	-calculation of frequency range of sensor output	183
63 Improvi	na linearity	183
6.3.1 Im	proved signal conditioning	183
Alla,	, <u> </u>	
		4

	6.3.1.1 Multiple function description	
	6.3.1.2 Custom function description	
6.3	.2 Mechanical alteration	
6.4	Context specific application	
6.4	1.1 Simplistic surface form measurement by interpolation	
6.4	2.2 Complex surface form measurement by 3D vector calculation	
6.5	Further tests and results on bending	
6.6	Development of a commercial device	
6.6	5.1 Mechanical design	
6.6	5.2 Electrical design	193
<b>6</b> .7	Other applications	
6.8	Conclusion	196
7 Co	nclusions	
7.1	Summary of thesis	
7.2	Achievements	197
7. <b>3</b>	Suggested further work	
8 Re	ferences	

Appendix A	Prototype Manufacture Information
Appendix B	Test Rig Development Information
Appendix C	Relevant Data Sheets
Appendix D	Associated Publications

, · · · ·

# **<u>1 A Problem of Access to Buried Infrastructure</u>**

#### 1.1 Introduction

This chapter provides a qualitative analysis of the issues involved when accessing buried infrastructure. The aim is to describe the undesirable effects of excavation in urban areas, to introduce the common reasons for requiring access to buried infrastructure, and to discuss the materials and techniques currently employed by the utility industries. In conclusion there will be a suggestion that there is justification for further analysis of the problems facing utility companies with a view to developing a generically applicable system which will reduce the undesirable factors of the current systems.

## 1.2 Accessing buried infrastructure, a typical scenario

Figure 1.1 below shows a typical arrangement of services buried beneath an urban roadway. For one of various reasons to be discussed a buried service may fail such that access is required to facilitate a repair.

Figure 1.1 Example of service distribution beneath roadways in urban areas.



Current methods of providing access normally require open excavation such that one or more people may work hands-on with the buried asset. This leads to various undesirable side-effects in urban areas such as:

• Road closure to allow direct open excavation, leading to congestion and disruption of other urban services.

- Pollution generated by spoil and materials heaps which often become somewhat distributed around the vicinity.
- Pollution caused by noise, which is a genuine problem for people working or residing long-term in the vicinity.
- Permanent damage to surface structures such as roads and pavements which often subside as a result of sub-standard reinstatement.
- There are other less quantifiable effects such as accidents caused by distraction.
- There is often third party involvement at various stages of any process.

**1.3** Types of failure requiring access to buried infrastructure Service failure requiring access for attention falls into two categories:

#### 1.3.1 Serviceable failure

Serviceable failure implies the effects of operational wear over time and also rapid deterioration of faulty structures. The causes and nature of this type of failure are specific to the service carried and the type of infrastructure involved. However, failure of this category can be divided into two distinct groups:

#### 1.3.1.1 Serviceable failure at a specific point

Specific points likely to lead to failure include joints, junctions and valves. The nature of and cause of point failures are somewhat specific to the service being carried. For example a bad joint in a power conductor will most likely vaporise a two metre stretch of cable resulting in total failure of service, whereas the equivalent bad joint in a water pipe may either obstruct flow or cause a leak, resulting in a service maintained at reduced levels. If a leaking pipe network carries water it may be acceptable to tolerate the reduction of service, if the pipe carries gas however, then for safety reasons it is important to effect a repair immediately. Point failure of a cable service whether power or communications is most likely to result in total failure. The most common causes of serviceable point failure in cable are from poor installation (e.g., bad joints), and the effects of ageing. The most common causes of services

occur from both external factors such as forcing of corroded valve taps and from internal conditions such as over-pressurisation, blockage, and corrosion.

#### 1.3.1.2 Serviceable failure over a region

Gradual degradation of a service may occur over time. For example cast iron water mains suffer corrosion both internally and externally (figure 1.2). This has a double effect of both mechanically weakening the pipe structure and obstructing the flow of water as the internal corrosion builds up. A result and symptom of this particular problem is discoloured and low-pressure supply to the consumer. A problem specific to concrete piping is the pH balance of the enclosed water, if the supply is allowed to become too acidic the concrete pipe is rapidly dissolved. Similarly cast iron gas pipes are affected externally but not internally. Power cables buried for several decades (up to 50 years is common) begin to lose the mechanical and electrical properties of their insulators, which may lead to earth leakage, brittleness, and corrosion of the conductors. Failure over a region of network implies attention over that region and possibly an increased area. In these particular cases of refurbishment and replacement there are some specific techniques allowing the avoidance of total excavation. These methods will be described subsequently.





## 1.3.2 Failure due to 3<sup>rd</sup> party intrusion

It is a common event that buried infrastructure can be damaged by a 3<sup>rd</sup> party. The type of activity leading to failure is often excavation for building or for access to another buried service. Some buried services are located fairly close to the surface (for example communications cable) which may be crushed by heavy vehicles mounting pavements. The damage caused by 3<sup>rd</sup> parties is usually restricted to an easily identified point, which has both the effect of allowing rapid diagnosis/repair and apportion of responsibility.

# 1.3.3 The obligation of buried infrastructure owners to maintain their assets.

In previous years the water distribution companies have been able to tolerate what is actually a high percentage of loss through network leakage. These leaks in the distribution network alone account for an average estimated 3514 million litres of water per day throughout England and Wales, which is about 15% of that entering the network [1], these figures are based on an average of 6 years up to the end of 1997. However, since privatisation of the utilities it has become a matter of public concern that the providers of basic services operate efficiently to minimise costs to the consumer. Particularly in the case of the water distributors this requirement has been reinforced during the mid 1990's because of several years of low rainfall and subsequent depletion of water reserves. Therefore less obvious leaks in the distribution network are now treated where they may have been left before.

#### **1.4 Materials and techniques used in the maintenance of buried infrastructure.**

The development of technology associated with buried infrastructure is largely industry driven. Many of the specific techniques in use are not described in published material and there is very little relevant academic material available. Much of the information presented in this introduction has been obtained through industrial contacts and utility support services such as the Water Research Council. The aim of this introduction is to clearly demonstrate the level and diversity of technology geared to access and repair of buried infrastructure, including where appropriate techniques of detection, diagnosis, repair/replacement, and minimisation of urban disruption. It is intended that this introduction may be of unique use to any subsequent academic research of related areas. The information subsequently presented in this section is most detailed for the water industry because this industry is most active in maintaining its infrastructure. There is also mention of gas, which largely parallels methods used in the water industry. Cable for communications and power distribution are also briefly covered.

#### 1.4.1 Maintaining water distribution networks.

There are many and varied methods in practice within the water industry. Many of the novel systems are developed specifically to upgrade/replace/refurbish complete lengths of pipe without requiring total excavation. These methods have been developed because of the nature of the network, which can be prone to rapid corrosion and blockage over large regions. Water networks are also prone to points of failure in an otherwise sound area, for which traditional open excavation techniques are employed.

#### 1.4.1.1 Diagnosing a water distribution problem.

In some instances it is obvious when the performance of a water distribution network has been compromised for example a burst water main can cause massive disruption both at the site of the burst and for the community affected by loss of supply (figure 1.3). However, the majority of water loss from the distribution network is less obvious. The consumer may typically notice low pressure and in instances of corroded iron pipe, a discolouration to the water. Water distributors are now increasingly making use of network distributed flow meters to determine likely regions in which a leak may be occurring. Flow rates may be compared with computer generated models to allow advanced analysis of network loading. This information may be used to predict areas in which existing infrastructure is of insufficient specification to hold the actual capacity.

Figure 1.3 A major distribution main leak outside Seattle.



#### 1.4.1.2 Locating leaks in the distribution network.

Low pressure leaks and low volume leaks may be identified by the techniques described 1.4.1.1. This allows them to be located within a geographical region of the network, but pinpointing the actual location for the purposes of repair is a challenging task. 'Small' leaks are often not evident on the surface for a long period of time. The longer a small leak remains undiscovered the more likely it is to eventually lead to destructive damage through erosion and subsidence.

There is one particular effect of any leak, which may be detected and analysed to provide a good location: acoustic activity. Any leak will cause vibration and there are three main ways in which this may be detected [2]:

#### 1.4.1.3 Pipe vibration:

Pipe vibration occurs in the vicinity of any leak. This vibration is transmitted along the pipe and may be heard by listening devices applied to near-by exposed points such as valves and hydrants. Frequency depends on the nature of the fault, on the material of the pipe, and on the volume and flow rate of the water, attenuation of the vibration along the pipe depends on the pipe material and the type and number of joints to the listening device. Typical vibration frequencies are in the range 500-800Hz. Several listening devices may be used in the vicinity of a leak and methods of correlation applied to pinpoint the exact location [3].

1.4.1.3.1 Soil Impact:

Soil impact occurs when the fluid flows from the fault in the pipe. The fluid strikes surrounding ground material setting up a resonance typically in the range

20Hz – 250Hz. Most ground types rapidly attenuate this signal and therefore a surface listening device may be used to accurately pinpoint the location of the fault.

#### 1.4.1.3.2 Fluid Circulation:

As escaped fluid impacts with surrounding ground it erodes a cavity. In this cavity the fluid circulates causing a resonance in the range 20-250Hz. The nature of this signal is similar to that for soil impact and would be detected at the same time as listening for soil impact noise.

# 1.4.1.4 Actions to repair/replace faulty water distribution infrastructure.

There are two main categories to the topic of actions taken when repairing/replacing substandard infrastructure. These take the form of operation on a specific point or part and operation over a complete section or region.

#### 1.4.1.4.1 Operations on a specific network point.

Operations on a specific point will always involve an open excavation directly from above. It is this type of operation which so often leads to urban disruption as described earlier in the chapter. Open excavation is required because it is necessary to provide human access to the affected part such that it may be manually attended. These operations are undesirable but necessary.

There are however, applicable tools emerging to assist in point excavation. These tools are designed to allow the excavation of a very narrow and deep hole such that disruption of surrounding structures is minimised. Vacuum excavation is a key player in this field. The process of vacuum excavation involves high pressure air or water lancing of compacted ground material to break the medium into constituent pieces. Spoil is removed from the excavation through a tube subjected to strong vacuum. Various systems are based on this process and leaders in the field of vacuum excavation include: General excavating [4], Slabach Enterprises [5], and Vacmasters [6]. A major advantage of fluid cutting and extraction over traditional mechanical methods is the proven reduced rate of damage caused during excavation.

#### 1.4.1.4.2 Operations on a complete section.

Operations on a complete section of pipe (for example a 100m length of 15cm water main) have the potential to generate the greatest urban disruption. Fortunately there has been much effort in recent years to reduce the urban effects from working on large sections of buried infrastructure, resulting in methods which allow 'trenchless' access in some instances. In a generic capacity these methods require an entrance pit at one end, an exit pit at the other end and perhaps one or more access pits between, depending on the method used and the length of the section worked on. The following is a list of methods and their applications:

#### 1.4.1.4.2.1 Slipline.

This is a process involving the insertion of a new main through the interior of an existing but depleted main. Typically this method is used to upgrade pipe where modelling demonstrates that the existing diameter is excessive and the insertion of a smaller diameter pipe will allow sufficient capacity to supply the demand. The existing main is prepared by cleaning out any encrustation. To do this a rotating scraper is pushed through. Then a cable is pushed through the section and used to winch the new polyethene pipe into place. Minor draw-backs of this method are based on the fact that the new pipe does not make a tight fit with the existing main so it has to be structurally capable of bearing the full operating pressure of the mains water. Secondly, if a leak should occur in the new section then location is made awkward by the fact that water runs along the annulus rather than breaking free at the point of fault. Slipline is a process which is simple to perform and widely applicable to diameters 20 mm - 1000 mm. More information is available from [7,8].

#### Figure 1.4 The slipline process.



#### 1.4.1.4.2.2Die draw.

Die drawing is a process developed in recent years to line an existing main. The process of preparation is the same as for sliplining. The application of the new pipe is under extreme tension so as to reduce its diameter during laying. This is achieved by winching the polyethylene pipe through a die before going into the existing main. Once the pipe is in place the tension is released and it returns to its original diameter making a tight fit inside the old main. This process may be used when the existing main is structurally sound, as the tight fit implies that the old pipe will bear load from the mains pressure. The advantage of this process is that lower specification polyethylene pipe may be used, and because there is little friction during application large diameters and long lengths may be processed. There are safety hazards posed by the high tensions required to draw the pipe through the die (up to 30 tonnes). Akkerman [9] in the US are an example of producers and providers of tensed insertion processes, in this case for sewers.

#### 1.4.1.4.2.3Pipe bursting.

Pipe bursting is a method replacing existing mains with higher or identical capacity mains without the need to remove the old service. There are two systems in use which are in principle the same. An expanding head is pushed into and through the existing main with rigid rods. Either of two mechanical actions is applied:

Percussive moling: The oversize pneumatic hammer head is winched through the existing main cracking it and pushing it into the surrounding ground. New pipe is pulled into the void immediately behind the head. This process is useful when replacing existing infrastructure with similar size pipe.





Hydraulic rod system: A hydraulic expanding head is pushed through the existing main on a series of rigid rods. The head is capable of extreme forces to crack the strongest of infrastructure and push it into the surrounding ground. Once the existing main is broken away the hydraulic head is pulled back through the void towing the new pipe behind it. This process is suitable for upsizing the existing main.

Limitations of both these processes are based on the potential for damage caused by ground heave. It is not possible to use these processes in ground where there are other buried services, or where the main is close to the surface.

#### 1.4.1.4.2.4Cement Mortar Lining.

Cement mortar lining is an established technology which has been used to renovate cast iron mains for over 50 years. The existing main is scraped clean in lengths of up to 150m. A device is inserted into the pipe which sprays a lining onto the inner wall as it is pulled through. The lining provides structural integrity and reduces encrustation. However, there are many drawbacks of the process which have lead to its decline:

- Quality control is difficult because it is not easy to determine the thickness of the lining applied.
- The life time of the lining at 30 years is relatively short.
- The process relies on the existing pipe to be structurally sound.
- The process does not actually reduce leakage, and the scraping process can actually increase it.
- In soft water areas the high pH value of the lining leads to unacceptable alkalinity of the water supply.
- The process is not predictable. It may be necessary to pass the lining machine through the pipe more than once to achieve a sufficient coating.
- There are many customer discomforts involved. Down time is significant because each application of the lining requires up to 16 hours curing time. The pipe requires disinfection once lined.
- Water back-flowing from connecting pipes can damage the lining before it has cured.
- The cement surface has poor hydraulic properties and hinders water flow.

#### Figure 1.6 Cement mortar lining process



Mixer, Pump and winch

There are still many companies [10,11]offering cement lining services because it is very effective when refurbishing large diameter pipes (greater than 1m).

#### 1.4.1.4.2.5Epoxy Resin Lining.

This process is similar to that of cement mortar lining. The main differences are that the lining is much thinner and smoother so that both flow and capacity are improved. This is not considered a long term solution to pipe rehabilitation as the lining will breakdown in time.

#### 1.4.1.4.2.6Narrow Trenching.

In the same way as conventional open cut trenching requires continuous excavation, so does narrow trenching. However, there are increasing numbers of trenching systems designed to cut a very narrow path, which is barely wider than the service to be installed or accessed. The advantage of this is that complicated paths can be easily laid (not always possible with systems such as percussive moling and sliplining) with minimal disruption to surrounding structures. The range of trench widths presently used are from 63mm – 127mm, and the range of cutting depths are typically to 600mm or in some cases 1000mm. Mechanisms employed to perform the cut are either carbide toothed wheels (often known as 'rockwheels'), or carbide toothed belts. There are many developers of this type of equipment, two of the leaders in the field are: CASE Construction Equipment [12], and John Deere Commerical Worksite Products [13].

#### 1.4.1.4.2.70ther area refurbishment processes

There are many other processes and variants on those described above when effecting a repair on buried infrastructure, in particular one of these is to insert a collapsed polyethylene tube into the prepared existing pipe and use a heat treat and curing process to expand it [14,15]. Other processes are described in varying detail by Evin [8].

#### 1.4.1.5 The bias of water distribution on this research project.

This project was conceived in the summer of 1995, a particularly bad year for drought throughout the UK. There was much criticism directed at water companies from OFWAT [16] and through the media as reserves ran low. This was compounded by the public revelation that although these privatised

companies were operating with excessive profits, their infrastructure was substandard and in worst cases some companies were losing half their water supply through distribution leakage. The issue of leakage became and remains a high profile topic.

It was the requirement to rapidly address any identified leakage which lead to the idea that an efficient system of access and repair would be beneficial. It was considered that any technology developed for the provision of access to buried infrastructure could be of use in a generic capacity to any industry with an interest, so including all the owners of buried infrastructure. As a result there is a brief description of gas, electricity and communications infrastructure given below.

#### 1.4.2 Maintaining gas distribution networks.

Gas distribution systems are similar in topography to those of water. A similar set of processes apply to the access and maintenance of gas pipes. As a matter of safety any gas leak must be dealt with as soon as possible, and these are generally caused by third parties because gas exerts very little operational wear on its distribution infrastructure. There are some additional tools available to this industry in the form of internal inspection devices known as PIGS. These devices were first used for internal inspection/maintenance of pipes in the 1940s by T. D. Williamson Inc [17], and they remain a leader in this field today.

#### 1.4.3 Maintaining cable networks.

Cable includes power distribution networks and communications networks (which may be both electrical and optical). These infrastructure types differ from those carrying fluids in three fundamental ways:

- Cable offers no internal access. If access is required to buried cable is required it is obligatory to use open excavation.
- Cable offers little or no mechanical resistance to the service it carries. Apart from the effects of the electrical resistance there is little operational wear on the cable networks. This increases the working

life of the infrastructure and reduces the frequency of refurbishment related access.

 Cable services don't suffer reduced performance through occurrences such as leaks, although the analogy can be drawn with the pipe blockage to demonstrate how they suffer from areas of increased electrical resistance. Such an area effects a potential gradient and therefore loss of energy occurs.

#### 1.4.3.1 Power distribution.

Electrical power distribution networks have been increasingly installed throughout the world over the past 100 years or so. In the UK, the National Grid has developed as the transmission system responsible for taking power from the generator companies and delivering it to the distribution companies. Being a power transmission power company it is necessary to maintain system efficiency by reducing transmission power loss to a minimum.

The efficiency factor concerned with power loss in a conductor is current. Over a unit length of conductor the power lost to electrical resistance is proportional to  $I^2$ . Now, because power is equal to the product of current and voltage it makes sense to maximise voltage and hence minimise the current flowing in a transmission line. The important point is that National Grid transmission lines are extremely high voltage, usually 275,000 volts or 400,000 volts. This makes them expensive to insulate safely. Hence the National Grid owns approximately 7000km of over ground cabling and just 600km of underground cable.

In contrast to the National Grid, localised distribution companies are not concerned with transmission of large powers over considerable distances. Further to this, distribution companies are required to deliver their product at the relatively low 240 volts or 415 volts r.m.s for the majority of their consumers. Hence they distribute electricity at the medium voltage of 38,000 to 10,000 volts to local transformers. Cable carrying electricity of this potential may be efficiently insulated and buried for practical purposes in towns and cities [18]. The result is an electricity distribution company owning far more buried than over head cable.

#### 1.4.3.2 Communications cable networks.

Buried communications increasingly take the form of high band-width optical cable allowing services such as cable television and internet access as well as telephone. New installations of optical and coaxial cable tend to be ducted quite close to the surface. Ducting is used so that the cables themselves do not require expensive armour and insulation from the hostile environment. The main problem for communications cable is subsidence crushing it, when for example, heavy vehicles mount the pavement.

#### 1.5 Key phases of the generic process to access and repair buried infrastructure.

It is possible to present any access and repair procedure as a logical series of key phases. It is helpful to do this so that any subsequent development of generic technology may target the specific phases which lead to greatest disruption, cost and inefficiency.

Diagnosis.

• Owners of buried facilities perform a diagnosis of the problem. This is based on information reported to them about localised performance of their infrastructure. Information may be provided internally by routine inspection procedures, or by external reports such as members of the public smelling gas. Typical diagnosis' are specific to the type of service involved.

Location.

• It is advantageous to define the location of the problem as accurately as possible. Various technologies specific to infrastructure type allow a degree of localisation, these can tie in with routine inspection procedures.

Access.

• The nature of the access method to be used depends on several factors: type of infrastructure, location of infrastructure relative to other buried facilities and surface structures, the prescribed treatment to the infrastructure, and the cost of the relevant options available. In many

cases the only option provided by current technology in the process requires open cut excavation.

Repair.

• There is no generic repair model. This part of the process is dependent on the exact nature of the infrastructure, and the type of the repair (replacement or refurbishment). It could be summarised that a general case involves either replacement of a section or some form of refurbishment of a section.

Extraction.

• Extraction is a term used to refer to the full reinstatement of underground and surface structure. This will almost always involve a degree of filling-in. Filling-in must be done to a high standard to ensure no future subsidence.

#### 1.6 Conclusion

#### 1.6.1 Summary of chapter 1

This chapter has introduced the issue of access to buried infrastructure, highlighting the disadvantages of the open excavation techniques in common use. The nature of buried infrastructure has been detailed and the common problems occurring which necessitate access have been discussed. A generic key phase has been identified which is applicable to any access and repair situation. This segmented process will be used in the analysis of an ideal system to reduce the undesirable of current practices.

#### 1.6.2 Research opportunity arising

As a result of investigating the access/repair processes for buried infrastructure, it has been possible to present a number of undesirable consequences arising from these operations (1.2). A generic analysis of typical processes (1.5) has allowed the identification of key phases, especially those which most contribute to the undesirable effects. As such the next chapter studies specifically the possibility of reducing undesirable effects with reference to the identified key phases, through the employment of a very novel and at this stage conceptual system. Chapter 2 concludes by specifying the technical research requirement of this project.

## 2 A novel concept on access and repair of buried infrastructure

#### 2.1 Introduction

The first chapter gave a view of some of industry's current methods to access and repair buried infrastructure, and significantly for this chapter, suggested the average generic process involved as key stages. The author entered this doctoral research programme in a position of holding novel ideas towards a solution to the problem posed. This chapter divulges these ideas in a way which describes a conceptual solution. Many technical challenges become apparent, any of which might be suited to academic and engineering research. As such a conceptual solution is presented around some fundamental parameters:

- The solution should be generically applicable to a broad range of situations, and different infrastructures.
- An important part of the solution lies in adaptation to new situations as they arise. (That is, a high degree of configurability and upgradability to encompass increasingly diverse underground repair operations).
- The author wanted to exploit the best features of an existing natural system which has been neither developed by man, patented or copyrighted by man, or owned by man. This system provides solutions to many of the conceived problems and gives good basis to develop sub-systems to address more situation-specific problems.

Chapter 2 begins by identifying an existing natural system which in many ways already addresses the problems that are likely to be encountered when embarking on a project of this specific nature (tunnelling and spoil processing). Some analysis of the existing system is presented in both academic and practical terms. Ways of adapting the desirable features into a synthetic mechatronic device are explored. The key technological challenges and philosophical quandries are identified with a brief suggestion as to how subsequent research may address each area. A decision process is described in which justification is given to the need to focus on one area of relevance and interest for the purposes of this thesis. The selected technology field is described in some more detail leading to the identification of a focused research challenge, the content of which will comprise the main body of material subsequently presented in this thesis.

#### 2.2 An existing system to exploit

The evolution process aims to encourage survival through adaptation to changing environments. Natural history, according to Darwinian Theory, has seen evolution provide organic solutions to the most demanding of survival problems. Today we see life in the air, the sea, on land, and underground too.

In this instance we are most interested in the mechanics of underground animals. Mobility is the main mechanical issue. For those animals living underground which are big enough to find the ground an obstacle there are two ways of moving between locations: displace matter in the intended path, or restrict movements to existing tunnels and cavities. One particular animal with the ability to move between locations in both these ways is the earthworm. The suitability of the earthworm as a mechanical vehicle for key hole surgery of buried assets is discussed in 'An introduction to robotic tunnelling with biological inspiration'[19], and 'Mechanical Earthworm to facilitate the access and repair of buried infrastructure'[20].

#### 2.2.1 Earthworm Locomotion Mechanics

The mechanics of worm locomotion are based on a hydraulic principle. The worm body consists of a fluid filled tube as shown in figure 2.1. The tube although continuous, is segmented by groups of circumferential and longitudinal muscles in the wall. The circumferential muscles constrict the tube to expel fluid out of a local area, or expand the tube to sink fluid into a local area. This process operates on a principle that the fluid is incompressible and so maintains a constant volume within the cavity. Using its circumferential muscles in this way various significant profile alterations are possible. Fig 2.1 The worm is a fluid filled tube



Figure 2.2 shows a typical region of segments with dilated circumferential muscles. To maintain a constant volume it is expected that as a segment increases its diameter it proportionally decreases its length. However, because segments are linked and fluid is allowed to flow between them it is possible that one dilating segment may maintain its length and hence increase its volume. In correspondence one or more other segments will decrease their volume (figure 2.3).





Figure 2.3 Variable volume of segments



Note: overall volume is maintained

This flexibility of segment configuration may be used to advantage in several ways. Dilated segments are used to create a fixture in a hole. The worm may then use this fixture as a purchase point when thrusting into the ground. Or the fixture point may be used as a basis for locomotion along an existing tunnel (figure 2.4).

#### Figure 2.4 Basic worm locomotion



Figure 2.4 represents a basic principle of locomotion without considering factors such as multiple fixture points and articulation. The real world process can be very complicated.

Locomotion through an existing tunnel is a relatively simple control process because the worm body is a very compliant structure [21,22]. The worm body will travel around bends in the path without need to apply deliberate articulation. Articulation is most important to locomotion when the worm is creating a new tunnel. In this situation the worm must make decisions about direction and take active measures to control the path of its journey. At this point the longitudinal muscles of each segment become relevant.

#### Figure 2.5 Use of longitudinal muscles in articulation



Unbalanced tension of longitudinal muscle causes bending.

Figure 2.5 demonstrates the principle of muscular operation for segment bending. For a real worm segment the process is more complicated because there are not three distinct longitudinal muscles as indicated but a complete covering all around the wall.

## 2.2.2 Passage through a solid

The earthworm is able to create new paths or tunnels in either of two ways: in loose soils it is possible to compact earth with a wriggling motion of the body segments (figure 2.6a) [21,22]. If the soil is hard or already highly compacted the worm may cut a path by 'eating' soil in its way, the soil passes through the worm and fills the path behind it (figure 2.6b). In the right conditions the passed soil is of lower density than the surrounding soil and a hydraulic pressure is created. This can result in worm casts being left on the surface of the ground.

#### Figure 2.6a-b Earthworm methods of handling soil



In summary the earthworm is an adaptable creature which possesses multiple capabilities to travel through the ground, by effective displacement of material and an efficient locomotion mechanism. The analysis of an earthworm as an underground vehicle leads to the question, in this instance, as to if it might be possible to replicate its best features in a man-made form? The feasibility of this question is explored in the subsequent section.

#### 2.3 Technical challenges of a man-made earthworm.

There are several key technical challenges envisaged in the development of a man-made earthworm. These are challenges both in the realisation of the basic vehicle, and also in the adaptation to the application of keyhole surgery of buried infrastructure. Key challenges are identified and briefly discussed below:

#### 2.3.1 Locomotion

Locomotion is key to the suitability of an earthworm as a vehicle for underground transportation. Quite naturally the earthworm possesses a very 'organic' mechanism in a collection of muscles used to achieve the desired motion in locomotion. The basic method of operation is described in the previous section 2.2.1. The question arises as to how this same set of actions could be attained through synthetic means. It is perceived that there are two distinct approaches to actuation:

- that which directly mimics the worms' locomotion muscle groups with corresponding mechanical actuators,
- and that which achieves the same actuation profile in time and space through a different actuator layout.

The benefit of copying the exact locomotion method employed by the earthworm is that the instructions are already there - as long as the control system for hundreds of actuators can be deciphered. The benefit of designing a custom earthworm is that the developer may select the best parts of different models to achieve the optimum result. Suggestions as to how each approach may be taken are given towards the end of this chapter.

#### 2.3.2 Location and positioning

It is necessary at any time during a tunnelling procedure to know exactly where all parts of both the vehicle and its excavation are located. Location would be with reference to some point on the surface, for example the supervisor's control point or the point of entry. This is necessary so that known obstacles can be avoided and the vehicle trajectory can be correctly maintained toward the target location. This is a challenge to existing technology such as G.P.S. which cannot work beneath the ground surface. It is proposed that there may be at least two different approaches dependent on the nature of the tunnelling operation. If an open tunnel is created through which there is a clear passage, then it may be possible to employ laser alignment through the tunnel [23].





If however, the vehicle maintains no direct physical link with the surface and is entirely autonomous then it would be necessary to consider such techniques as acoustical correlation. In such a situation the vehicle might act as a beacon emitting an acoustic signal which would be detected with varying amplitude by different listening devices at known locations.

Figure 2.8 Locate vehicle through ground via acoustic correlation



Proportional amplitude of signal at each pick-up determines relative position of worm

### 2.3.3 Ground cutting

It is envisaged that an urban-burrowing robot would be expected to displace allmanner of ground material. From hardcore and building rubble to thick clay and sand, there are many typical types of solid material on which pavements and to a certain extent roads are founded. There are different optimised processes for removing and handling each type of material, a high-pressure fluid jet might be most effective for one material whilst a mechanical drill might be more efficient on another. It is not the intention at this stage to profile each ground type and match it with a set of cutting technologies. The intention is to demonstrate that there will be a need to either carry more than one cutting system, or for the vehicle to hold capacity for interchanging of cutting tools depending on the ground type experienced.

This suggestion of a tool magazine in the worm head can be useful in the conceptual handling of tools for a repair process. It is envisaged that there may be tools required to cut and weld, and those which would be used to position materials and resins. There are other potential users of a tool magazine which will be discussed in other areas of the key technological challenges.
#### Diagram of interchangeable tool magazine 2.9



## 2.3.4 Spoil processing

Cutting a hole in ground material is one problem. Moving ground material away from the cutting face is another problem. It has already been demonstrated how the earthworm may manage the spoil handling problem in two different ways (2.2.2). However, the analogy to a tunnelling earthworm becomes slightly tenuous at this point because the nature of urban ground is rather different from that of garden soil. Garden soil is reasonably homogenous to the earthworm, which if it does find an obstacle in its path, who is to say that it is bothered about deviating, perhaps it will alter its course and continue tunnelling through soil until it hits the next obstacle (figure 2.10).



On the other-hand, a mechanical device inserted into the ground has a distinct objective and target destination. Undoubtedly the mechanical device will uncover unforeseen obstacles which must be circumnavigated at worst, or if appropriate forcefully removed. Further to this, it is expected that urban ground is not very homogenous to a mechanical earthworm, which will be required to 'digest' large lumps of solid material as well as those smaller particles which pose less problems.

Accepting that there is an upper limit on the size of particles which may be passed through any spoil processing system, there are two philosophies on handling those parts which are too big:

The first philosophy is to ensure that whichever cutting technique is employed, its effect on large particles is to cut, abrase, or crush them in the action of removal from the cutting face. The second philosophy is to accept that not every cutting method employed will be able to reduce particle size, and that large particles entering the worm 'mouth' should be crushed in hydraulically powered jaws. If large particles are removed which are too big even for the jaws, and too hard to be drilled (for example a granite boulder 150mm diameter) these must be displaced into the tunnel floor such that the vehicle may pass over them (Figure 2.11).



Figure 2.11 Displacement of boulders to tunnel floor

If a very large obstacle is uncovered and found to be firmly embedded then it will be necessary for the vehicle to deviate from its intended path such that it may

31

reach its destination by circumnavigation. It is anticipated that large obstacles might include other buried infrastructure and foundations as well as buried rubbish and natural ground features such as surface rock.



## Figure 2.12 Circumnavigation of large obstacles

## 2.3.5 Underground vision/obstacle identification

There are two concepts to the term 'underground vision'.

## 2.3.5.1 Global vision:

The vehicle is moving through a solid ether in-which there may be buried a variety of different structures and objects. Some of these buried things may be in use and must not be disturbed, others may be obsolete structures and rubbish, which may be disturbed if necessary. It is important to have some idea of what structures and objects exist in the operating locality so that the vehicle path may be planned around them. It may be possible to estimate the location of buried services through records such that the operator is 'aware' of 'dangerous' areas. However this does not address the issue of planning around unrecorded items such as obsolete infrastructure and rubbish. To confirm the suggestion of records and to help identify other relevant items it would be most useful to have some method of 'seeing' through the ground.

There are various techniques in use to detect buried objects, they have applications such as archaeological surveying [24] and buried infrastructure surveying. These techniques are discussed in detail elsewhere. There are limitations to the capability of these systems to determine accurate information such as depth of anomaly and nature of anomaly.

In summary, the global vision concept is one of mapping the underground environment in three dimensions before the excavation operation begins. This map may be digitised for use in an automated system or may be presented visually to a manual operator of tunnelling equipment. In this sense the vision requirement is one of relative spatial awareness.

#### 2.3.5.2 Local vision

In addition to knowledge of the 3D environment through a global vision system it is necessary to provide comprehensive feed-back on processes and status at the cutting face/repair site. The nature of feed-back may in some ways mimic the sensory capacity of humans, not just in the visual sense:

#### Eyes

- Perhaps the most intuitive and applicable of the senses, a simple monochrome vision system or a stereoscopic colour vision system sensitive to an extended range of wavelengths. The right vision system at the head of the worm would provide the most immediately useful information to a remote operator of both the tunnelling procedure and the repair procedure.
- Problems posed by this suggestion include at the very least an issue of illumination in an otherwise completely dark environment, and an issue of maintaining quality of vision in a dirty and wet environment.

Ears

 An ability to hear what is happening at the worm head may be useful.
It is anticipated that changing ground qualities may be detected through changing acoustics of the cutting process. There may be other benefits not yet imagined.

#### Touch

• There are two facets of interest in this instance, temperature and load. In an environment where many abrasive actions are taking place, as well as a degree of materials processing (when effecting a repair) there

33

may be generation of excessive heat. A monitor of heat would be beneficial in controlling the process. With respect to load, the interest is in resistance to cutting and drilling of ground material and obstacles. The benefits of load feed-back are many in the field of process efficiency and real-time feasibility evaluation of achieving short-term objectives (for example is it worth trying to drill through an old brick, or would it be better to circumnavigate it?).

Taste and Smell

• The idea of taste and smell as sensory terms refers to an ability to sample the chemical constituents of materials and the ability to analyse fumes for their chemical content. These facilities may or may not be of use to the tunnelling system in the operators efforts to determine information such as the nature of the obstacle or the stage of the repair process. However, extending the concept to encompass any method of material identification, the inclusion of a hall-effect device might be useful to determine if an object were ferrite.

# 2.3.6 Power supply

Power is an issue to the operation of the vehicle. Major consumers of power are expected to be the cutting process, the locomotion process, and the repair process. It will be indicated (2.6.1) that there are two possible configurations of the whole system, either entirely remote, or tethered through the entrance tunnel to the surface. The nature of the power requirement may be one of the more significant factors in determining which configuration is more feasible. To maintain the simplicity and relevance of this discussion subsequent analysis is presented qualitatively.

If the vehicle were entirely remote it would need to carry its own power supply. It may be possible to fill the void (equivalent to a fluid sack) with an electrolyte solution so as to devise for example a custom lead acid battery. However, the safety of doing this may be questionable from the point of view of damage leading to short circuit and fire, or leakage of electrolyte into the surrounding environment. It would also be questionable as to whether the battery could store the required energy for an operation of average duration and complexity.

The cutting and crushing of ground materials is a power intensive process, a process which naturally lends itself to hydraulic application. It would not be feasible for a self-contained vehicle the size and shape of the robotic worm to hold a hydraulic pump or the energy required to operate it (see section 2.7 for size estimates).

In summary, it may be possible to provide all power requirements from on-board reserves. There would be significant challenges and limitations in doing so. It may be preferable to locate the energy reserve outside the vehicle and provide a physical link between the two. This would allow greater flexibility in selection of the ideal energy medium and remove the theoretical limitation on total energy available for one operation.

## 2.3.7 Control

The question arises as to how to control a vehicle of such complexity and unusual application. There are three key challenges:

#### 2.3.7.1 Locomotion control

At the lowest level it is necessary to know the status of each locomotion actuator. This status information might be absolute position. One level up it is necessary to evaluate if each actuator is in the status expected. The next level builds a map of the vehicle profile determined by the status of each actuator, this is compared to the map expected. A decision must be made about what profile to arrange next. This decision is based on higher level objectives such as 'get to target location', and is dependent on high level questions such as, is the current locomotion strategy effective? The high levels of the process may be entirely automated through a knowledge base and inference engine, or may be supervised by a human operator, or entirely manual. However, it would not be feasible for example to expect a human operator to manually set and read each locomotion actuator.

# 2.3.7.2 Ground cutting and spoil handling

From the mechanism employed to select and present the requested tool to the measurement of how efficient the cut is and how effective the spoil removal system is various levels of control are required. Decisions about the best tool to use may be automated depending on a fixed or adapting knowledge base, or the decision may be made by the operator depending on factors such as progress rate and ground type. A low level control system may be employed to actively pump spoil through the worm body in a peristaltic process. The mouth of this debrisingesting device might be inanimate or may be actuated to allow active positioning.

#### 2.3.7.3 Repair process

A process of teleoperation or automation. It will be important to allow dextrous use of the repair process tools such that they may effectively replace human hands. Quite independent of issues of automation and manual control it will be necessary to provide detailed feedback of tool status and work piece status. It is envisaged that the majority of procedures will be performed manually by teleoperation, perhaps making use of advanced virtual-immersion technology to provide visual feed-back and tactile feed-back. In such a way tools may be used reasonably intuitively. Specifying manual control of repair process tools would remove complexity of automated control.

This section has highlighted the complex nature of the challenge to develop a synthetic earthworm. It has in no way covered all the issues which will require considerable effort to succeed, but it does give a broad indication of the diversity of technological fields which may be applied to a refined system. The next part of this chapter considers the synthetic earthworm as a mechatronic system, making use of graphical presentation to clarify the inter-relation of various concepts.

#### 2.4 The mechatronic device – some applications

The earthworm provides an interesting base vehicle upon which to apply mechatronic principles. The natural worm is in segment form, most segments being mechanically similar except for certain segments holding specific functions such as front and rear segments and reproductive segments. The repetitive modularization of the earthworm body lends itself to applications of distributed control/hierarchical control, redundancy engineering, case specific configurability, and sensor and actuator networks.

#### Figure 2.13 Segmentation of the earthworm



Configurability is of immediate importance because it is fact that the repair process (for example) for an impact damaged power cable will be different from that of an impact damaged water main. To allow the system to be useful in both scenarios the earthworm vehicle must be able to either hold all the tools and materials it would ever need for any job (unrealistic) or hold a changeable magazine of tools and materials suited to a specific job.

Actuation modelling is an important aspect of proving an understanding of earthworm locomotion and will be of benefit when applying the natural process to a synthetic mechanism. As will be discussed further in this chapter, there are many options and many restrictions on how robotic earthworm-like locomotion may be achieved. Figure 2.15 demonstrates the simplest actuator model with its spatial relationships and degrees of actuation for a single segment of the vehicle.



e.g., power cable tool head

It is envisaged that there may be several appropriate control strategies to successfully control individual actuators within individual segments within a vehicle which may be tailored to specific needs (i.e., the number and complexity of segments may vary from job to job). These various strategies must be investigated for their overall impact on system robustness, efficiency, and complexity. Strategies of interest additional to conventional monolithic centralised control include:

- top-down hierarchical control, employing various levels of subcontrollers to perform and monitor increasingly lower level tasks.
- Communal distributed control via a networked multi node environment, where each 'object' within the system affords a degree of intelligence and many nodes are able to take system control when appropriate.
- Stimulus and response control, where system elements (mechanisms) are clustered such that one cluster's actions are perhaps partially or wholly determined by those of its neighbours. In this process many actuation and other operations are pre-programmed as sequences which may be easily initiated by a central controller.
- Sequential control such that a simple trigger from any controlling unit may initiate a series of actions. In analogy to a neural network, operations may be tuned in real-time by various feed-backs.

38

Figure 2.15 Segment degrees of actuation



Actuator compound function (e.g.,)



angle of articulation and length as arc



## 2.5 The selection of a suitable research problem

It has been shown throughout this chapter that the project is very broad and has identified many different technological challenges which must be addressed before a fully functioning product is forthcoming. It is clear that each of these technological challenges in its own right is worthy of significant study to PhD level. As such it was necessary to make a decision about where to take the project for the new research phase. There were two main options open. The first option was to maintain the project at a broadly conceptual level with a large emphasis on modelling. This route would lead to a reasonably refined theoretical proposition of the intended system, and would be most useful to subsequent researchers in the field. However, this route would not allow time for specific focus on a particular technical problem in great depth. The second option was to leave the initial investigative and analytical stage as it has been presented, with the intention of concentrating efforts on the in-depth research into problems posed by a particular relevant field of interest. This would provide the opportunity to focus research and thesis into a technical investigation with the intention of presenting at least a prototype of some part or system which was novel.

The decision was taken to opt for a focus into one particular area of interest. The reason for this decision was the desire to move beyond the conceptual stage and rapidly develop a working system of some description.

The technical area selected for development was that of locomotion. As the project progressed it became clear that this in itself is multi-faceted topic and as such there has been further refinement into an even more focused field. The remainder of this chapter will expand on the field of earthworm locomotion indicating the key components and the options available. The chapter will conclude by highlighting the exact topic for subsequent research.

#### 2.6 An expansion on locomotion.

To begin this section on synthetic-earthworm locomotion, it is necessary to understand the method(s) of creating the passage through which the vehicle will travel (that is to say, spoil handling), as this may have a significant effect on the way in which locomotion should be achieved.

It was considered possible to develop a tunnelling system, with a mechanical worm as the tunnelling vehicle, in either of two ways: that which merely displaces spoil within the ground, or that which extracts spoil out of the ground. These methods have been described in detail below. However, a third case involving soil compression to create tunnels, as used by the earthworm animal (figure 2.6) will not be considered because this method causes structural alteration to the ground, and leaves no material to back-fill the tunnel.

The danger of the process described in figure 2.6 is that compression of surrounding ground material may cause damage to nearby buried services and perhaps ground heave which damages surface structures. These effects are noted by some pipe moling processes which are based on expansion of surrounding ground; these have been described in chapter 1. The sections below describe the two processes of spoil displacement and spoil extraction.

## 2.6.1 Spoil displacement

This system involves removing spoil from the path of the vehicle, passing it through the vehicle and packing it out the rear end to fill in the tunnel (Figure 2.16).





In this system the vehicle would need to be self-sufficient as there would be no physical link other than wireless communications. Advantages are that the vehicle need not extract along the entry path, and that structural quality of the ground is always maintained. Disadvantages are that communications through the ground might be difficult, the vehicle would be limited by the amount of energy it could store, the vehicle must carry all materials and equipment required to complete a job, and it may also prove difficult to extract if a failure were to occur.

## 2.6.2 Spoil extraction

Spoil extraction involves the creation of a continuous tunnel from entry point to cutting vehicle. The soil cut from the vehicle path is passed through the tunnel to a temporary store outside (Figure 2.17).



#### Figure 2.17 Spoil handling by temporary extraction to the surface

On extraction the soil is passed back down the tunnel to the cutting vehicle. The vehicle reverses out of the tunnel packing the soil behind it. The main advantage of this system is that a continuous physical link can be maintained allowing: the power source to be kept out of the vehicle, hardwired communications, relatively easy extraction in the case of vehicle failure, passing of repair materials from outside the tunnel to the vehicle tools, and potentially easier positioning and alignment. Complications are created by the need to provide: a temporary tunnel

lining, and an extra system component to handle pumping of spoil through the tunnel. In this system the vehicle must extract along its entry path.

It can be seen that the main difference between the two cases presented above is that one can make use of pressure differential to assist motion and the other cannot. If spoil is moved from the head of the worm and released at its tail, then when enough material is built up so that the worm is enclosed any subsequent displacement of spoil to the rear of the worm will tend to push the vehicle into the void it is creating.

#### 2.6.3 Compliance:

The earthworm is a highly elastic structure. It is possible for the earthworm to deform along both active and passive axes if required. The earthworm may also choose to resist deformation if required. This ability is an important feature of the earthworm's locomotion mechanics. It allows the earthworm to pass through small holes, and to assume the profile of non-circular holes. This can be done without conscious effort as the body simply relaxes to make the best fit (refer to figure 2.11).

If this ability were translated to a synthetic version of the worm, then in principle the actuation complexity and the control complexity when following an existing tunnel could be reduced; even if the tunnel were not straight. A non-compliant vehicle would be required to sense deviation in the path and accordingly control the motion actuators to steer the vehicle round a bend. A compliant vehicle would not need to know that its path was not linear, the elasticity of the body would tolerate a high degree of deformation caused by the tunnel walls guiding it round a bend (analogous to trains running on railway track).

However, the key factor is in choosing when and when not to allow compliance. The worm must be forceful against resistance to make progress, further, if as in the proposed application, the worm is required to cut its own path through solid ground then compliance must be restricted along the active axis. An inherently compliant actuation material will require closed loop control to determine the relationship between extension along the active axis and load on the active axis for a given power input. If there is no feedback from the actuator and the power input is flat then the following static results could be expected by Newtonian Theory (from standard design/engineering texts):



note divergeance from expected actuation

As indicated in the graph above, open loop actuation is affected by load. Feedback is required to compare the expected actuation with that achieved, and the power supply adjusted accordingly. In this way, simple actuators can be combined with simple controllers to produce reasonably load-independent actuation.

# 2.6.3.1 Closed loop control of actuators

The charts illustrated below indicate how feedback of actuator attainment may be used by a controller to vary the power to the actuator such that the desired actuation profile may be achieved regardless of load (within certain limitations).



Controlled power profile under minimal load



It can be seen in each case above that identical actuation profiles were obtained under different loading conditions by employing some form of attainment feedback to a power controller. In each case the instruction was to obtain an actuation of 8 units in a linear manner. Feedback was used to determine the resistance to actuation and the power delivery adjusted accordingly.

Examples of feedback sensors and their accompanying power controllers may be linear potentiometers for linear displacement actuators, and circular potentiometers for rotary actuators. These systems may encode either an absolute state or return incremental states. A typical control loop might be as indicated in the figure 2.18.

Figure 2.18 A typical control loop example



## 2.6.4 Methods of achieving worm-like actuation

It is proposed that there are two feasible ways of achieving worm-like actuation. The mechanical parts employed in each of the two ways may not be very different, but the philosophy behind their layout and control is unique to each.

## 2.6.4.1 Natural worm actuator topography

This philosophy is based on the supposition that nature is a better designer than the best human minds, whether this be consciously or by some process of evolution. As such the ideal model on which to base the synthetic earthworm locomotion system already exists and needs only to be understood, copied and applied to the man-made system.

The basic earthworm locomotion model has been described earlier in this chapter and the requirement is one of translating the two distinct muscle groups into mechanical actuators. Figure 2.19 illustrates the function of longitudinal muscles for bending within the earthworm.

### Figure 2.19 Longitudinal muscle groups



Length-ways muscle groups as used by earthworm animal to achieve bending and segment length change

### Figure 2.20 Circumferential muscle groups



Circumferential muscle groups as used by earthworm animal to achieve dilation and constriction of segments

There are problems associated with the attempt to match a suitable actuator technology to the requirement demonstrated in the two figures above. Some starting points are discussed towards the end of this chapter.

# 2.6.4.2 Custom-designed worm actuator topography

This philosophy accepts the successful operation of the earthworm locomotion system as presented by nature. However, it also caters for the limitations of current engineering technology the possible differences in specific purpose and requirements between the earthworm animal and the earthworm robot. The intention of this philosophy is to arrange a compromise so that the objective is fulfilled efficiently. One example of how a custom-designed actuator topography might be arranged, is based on a two dimensional grid within an elastic material.



#### Figure 2.21a,b,c Actuator array

Each element of the grid is an actuator in either the x or the y axis, which are perpendicular to each other. If every actuator is simultaneously effected to the same extension then a 2 dimensional scaling of the material will result. This is illustrated in figure 2.22a. If however, actuation is restricted to those actuators in one orientation then the result might be illustrated by figure 2.22b.

Figure 2.22a,b Basic deformations caused by integral actuator grid arraya) 2D deformationsb) 1D deformations



Having established a compound sheet material (with embedded actuators arranged as above) it may now be possible to consider deformation of the sheet form so that it is no longer flat. If the sheet were to be formed into a cylinder such that there were closed rings of actuators around the circumference and those perpendicular to the circumferential actuators were running along the length of the cylinder, it could be possible to attain worm-like locomotion through appropriate control. Figure 2.23 demonstrates the roll-up principle.

### Figure 2.23 Roll-up of 2D actuator grid into cylinder



# 2.6.5 Examples of appropriate actuation technology

The earthworm animal achieves its mechanical and locomotion properties through high integration of muscles within its exterior wall. Further to this, the nature of muscle is to be soft and compliant in its relaxed phase, but hard and less compliant in its contracted phase. Muscle may be thought of as a structural tie, as such it exhibits tensile strength rather than compressive strength. Tensile loading is achieved by attempting to shorten the length of the tie (muscle contraction). Muscles do not actively extend, they must elongate after contraction through a combination of interior elastic memory and external load. Muscles are often arranged in organic systems such that they are paired antagonistically. This allows the contraction of one muscle to facilitate the elongation of an opposing muscle, which is analogous to a double acting piston.

In the earthworm animal the longitudinal and circumferential muscle groups act as antagonistic pairs with the constant volume fluid sack providing a mechanical couple (see figure 2.20). The challenge to a designer wishing to mimic the properties of natural muscle are great. Conventional actuation does not lend itself readily to the compliant nature of organic muscle. With a few exceptions it was necessary to consider the nature of actuation systems, including coupling devices, which would be a best compromise in providing earthworm-like properties:

There are those systems which may be said to directly mimic natural muscle in-asmuch as they exhibit tensile strength, and that they will tend to contract on request. There are other actuator systems reliant on coupling mechanisms to achieve the required exertion profile.

# 2.6.5.1 Polymer gel.

The possibility for certain polymers to act as muscle material was first published by Katchalsky [25] and Khun [26] in 1949. However, until recent years the technicalities involved have prevented practical application of polymer gels in actuation. The science of polymer gels is specific to chemistry, technical information may be found elsewhere [27-30]. Polymer gels can be made to swell or shrink so effecting actuation by volume change. Or gels may be bent or deformed without undergoing significant volume change. Volume change is of greatest interest to the application and will be considered first.

Initial experimentation in the 1950s and 1960s by Khun, Kalchasky and others shows how polymer hydrogels can exhibit large reversible volume changes in response to PH variation of their environment. Brock and Lee [30] describe an actuator system developed about this principle. They use PAN hydrogel fibres (which are described by Shahinpoor [29]) as tendons in a cylindrical chamber. The chamber was irrigated by two sets of perforated teflon tubes, one carrying acid, the other, base (figure 2.23).

# Figure 2.23 Polymer gel actuator using acid and base irrigation



PAN fibre irrigation

The gel fibres were connected to an external spectra tendon which in turn was mechanically loaded. This experiment demonstrated how the system could perform work by shrinking the gel fibres on addition of acid and swelling them on addition of base. However, the maximum tension generated by the gel actuator was only 0.35N. Unbalanced irrigation across the fibre mass was observed to be a problem.

Shahinpoor [29] describes similar systems but demonstrates robust mechanical design and a strength performance equivalent to human muscle of 20kg/cm<sup>2</sup>. Descriptions of the practical preparation and assembly of these systems are given.

The advantages of these new generation muscle like actuators are:

- muscle-like physical properties such as tensile strength, good power per unit volume, flexibility and elasticity in non-actuating planes, and good life-cycle.
- there is a reasonably established preparation and manufacture process,
- the necessary hydrogel fibres (Polyacrylinitrile) are commercially available.

The disadvantages are:

- complex manufacture process requiring experience to perfect,
- corrosive acid and base fluids required during operation,
- even irrigation across the active fibres is difficult to achieve, this results in an unbalanced response and internal stressing,
- response delays vary according to the efficiency of the irrigation system,
- supporting components such as acid and base containers are bulky and preclude mobile applications.
- the acid/base solution is not easy to reuse.

Similar performance to irrigation of PAN fibres with acid and base has been achieved through application of an electric field. In this system the PAN fibres are irrigated in a sealed chamber of an electrolyte solution. According to Shahinpoor [29] a field of just 'a few volts' per centimetre is required to control the pH of the electrolyte. This method of controlling the environmental pH of the PAN fibres affects their volume just as described above. There are several advantages to this system:

- there is no need for a complex irrigation system and the associated control systems,
- requires electrical rather than chemical power source,
- response is better proportioned to stimulus (no unbalanced irrigation problems, no fluid flow propagation delays etc.).

### Gel bending:

Cantilever actuators have been manufactured from polymer gel strips [31-34]. These devices have been developed on a millimetre scale with a view to application in keyhole surgery for example. Experimental systems include actuators dependent on an electrolytic environment for operation, and integrated actuators able to operate in air or free space. The power source is electricity and the response is indirectly proportional to an analogue input. The main disadvantage of electrical drive is that a pd of several hundred volts per centimetre is required across the gel, which has practical safety implications.





# 2.6.5.2 Conventional Chemistry actuation: The process of many chemical reactions is one of:

 $X + Y \rightarrow a$ 

The conversion of the agents to the product will involve a change in entropy. Combustion of hydrogen is a simple and explosive example where much heat is generated as water is produced:

 $2H_{2(g)} + O_{2(g)} \leftrightarrow 2H_2O_{(l)}$ 

Some chemical reactions are more complicated and produce more than one product, or may require more than two chemicals to initiate. The important effect on matter involved, with regard to actuation, is one of phase change. Most chemicals can exist in either of three physical phases: solid, liquid, or gas. The volume held by a lump of matter increases as it passes from solid through to gas. Useful actuation can be done from any reaction which results in a phase change from liquid to gas. For example:

 $Mg_{(s)} + 2HCl_{(aq)} \rightarrow H_{2(g)} + MgCl_{2(aq)}$ 

It is common for compounds to exist in solution with a solvent as indicated in the equation above, where hydrogen chloride gas is dissolved in water [35]. This is often necessary to allow reactions to initiate. The liquid phase is also of mechanical importance as an actuator power source because it is easier to move than a solid. In the example above solid magnesium dissolves in the hydrochloric acid solution to form hydrogen gas and solid magnesium chloride. The magnesium chloride will readily dissolve in the solution as would table salt in water, hence resulting in just gas and liquid. The problem with this example is that it is difficult to inject a solid to the reaction chamber. A better example combines two solutions to give a gas and one solution:

 $Na_2CO_{3(aq)} + HCl_{(aq)} \rightarrow CO_{2(g)} + NaCl_{(aq)} + H_2O_{(1)}$ 

This reaction combines aqueous sodium bicarbonate and hydrochloric acid to evolve carbon dioxide, brine, and water. The reaction products are relatively harmless except that there is popular objection to unnecessary release of carbon dioxide to the atmosphere. Unfortunately it is not feasible to reverse this reaction so the products are disposable. The internal combustion engine is an example of chemical reaction based actuation where oxygen and hydrocarbon combine to evolve much gas. However, in this instance a spark is required to initiate the reaction. This type of reaction is limited in application (except where considerable evolution and development have refined the process) because of the requirement to supply 'fuel' and to remove 'waste', not to mention the considerable side effect of heat generation.

A rather more preferable chemical reaction would be one involving a system of chemicals which has two stable phases, one liquid and one gas. A transition between states could be effected on request through the introduction of a suitable stimulus (e.g., raising of energy or lowering of energy). There are examples of such 'closed' and reversible systems emerging in the field of actuation [36]. These special chemical reactions are useful because they can be initiated and controlled quite accurately. They can also be theoretically reversed. The reaction process requires electrolysis of ionic solutions to generate gas. In principle the electrolysing potential may be reversed to re-adsorb the evolved gas into the original chemical form. The most basic electrolysis cell format is:

Figure 2.25 Basic electrolysis cell



Pure water is the most basic of electrolytes. Water readily dissociates into its ionic components:

 $H_2O_{(1)} \leftrightarrow H^+_{(aq)} + OH^-_{(aq)}$ 

At any positive cell potential the cathode will generate hydrogen gas as electrons combine with the hydrogen ion. The production of oxygen gas at the anode requires specific conditions: the cell potential must be above that required to free an electron from the peroxide ion (i.e., at least 1.23v [37]). Further to this the potential must not be so high that the oxygen generated has sufficient energy to

oxidise the anode material; the lower the reactivity of the electrode material the higher the cell potential may be. Selection of the right electrode material minimises this problem, platinum and gold are suitable electrode materials for this purpose [36].

The current through the electrolyte is proportional to the gas evolution rate. High currents generate more gas. Increasing the voltage will increase the current flowing according to Ohms law and therefore generate gas more quickly. The problem is that water is not a good conductor of electricity and actuation resulting from gas evolution would be slow.

A more appropriate and commonly used [36,38,39] electrolyte solution is copper sulphate. This may be electrolysed to generate oxygen at the anode. Gas production is substantially greater for a given cell potential over a water electrolysis cell. This is because copper sulphate is a good electrical conductor and supports a high current flow for a given voltage relative to water. The basic process is:

at the cathode:  $2Cu^{2+}_{(aq)} + 4e^{-} \leftrightarrow 2Cu_{(s)}$ the copper ion reduces to copper and combines with the copper electrode,

at the anode:  $2H_2O_{(1)} \leftrightarrow O_{2(g)} + 4H^+_{(aq)} + 4e^-$ 

water gives up electrons to form oxygen gas and hydrogen ions. The cell potential, according to Neagu [36], must be maintained above 0.89v to allow oxygen evolution. If the cell potential is allowed to exceed 1.23v however, hydrogen evolution may occur at the cathode which would restrict copper depositing on the electrode. This would reduce the cell life [36]. The type of material used as the cathode also affects the conditions under which hydrogen evolution may occur, copper being a better inhibitor than platinum [36].

Oxygen is slightly soluble in water. However, saturation is quickly achieved so this property of the system does not seriously affect the efficiency of the actuator. However if the dissolved oxygen is allowed to reach the cathode it will take up the free electrons over the copper ions arriving from the copper sulphate solution. If this reaction is allowed to proceed it will affect actuator efficiency. To restrict the oxygen molecule's access to the cathode a special barrier may be used which allows passage of positive particles but not negative or neutral particles. Negau demonstrates [36] how a NAFION [40] coating may be applied to the cathode to prevent contamination by dissolved oxygen. The coating allows the copper ions to combine with the arriving electrons whilst the oxygen molecules are held back. Since the oxygen remains in solution, saturation occurs and gas evolution proceeds.

This reaction can be reversed. The cell polarity is reversed so that the platinum anode becomes the cathode. The new cathode absorbs oxygen to form hydrogen peroxide ions and water:

 $O_{2(g)} + 2H^{+}_{(aq)} + 4e^{-} \leftrightarrow 2OH^{-}_{(aq)}$ 

$$H^+_{(aq)} + OH^-_{(aq)} \longleftrightarrow H_2O_{(1)}$$

The copper electrode becomes the anode and copper oxidises to form copper ions:

 $Cu \leftrightarrow Cu^{2+} + 2e^{-}$ 

This reaction is 100% reversible in theory. Factors which may affect its cycle span are:

- Isolation the cell must be chemically sealed to ensure no contamination from external substances, and to prevent leakage of internal substances (particularly oxygen gas).
- Pollutants whilst it is possible to seal the cell it is not possible to ensure that the electrolyte and the electrodes or the cell walls are entirely clean (during manufacture) from compounds which may promote parasitic side reactions.
- The integrity of the NAFION coating on the copper electrode is known to be not entirely non-porous to negative and neutral particles [36]. In time copper particles will pollute the electrolyte as hydrogen evolution prevents their adherence to the copper electrode.
- The NAFION coating is acknowledged as a problem to the static performance and life cycle of the chemical actuator [36].

Kempe and Schaper [39] describe an alternative and earlier electrolysis actuator. Their system is based on aqueous potassium peroxide as an electrolyte and silver and platinum electrodes. The process is designed to evolve hydrogen gas in one direction and adsorb it if the polarity is reversed. At the silver electrode:

 $Ag_{(s)} + 2OH_{(aq)} \leftrightarrow AgO_{(s)} + H_2O_{(l)} + 2e^{-1}$ 

and at the platinum electrode:

 $2H^{+}_{(aq)} + 2e^{-} \leftrightarrow H_{2(g)}$ 

The researchers describe load capability of between 1000N and 2000N. However, for both systems detailed there is an issue of time scale. Negau [36] demonstrates how the copper sulphate micro-actuator was developed to deflect a miniature membrane. The paper indicates that good inflation, or the 'out-stroke', may be achieved within about 2 seconds of applying current. The 'return-stroke' may take as long as 7 minutes by applying a reversed current. Kempe [38] reports 'out-stroke' and 'return-stroke' to be between 10 seconds and 60 seconds each, and points out that this time may be tuned by cell topography.

Both these researchers have been developing valve actuators with good static performance and linear controllability. In neither application was time scale considered to be particularly critical. It is perhaps possible that fast acting actuators with good linear characteristics may be developed from the electrolysis cell principle. However, current technology appears to be not appropriate for highly dynamic actuators.

## 2.6.5.3 Shape memory alloy.

Shape memory in alloys is an effect which has been known of for over 50 years. The effect results from a change in the arrangement of atomic arrays from a highly arranged and densely packed structure (known as the parent phase) to a less densely packed structure (known as the martensite phase). A transition between these phases is caused by application of heat after deformation (for a one-way alloy), or heating and cooling (for two-way alloys). A good description of the shape memory process is given by Friend [41].

To make use of the shape memory phenomenon in actuation consideration must be given to the martensite and the parent shape. Waram [42] and Tautzenberger [43] gives detailed discussion on how effective reciprocal actuation may be obtained from shape memory alloy formed in the desired manner. Conventional forms employed in the application of SMA are commonly those used for springs. In these cases the alloy may be programmed to exhibit 'contraction' on heating and 'extension' on cooling, or visa versa if desired. This is illustrated in the figure below.

# 2.6.5.4 Coupled mechanisms.

Using the appropriate coupling system almost any actuation profile may be achieved with any actuator, in theory. In implementation terms there are practical limitations, for example a micron displacement piezo actuator would not be used in place of a car's petrol engine. However, it is very conceivable that mechanical ties might be used in place of direct muscles in a way that the tie exhibited most of the muscle properties such as flexibility with tensile strength. A conventional actuator may then be employed outside the vicinity of the 'muscle' location. This is represented for the circumferential muscle group of the earthworm vehicle in figure 2.26.

Osada [44] demonstrates the use of coupling mechanisms in very relevant 'wormlike' manipulators. These manipulators make use of external winching actuators, cable ties and compression springs to effect displaced actuation in the place and manner desired.

# Figure 2.26 Example, use of coil and lead screw mechanism in actuation

Continuous coil with single actuator



# 2.6.6 Positional feedback for a mechatronic earthworm segment

The previous section demonstrated how the synthetic earthworm would in certain circumstances require positive actuation and in other circumstances need to

submit to the resistance of obstructions. To understand at any time the nature of external loading and loading caused by actuation there is need for a system feedback. This feedback would allow a profile of the worm shape to be formed. Taking one worm segment in to consideration in isolation it is possible to apply positional feedback in either of two ways.





# 2.6.6.1Worm shape by direct actuator state feedback

Measuring the direct actuation of an actuator is an established practice and there are methods applicable to many of the different types of actuator. Direct feedback systems usually rely on some form of mechanical linkage from the actuator output to the feedback transducer. There are also those whose state may be probed internally, such as electric motors, whose armature voltage is proportional to velocity.

If this principle were adopted for the synthetic earthworm then every actuator within each segment would require its own feedback and control loop. A system whose locomotion actuators are many may find difficulty in powering and controlling such a network.

### 2.6.6.2 Indirect feedback

In a complex mechanism where there are many actuators contributing to the system status, specific information about individual actuators might be less useful than data provided from measuring particular *components* of the structure. The worm segment structure is an ideal example where this could be the case: A segment based on a grid array of actuating elements may total more than 100 individual components. This combined with position and orientation of each element being dependent on its neighbouring elements suggests that element specific information is fairly meaningless without simultaneous consideration of position within the hierarchy. This is in contrast to placing sensors within the material which is being actuated (i.e., within the earthworm skin), a sensor used in this manner would provide information about the resultant of many or all actuating elements.

## 2.6.7 The Sensor Mesh

The mechatronic worm provides an interesting challenge for positional status feedback. First it is necessary to identify the fields which are to be measured, and then to establish factors such as: measured relative to what, accuracy required, frequency of sampling etc. Considering the skin of one segment it would be useful to know:

- the circumference of the segment at several points along its length,
- and the length of the segment at several points around the circumference.

If this data were available in some useful metric such as millimetres it should be possible to map the segment topography in 3 dimensions with an accuracy proportional to the density of sensors. Figure 2.28 shows a conceptual arrangement of strain sensors which may be incorporated within the vehicle skin.



Each element measures longitudinal or circumferential strain in its location

# 2.6.7.1 High strain transducers

The nature of the worm implies that actuation leads to stretching and bending of the segment surface. Suitable sensors are therefore those based on strain measurement. A design problem lies in the high static strain range in which the sensor must operate. Conventional metal leaf strain gauges operate at up to 4% strain range. It is likely that the surface skin of the worm segment might be required to strain to a maximum of 100% greater than original length and 50% less than original length at the other extreme. Brief investigation suggested there to be no suitable strain measuring devices off the shelf to handle these requirements. However it was supposed that there are several possible systems which could be investigated to develop a strain gauge to operate within the required extremes. These might include continuous surface contact systems which deform with the surface to be measured, or they might be devices with two points of contact and a system to measure the distance between those points of contact.

## 2.6.7.1.1 Fluid filled tubes

Fluid filled tubes are used as electrical resistors in a bridge circuit. The fluid in the tube is a reasonably high resistance conductor such as saline solution. If the tube is stretched the sectional area of the fluid in the tube decreases, this increases the resistance. As the tube relaxes it reforms to its original dimensions and the sectional area increases. This type of device may be continuously bonded to the object surface.

## 2.6.7.1.2 Conductive rubber

Conductive rubber is widely used in products having rubberised key pads such as miniature computers and hand held electronic games. It is possible to stretch this rubber typically to 150% extension. Straining of conductive rubber is known to induce a resistance variation, although this is not normally proportional to elongation.

2.6.7.1.3 Adaptation of line displacement sensor.

An ideal solution to the problem for a fixed plane 2 dimensional surface would be displacement measurement between points on the surface. However, the surface of a robotic worm segment would be cylindrical and not a flat plane. A displacement sensor developed to measure the surface distance between points on a non-flat surface, which may vary its undulations, would be required to deform and conform to its parent object whilst maintaining the ability to measure. An interesting challenge.

# 2.7 Estimation of some operational requirements

It is understood that the locomotion requirements are known in a qualitative way. On a hierarchical scale: the entire worm vehicle must be able to:

- both resist external forces and give to them at the discretion of the controlling system,
- exert forces required to effect transportation through the void created by the cutting head,
- and exert forces required to steer the cutting head along the desired path.
- Each segment of the vehicle must be able to:
- both resist external forces and give to them at the discretion of the controlling system,
- and assume the articulation and dilation profile requested by the controlling system regardless of the resistance experienced.
- Each actuator element of a segment must be able to:

- both resist external forces along the active axis and give to them at the discretion of the controlling system,
- and comply with external loading not along the active axis by effecting minimal resistance.

The keywords below, as derived from the statements above require eventual quantification:

- Dilation the limiting parameters for segment diameter (e.g., 80mm<Ø<150mm).</li>
- Length the limiting parameters for segment length (e.g., 100mm<L<200mm).
- Articulation the limiting parameters for segment bending (e.g.,-45°<θ<+45°).</li>
- Turning radius the constant defining the tightest radius through which the vehicle may turn.
- Total length the limiting parameters for entire vehicle length (e.g., lm<L<2m).</li>
- Volume the constant defining the vehicular volume.

The variables and constants detailed below give important performance specifications, although they are ambiguous qualities derived from overall vehicle performance. They are also very dependent on ground properties:

With respect to the cutting head and tool chuck:

- Frequency response the excitation response of these parts to a range of mechanical vibrations caused by machining and operations.
- Thrust the maximum load (either static or dynamic) applicable before vehicle slippage occurs.
- Torque the maximum torsion load applicable before vehicle slippage occurs.
- Stiffness a resistance to deflection due to oblique loading (both static and dynamic).

With respect to the whole vehicle:

- Ground pressure the pressure exerted by a ground type exhibiting fluid properties (e.g., wet clay flows as a viscous fluid in contrast to dry clay which is structurally rigid).
- Ground friction the mechanical resistance of a ground type to the vehicle skin sliding against it - to be defined over a range of contact pressures.
- Range the distance, affected by a broad range of factors, that the vehicle may travel underground. (Factors depend on tunneling mechanism soil displacement: self-sufficiency, communications capability, ground conditions, equipment carried, operation to be performed. soil extraction: tunnel lining system, length of umbilical).
- Depth the depth at which the vehicle may safely operate affected by factors such as those listed for range.

The actuating elements of both the worm based topography and the 2D scaleable flexible sheet system combined with the rubber material they are to be embedded within, would be likely to influence many of the performance criteria listed above. Because of this the actuation technology to be employed in either system is critical.

It is possible to effect actuation in either of two ways as discussed: direct actuation, or indirect actuation through linkages (for example cable). The systems outlined above show examples of both - a spiralled cable is a link to a linear or rotary actuator, individual actuation elements in the 2 dimensional grid perform direct actuation. There are advantages and disadvantages for each in this application:

A linked system is able to perform complex actuation as in the example of controlling segment diameter with a spiralled cable. The power conversion device is located away from the area of direct actuation. This allows for miniaturisation of moving parts but increases overall volume.

A direct actuation system is usually limited in controllability to simple movements. However, because the power conversion device is part of the actuator, volume requirement is reduced and the system has greater integration.

## 2.8 Conclusion

This chapter has presented the concept of a robotic earthworm which may be employed as the vehicle on which to base a key-hole underground access/repair product, for use by various utility industries. The concept has been tailored to address the criteria identified in chapter 1 as being disadvantages of the current methods used in industry.

This chapter has broadly identified the product and application setting for a major design project. The design project to develop the conceptual robotic worm would include many facets suited to academic and industrial research; these facets have been identified during the course of the chapter. To allow detailed and technical development work it was decided that a focus would be made on the locomotion mechanism involved in this device. The second part of this chapter (from 2.6 onwards) identified the two key components for research involved in locomotion, that of actuation, and that of positional/movment feedback. At this stage a decision was taken to focus specifically on the issue of feedback because the nature of the problem was well suited to the author's interests and abilities. Chapter 3 begins the phase of investigation into relevant work on sensing technology. An existing transducer principle is described and adapted from fundamental physical principles to provide a mechanism suited to measuring in the required manner. Suitable sensor circuitry is discussed, and a prototype is designed and manufactured for testing.
# 3 Sensor Technology and Development of a Novel Surface Measurement System

#### 3.1 Introduction

Chapter 3 builds on the context of the opening chapters to provide a technical background for the proposal of novel sensing instruments and systems. The chapter examines parameters of measurement and attempts to unify associated terminology. The fundamental physical principles of various categories of measurement instrument are identified, this is mostly specific to the perceived solution although there is a general basing. The limitations and potential of existing sensor systems are discussed where appropriate. A proposal is made that a particular category of sensing devices is more suited in this instance to the development of a feasible instrument. The selected category is examined from first principles and a theoretical design is made. The chapter concludes with detail on how a prototype was constructed.

## 3.2 Measurement and Instrumentation fundamentals

This section provides background academic and industrial information on the principles of measurement and a discussion on terms and terminology related to measurement systems.

## 3.2.1 Sensor definitions

Brindley [45] suggests that technically a sensor is a passive transducer, and that a transducer is a device able to convert one physical quantity (measurand) into another physical quantity (often electrical) which is a signal. A passive transducer is a proper classic transducer (dictionary defined) in that the output signal is derived directly from the input measurand. An active transducer requires external power to provide the response of output to input.

Morris [46] describes measuring systems as instruments. He suggests that within each measuring instrument lies the primary transducer, which has an output as a function of the measurand. In many instances the output of the primary transducer is not a form appropriate for use and subsequent stages of conditioning may be included. Morris makes no reference to the term sensor, however the same distinction between active and passive transducers is made (as elements of the measurement instrument) as made by Brindley.

Boyle [47] hints of a subtlety in meaning between terms, but the first sentence of his text chooses to ignore the differences and allow interchangeability of sensors, measuring instruments, and transducers. Smith [48] says, 'A transducer is a device that converts one form of energy into another'. The Oxford English Dictionary would concur. He clearly suggests that a subset of transducers are those with an electrical output and a purpose to drive subsequent electrical or electronic circuitry. These transducers are, 'often described as electrical sensors'. Smith expands on Morris' encapsulating modular term of *measuring instrument* by stating that it is often necessary to include intermediate transducing steps when translating the measurand to the final useful output.

Norton [49] makes the classical definition of transducer and shows how it is somewhat inappropriate to take this meaning literally in common use, because, as he demonstrates a huge number and variety of artefacts can be defined as transducers: 'a piston or crankshaft in an automobile engine, a valve in a steam pipe, a typewriter, a violin, and a frying pan'. Norton goes on to make a special case definition of transducers used for measurement, 'A transducer is a device which provides a useful output in response to a specific measurand. (The measurand is the physical quantity, property, or condition which is measured).' He finally states that in this special class of transducer it is appropriate to describe the device as a sensor.

Ruocco [50] makes a clear and useful distinction between the terms transducer and sensor in the context of robotics. He states that a transducer is an elementary device with the capability of converting a physical non-electrical input quantity into an electrical output quantity (In this definition there is a deviation from the classical meaning of the word). The transducer is as simple as this. A sensor on the other-hand is not an elementary device, it is usually based upon a transducer as described but in addition it includes the capability of processing the transducer output signal in accordance with a given algorithm to provide an output suitable for interfacing to a process control system. Remaining in a robotic context, Stadler [51] apparently demonstrates questionable distinction in his use of terms, because a chapter is entitled 'Sensors and Instrumentation', and the text within describes whole systems as measuring instruments, and generally 'the primary sensing device such as a transducer of some sort' is included within it. An interchangeability has been made between sensors and transducers, but a transducer is clearly presented as an element of the whole instrument, never as a term to describe the instrument itself. The contradiction comes in the title of the chapter and subsequent sections of the chapter where headings refer to whole measuring systems as sensors therefore suggesting the term is interchangeable with the term measuring instrument. The intention is not to criticise a very useful academic text on the grounds of terminology, the intention is to demonstrate how there is a reflection on the unclear academic and industrial ethos for terminology in the field of measurement.

From this brief inspection of academic and engineering texts it can be seen that discrepancies or maybe misunderstandings exist in terminology. It can be argued by the likes of Boyle who's handbook errs on the practical industrial bent, that the argument is of no relevance because a device is specified to 'do' what it 'does' and as long as it does what the specification says it should then why is there a problem. Even Tse's [52] text which is of a very theoretical and analytical nature chooses to 'conform to common usage' such that the terms: sensors, transducers, instruments, and systems are used interchangeably.

It is the authors belief of these terms and others to be discussed (such as accuracy and resolution) that terminology needs to be consistent and needs to be clarified so that understanding of specifications is universal and subsequent misinterpretation of information does not occur. Furthermore, as the technical element of this thesis is to focus upon measurement it is necessary to present a standard for terms as they will be used throughout the text.

## 3.2.2 Definition of measurement terms as used in this text

A hierarchical relationship between mentioned 'device' terms has been established for use throughout the technical section of this thesis:

## 3.2.2.1 Transducer

Transducer is considered in its literal sense to be a device which converts energy from one form to another. This being the case, a transducer is always involved at the elemental level of a measurement system. It also follows that a literal transducer is a passive or self-generating device, because it uses the energy of the measurand to directly generate the signal output.

#### Figure 3.1 The transducer





## 3.2.2.2 Primary Transducer

A primary transducer is the sensing device which is either directly or indirectly linked to the input measurand. Any other transducers in the system would make use of the output from the primary transducer or other subsequent stages.

## 3.2.2.3 Active Transducer

The term 'active transducer' refers to a system or sub-system which includes a literal transducer (or maybe several), but the primary transducer is not directly linked to the measurand. The input to the primary transducer may be a supply voltage and the output from the primary may be an electrical signal which differs according to a variable state property of the device. This variable state property may then be linked directly to the input measurand. The Measurand is said to modulate the transducer input.

When this system is used it is not necessary for the input measurand to continuously (or ever) supply energy for the primary transducer to operate. It is however, necessary for the active transducer to be supplied with a continuous power at its energy input.

### Figure 3.2 The Active transducer



# 3.2.2.4 Sensor

A sensor includes a primary transducer and subsequent capability to condition the output signal such that it is calibrated in some manner for transmission or use. A passive transducer would rarely qualify as a sensor in its own right, an active transducer would in some instances qualify as a sensor in its own right.

# 3.2.2.5 Instrument

An instrument used for measurement (and so including a sensing device) is dependent upon a sensor and will include a stage to indicate or record measurements.

# 3.2.2.6 System

A measuring system has connotations to generic capacity and as such may be expected to include a degree of modularization and adaptability.





It may also be expected to assert control over sampling parameters. In all other respects the system is based upon the measuring instrument.

## 3.2.3 Quantification and qualification of measurement

It has been established that transducers whether active or passive, generate an output signal which is in some way related to the state of an input measurand. It is important to understand the nature of this output signal and indeed the nature of the input measurand to know what the relationship is between the two. Only when this is done can truthful and meaningful information be gleaned from a measurement instrument. The following is an explanation of the common terms classified to quantify the important factors affecting the relationship between input and output.

#### 3.2.3.1 Accuracy

Hansman [53] and Brindley [45] define accuracy as the difference between the true value of the measurand and the measured value as indicated by the instrument. It is often the case that accuracy is declared in terms of the error induced in the worst case such that a guaranteed accuracy is given or a maximum error is given along with an average error. Many factors affect the accuracy of an instrument and these are broadly categorised as either systematic or random. Systematic errors are 'features' of a measuring instrument, their effects, once noted, are predictable. Random errors cannot be definitely predicted although it is possible to quantise their effects through probabilistic methods. Measurable error sources of an instrument are described below.

## 3.2.3.2 Sensitivity

Sensitivity is generally defined as the ratio of output signal range to input measurand range. In generic terms an assumption is made that the relationship of input to output is linear, although it is often necessary to quote the average sensitivity or the range of sensitivity for a commercial device.

## 3.2.3.3 Linearity

Linearity is defined by the 'constantness' of gradient of the plot of output against input. Very few instruments have a truly constant gradient, and as such exhibit a non-linear response of output to input. As such, linearity is usually quoted with reference to deviation from some line of best fit, or over an optimal operating range of input measurand.

#### **Figure 3.4 Practical limitations of Linearity**



Input measurand

The effects of non-linearity if not compensated for, add to the overall systematic error in the instrument. Non-linearity need not be an issue in modern signal processing built into measurement instruments, because the effect is systematic and absolutely repeatable; therefore it can be profiled.

#### 3.2.3.4 Hysteresis

Hysteresis causes a difference in the output of a sensor when the direction of the input has been reversed. This phenomena may be caused by mechanical effects such as strain and backlash, physical effects such as magnetic memory, or by electrical components within circuitry for example operational amplifiers. Hysteresis is a real problem in terms of defining repeatability and hence overall accuracy. With reference to the chart in figure 3.5 it can be seen that any quantisable value enclosed within the response curves is possible, and that for any given input measurand a range of output signal is possible purely dependent on the previous and instantaneous physical state of the input measurand effector.

#### Figure 3.5 The hysteresis phenomenon



Input measurand

The effects of hysteresis are systematic in that it is possible to map the effect. However, the effects of hysteresis are based not only on the instantaneous state of the system, but also on how the input measurand has been changing in recent history, because of this a very complicated algorithm or a sizeable database would be required to describe the interpretation. It is preferable to design a system such that the hysteresis phenomenon is negligible in its contribution to overall error statistics.

## 3.2.3.5 Resolution

Resolution is the ability to distinguish between two different input values. The closer together the input values the greater or higher the resolution. Many analogue devices exhibit a continuous scale between maximum and minimum for both input and output. In these instances it is often found that a reference to infinite resolution is made, which in theory is potentially attainable. However, resolution is ultimately affected by other factors such as noise. Resolution is not a measure of absolute accuracy (because of full-scale errors) although it is a good measure of incremental accuracy.

## 3.2.3.6 Drift

Drift is the movement of the output signal when the input measurand is held constant over a period of time. Drift may be both systematic and random. Drift becomes systematic when its effect may be proportioned to some secondary input source. In many cases transducers are susceptible to not only the input measurand but also to other external physical quantities such as temperature, or moisture. Systematic drift may be profiled and the effects removed from the output signal, random or non-profiled drift is a contributor to overall error.

Random drift (or noise) of an average frequency two or more times higher than the sampling rate of the instrument may generate output errors such that there is little correlation between successive readings, even if the measurand has not changed. The effects of this relatively high frequency drift may be filtered out at the expense of dynamic response.

#### Figure 3.6 Example of signal drift.



to remove



## 3.2.3.7 Dynamic response

because of noise

This is related to the delay between a change in input measurand and the corresponding change in output signal. It is an issue when a relatively instantaneous change in input does not accompany a simultaneous change in output. This response time generally leads to an upper limit at which the input measurand may vary and the corresponding output be guaranteed within specification.

#### 3.2.3.8 Cross sensitivity

This term is variously described as any external physical quantity which may lead to an unwanted influence on the transducer output signal. In classic terms it is a factor which comes about because of the transducer's sensitivity to more input types than just the selected input measurand (for example a displacement sensor may also be sensitive to variations in temperature and to variations in supply In practical terms it is more often used to relate specifically to energy). phenomena occurring in environments where there are many instruments and many systems generating signals which in some way interfere with the output signal of the sensor in question.

In addition to multi-device environments, a single sensor may be open to transverse sensitivity, where for example an axial accelerometer also registers on its output off-axis forces.

## 3.2.3.9 Operational range

This is quite simply the input measurand range of values for which the instrument, sensor, or transducer is designed to measure. Operational range is important because attempts to operate outside the limits may result in large error readings or even mechanical damage.

## 3.2.3.10 Repeatability

This term is used to refer to the ability of a measuring instrument to return the same output signal for a given input measurand over a number of cycles. This value may be quantified in various ways such as a probability that the reading will be within a range of position about the true position (or a probability that the reading is within tolerance), or that it is guaranteed 100% repeatable to within a given tolerance, or that the signal is true within a certain percentage of the full scale. Precision is another term used to describe repeatability. Reproducibility is a similar term although it refers to the same test conducted with the measurement instrument being entirely removed and repositioned between cycles.

This list of terms and factors affecting error in output signals is not exhaustive and probably does not cover the full diversity of terminology, especially those terms with similar meanings. However, an effort has been made to mention in a qualitative fashion the key factors which affect measuring instruments. The sources for information in this section are from standard instrumentation and measurement texts [45-53] and a good source for further reading on terminology would be Elgar [54], and on error analysis for measurement in general Turner [55].

## 3.2.4 The sensor groups by measurand

There are instruments to measure all manner of quantities, ranging from liquid levels in a container to iron density in asteroids. However, all these diverse and specialised instruments are based on the measurement of just a few physical phenomena. Ristic [56] and Usher [57] independently highlight six distinct groups of sensor by measurand:

- Thermal, e.g., temperature, heat, heat flow,
- Mechanical, e.g., force, velocity, position,
- Chemical, e.g., concentration, composition, reaction rate,
- Magnetic, e.g., field intensity, flux density, magnetization,
- Radiant, e.g., e.m.s. intensity, wavelength, polarization, phase,
- Electrical, e.g., voltage, current, charge.

For example, the problem of measuring liquid level in a container is a mechanical issue which typically may be measured through force due to gravity or position of a float. There are various ways in which iron content of bodies may be inferred, perhaps magnetisation, resistivity, reflected spectra, or transmitted spectra (for small pieces). Spatial measurands are typically linked to electrical, magnetic, or optical transducers.

#### 3.2.4.1 Sensor group for surface measurement

Development of a concept in chapter 2 described how the skin of the earthworm vehicle might be treated as a surface, and that measurements of surface form might provide information on the changing mechanical form of the whole body.

It is the author's belief that surface form measurement must be the quantification of position information of discrete and identifiable points in one or more dimensions relative to some known (absolute) point or relative to other discrete (incremental) points. In review of previously published academic material of a relevant nature surface measurement has been a subject of study in three key areas: geology/earth science, aircraft location, and machine tool manufacture. It will be seen that there is a degree of overlap from geological requirements to aviation requirements, and that surface form measurement in machining is very much different. On the one hand geological spatial measurement is large scale to global, and on the other hand spatial measurement in machining is macroscopic to microscopic. In the application presented by this research the measurement requirement is orders of magnitude below geological requirements and at the upper limits of measurement requirements in machining. However it was noted that the principles of measurement from these established fields could be of benefit in solution to the problem posed. Other than these fields which are described below, there are conventional sensors which may be arranged suitably to perform the task required, these will be discussed subsequently.

#### 3.2.4.1.1 Geological Surveying

Since the late 1950s airborne and subsequently orbiting radar-scatterometer systems have been employed to 'observe' physical properties of the earth's surface [58]. One of the main applications of this early technology was to assess parameters of cultivated vegetation such as geometry, density, pattern, and height. However, of more interest to this work is discussion of the altimeter application of radar. Munday [59] describes how radar altimeters were originally designed for air and space craft (see figure 3.7) to measure their height above ground and sea, particularly for aircraft flying in poor visibility.

Figure 3.7 The satellite radar altimeter.



centre of orbit

reference ellipsoid

The technique has since been applied to oceanography, polar region research, and geology. Munday reports that an altimeter flown in 1978 on the Seasat satellite was able to measure altitude to 10cm r.m.s. with an application to measure seasurface profile. The same instrument was reported by Jay Zwallag [60] to

measure surface elevation over Greenland and Antarctic ice sheets. More recent technology employs laser scanning, which will be described in the next section.

3.2.4.1.2 Aircraft altimetry

NASA describes various methods [62] of obtaining altitude information and highlights the principle categories of these systems. There are three categories which will become clearly distinctive:

Measurement of height above terrain,

Measurement of altitude (pressure based) above sea level,

Measurement of height above sea level.

In the context of aircraft altitude these categories were devised to apply to low altitude, through normal altitude, to high altitude flying respectively. The principle of measurement above terrain is different from the principle of measurement from sea level. 'Terrain' is not a constant and hence provides only distance information relative to the ground. Examples of techniques based on this principle are:

Radio and radar altimeters



The time taken for pulses to reflect, or phase shift between transmission and reflection of continuous waves.

Laser altimeter



The time taken for short pulses to reflect. Higher resolution and more accurate than radar

#### **Capacitance** Altimeter



Isolated parallel plates act as a capacitor between wing tips or from nose to tail. Earth acts as a third conductor which modulates capacitance. Good close to ground (<100m) Reference Watton [63]

Sonic Altimeter

This is as for radio/radar system but for low speeds and low altitude.

'Sea level' is a global reference for zero altitude measurement. Sea level is defined by barometric techniques, and as such requires regular calibration due to the effects of weather. Altitude therefore, is measured as a function of atmospheric pressure, in this case. Typical devices which are responsive to air pressure are the aneroid, in which a chamber deflects as a function of pressure, density in which the breakdown potential of air might be measured, hypsometry, in which the boiling point of a pure liquid is measured. These systems are typically intended to measure altitude above sea level in the range 3000ft to 60,000ft. Higher altitude devices may not use the principle of air pressure as this factor drops off exponentially with increased height. Therefore a number of techniques have been developed about gravity changes, magnetism, and cosmic ray intensity.

Measurement of aircraft altitude is a broad subject where some of the principles are of relevance to generic spatial distance sensing, although it has been demonstrated that there are specialised techniques not transferable to the application presented by the work in this thesis. As such it was deemed as being of little reward to expand further in this area.

#### 3.2.4.1.3 Surface Measurement in Manufacturing

Surface measurement is increasingly important to quality control in automated production, rapid reverse engineering, and to robot manipulation in automated processes. Hence a host of academic material covers the diversity of this subject. The topic can be divided into two categories: macroscopic/microscopic measurement of roughness and texture, which is of little relevance to this work, and object level shape and form, which *is* of interest to this work.

In this instance there is an assumption that object level is in the order of millimetres  $\pm$  two orders of magnitude. As such categorisation leads to the group of contacting devices and the group of non-contacting devices.

#### 3.2.4.1.4 Contact surface measurement devices

Renishaw Plc lays claim to the original kinematic probe [64] for a coordinate measuring machine (CMM). This device was first produced in 1973, and presently there exist stylus based probes boasting a repeatability of approximately 0.3 microns over a measuring cell range of  $1m^3$ . The principle of operation is that the probe is systematically offered up to the surface for measuring until such time as contact is made and an opposing load is detected in the probe. To avoid damage to the stylus or the surface there is a spring deflection. At the moment of contact detection the current position of the stylus is recorded in 3 dimensions.

There are limitations to the process in its basic form, which include: difficulty in measuring on steep slopes, inability to measure occluded forms, potential to mark the surface or damage the stylus, the process is slow. There are CMMs which are able to address the difficulty of profiling complex surfaces through an interchangeable magazine of different probes, but the process remains slow. The main applications would be reverse engineering and post-production quality assessment.

Ossana [65] confirms the suitability of the stylus probe to millimetre range surface measurement by demonstrating the digitisation of human body surfaces.

#### 3.2.4.1.5 Non-contact surface measurement devices

There is greater diversity to this sub-field of surface measurement than for contacting devices. Applications range from machine vision to surface inspection, and object recognition.

80

Ip [66] reports on two non-contact optical surface measurement systems. The first system (a laser displacement sensor) employs laser scanning, where a spot is projected onto the object surface and the deflection of the reflected spot is focused onto a CCD. Position information is obtained through triangulation (see Besl [67] for detail). Interpolation is used to approximate the surface between measurement points. The second system requires a grey-scale image of the object surface and derives form information from shading levels. This work remains experimental and further details may be found from Ip [68,69].

Stout [70] reviews a range of systems for microscopic and macroscopic measurement, which includes the stylus as previously mentioned. He also describes the Focus Detection System which may operate into the millimetre displacement range. This system focuses a low power laser spot onto the object surface. As the surface height varies (either the spot or the object surface is moved) the focusing lens is automatically adjusted to maintain focus, and hence the focus lens displacement is proportional to the surface height variance. Like all probes, the basic system struggles with accuracy on steep slopes.

Other non-contacting 3D profiling technology is based on optical methods. There are numerous methods both novel and established depending on the scale, precision, and complexity of measurement required. A summary of current techniques is given by Chen et al [71], this summary includes: time in flight [72], laser scanning [73], Moiré [74], Laser speckle pattern sectioning [75], Interferometry [76], Photogrammetry [77], Laser tracking system [78], and structured light [79].

This section has thus far described existing technology for surface measurement of an optical and mechanical nature. For the purposes of this research all these systems have a fatal common condition which is that they are external to the surface being measured. It is believed to be a unique occurrence of the need to measure a surface which is entirely surrounded by solid material, hence any form of external probing, whether mechanical, optical, or other is not possible. As such the remainder of this section considers existing technology which is surface mounted or embedded within the object it is to take measurements from.

81

# 3.3 Surface mounted/embedded systems for surface measurement

## 3.3.1 Principle of employment

The options available to measure surface profile from within the surface are limited. It is supposed that there are only two obvious techniques applicable in this instance. The first is a measurement of distance between nodes located on the surface. As the surface deforms the distance between nodes could be expected to change. The second is to measure strain within the surface. The author is not aware of experimental or commercial systems operating in this manner. Either way the practical approach could be common. The surface must be regarded as a series of discrete 2-D cells perhaps as shown in figure 3.8a. Suppose an external load displaces any node vertically down such that the cell becomes 3 dimensional as in figure 3.8b.

## Figure 3.8a,b a cell partitioned surface



In diagram 3.8b it is shown that a flat surface cell has taken on a 3 dimensional existence and is in fact a tetrahedron. Solution of this tetrahedron will provide information on the relative position of its 4 nodes in 3-space. Basic trigonometry and 3-D vector techniques may be used to analyse the geometry (Gasson [80]). Figure 3.9 expands the 3-space geometry for any tetrahedron. In figure 3.9 triangle PRO is taken to be in the reference plane and point O is taken to be the reference origin. The six edges of the form  $(\overrightarrow{OR}, \overrightarrow{OP}, \overrightarrow{RP}, \overrightarrow{OQ}, \overrightarrow{RQ}, \overrightarrow{PQ})$  are known by measurement hence all angles of the triangular faces may be determined. In this instance the cartesian map is oriented such that the X axis is in line with the

elevation of triangle PRO. First the elevations of PRO and PQR must be determined such that their intersection's with axis  $\overrightarrow{PR}$ , s and t respectively are known. This is simple trigonometry  $\overrightarrow{Os} = \overrightarrow{OR}$ .Sin $\beta$  and  $\overrightarrow{Qt} = \overrightarrow{RQ}$ .Cos $\chi$ . Knowing the length  $\overrightarrow{PR}$  the positions in X and Y of s and t may be deduced through pythagoras' theorem. Since triangle PRO is in the base plane, s and t have a zero Z value. Now point R can be determined as  $R(\overrightarrow{Os}, \overrightarrow{sR})$  and P as  $P(\overrightarrow{Os}, \overrightarrow{sP})$ .





To determine the cartesian position of point Q the following procedure may be used. A projection of line  $\overrightarrow{OQ}$  should be made onto the plane parallel with XZ inwhich imaginary triangle Qtv lies (refer to figure 3.9). This projection exactly matches line  $\overrightarrow{Qv}$  and may be calculated using pythagoras:  $\overrightarrow{Qv} = \sqrt{\overrightarrow{OQ}^2 - \overrightarrow{vO}^2}$ . Now triangle Qtv may be resolved such that angle  $\alpha$  is known (using the cosine rule to find the angles). The determination of angle  $\alpha$  allows imaginary triangle Qut to be resolved such that the Z axis value of Q is known and the formula for the X axis position of Q is known. The Y axis value of Q is simply the difference in position between points t and s. The solution to the generic tetrahedron when O is the origin is thus:

For point R:  $R_{(x)} = \overrightarrow{Os} = \overrightarrow{OR}.Sin\beta$   $R_{(y)} = \overrightarrow{sR} = \overrightarrow{OR}.Cos\beta$ For point P:  $P_{(x)} = \overrightarrow{Os} = \overrightarrow{OR}.Sin\beta$  $P_{(y)} = \overrightarrow{Ps} = \overrightarrow{PR} - \overrightarrow{sR}$ 

83

For point Q:  $Q_{(x)} = \overrightarrow{Os} + \overrightarrow{tQ}.\cos\alpha$ 

$$Q_{(x)} = \overrightarrow{os} + \overrightarrow{iQ.Cosu}$$
$$Q_{(y)} = \overrightarrow{st} = \overrightarrow{PR} - \overrightarrow{Os} - \overrightarrow{Pt}$$
$$Q_{(z)} = \overrightarrow{tQ.Sin\alpha}$$

To find  $\alpha$ :

$$\alpha = 180 - \cos^{-1}\left(\frac{2\overrightarrow{Os}^2 + \overrightarrow{Qt}^2 - \overrightarrow{OQ}^2}{2\overrightarrow{Os}.\overrightarrow{Qt}}\right)$$

This process applies to any form of tetrahedron whether totally symmetrical or not. Having formulated a resolution for a single cell the next step is to define a spatial relationship between adjoining cells (figure 3.10) such that surface profile could be determined. Application of this and other issues in surface/membrane meshing are discussed in briefly in chapter 6.

#### Figure 3.10 Network of cells to measure surface shape.



Shown with no surface deformation

#### 3.3.2 Application to specific technology

Having established a theoretical manner in which surface mounted displacement sensors could be employed, the next step is to consider suitable existing or conceptual devices, which could perform the sensing task. These as mentioned earlier fall into the category of positional sensing and strain sensing.

In the instance of positional sensing, measurement instruments are employed to quantify absolute distance between reference nodes within the surface, thus enabling inference of relative node location. There are a variety of devices suited to absolute distance measurement which will be described in any good industrial sensor handbook such as those by Brindley [45] and Boyle [47]. The current concern is with contacting devices and some of relevance are briefly described below.

## 3.3.2.1 Resistive displacement devices.

At its most simple a resistive displacement sensor may be constructed from a linear or rotary potential divider. In this instance the potential divider is an active transducer with a power supply input a signal voltage output and the mechanical movement of the wiper which modulates the ratio of the divider. These systems are dependent on good and consistent electrical contact between the wiper and the resistive element, and are affected by temperature.





Figure 3.11 demonstrates a linear resistance transducer (which is active because a mechanical linkage modulates the division of supplied voltage), and a simple buffered circuit, which will allow some stability under loading.

#### 3.3.2.2 Inductive displacement devices.

The principle of operation in an inductive sensing device lies in the measurement of self-inductance of a coil as the magnetic permeability of space around it varies. In a proximity sensor, this typically relies on a permeable body external to the device, and in a positional or displacement sensor this typically relies on a moving core within the principal coil (the core being mechanically linked to the object of displacement). The self inductance of a coil with a high permeability core (relative to air) is approximated in general [81-83] to:

$$L = \frac{\mu_0 \mu_r N^2 A}{l} \tag{1}$$

If however the core of the coil is air or there is surrounding material in close proximity of a high magnetic permeability then equation 1 will not give an accurate approximation. The effect on self inductance of an external body is complex and beyond definition by a linear equation. This is because the magnetic field generated by a solenoid is not confined to the volume of space inside the coil, although it is strongest at this point.

85

#### Figure 3.12 Magnetic field of a solenoid



Typical field lines for a solenoid are as illustrated in figure 3.12. The field strength at any point on any field line is a sum of all the parts of the conductor coil generating the field, which is not necessarily difficult to solve, but very time consuming.

Each point of the magnetic field generated can be thought to facilitate energy storage dependent on the magnetic permeability of the medium in that region (and as such energy stored is proportional to the flux density B). To maintain a qualitative approach it can be observed from figure 3.12 that regions of close packed field lines have greater intensity of magnetic field than regions of less or no field lines. The placing of high permeability material in intense field regions (e.g., within the coil) therefore generates intense flux density, whereas the same material placed in regions of less intense magnetic field generate a lower (but quantifiably contributable) flux density. As a result, the overall induction of a coil is affected by the magnetic permeability of surrounding material, but the measurable effect (e.g., contribution to noise and error) is only observed when the externally produced flux density is measurable relative to the internally produced (i.e., inside the coil) flux density. This is most likely to give rise to deviation from the solenoid approximation equation (1) when there is an air core and bulk material of high permeability in the proximity. The effect of surrounding material diminishes rapidly as the field strength is inverse square proportional to distance.

Analytical methods are available to resolve magnetic fields and flux density, Jiles [83] and others demonstrate the depth of the problem. It is not appropriate to investigate self inductance in greater detail at this point.

Figure 3.13 Inductive displacement devices



Figure 3.13 indicates the two main arrangements for displacement and proximity measurements. The accompanying ac bridge circuit makes use of a second inductor to assist with noise rejection (more detail from Wobschall [31]). Measurement of inductance requires changing current, hence all instrumentation circuitry for these devices is based on oscillating voltage. Some circuitry presents signals in an a.c. format where amplitude is proportional to measurand (e.g., as for circuit in figure 3.12), others present an a.c. signal where frequency is proportional to measurand, and still others present a d.c. signal proportional to measurand. A range of instrumentation circuitry is discussed by Wobschall [84], and Cirovic [85].

## 3.3.2.3 Reluctive based displacement devices.

There are many variations on the principle of operation which is mutual induction of current within one coil by excitation from another coil. The degree of reluctance within the magnetic circuit between the two coils is typically varied by a moveable iron core such that the mutual induction of the secondary coil is proportional to the displacement of the core.





A common and reliable device is the linear variable differential transformer (LVDT) which makes use of common mode rejection in opposing secondaries to reduce noise and to amplify the reluctance effect.

## 3.3.2.4 Capacitance based displacement devices.

The operation of capacitance displacement devices is based on either overlapping area of parallel plates, separation of parallel plates, or change of dielectric separator. Devices exist based on each of these principles. The relationship of factors in capacitance is:

$$C = \frac{\varepsilon A}{d}$$

where  $\varepsilon$  is the permittivity of separating material, A is the overlapping area, and d is the separating distance.

## 3.3.3 The suitability of contact sensing devices to flexible surface

#### measurements.

As discussed resistive devices have a weakness in the wiper contact with the resistive element. This moving contact is known to induce wear and to introduce an element of inconsistency in electrical contact. Self inductive devices are conventionally rigid devices with a rigid and brittle ferrite core which must be at least of the length of displacement to be measured. These factors each are incompatible with the requirement for a flexible device. Reluctance or mutual inductance devices require that primary and secondary coils are maintained in the same axis otherwise flux linkage is non-linear and unpredictable, also, off-axis coils loose any benefit of common mode rejection (as employed by the LVDT). Capacitance transducers are difficult to work with because capacitance values and ranges are small, and there is little tolerance in terms of mechanical factors such as the magnetostrictive displacement sensor as described by Boyle [47] cannot be physically operated other than in a straight line and they are too intricate to be rugged instruments.

It was possible to reject all existing contact displacement sensors for one or more reasons. However, the intention was to investigate the use of conventional fundamental principles in a new way, to determine the feasibility of adapting established technology to an unusual requirement. The remainder of this chapter works through the concept, theory, and production of a prototype sensor developed to address the specific requirement of measuring distance between points on a varying surface (i.e., conforming to both extension and articulation).

# 3.4 The development of a flexible inductive displacement sensor

Inductive principles were selected in preference to resistive and capacitive technology for three key reasons:

Prototype capacitive systems are particularly prone to deleterious effects of stray capacitance from rapidly made circuits on breadboards and PCBs.

Resistance systems require a wearing contact, which may exhibit inconsistent electrical properties within a flexible device.

A particular solution was envisaged based on the use of self-inductance of a coil.

## 3.4.1 Considerations of coil geometry

As can be seen from equation 1 earlier in this chapter, the inductance (approximate) of a coil (solenoid) is dependent on the magnetic permeability of the material which comprises its core. It is not easy to calculate the exact inductance as this depends on the flux distribution in and around the coil. Equation 1 assumes that the flux density is consistent inside the coil and negligible outside the coil. It is generally accepted that flux density varies over the cross-section and along the length inside the coil. This implies difficulty in analysis of the tuneable inductor, because not only does it not have a core of consistent magnetic permeability, it also requires the content of the core to change in its magnetic permeability during operation (i.e., the solid core slides in and out of the coil). However, it is seen that the magnitude of L from equation 1 can be built up in summed parts according to the component:

$$\Delta L \propto \Delta \frac{N}{l}$$

Now it is possible to treat a solenoid with two different concurrent core components as two or more smaller solenoids, and for the purposes of approximating the overall inductance L, the inductance may be calculated for each component solenoid and then summed (see figure 3.15). Figure 3.15 shows a coil of length 1. A high permeability core is exactly half way inserted within the coil, the other half of the coil encloses air.

89

#### Figure 3.15 Axial section through solenoid



To determine the approximate value of inductance L for this system two components should be calculated (from equation 1) separately and the result summed.

First some parameters must be defined:

Radial area of coil is A Permeability of air is  $\mu 0$ Relative permeability  $\mu R$  of solid core is 1000 Number of coil turns N is 1000

Length of coil is 1

Component 1. The half of coil enclosing the solid core

$$L_1 = \frac{\mu R \mu 0 \left(\frac{N}{2}\right)^2 A}{1/2} = \frac{500 \mu 0 N^2 A}{1}$$

Component 2. The half of coil enclosing air

$$L_2 = \frac{\mu 0 \left(\frac{N_2}{2}\right)^2 A}{\frac{1}{2}} = \frac{0.5 \mu 0 N^2 A}{1}$$

As can be seen there is a factor of 1000 times difference between the contributory elements of the core thus implying that the permeability of air relative to a ferrite core (for example) is negligible. The sum of parts may be used for a core in any axial position. Figure 3.15 shows a coil of finite length 1, and it happens that this length 1 approximately determines the operational sensing range for a displacement transducer of this type. This is because displacement is determined by measurement of the inductance of the coil as the cylindrical core is moved, the permeability of air being negligible as discussed, a more or less linear function is

derived between core displacement and inductance. The limitation of this system with reference to our target specification is that at the very least neither the core or the coil are flexible. A coil could not be made flexible because its geometric properties are contributory to its inductance. It was however, proposed that a very thin coil could have negligible effect on the mechanical properties of a flexible device, this is graphically represented in figure 3.16.

#### Figure 3.16 Overcoming coil rigidity



# 3.4.2 Consideration of core design

So a thin coil improves the flexibility of a device as a whole, there are remaining issues such as skewing of the coil but these may be considered subsequently if the effect is significant. However, until this time the sensing range has been a function of coil length and this may no-longer be the case because the coil is required to have no significant length. As such it becomes necessary to consider an alternative method of varying the magnetic permeability of the enclosed material. Figure 3.17 shows a 2-dimensional section of two solenoid/core arrangements. The first is of the form shown in figure 3.15. It is clear that the quantity (in 2D) of ferrite material enclosed in this case is a function of two variables, w and  $\Delta l$ , where w is constant for a cylindrical core, and  $\Delta l$  is by definition equal to or smaller than l the overall length of the coil.

#### Figure 3.17 Variables for core geometry



The range limitation of this form is apparent. The second arrangement indicates a tapered ferrite core whose quantity of enclosed material (in 2D) is a function of w, l, and  $\theta$ . By trigonometry the function is:

$$A = l \left( w - \frac{1}{\tan \theta} \right)$$

An assumption is made that in the case of the tapered ferrite core, the apex and the base are situated such that they lie beyond the confines of the coil enclosure. It is seen in this case that 1 is no-longer a constant with limiting parameters, but actually a variable within the equation. As suggested this variable is made constant by a coil of short but definite length. Now the angle of taper is held constant (i.e., a straight sided cone) allowing the single remaining variable w to define the area A. It can be envisaged that w would be a linear function of linear displacement should the tapered core be continuous with straight sides.

To determine the mass of ferrite material within the coil and hence an approximation of its inductance, consideration in 3 dimensions is required. The form of the core within the coil would be a truncated cone for which the volume is given by:

$$V = \frac{1}{3}\pi r_b^2 h - \frac{1}{3}\pi r_a^2 (h - l)$$

where:  $r_b$  is the larger radius of the cone

ra is the smaller radius of the cone

h is the height of the apex from  $r_b$ 

l is the height of the apex from  $r_a$ 

Having established the ferrite volume, an approximation is made by either of two methods:

Average the known ferrite permeability across the total volume (i.e., equate to a fully inserted cylindrical core of permeability less than that of the tapered ferrite core) or equate the ferrite volume from a conical form to a cylindrical form whereby the permeability is maintained such that the representation is of a partially inserted cylindrical core, the inductance for which may be determined by the method described previously.



Figure 3.18 Conical core equivalent to cylindrical core in two ways

In theory these methods imply a linear relationship between core displacement and variation of inductance. However, the non-uniform field of the solenoid would be expected to introduce significant digression as the core becomes nearer to total extraction (figure 3.19).

#### Figure 3.19 Nonlinearity of inductance against core insertion



#### 3.4.2.1 A flexible ferrite core

In order that it becomes possible to flex a ferrite core some consideration of materials is necessary. A simple solution to the problem was found in a suspension of ferrite powder set-up in the liquid phase of a cold-curing silicone rubber compound. Conventional ferrite core was crushed to a fine powder (typically less than 10microns) and mixed by mass with catalyst curing compound

(see figure 3.20 for core form). In order that the relationship between relative density of ferrite material and effect on inductance could be understood a range of densities were manufactured and the results are presented in chapter 5.

Figure 3.20 Sectional view through manufactured rubber core



Ferrite doped silicone rubber compound

# 3.4.3 Design and manufacture of inductor

The constraint of variables such as coil geometry and core geometry was necessary at this stage. The selection of values as presented in this subsection may be seen as somewhat arbitrary. However, it was necessary to ensure that the electrical function of the sensor circuitry was considered and that it would be possible to achieve a useful output in terms of level and range. Further to this it was necessary to consider the desired operational range of the displacement sensor and the effect that this constraint would have on the selection of values. Where a variable value was not of significant importance to the overall operation of the sensor system (assuming the appropriate magnitude had been selected) then a value was selected to assist in computation. For example the number of turns on the sensing coil was required to be in the order of 50 turns to provide an output frequency of the sensor in the 100KHzs range, such that 49 or 51 turns might have made little difference other than to complicate the model.

## 3.4.3.1 The Coil

The coil was formed on a perspex former as shown in figure 3.21. The number of turns was 50.





# 3.4.3.2 The Tapered core

The core was manufactured integrally with the housing. The dimensions of the core made for testing were as shown in figure 3.22.

#### Figure 3.22 tapered core geometry for prototype modulator



# 3.4.3.3 Modulator housing

As from the discussion at the start of the chapter the variable inductor is neither the measuring transducer nor is it a sensor in itself, it is a component of the primary transducer which tends to modulate its output. As such the variable inductor is referred to as a modulator. The housing was designed of the form which would measure displacement of two points, one relative to the other.

## Figure 3.23 Two parts of prototype modulator



Two parts were required, one to hold the coil in place and the other to hold the tapered core in place.

#### Figure 3.24 Manufacture of part to hold coil



Figure 3.25 Manufacture of part with tapered ferrite core



Each part was designed to integrate a compatible mounting system which would be used for attaching to experimental apparatus (see chapter 4). The whole

volume of each part was manufactured in moulded and cold cured silicone rubber to afford the properties of flexibility and elasticity. Assemblies of the mould tools used are shown in figures 3.23 - 3.25. Detailed design drawings and part specifications may be found in appendix A.

#### 3.4.3.4 Circuit Design

There were various options for circuit design. These were centred on the measurement of frequency or the measurement of phase change. To establish which would be more appropriate in this instance it was necessary to estimate the foreseeable frequency range which would be generated. In a typical LC oscillating system the frequency of oscillation is given by the form of:

$$f = \frac{1}{2\pi\sqrt{LC}}$$
(2)

Now if L were to vary by a factor of 100 between extremes (which might be reasonable for a doped silicone rubber core) then it may be seen from equation 2 that the frequency could be expected to vary by a factor of 10. This shows that the instrumentation circuitry should be designed to generate and measure frequency change rather than phase shift (very small frequency change would be measured as phase shift). A suitable frequency generating circuit was found in the Colpitts oscillator which is described in figure 3.26. The output of the Colpitts oscillator was conditioned in subsequent stages such that it was TTL compatible.

#### Figure 3.26 The Colpitts Oscillator



## 3.4.4 Anticipated performance of prototype sensor

The calculations presented below suggest the expected performance of the Colpitts oscillator and its prototype displacement modulator, in terms of frequency range. The calculation is based on the approximation equation for a solenoid (1), and therefore it is expected that subsequent experimentation may show marked deviation from the findings presented here.

Assuming the core is fully inserted within the coil, the geometry would appear as figure 3.27.





In this instance the base diameter of the cone is known to be 8mm. For a generic case where the base has been axially displaced some distance x the term  $r_b$  must be derived from the knowledge of x (or in the case of taking a measurement x is determined from knowledge of ferrite volume) so that:

$$r_b = 4 - x \tan \theta$$

where 4 is the base radius of the cone. Similarly the term  $r_a$  is derived from the same equation where x is increased by the length 1 of the coil.

To determine the volume of ferrite material within the confine of the coil the standard truncated cone formula is employed:

$$V = \frac{1}{3}\pi r_{b}^{2}h - \frac{1}{3}\pi r_{a}^{2}(h-1)$$

so for figure 3.26:

$$r_a = 4 - 1 \tan \theta = 4 - 4 \times \frac{4}{28} = 3\frac{3}{7}$$
  
V =  $\frac{1}{3}\pi$ .16.28 -  $\frac{1}{3}\pi$ . $\left(3\frac{3}{7}\right)^2$ .(28 - 4)≈174 mm<sup>3</sup>

Now the equivalent cylindrical volume of ferrite material is given by:

$$\mathbf{V}=\pi\mathbf{r}^{2}\Delta\mathbf{l}$$

where  $\Delta l$  is the height formed by a cylinder with equal volume to the conical section and radius equal to the base radius of 4mm. Therefore:

$$\Delta l = \frac{174}{16\pi} = 3.46 \text{mm}$$

Earlier in this chapter we accepted that  $\Delta l$  contributes proportionally to the total inductance L of the coil:

$$\Delta L \propto \Delta \frac{N^2}{1} \therefore \frac{50}{4} = \frac{N_e}{3.46} \text{ and } N_e = 43 \text{ turns}$$
$$L = \frac{\mu_r \mu_0 N^2 A}{1} = \frac{\mu_r \mu_0 \times 43^2 \pi \times 0.004^2}{0.00346} = \mu_r \mu_0 26.8 \text{H}$$

The value of  $\mu_0$  is known to be  $4\pi \times 10^{-7}$  Hm<sup>-1</sup> and one might estimate the relative permeability of the manufactured doped silicone rubber core to be in the order of 100 times (although this will be calculated for samples used in testing). Therefore it is possible to produce a quantitative estimate of the inductance of the coil when the tapered core is fully inserted:

 $L = 4\pi \times 10^{-7} \times 100 \times 26.8 = 3.37 \text{mH}$ 

and this inductance allows calculation of frequency generated by the Colpitts oscillator

$$f_{osc} = \frac{1}{2\pi\sqrt{LC}} = \frac{1}{2\pi\sqrt{3.37 \times 10^{-3} \times 3.2 \times 10^{-9}}} \approx 48.98 \text{KHz}$$

where the value of C is derived from fixed stable capacitors which are detailed in circuitry included with appendix A. We now have an idea of the base frequency of the oscillator, the next step is to determine by how much the frequency will rise as the tapered core is extracted, both in terms of overall range and also in terms of frequency change per unit displacement of the core. Figure 3.28 shows the configuration of the modulator at the point of maximum extraction (in the linear taper range).

The same procedures apply as for calculation of base frequency such that the inductance of the coil in this state is calculated to be  $2.7 \times 10^{-5}$  H, which is very small, so small that the permeability of air in the remainder of the coil might be significant. The inductance due to the air core is approximately  $3.9 \times 10^{-5}$  H so

that the total inductance of the coil is  $6.6 \times 10^{-5}$  H. The resulting frequency of oscillation in the colpitts circuit would be 346KHz. This implies a frequency swing of a factor 7 times. However L is root-inversely proportional to f so that one could not expect a linear function between core displacement and output frequency, without careful selection of coil/core geometry and associated accurate field calculation. In principle the calculations provided in this section suggest that the relationship between input displacement and output frequency for the sensor as a whole should be as indicated by figure 3.29, below.





Figure 3.29 Predicted response curve of frequency output to displacement input



Subsequent chapters will define the actual response characteristics observed during experimentation, and if necessary explore more in-depth theory of coil inductance to provide a better model.

## 3.4.4.1 Usefulness of the output signal

As is made clear by the preceding section, the intended output signal from the sensor is a frequency dependent upon core displacement. There is no reason to assume that a signal conversion could not be provided to perhaps generate a d.c. signal or any other form of signal as required. An ideal sensor has its output signal directly proportional to the amplitude of its measurand, when this is so then it is known that device sensitivity is uniform over the operational range. In the past non-linear response of a device was often unwanted and difficult to work with, and this remains the case today. However, modern techniques allow 'clever' manipulation to condition any systematic and repeatable signal into any form desired (i.e., a non-linear response may be massaged into a linear response). As a result of this it is not unreasonable to suggest that the response curve indicated by figure 3.29 may be a useful source. The following two chapters will provide some detail on how the output signal was processed.

#### 3.5 Conclusion

This chapter has taken an overview of existing surface form measurement and displacement measurement systems. It has been shown that the measurement problem presented by the tunnelling earthworm is unique and that no existing system can cater for the needs. As such an existing and well utilised physical principle (inductive sensor) has been adapted from its fundamentals into a theoretical device which may be used to modulate the output of an electronic oscillator. The theory has been applied to a mechanical design so as to produce a functioning prototype which in itself is suitable for testing. The next chapter discusses the test methodology and presents the implementation of suitable testing apparatus.

101
# 4 Experimental design

#### 4.1 Introduction

Chapter 4 identifies the test criteria relevant to the prototype sensor presented in chapter 3. A methodology is given as to how best to obtain useful information about the identified test criteria. It becomes clear that a dedicated test rig is required to perform consistent testing over long periods of time and for prototypes of variable configuration. The test rig as used is developed through chapter 4. Chapter 4 concludes by stating how the test rig should be used to obtain optimal data from the prototype sensor.

#### 4.2 Background to experimental design

Hicks [86] describes experimentation as the testing of a hypothesis such that it may be statistically validated or invalidated. In the situation presented by this work there has so far been little mention of hypothesis as such. The approach has been to loosely define a specific requirement with a view to an application, followed by a certain amount of technical development to provide a 'device' suitable for concept testing. In the development of a prototype certain constraints have been enforced and a simple model has allowed a theoretical prediction of some of the required performance criteria. However, on arrival at this chapter the author's vision for testing has been not so much a focus on success and failure within individual tests; rather, the anticipated aim of experimentation was to ascertain a measure of performance across a broad range of test criteria. From this series of observations it was hoped that refinements would present themselves, as indeed they did.

The difference between this observational assessment and a classic test of the hypothesis is that there is tolerance of undefined elements within the system. For example, the simple model of performance given at the end of chapter 3 would not be sufficient in itself as a basis for reliable comparison with measured results. This is because the model is based upon simplified and inherently inaccurate descriptions of the definition of self-inductance. An experimental observation is not to judge success or failure or even to necessarily compare expected results with actual results, it is to record the cause and effect for inputs and outputs. From

this record it is hoped that reliable relationships between input and output can be established in mathematical terms. It is also anticipated that results which exhibit a divergence from the desired property may indicate some variable in the system which has previously been overlooked or considered insignificant. In this way it was envisaged that there might be a degree of evolution as the initial prototype demonstrated its weakness for the second prototype to build on.

The following section details briefly the stages in evolution from the initial experimental work through to the presentation of a prototype suitable for thorough analysis.

#### 4.2.1 Predecessors to the prototype

It is necessary to discuss predecessors at this point because observations on their performance under limited tests were critical in establishing both a suitably refined product for testing and for the design of a suitable testing apparatus.

In the beginning there was a theorem that self-inductance of a coil is a variable dependent on several properties, including magnetic permeability of the space within and around it (see chapter 3). This theorem was accepted and established as the basis for the principle concept design of a novel transducer. The novelty of the transducer was about its mechanical topography, that is, the flexible elastic requirement of the device as a whole. This implied the requirement for a flexible elastic ferrite core, and a rigid coil of minimal spatial dimensions. Since it was envisaged that a tapered ferrite core would be required (see chapter 3) it was found that the most appropriate method of obtaining a part with such properties would be to manufacture it. The preliminary experimentation then was in establishing the suitability of selected materials and methods in manufacture. Design and testing was in no way extensive or exhaustive, the intention was to find and refine a suitable method to a degree that its magnetic and mechanical properties would allow satisfactory utilization in subsequent experiments. The process involved manufacture of controlled samples with differing levels of ferrite density. The objective was to ascertain the following: degree of flexibility/elasticity compared to undoped base material, confirmation that the

manufactured objects exhibit the desired magnetic property, and to establish a relationship between level of ferrite density and magnetic permeability. These tests were conducted on a small batch of manufactured parts under reasonably consistent conditions (i.e., temperature variation was minimal and measuring apparatus was consistent). The experiments are presented with their findings in chapter 5.

The second initial test was to establish the feasibility of the tapered core design. Modelling in chapter 3 indicated that the principle is sound, however, some comment was made about the limitations of the model and it was desirable to explore the practical effectiveness of the taper. A true circular taper form was difficult to produce because of the die machining requirement. In this initial phase it was deemed satisfactory to approximate a circular cone by a more elliptical conical form as shown in figure 4.1.

Figure 4.1 The elliptical tapered core Circular taper Elliptical taper





Ideal, difficult to manufacture Not ideal, made from two halves

The taper was cast from a negative die and the cured product was then offered to the sample coil which was as described in chapter 3. At this time the taper was not encased within undoped silicone rubber. The aim of this test was not exhaustive or quantified. Testing was to confirm or otherwise expectation by observation that the inductance of a coil would be altered by the axial movement of a tapered core within its confine. The results of this experiment are presented in chapter 5.

The next phase of development was in the encasement of the tapered core within undopped silicone rubber, such that the alignment of the core could be maintained. The challenge was in manufacture, the tests were concerned with mechanical integrity (i.e., could the two parts be satisfactorily bonded?), and in the retaining of magnetic properties. Again, the tests were not intended to be exhaustive and quantified, the design was to generate an observed impression of performance. The tests include tension, compression, and torsion across the joint, as well as offering the part up to the induction coil. The results of these experiments are presented in chapter 5.

Finally it was necessary to give consideration to mechanical linking between the sensor device and the surface or apparatus it was to be bonded to. The consideration in this instance was purely for compatibility with testing apparatus. Earlier testing had demonstrated a degree of significant sample movement and as such the design was modified to its final state (presented in chapter 3) to clamp the parts securely. At this point a prototype had been developed which was mechanically independent of the test apparatus, but fully compatible should it be necessary to attach it.

#### 4.3 Identification of test criteria

#### 4.3.1 What is the purpose of testing the prototype?

This is an important question to ask at the outset of any experimental design process. According to Hicks [86] and Kempthorne [87] it is most important that experimental design is based upon the following key criteria:

- Appropriate dependent variables must be selected for measurement,
- The measurement method should be of higher resolution and accuracy than the anticipated output resolution and accuracy of the variable,
- The setting of control variables (e.g., displacement positioning of tapered core by test rig) should be of higher resolution and greater accuracy than the anticipated tolerances in normal operation.
- Every effort should be made to eliminate the effect of significant secondary inputs.
- Randomization should be introduced to experimental procedure to average over time, the effects of unidentified inputs.

If these guidelines are followed there may be a reasonable assumption that the data collected are representative of the true relationship between controlling and

depending variables. They both give much discussion on the meaning and manipulation of sample data, some of which is relevant to the results collected in this thesis. However, it is considered that analysis of data is a task for inclusion with the discussion of results obtained, given in chapter 5.

The purpose of testing the prototype, specific to this research work was:

- To prove or otherwise that the prototype as implemented, in principle could be used as a displacement sensing device when coupled to suitable electronic driving circuitry.
- To demonstrate the potential for measuring distance between points along a curved surface.
- To ascertain a specific measure of performance and capability for the prototype in terms of popular categories such as accuracy, resolution, and range.
- To identify the cause of unexpected observations with a view to understanding the interaction of theory with application.

## 4.3.2 Relevant performance categories

An explanation of performance categories was presented in chapter 3. In respect of the prototype specifically discussed here the main categories of most initial interest were: range, resolution, sensitivity, drift, temperature coefficient, hysteresis, repeatability, and accuracy. Dynamic response, although an important property definition for any sensor has not been included here because it is a considerable field which might require entirely different experimental design, and it was thought that this would be supplementary research on completion of the more static work.

## 4.4 Experimental design – Apparatus

It was determined that all the static tests, and those without dynamic influence (low gradient measurand) should be conducted about one set of apparatus. The requirement of the apparatus was to provide an accurate and absolute primary input measurand, whilst maintaining or monitoring all other secondary inputs. The apparatus was required to be able to reliably hold the prototype in a truly static condition for long periods of time (several weeks at a time). In addition to these basic requirements there were also other higher-level procedures to determine properties such as hysteresis and repeatability, which made the necessity for some suitable drive and control for the input measurand. There was a need to record data in a flexible manner such that appropriate samples could be taken depending on the test being conducted. In its simplest terms the test apparatus was as indicated in figure 4.2 below.





With regard to the mechanical flexibility of the prototype, and the need to identify a credible test to measure the 'performance' of this factor, it was determined that a mechanical addition to the initial experimental apparatus would allow this function. The reason for design in this manner was to allow a profile of the prototype to be built up in direct comparison to conventional displacement sensors (i.e., those which measure in a straight line) in terms of classical properties. In respect to this, the mechanical flexibility of the prototype is of secondary importance, and it was deemed that apparatus to 'test' this property and its related variables would be complicated. An example of an add-on for these purposes is given in this chapter, and discussed in more detail in chapter 6.

The requirements of the testing procedure were such that a bespoke system was developed. The entire system incorporating mechanical, electrical, electronic, software, and control/data processing interface was designed manufactured and assembled by the author. The remainder of this section expands on the design and manufacture of the experimentation apparatus with a particular view to achieving the requirements presented earlier in this chapter.

#### 4.4.1 Measurand drive

The measurand drive is a mechanism, which offers both a means to securely hold the prototype, and a means to vary the input measurand in a quantified manner. Key features of this mechanism include: the fabrication material, the drive screw, the servo motor, the prototype's clamping options, and sub sensors and indicators. Various forms of the drive were considered based on the requirements specified, however there were limited options and it was not a difficult task to realise the most efficient solution available. The most significant constraint on design was funds for new equipment such that much of the design was based on 'found' materials and components.

## 4.4.1.1 Fabrication material

The selection of a suitable material was critical in the design of the measurand drive. It was known from experience and from chapter 3 that the inductance of a coil is affected by the magnetic permeability of surrounding material. As such, it was important to be able to state either that surrounding material had magnetic permeability approximately equal to that of air (and because of this had no significant effect on the properties of the coil), or else that the magnetic permeability of the surrounding material was known and that its effects were quantified and could be filtered from the logged data. The material was selected such that its magnetic permeability was insignificant.

A suitable material was selected in glass-filled nylon. This material offers the required properties for a structurally precise fabrication. It exhibits good tensile and compressive strength, good rigidity, and a low thermal expansion coefficient (2.3X10<sup>-5</sup>/°C). This material was also found to offer excellent machining properties such that it could be treated as a conventional machining material. The entirety of the measurand drive was constructed from machined parts in glass-filled nylon and nylon bolts were used to join the parts.

#### 4.4.1.2 Topographic design

The design requirement was to allow appropriate clamping of the prototype, and subsequent displacement of the core part within the coil part. The prototype part had been constructed such that its length including the clamps and the core was approximately 100mm, its sensing displacement range was approximately 30mm. However, it was envisaged that further prototypes of either larger or lesser design might be implemented as a result of subsequent findings, and in particular it was envisaged that there might be a requirement to mount a prototype upon an elastic surface on the drive bed. It was therefore determined that the topography should allow for this future work, and the maximum length bed possible (determined by raw material dimensions and machining limitations) was constructed. Figure 4.3 shows the measurand drive in operation. It was constructed with a single axis travel of approximately 300mm. A screw thread of 2.5mm pitch was turned on a nylon bar of 60mm diameter and this was geared to a stepper motor via a long HTD toothed belt. Technical drawings of this apparatus are given in appendix B. A Vernier scale was attached to the side of the drive as indicated in the figure, this was graduated to allow determination of absolute displacement by interpolation to 0.01mm.



This, of course is a manual measurement instrument and the intention was that it should be used for calibration and for correlation checks during experimentation. An LVDT was provided as an additional reference displacement sensing device.

The advantage of the LVDT was its analogue signal out, which was used in automated control modes to provide a correlation for each data point collected.

The LVDT featured variable range and sensitivity selection which could prove useful in the adaptability of the apparatus for different tests. A digital temperature sensor was also mounted upon the measurand drive in close proximity to the prototype's inductive coil. Figure 4.4 indicates the key components of the measurand drive. Various components such as the transmission and the temperature sensor are described as necessary in subsequent sections. Due to the construction material selected for this piece of equipment it was realised that there would be a limited operational life of bearing surfaces and in particular the leadscrew.





Great care was taken in manufacture to construct a tight thread so that there would be minimal backlash. It was anticipated that despite the limited lifetime of the equipment it should be of sufficient robustness to last through the main test phases of this research. With the addition of dry lubrication the measurand drive has remained in undiminished form for over three years.

# 4.4.2 Data logging

As has been indicated by discussion within the measurand drive section, there were multiple sources of data to be logged. These sources included: two temperature sensors, one reference LVDT signal, and the prototype sensor signal.

The intention was that for every prototype signal sample, each of the other sensors should be sampled at the same time. During tests where the input measurand was made to change the sample rate could be set relatively high (to 1KHz) such that the logging equipment required bandwidth sufficient to handle the maximum data flow. Conversely, during long-term drift tests lasting several weeks it was necessary to have the ability to sample at a lower sample rate, typically once every 10 seconds. In this instance bandwidth was not the issue, but reliability was.

The requirement was therefore one of regular sampling at defined time intervals from 1ms to 10s. Hence it was determined that a real-time system should be responsible for handling the time-based activity. The nature of the signals to be sampled were such that a system with multiple configuration capability would be of most benefit, and before the real-time controller and its interfacing is discussed further some comment will be made on the selection and the nature of the sensors involved.

#### 4.4.2.1 Prototype sensor

In chapter 3 it was indicated that the prototype sensor would be based upon a Colpitts oscillator. The output of the classic oscillator is a pure sinusoidal signal. Some manipulation of supply rails to the oscillator allowed oscillation about +2.5v with 0v and 5v being the peak voltages. This manipulation was performed such that interface with a TTL level control/logging circuit would be simple. Figure 4.5, demonstrates the primary sensor circuitry with power provision. Therefore, the prototype sensor was designed to output a 0-5v sinusoidal signal with frequency proportional to input measurand. It should be noted that a special high bandwidth, low voltage, low temperature coefficient operational amplifier was employed within the oscillator circuit. The amplifier came in a dual package so that the spare amplifier could be used as a false ground source buffer as shown (a stable reference was required at half supply voltage such that the oscillator amplifier could treat the single supply as a dual supply). Several different amplifiers were employed during testing, and their characteristics were shown to affect the results as reported in chapter 5. Figure 4.6 indicates the expected output signal in terms of the circuit shown in figure 4.5 and in terms of the modeling presented in chapter 3.



#### Figure 4.5 5v swing Colpitts circuit from 5v single rail supply False ground generation

Figure 4.6 Expected signal out from prototype sensor



# 4.4.2.2 Reference LVDT sensor

The LVDT was used as an 'off the shelf' device complete with its own amplifier. The amplifier was designed to offer a d.c. output signal proportional to input measurand in a selectable range between  $\pm 15v$ . As both the range and the offset of the amplifier were adjustable it was possible vary the sensitivity of the LVDT over a reduced displacement range. The LVDT maximum displacement range was 50mm.





0<Vrange<30 -10<Voffset<+10

## 4.4.2.3 Reference temperature sensors

It was established during early experimentation (see chapter 5) that there might be a considerable temperature effect on the performance of the prototype sensor. It was therefore determined that temperature sensors should be installed about the experimental apparatus to monitor the effect. Initially it was supposed that three points should be monitored, proximal to the coil, proximal to the Colpitts oscillator's amplifier, and a separate sensor for ambient environment measurement. It was decided that in consideration of convenience in terms of setup, sampling, and calibration, that packaged digital sensors should be employed on a distributed network. Two different devices were used, based on the I<sup>2</sup>C serial network. I<sup>2</sup>C is a half-duplex 2 wire open collector bus operating at 1Mbit/sec, and is intended for logic-level communications over short distances (typically 2m). As such the signal out from these temperature sensors was incorporated within the communications protocol of the I<sup>2</sup>C bus.

# Figure 4.8 Photograph of LM75 digital temperature sensor network cable



The two devices used were a 9bit output device from National Semiconductor (LM75), and a 13bit output device from Dallas Semiconductor (DS1624). Both

were packaged in 8 pin dil plastic cases. The general descriptions may be found in Appendix B, along with a general description of the  $I^2C$  protocol.

# 4.4.2.4 The real-time controller for data logging

This sub-system of the experimental apparatus is quite separate from the experiment controller, which is discussed in the next section. The purpose of the real-time controller was to ensure that sensor sampling and data recording occurred consistently and at the time interval requested. The only way this could be guaranteed was to dedicate a processor specifically to trigger acquisition phases from a timer interrupt. In a logical sense the operation of the real-time controller was as shown in figure 4.9.

It can be seen from figure 4.9 that acquired data is moved into a buffer. Another asynchronous function of the real-time controller was to communicate with the experiment controller to collect configuration information and also to return sampled data on request. The buffer was designed to store several samples in a FIFO arrangement incase the experiment controller was busy on other tasks during multiple sample times. The communications function of the real-time controller, and its other duties will be detailed in subsequent sections.





The hardware selected upon which to base the real-time controller was an Arizona Microchip Peripheral Interface Controller. The advantage of this manufacturer was that there is a large family of 'Microchip PIC' microcontrollers to choose from and that availability was found to be good, as was support in terms of development tools. The specific device a P17C756 was selected for the following key reasons:

- 200ns instruction cycle
- Expanded reduced instruction set
- Large programme memory
- Large data memory
- 50 I/O pins
- Assembler and C level programming options
- Large range of peripheral features including
- Multi-channel 10bit A/D convertors
- Universal serial ports including RS232 and I<sup>2</sup>C
- 3 16 bit timers
- Prioritized interrupts on timers, selected external pins, and other peripherals
- Emulation support
- UV erasable parts for development

## 4.4.2.5 Interfacing sensors to the real-time controller

The selected microcontroller was based on TTL compatible logic levels, and for the analogue ports 0-5v only. It was therefore necessary to ensure that each of the input signals from the microcontroller complied with either of these requirements. Subsections expand on detail in terms of modelling and implementation.

## 4.4.2.5.1 Prototype sensor interface.

In figure 4.6 it was shown that the output of the prototype sensor was frequency based, and that the voltage swing was confined within TTL limits. For full compatibility with digital circuits it was necessary to pass the signal through a schmitt trigger which ensured the signal was definitely either high or low, and therefore a pulse train. It is already known that the frequency was expected to vary between approximately 50KHz and 350KHz (chapter 3) so the question posed was how best to measure this frequency.

The P17C756 microcontroller offered two methods of frequency measurement. The first was an option to use the internal clock to time the duration of a defined number of cycles on the incoming signal. The second option was to count the number of incoming signals during a defined period of time (again using the internal clock). It was necessary to investigate these options further to understand the quantisation error each would introduce for the given frequency range.

At 50KHz the cycle time of the sensor signal would be  $20\mu s$ . If the microcontroller's internal clock were used to time a cycle of this period it would take 100 clock counts at 200ns. However, because the sensor signal is an input variable and not definitively quantified, a return of 100 counts can only suggest a value for the sensor signal within a certain error range. Figure 4.10 demonstrates how the error or uncertainty comes about, and demonstrates how the error may be quantified.



Accuracy of measurement is  $t_{sig} \pm 0.5 t_{ck}$ 

Error of measurement is  $t_{ck}/t_{sig}$ 

With reference to figure 4.10, the shaded region indicates the amount of uncertainty in any measurement of the signal frequency. This error is due to the real time between increments of the timer clock during which the sensor signal's cycle is completed. It is clear that the accuracy and error factors improve as  $t_{ck}$  becomes small relative to  $t_{sig}$ . Resolution in terms of time is also limited to the

period of  $t_{ck}$ , such that the range of variance in the sensor signal divided by  $t_{ck}$  and rounded down to the nearest integer determines the number of discrete steps which may be indicated by the measurement system. For example, suppose that the timer clock is incrementing at 200ns intervals, and that the sensor signal range is 50KHz to 350KHz (these figures are based on the actual model used thus far). First predict the error and accuracy at the low-end frequency:

- T<sub>sig</sub>=20µs and accuracy of measurement is 99%
- Error is ±0.5%
- Next predict the error and accuracy at the high-end frequency:
- $T_{sig}=3\mu s$  and accuracy of measurement is 93%
- Error is 3.3%

Finally to predict the measurement resolution in terms of discrete timer increments:

Output range = 
$$\frac{t_{sig-low} - t_{sig-high}}{t_{ck}} = 85$$

And in terms of frequency, one increment on the measurement output is equivalent to a change of approximately 3.5KHz on the sensor signal. In terms of linearity in this response the chart in figure 4.11 demonstrates that the counts per signal cycle increase exponentially as the signal frequency drops.



Figure 4.11 Frequency of output varies exponentially with displacement

This might appear to be an undesirable response curve, however, it should be recalled at this point that the frequency generated by the Colpitts oscillator is not linearly proportional to mechanical displacement of the tapered core. It may be interesting to predict how the count rate responds to core displacement:

$$count = \frac{1}{f.t_{cyc}}$$
(4.1)

where f is signal frequency and  $t_{cyc}$  is the counter time period

$$f = \frac{1}{2\pi\sqrt{LC}}$$
 from the Colpitts Oscillator (4.2)

$$L = \frac{\mu_0 \mu_r N^2 A}{l}$$
 from the inductance of a coil (4.3)

This equation was adapted in chapter 3 to approximate inductance from a tapered ferrite core, where two parts contribute to the total inductance:

$$L = L_{c} + L_{a}$$

where  $L_c$  is the inductance contribution due to the ferrite volume enclosed, and  $L_a$  is the contribution due to the air volume enclosed, such that:

$$L = \frac{\mu_0 \mu_r N^2 V_c}{l^2} + \frac{\mu_0 N^2 (V_{max} - V_c)}{l^2} = \frac{\mu_0 N^2 (V_c (\mu_r - 1) + V_{max})}{l^2} \quad (4.4)$$

where  $V_c$  is the volume of ferrite core enclosed within the core and  $V_{max}$  is the total cylindrical volume enclosed by the coil. The volume terms are derived from geometric parameters of the system so that when a prototype coil was designed as specified in figure 3.20  $V_{max}$  was determined to be 2.01e-7m<sup>3</sup>. The standard formula for volume of a cone was used to determine the instantaneous value of  $V_c$ . The equation below expresses the value of  $V_c$  in terms of displacement d and specific parameters and may be analysed with reference to figure 3.26:

$$V_{c} = \frac{\pi}{3} \left( \left( (r_{max} - d \tan \theta)^{2} (d_{t} - d) \right) - \left( (r_{max} - (1 + d) \tan \theta)^{2} (d_{t} - 1 - d) \right) \right)$$
(4.5)

where  $r_{max}$  is the maximum base radius of the conical section enclosed within the coil,  $\tan\theta$  is the parameter defining the angle of taper,  $d_t$  is the total displacement range, and l is the length of the coil (and hence the height of the truncated conical core). This is a most useful expression as all parameters are constrained by design geometry allowing volume to be directly related to displacement. It was found that expansion and subsequent simplification in terms of d gave an efficient form

of the equation, but only once the parameters had been quantified (this is annotated below). A complete expanded equation relating count time to displacement is:

$$count = \frac{2\pi \sqrt{\frac{C\mu_0 N^2 (\pi/3 (((r_{max} - d \tan \theta)^2 (d_t - d)) - ((r_{max} - (1 + d) \tan \theta)^2 (d_t - 1 - d))))(\mu_r - 1) + V_{max})}{l^2}}{t_{cyc}}$$

(4.6)

From this point it is necessary to quantify the parameters as used during mainstream prototype testing:

- C (capacitance) was constrained at 3.2e-9F
- $\mu_0$  (magnetic permeability of free space) is universally taken as  $4\pi e^{-7}$  Hm<sup>-1</sup>
- μ<sub>r</sub> (relative permeability of ferrite core) is set as 100 (as discussed in ch
   3)
- N (number of turns on coil) was constrained at 50
- l (length or thickness of coil) was constrained at 4mm
- r<sub>max</sub> (maximum radius of core) was constrained at 4mm
- $tan\theta$  (gradient of taper) was constrained at 1/7
- dt (total displacement range) was constrained at 28mm
- V<sub>max</sub> (total volume enclosed by coil) was constrained at 2.01e-7m<sup>3</sup>
- T<sub>cyc</sub> (counter time period) was constrained at 200ns

The taper volume expression was reduced as follows:

$$V_{c} = \frac{\pi}{3} \left( (0.004 - d/7)^{2} (0.028 - d) \right) - ((0.004 - (0.004 + d)/7)^{2} (0.024 - d)) \right)$$

$$= \frac{\pi}{147} \left( 8.128 \times 10^{-6} - 6.24 d \times 10^{-4} + 1.2 d^{2} \times 10^{-2} \right)$$
(4.7)

and subsequently the full expression for counter increments in terms of displacement was reduced to:



 $=9.8696 \times 10^7 \sqrt{1.1075 \times 10^{-12} - 8.40656 d \times 10^{-11} + 1.61628 d^2 \times 10^{-9}}$ 

(4.8)

Now in the case of the prototype described in chapter 3 it is possible to predict the relationship of displacement of core (input measurand) against the timer's increment value (sensor signal out). Figure 4.12 illustrates this relationship.

The chart in figure 4.12 predicts a near linear relationship between input and output of the sensor as prototyped. This was slightly surprising to the author as it was noted that the defining equation includes a quadratic component.





In order that this could be better understood the quadratic expression in equation 4.7 was plotted over an imaginary displacement range, as was the full definition in equation 4.8.

It is clear from figure 4.13 that although the quadratic component of equation 4.8 plots a typical quadratic curve, the effect of the full expression (in particular the square-root of the quadratic) is to more or less linearise the response curve with the exception of the region around the minima. As such it is shown that the principle of design in terms of signal handling is appropriate.



#### Figure 4.13 Defining expressions over an imaginary range

However there is an issue in terms of resolution. This arises because the frequency of the sinusoidal signal from the Colpitts oscillator must be quantised in time segments limited to the period of  $t_{cyc}$ . As in indication of the implications of this the chart in figure 4.10 has been re-plotted to show the quantisation effect, this is shown in figure 4.14.



In fact, the model indicated that the counter intervals would range from a maximum of 103 complete counts at the lowest output frequency to 14 complete intervals for the highest output frequency, giving and output range of 89 discrete

steps, and hence a resolution of  $\pm 0.13$ mm (approximately). This resolution may be suitable for some applications, although it is often preferable to ensure resolution is better than 1% of the full input range or in terms of binary bits at least 7bit resolution. The resolution of this system may be improved by either of two methods:

- Increase the measurement period, typically by timing multiple signal cycle periods or by dividing the signal frequency before timing the cycle period.
- Reduce the period of t<sub>cyc</sub> the timer interval.

It was by the nature of the microcontroller employed not feasible to consider significant reduction in the timer interval, so that a decision was made to count multiple signal periods. Various settings were used based on timing the period of 16 complete oscillator cycles whose output had first been divided by a  $2^x$  where index x was typically 2-8. This allowed the generation of higher resolution output, often around 1000 discrete steps over the complete displacement range (which could be expressed as roughly 10 bit resolution or resolution to  $\pm 14\mu$ m). The response pattern of the prototype sensor is presented and discussed in detail in chapter 5.



Measurement of sensor signal frequency was accomplished by use of hardware and software functions of the P17C756 microcontroller. Figure 4.15 shows how a TTL signal which has been pre-scaled to increase resolution was applied to an I/O port of the microcontroller. The I/O port had been configured such that internal hardware would use the internal clock to time a duration of 16 rising edges. On completion, the timing count stored in a CAPture register would be made available to higher level functions for processing. It is appreciated that extending the duration of a frequency capture process may result in undesirable effects if the input measurand were to be changing rapidly. This potential problem is described and addressed in chapter 6.

#### 4.4.2.5.2 LVDT sensor interface

The LVDT external amplifier was configured such that it's analogue d.c. output was compatible with the analogue ports of the pic microcontroller (i.e., 0-5V range). The analogue capability of the microcontroller in use was limited to input only, and via a 10 bit successive approximation method to digitise the signal. The analogue peripheral was multiplexed to 16 I/O ports, although only one was required in this instance. Sampling time was dependent on the microcontroller clock speed and was approximately 1.6µs. A sample was initiated at the same time as other sensors were triggered to provide data, and an interrupt routine moved the sample into a safe register for processing.

#### 4.4.2.5.3 Temperature sensor interface

As previously discussed the temperature sensors used in this application were interfaced directly to an  $I^2C$  serial bus. The connections made between devices and the microcontroller were as shown in figure 4.16.



Figure 4.16 Use of I<sup>2</sup>C serial bus

Digital temperature sensors

# 4.4.3 Experiment Controller

The experiment controller was designed as its title suggests, to manage all aspects of experimental procedure. Whilst the main component of the controller is its software based logical flow, all that encompasses experiment control has not been contained only within a computer programme. Significant duties were delegated to peripheral components for their particular ability, these include a microcontroller to handle real-time critical functions, a motor controller to power the measurand drive, and various sensors to determine status and conditions. The experiment controller can be divided into parts of specific function which are detailed below.

# 4.4.3.1 User Interface

The user interface was designed and implemented using the LabView<sup>™</sup> PC based development environment to allow interaction with the test rig for the following purposes:

- Configuration of test parameters
- Observation/processing of 'real-time' data
- Observation/processing of stored data
- Ultimate system control including automation override

# 4.4.3.1.1 Configuration of test parameters

The setting of parameters allowed the user to configure the test system in an optimal way for the type of information required. This includes settings within two different facets, the first being calibration of system peripherals, the second being setting of high level control strategies for the required type of test.

# 4.4.3.1.2 Observation/processing of 'real-time' data

The software was designed such that a degree of sampled data could be indicated both graphically and in numeric form. This function was most useful in determining system reliability during extended test periods. Figure 4.17 indicates a typical continuous plot of data as it is sampled.

Figure 4.17 Real-time data display



The display charts were able to plot up to 1024 points at any time for a rolling graph (as in the LVDT Reading, and Sensor Reading charts), and an unlimited number of points for a scatter chart (as in the plot of output / LVDT). A degree of real-time data processing was available in the form of successive transfer functions to compensate for systematic effects which become apparent in discussion of the results (chapter 5). In Figure 4.17 it can be seen that linear relationships result from a plot of temperature against sensor output. The temperature coefficient could if necessary be calculated from this historic data to provide a compensation in the plot of sensor output/LVDT (where the LVDT is a relatively temperature stable reference). The nature of the LabView software allowed easy user configuration of 'real-time' transfer functions. An important utility of real-time data processing was storing to disk such that the data was available for subsequent analysis in a generic format.

## 4.4.3.1.3 Observation/processing of stored data

Stored data was typically retrieved for the purposes of condensing (reduction of huge data sets from extended test periods, typically 100K samples of 30 bytes each) or for the purpose of performing an off-line calibration check (update of transfer functions via curve-fitting).

4.4.3.1.4 Ultimate system control and automation override.





The user interface also provided standard functions such as start, stop, emergency stop, reset, and file handling. Figure 4.18 shows a screen capture of a typical user interface.

# 4.4.3.2 Logical flow

The user interface also held the underlying responsibility for programme flow. The programme was designed such that it was event driven from an encapsulating loop. A certain sequence of register tests were performed in the loop to determine status and actions. The main structure of the PC based LabView programme was designed around a communications protocol with an external microcontroller (previously described). The function of the communications protocol is described in the next sub-section. An important requirement of the logical flow of the user interface software was to manage incoming data in such a way that it was appropriately extracted and labeled according to its source.

# 4.4.3.3 Communication

In each loop of the main programme software a message was sent to the microcontroller via the communications system (based on RS232). The message sent was either a control or configuration instruction, or else it was a poll for any data to return. The form of the message sent was as follows:

HEADER DATA NULL

where HEADER is some 3 character (ASCII) sequence to identify the message, DATA is any data (ASCII or binary depending on message type), and NULL is a universal message terminator (required especially if binary data is encapsulated). For example, the peripheral microcontroller was programmed among other functions to control a stepper motor for the measurand drive, the message format to access motor control was as follows:

Motor control instruction Energise Direction Speed

an example of its use is:

Mot0,0,1200null

This string instructs the microcontroller to: not-energise the motor, direction is forward, and speed setting is 1200. If the energise variable is zero the subsequent variable settings will be written to the microcontroller registers, but will not energise the motor. The speed setting is a full scale 16bit number.

A full listing of all the instructions used in this protocol are given in appendix c. In all instances the message sent out to the microcontroller generated an a response. This would either be an ACKnowledgement of meaningful message received or else report of a BAD message received. A bad message would imply that something was transmitted and something was received, but it made no sense to the receiver, in this case the user interface software would normally resend the message. In some situations no response would come from the micrcontroller (e.g., if it was powered down), in this case an indicator on the user interface would be enabled and the message would be retried until intervention or success.

To maintain reliability in an asynchronous two way protocol, the user interface software (on host PC) was defined as the master, and the microcontroller was defined as the slave. The slave was programmed such that it would not speak unless spoken to, and it would only answer the questions asked of it. In this way the master was able to handle communications in sequence, only requesting data when ready for it. It was for this reason of awaiting the master that the microcontroller required substancial data buffers for messages to return.

127

The flow chart shown in figure 4.19 demonstrates in simplistic terms how the microcontroller handles messages from the host computer.



Figure 4.19 Message-in handling of microcontroller

Data received at the RS232 serial port of the microcontroller triggers an interrupt which moves the data to an appropriate register before setting a flag for the main programme loop to test. When the flag indicates a 3 character header has been received the microcontroller then interprets the header to determine what if any data is to follow and where the data should be placed. Interpretation of the header also allows flags to be set for the main programme loop to test to determine what if any action is required.

In some instances a message from the host computer would require a response from the microcontroller, for example a request for motor status. The protocol structure was designed such that there are two types of response, the first is generic acknowledgement or rejection of the latest message received, this is an automatic 3 byte return as already discussed. If further data is to be sent such as sensor data, then a return message is compiled with a DAT for data header. The data is compiled in binary format. The return message is appended to a software FIFO transmit queue. The transmit queue is picked off by the appropriate message instruction from the host computer. A message counter is associated with the transmit queue so that if no messages are waiting for transmission this may be conversed to the host. Messages other than sensor data may be returned to the host computer. These other messages include status reports and reports on asynchronous peripheral events such as triggering of the safety cut-out mechanism. In this manner full two-way conversation between master and slave has been achieved. The simplicity of the protocol to eliminate simultaneous transmission errors ensures data integrity. The early results in chapter 5 indicate how the protocol was optimised for high speed operation.

# 4.4.3.4 Data processing

The nature of the development tools for the user interface software allowed easy integration of complex data processing functions. In general the nature of data processing applicable to this system was filtering, calibration, and linearisation. These functions were partially embedded within the microcontroller and mostly incorporated within the user interface software.

# 4.4.3.5 Filtering

Low pass filters were optionally applied to raw data to remove spurious points. This was typically performed by the microcontroller on a moving average basis. Other filters were applied to data in 'real-time' and off-line processing depending on the information required at the time. A substantial number of filters were available in software for the user interface.

# 4.4.3.5.1 Calibration

Calibration was employed both within the microcontroller and within the user interface software to convert raw data into meaningful and useful ranges for observation and subsequent processing.

# 4.4.3.5.2 Linearisation

A substantial element of data processing was to utilize functions for the purpose of observing non-linear but systematic characteristics as reliable linear characteristics. This was achieved mostly through the application of polynomial functions fitted to response curves, although other methods were considered such as conversion tables and methods of interpolation.

# 4.4.3.6 Safety

Safety became an important aspect of control within the experimental rig. A large motor was employed on the measurand drive such that gears and end stops were hazardous parts of the mechanism. A safety cut-out system was implemented around the hazardous points with an additional user operated switch in case of difficulties. The normally closed safety circuits were designed to open on triggering such that a hardwire link to the motor driver would cause deenergisation. Further to this the safety system was linked to the microcontroller which would assess the cause of the stop and relay this information to the host user interface via the communications protocol.

## 4.4.4 The integrated test-rig

The diagram in figure 4.20 is a schematic representation of the electrical and mechanical integration of the various components which go to make up the entire test-rig.



Figure 4.20 Schematic representation of physical test rig

The diagram shown in figure 4.21 represents the interaction of core modules of the software system. As can be seen from figure 21, the software is hosted on two independent platforms, which are linked by a communication protocol previously discussed. The diagram shown in figure 4.21 represents the interaction of core modules of the software system. As can be seen from the diagram, the software is hosted on two independent platforms, which are linked by a communication protocol previously discussed.



#### Figure 4.21 Schematic of software integration across control system

## 4.5 Tests conducted

The tests conducted fall into two distinct groups. There was one group concerned with the stability of the sensor system over time. There was a second group concerned with performance in active use. These two groups are expanded in the sub-sections below.

#### 4.5.1 Testing sensor system stability

It was critical to understand the consistency of performance over time, particularly with a view to identifying systematic variations distinct from un-attributable noise. Tests were designed to hold the system in a steady state so that variables were reliably fixed or else reliably monitored. The test was conducted so that regular samples were taken over a long period of time, in some cases up to 3 weeks. Chapter 5 details how systematic variations were identified and compensated for. The stability tests were not exhaustively conducted with tests for every variable combination because time would not allow this. A range of random tests were conducted, together with several tests at specific variable settings (for example with the minimum and maximum measurand settings).

#### 4.5.2 Testing active performance of sensor system

A variety of tests were designed to optimise data obtained for characterisation of specific performance criteria. First and simplest was the determination of output range for a given input range. This test was conducted for different frequency sample times as discussed in section 4.4.2.5.1. The test involved a simple sweep of the input measurand (via the measurand drive) over its full range, to record both the reference LVDT reading and the prototype sensor reading. From this information it was possible to interpret information on displacement resolution including maximum resolution average resolution and instantaneous resolution. Further to this it was possible to infer hysteresis between a rising input measurand and a falling input measurand. The second test was designed to expose error due to repeatability. A high level function of the user interface was designed specifically to allow repeatability testing. The process involved driving the input measurand to some random displacement within the valid range. The drive was then required to return to some specified displacement based only on the prototype sensor's series of readings. The difference between position achieved and position requested gave rise to a reading of repeatability. The absolute positions were measured both by the reference LVDT and by the operator using the vernier scale. It was possible to infer other characteristics such as overall error and accuracy from a considered combination of all test data obtained. The results of these tests are presented in chapter 5.

#### 4.6 Conclusion

This chapter has presented a test methodology to characterise typical performance criteria for the prototype sensor developed in chapter 3. The test methodology has required in its implementation the development of a specific experimentation rig, which has been described in this chapter. This chapter has concluded by stating how the test rig was used to measure relevant performance criteria for the prototype sensor. Chapter 5 gives a presentation of the results obtained using this equipment and results from earlier tests prior to the use of the main test rig.

132

# 5 Results and discussion

## 5.1 Introduction

The results chapter presents a selection of results from experimentation which are representative of all the results taken. Each presentation is accompanied by a discussion where appropriate so that a logical and chronological progression may be seen between results as the chapter progresses. The chapter has been divided into three main parts. The first part is concerned with initial feasibility studies before the main prototype was developed, this section includes analysis of the manufactured ferrite core. The second part is concerned with the main results. This section demonstrates a degree of evolution as the prototype and the experimental environment were refined. The second part is the major part of this chapter. The third part is concerned with preliminary testing of the prototype under bending, for which there are no bench-marks. Some model development is included both for the prototype sensor and for the test rig.

## 5.2 Preliminary Results

## 5.2.1 Concerning properties of elastic ferrite core

Tests were conducted to ascertain the effect of doping silicone rubber with ferrite particles as described in chapter 3. The tests were concerned primarily with magnetic properties, although some consideration was given to the effect on mechanical properties

## 5.2.1.1 Magnetic properties of elastic ferrite core

This test concerned the magnetic permeability parameter of the ferrite core. The objective of testing was to quantify relative magnetic permeability according to doping levels by mass. Core samples were prepared in batches from zero doping through to approximately 30% doping (at which point mechanical properties of compound begin to deteriorate significantly). Figure 5.1 shows a typical batch of ferrite doped silicone rubber cores. They were produced so that they would provide a free sliding fit within the prototype sensor's coil (as described in chapter 3). The table in figure 5.2 quantifies the sample data plotted within figure 5.3.

Figure 5.1 Core samples used in permeability testing



Figure 5.2 Core sample data

C.L.		
Exact doping levels by percentage mass		
	Batch 1	Batch 2
Core 1	9.26%	9.15%
Core 2	17.48%	17.44%
Core 3	25.05%	24.85%
Core 4	31.98%	31.46%

The Y axis data shown in figure 5.3 indicates the Colpitts oscillator time period in terms of microcontroller cycle periods. The important information from this plot however, is that the level of doping (and hence magnitude of magnetic permeability) is reasonably proportional to the frequency of oscillation in the Colpitts oscillator. This is true down to zero doping level.



Figure 5.3 Relationship of ferrite density with inductor performance

# 5.2.1.2 Mechanical properties of elastic ferrite core

It was considered not entirely relevant at this time to perform mechanical testing on the samples manufactured. The reason for this is that it was understood that silicone rubber was a convenient

# 5.3 Main Results

# 5.3.1 Results to determine optimal configuration

The initial test was a comparison of the modelled sensor output range against the actual output range observed. From chapter 4 it will be recalled that an approximate model suggested that the Colpitts oscillator (as designed in chapter 3) frequency should vary in the range 50KHz to 350KHz. The Colpitts oscillator was tested under normal operating conditions (i.e., the output stage was loaded by a microcontroller I/O port) to observe the frequency range actually achieved. Figure 5.4 shows traces of the output signal post-digitisation for a) the ferrite core fully extracted, and b) the ferrite core fully inserted.





It should be noted that the core used in this test was a 30% doped silicone rubber ferrite core type. The time base of the traces shown was 5µsec in each case such that the measured frequency extremes were: 292KHz (fully extracted), and 260KHz (fully inserted. The result clearly indicates that although the frequency range achieved lies within the predicted region, it is a very much reduced range. The reason for this mis-correlation with the expected result is perhaps as previously highlighted, that the complexity of the model employed in chapter 3 was insufficient. However, further results presented in this chapter indicate that it was feasible to proceed with testing despite the disappointing coil response to

displacement. Chapter 6 is partially devoted to re-examining the inductance model with a view to understanding the results obtained.

With respect to the microcontroller frequency sampling function, the signal was digitised such that signal frequency became a function of 'number of oscillator cycle periods' as described in chapters 3 and 4. The chart shown in figure 5.5 graphically represents the range of sample values for the microcontroller as the sensor is allowed to measure full displacement.



Figure 5.5 Sensor output range in terms of microcontroller time cycles

With reference to the data shown in figure 5.5 it can be seen that the microcontroller counted 216 cycles for the highest frequency generated, and 246 cycles for the lowest frequency generated. This is a range of just 30 discrete steps (equivalent to approximately 1mm per increment). This suggests that the microcontroller was detecting a frequency range of:

$$f = \frac{16}{n_{cyc} \times t_{cyc}}$$
  
$$f_{l} = \frac{16}{246 \times 250e - 9} = 260 \text{KHz}$$
  
$$f_{h} = \frac{16}{216 \times 250e - 9} = 296 \text{KHz}$$

This shows that the microcontroller was functioning correctly in terms of frequency measurement (this was also confirmed with a fixed frequency generator

source). Note that the factor of 16 was included because the microcontroller hardware was configured to pre-scale the sensor frequency by 16.

The sensor system was then configured to measure signal stability over time as the measurand was held static on the experimental rig (mechanical statics were subject to thermal expansion as mentioned in chapter 4). The initial test was conducted over a period of approximately 24 hours with sampling every 10 seconds. The chart shown in figure 5.6 indicates how the microcontroller recorded the prototype sensor's frequency output over this period.





The reliability of the sensor signal appears to be good, based on approximately 5000 samples shown in figure 5.6 only 18 could be identified as spurious. It is clear that the measured frequency was consistently within the interval between two quantised steps (216 - 217) which amounts to a frequency in the range 296.3KHz - 294.9KHz.

It was clear from these first findings that the resolution of the system was being limited by the frequency measurement function because it was restricted to a time base of 250ns (microcontroller clock period). 3 solutions were considered to address this issue, alone or in combination:
## 5.3.1.1 Increase the microcontroller clock speed.

It was possible in principle to increase the clock speed of the processor to 33MHz and hence achieve a 121ns time base. This would be a favourable solution to increase performance slightly and theoretically double the system resolution. However, it was found that the microcontroller's secondary functions such as communications and motor control would be adversely affected by a change to the time base, and too much time would be required to modify the code to compensate. In a revised system it would be prudent to maximise the built-in capability of the system clock.

An alternative to using the internal clock would be to make use of an external very high speed clock in combination with an external dedicated logic counter. In this way clock speeds of 100MHz and hence time bases of 10ns would be possible. The draw-back of this solution would be the increased complexity of hardware and interfacing between peripherals and the microcontroller. However, this solution would be considered in a subsequent revised system.

# 5.3.1.2 Count more than 16 cycle periods from the sensor output signal.

It was determined that the simplest manner in which resolution enhancement could be achieved would be to measure frequency over a longer time period, by increasing the prescale function of divide by 16, to higher powers of 2. It was found, as will be reported below, that scaling of 64 and 128 was not unreasonable. However, it was realised that an increase in the sample time period may introduce undesirable errors during rapid changes of the input measurand. It was decided that this method would be suitable for static and slow speed testing of the prototype sensor. This solution would not be favoured in a subsequent redevelopment of the system.

## 5.3.1.3 Increase the sensor output signal frequency range.

As reported, it was expected that the sensor system's frequency output range would be much larger than was actually observed. It is clear that if design optimisation could be employed, a larger range of induction from the coil might be achieved and hence an improved frequency swing for a given displacement. To highlight the potential for optimisation the tapered silicone rubber core was temporarily replaced with a conventional solid cylindrical ferrite core to observe the frequency difference between full insertion and full extraction. Figure 5.7 graphically represents the frequency swing between full insertion and full extraction and full extraction, where the measured range was 140KHz - 290KHz.



Figure 5.7 Frequency swing obtained from traditional solid ferrite core.

## 5.3.2 Main-stream results

Main-stream experimentation was conducted over a period of some 7 months which includes time for a degree of alteration and improvement based on successive findings. As such, it should be evident in this section that the performance of the prototype improved in time as a result of mechanical, electrical, and software enhancement of the experimental object and the experimental rig.

## 5.3.2.1 Test results for first version of prototype

#### 5.3.2.1.1 Steady state drift test

A drift test was conducted over a period of days to determine the effects of the environment on the stability of the prototype. The input measurand was held static by the measurand drive whilst frequency samples were taken every 10 seconds by the automated system. A sample was also collected from the reference LVDT sensor to monitor the stability of the test rig. Figure 5.8 shows how the sensor's output frequency varied as recorded by the experimental rig over a period of 10 days.

Figure 5.8 Steady-state drift



From this trace it can be seen that increased prescaling has been employed to increase the resolution of frequency measurement (as described earlier in this chapter). As a result of this, the recorded oscillator cycle count for each successive pulse train from the sensor is in the order of 9000 as opposed to 200-300 as observed earlier. From figure 5.8 it can be deduced that there is a considerable drift in the recorded frequency and that there is a periodic component which has a 24 hour cycle. There is also evident spurious signals which tended in all but one case to record a 'low' result. Although the timing of these events appears random the fact that they tended to be low might suggest that there is a systematic solution to their existance. This will be addressed subsequently. Regarding the periodic drift it was natural to assume this might be due to temperature variation and as such the system was rapidly modified to include the temperature sensing sub-system which has already been described. There were two options available to characterise the effects of temperature, the first was to conduct experimentation within a controlled environment such that temperature could be a controlled variable. The second option was to conduct experimentation within normal 'room' conditions such that temperature would vary independently and freely, but the temperature was recorded and its effect on the experimental output was also recorded. Either method would in time allow a correlation between experimental performance and temperature to be observed. It was necessary to employ the second option because of the expense involved with

purchase and use of an environmental chamber. The effects of temperature are quantified subsequently.

## 5.3.2.1.2 Variable measurand tests

It was noted that the frequency drift amounted to a range of measured output of 30 discrete incremental steps. In order to understand the error magnitude contributed by this range in output for a fixed input, it was necessary to measure the complete working range of output for the complete working range of input. The chart shown in figure 5.9 is the recorded output from the test sensor as the input measurand was driven through 3 complete displacement cycles.





From figure 5.9 it is possible to determine that the effective output range is from 9190 to 8680 oscillator cycle counts and therefore 510 discrete incremental steps. This being the case, the error contribution from a drift of 30 steps is 5.9%, which is rather high.

Other useful information was extracted from the range test data these have been compiled into a list of results as presented below.

## 5.3.2.1.3 Linearity

To determine linearity of the output it was necessary to make a correlation with the reference LVDT sensor. The chart in figure 5.10 shows how the two data sets relate when mapped onto each other. It can be seen that there is a non-linearity between the prototype sensor and the reference LVDT sensor data (the LVDT is assumed to be linear 1%). This non-linearity is better observed in figure 5.11, which shows a plot of prototype sensor data against the reference data.





Figure 5.11 Linearity of prototype sensor



In terms of quantification the maximum divergance from a line of best fit occurs towards the high-end frequency of the prototype sensor, where the error attributable is approximately 20%. It is clear that a necessity exists to examine the response curve of the prototype sensor to changing input.

## 5.3.2.1.4 Hysteresis

With reference to figure 5.11 there was no observable hysteresis beyond the amplitude of random noise and quantisation noise. This is an important finding because any significant hysteresis would irretrievably diminish the performance of the sensor in normal operation. Figure 5.12 zooms in on a section of the graph shown in figure 5.11to demonstrate that hysteresis is at a level more or less masked by noise.





#### 5.3.2.1.5 Resolution

With reference to figure 5.9 it was found that when using a pre-scale divider of  $2^9$  an output range of 510 discrete incremental steps was achieved. It is known that the input displacement range was 28mm. Therefore the theoretical resolution of this system as configured was to 55µm on average.

## 5.3.2.1.6 Repeatability

At this stage repeatability was not tested as it was found desirable to improve performance in terms of resolution, drift, and stability.

On conclusion of the first main test it was clear that there existed a potential for comparable performance with conventional sensing. However, it was deemed that refinements to the system were required to optimise the performance to a level which would best demonstrate the capability of the novel prototype sensor.

#### 5.3.2.2

Prior to any modification directly involving the sensor mechanics or electronics temperature sensors were placed both at the modulator site and at the oscillator site (specifically the temperature of the op-amp). The temperature sensors employed were as described in chapter 4. Their specific signal information was in the range of  $-25^{\circ}$ c to  $+125^{\circ}$ c and intervals of 0.5°c.



Figure 5.13 Prototype sensor output drift over a period of time

The chart shown in figure 5.13 indicates how the prototype sensor output drifted over a period of days, once again following a 24 hour cycle. The data shows how the signal drifted from its average value in terms of oscillator cycle counts. The significance of this data becomes apparent when compared to the temperature data recorded during this test which is plotted as shown in figure 5.14 below. It is clear

that the periodic drift in frequency exhibited by the prototype sensor was largely due to environmental temperature variation. The offset observed between the two temperature points recorded may be explained by the oscillator's location in an enclosed chamber with other electrical systems.





The chart shown in figure 5.15 plots sensor drift against oscillator temperature. A linear trendline has also been included to show where the line of highest density lies.



Figure 5.15 Deduction of temperature coefficient

This chart is slightly unclear because of spurious data points (discussed and addressed in section 5.3.2.4.1). A software filter was applied to remove these points, the revised data is presented in figure 5.16.



Figure 5.16 Filtered data to show temperature coefficient

After filtering the data set it became clear that two quite distinct groups existed each with its own temperature offset, this is highlighted by figure 5.16. Although there appeared to be a shift in temperature effect at some point during the experiment, it was important to note that the temperature gradient remained reasonably constant.



Figure 5.17 Deivance of temperature from sensor output

The reason for the observed function shift was at this time unclear, although it was supposed that there might have been some mechanical adjustment in the system, perhaps settling of the prototype. The chart in figure 5.17 highlights the point in time when the shift in temperature response occurs. It can be estimated that the anomaly occurred at some point during the  $5^{th}$  day of recording, and notably at a time when the daytime temperature reached a new high. Other tests revealed a tendency for sudden and unpredictable shifts in output when operating over a period of days.

It was assumed that further investigation would be able to account for shifts in temperature relationship, and it was determined that it would be useful to attempt to use the data collected to generate and apply a temperature compensation function. With reference to figure 5.16 the linear function from the larger data set: y=7.3x+9061.5 was applied to form the specific function y=D-7.3t-9061.5. In this function two inputs are applied, D the prototype sensor reading, and t the corresponding temperature of the Colpitts oscillator. The output y is the temperature compensated value of the sensor reading. Figure 5.18 demonstrates this function applied to the prototype sensor data set.





It can be seen that drift is reduced from a range of approximately 30 incremental steps to approximately 10 steps, even accounting for the anomaly which caused the shift in temperature response. Assuming it will be possible to remove this anomaly the resultant drift might be only 4 incremental steps on average, this being an error of 0.8%. Subsequent testing was designed to allow an 'on-line' temperature compensation function.

## 5.3.2.3 Test results for the third version of the prototype

It was anticipated that an improvement in temperature drift might be achieved if a more suitable op-amp were selected to operate within the Colpitts oscillator. This was considered because it was found that the lf353 as used to this point exhibited a temperature coefficient of typically  $10\mu$ V/°C. An op285 op-amp was selected as a direct replacement as it was compatible in every necessary way (mechanically and electrically) and its published temperature coefficient is 10 times better than the previous device at just  $1\mu$ V/°C. These devices are featured in appendix D.

A similar test series was conducted with the modified prototype, with a particular interest in the drift performance.

#### 5.3.2.3.1 Steady state drift test

A steady state drift test was conducted over a period of days much the same as before. The chart in figure 5.19 shows the prototype sensor output over a period of days. Most notable was the periodic feature as observed in previous experimentation, and also the random spurious data points.



Figure 5.19 Prototype sensor output over 4 days

It is also evident from figure 5.19 that the drift pattern appears to be of lesser amplitude than in previous experiments with the lf353 op-amp, the range of drift over 4 days being approximately 10 incremental steps. This finding was as expected and the improvement was a factor of 3. Figures 5.20 to 5.22 may be used to confirm that the drift pattern is due to temperature.



Figure 5.21 Relationship of oscillator temperature with drift



The observation that modulator and oscillator temperature, although offset from each other, tend to vary over time at the same rate, indicates that the temperature effect was due to environmental temperature changes rather than, for example, varying power dissipation in electrical components of the sensor circuit. This

149

result indicates that the electrical circuit is reasonably stable, although there would be a period of warm-up from a cold start.



Figure 5.22 Relationship of modulator temperature with drift

A calibration of the data output was performed with reference to the oscillator temperature. The chart in figure 5.23 indicates how the signal drift has been improved from a pre-calibrated condition.



Figure 5.23 The effect of temperature compensation

Prior to compensation the effect of drift was to add a 10 increment uncertainty to the output signal over a 4 day period. With reference to figure 5.23 it is clear that the 10 increment drift has been reduced to approximately 3 steps over 4 days (excluding spurious data points). Analysis of the accompanying range information will allow quantification of the error percentages due to these drifts.

## 5.3.2.3.2 Range test

A range test was conducted for the modified prototype sensor. The chart in figure 5.24 indicates that the input measurand was driven over its full displacement range approximately 3 times.



Figure 5.24 Output range of prototype sensor

The output range recorded for the OP285 version of the prototype was 170 discrete steps. This implies that the error due to drift was 14% before compensation and 1.8% after compensation. Clearly the reduced output range of this configuration has resulted in a poor performance in spite of the seemingly superior characteristics of the op-amp selected. It was apparent that the OP285 was an unsuitable device in this application, further compounded by exhibition of a significant hysteresis (approximately 3% uncertainty) on the output signal. This hysteresis is demonstrated by the chart in figure 5.25. As a result of problems encountered when attempting to base the Colpitts oscillator on a conventional dual supply amplifier, it was deemed appropriate to redesign this circuit with a view to increasing logic compatibility and maximising the performance of the oscillator in this application. The fourth test series presents the alterations made and accompanying new results.



Figure 5.25 Evidence of hysteresis due to circuit design

## 5.3.2.4 Test results for the fourth version of the prototype

The main alteration in the system between the second and the third test series was a change of operational amplifier in the Colpitts oscillator. There was also an investigation into the quantity of spurious data points recorded during previous experimentation.

The selection of operational amplifier coincided with a redesign of the power supply for the Colpitts oscillator. In initial versions of the prototype sensor a dual supply was used at  $\pm 12V$ . This was required to allow the lf353 and op285 opamps to produce sufficient output swing for subsequent pulse train digitisation. Conversion of a dual supply generated signal into a TTL compatible signal involved various stages which it was preferable to eliminate. A high performance single supply op-amp having a rail-to-rail output swing capability was found in the lm6142. This was a vast improvement over the previous component as no special supply was required other than to provide a 2.5v reference, and furthermore the oscillator output required only one stage of amplification to saturate the sine wave into a full-swing pulse train having automatic TTL compatibility.

It was found that modification of the Colpitts oscillator as described produced a very slight rise in output frequency due to the different characteristics of the opamp and the discrete components operating at different voltage levels. Whilst it may be possible to evaluate these differences mathematically it was deemed unnecessary at this time. The performance of this circuitry revision is presented subequently.

#### 5.3.2.4.1 Correction of spurious data points

It was necessary to diagnose the cause of spurious data points in order that a solution could be found. It was supposed that the following components of the system could be responsible for data errors: noise affecting the signal quality of every sensor in the experimental rig, poor physical communications between host PC and microcontroller, programme problem within host PC, or programme problem within microcontroller.



Figure 5.26 Testing of frequency measurement function

It was observed that the most obvious spurious signals were present on the prototype sensor signal trace, and there were occasional spurious points on the signals from the temperature sensors, although the errors were not synchronised in any way. In order that the modulator of the prototype sensor could be eliminated as a source of spurious signals the Colpitts oscillator was replaced with a crystal derived pulse train which was known to be stable and free from noise. The signal source was divided to approximately 300KHz before being applied to the capture facility of the microcontroller. Figure 5.26 shows the frequency information captured over a period of approximately 2 days.

From figure 5.26 it is clear that spurious points remain despite the stable frequency source. As such this demonstrated that the error was not in the incomming signal, but in subsequent digital processing (i.e., in the microcontroller software, PC software, or communications link between the two). A check was made on the integrity of data transmitted between microcontroller and PC. It was observed that the message structure was intact and erroneous entries occasionally occurred within the message structure resulting in spurious data. Therefore it was revealed that the problem lay somewhere within the microcontroller code.

The problem was rapidly identified as data preparation before embedding into message for sending. The reason for this was that the message structure was always correct but the data held within it was not always correct, suggesting that at the time the message is assembled invalid data is accessed. The cause of this could either be down to reading the wrong data registers (pointer error), or reading the correct data registers, but finding them to hold invalid data. Debugging indicated that the correct registers were accessed but that they occasionally held invalid data. The problem was attributed to the real-time software running out of time to complete data conversions from binary to ascii format (at this stage all data was converted to ascii before being sent to the PC). Time was limited because there were various interrupts servicing the multiple functions of the microcontroller and data preparation (a lengthy and iterative process) had to be completed between interrupts and before the next data sample arrived from the sensors (maybe as frequently as every 2ms). The solution was an all-round efficiency improvement of the data processing, message preparation, and communications format of the microcontroller code. Code was modified so that there was no requirement to convert binary data into ascii data before sending to the PC. This had the advantage of reducing the 'idle' time load of the real-time process (no lengthy conversions to be made), and reducing the byte length of each message to be sent (e.g., a hex value of hFFFF was now sent as two hex bytes, (FF)(FF), instead of its five byte (five character) ascii equivalent of 65535). As a result of this improvement the incidence of corrupt data was reduced to negligible.

## 5.3.2.4.2 Static drift test.

The temperature coefficient of the lm6142 is rated at  $3\mu V/^{\circ}C$  and based on previous experimentation it was expected that there would be a significant temperature effect on the Colpitts oscillator. Figure 5.26 shows the recorded output of the prototype sensor over an 18 day period.









As expected there is a significant periodic drift which is based on a daily cycle, although a significant drop off was recorded over the 3<sup>rd</sup> and 4<sup>th</sup> days. This drift

pattern was closely matched by temperature variation, which is plotted in figure 5.27.

The temperature chart indicates that environmental temperature significantly dropped off to explain the sensor output drift pattern. The correlation of sensor drift to temperature variation is expressed in figure 5.28. It can be seen from the best fit function that the relationship is approximately linear. Therefore this function was applied to the data set to afford temperature compensation. The chart in figure 5.29 indicates how drift has been improved from before compensation to after compensation.



Figure 5.28 Temperature response of prototype output

Figure 5.29 shows how drift has been improved from a range of 36 steps over a 4 day period to just 3 steps when applying temperature compensation. The accompanying range test for this configuration will allow the error contributions of drift to be determined.

Figure 5.29 Application of temperature compensation



## 5.3.2.4.3 Range test

The input measurand was cycled through its entire displacement range producing the output response from the prototype sensor as shown in figure 5.30.



Figure 5.30 Prototype sensor response to input cycle

The range of the sensor output during this test was approximately 900 incremental steps. Therefore the resolution was on average 31µm per step. The error due to uncompensated drift was 4%, and after temperature compensation 0.3%.

#### 5.3.2.4.4 Linearity

The chart in figure 5.31 shows the linearity of the prototype sensor with respect to the reference LVDT. From the information presented it is clear that the response of the prototype sensor to displacement is not linear. The response curve is of the typical cumulative form where sensitivity diminishes to zero towards each extreme, and is at a maximum towards the centre of the data set.





#### 5.3.2.4.5 Hysteresis

The level of hysteresis can be observed from the chart in figure 5.31. It is clear from this figure that hysteresis is at a level of quantisation noise. Figure 5.32 shows a close-up of a mid-portion of the data shown in figure 5.31. This chart proves the lack of hysteresis in the system (i.e. it is low enough to be masked by noise), which is favourable.



Figure 5.32 Close-up of response curve to observe hysteresis

## 5.3.2.4.6 Repeatability

The repeatability of the system was conducted under a specific test as described in chapter 4. Initially a post-temperature compensation repeatability test was conducted with uncalibrated data, (i.e., the effects of temperature were removed but the displacement response was not linearised).



Figure 5.33 Chart of positioning errors about a single homing point

The temperature compensation function applied was similar to the linear functions described earlier in this section. The chart in figure 5.33 indicates how the absolute resting point of the driven slide finished when the prototype sensor was

employed as feedback to return the slider to the same point repeatedly. As can be seen, the range of final positions varied over 0.18mm when recording 10 repeats of the test.

The result shown in figure 5.33 was typical of positioning errors over the full displacement range. It may be observed that the data of figure 5.33 falls into two categories, high reading and low reading. The four high reading data points were obtained when driving the input measurand in the opposite direction from when the low reading data points were recorded. It was supposed that this effect could be the product of mechanical hysteresis (which has already been observed as insignificant), or else that it could be the product of quantisation in data processing. Raw data from the digital stage of the sensor (i.e., the microcontroller) was presented in integer number format. The temperature compensation function includes coefficients which are real numbers, therefore it was expected that the temperature compensated output would be a real number and not an integer. During this interim stage of repeatability testing before application of response linearisation it was preferable to display integer values for sensor output, such that the data was either rounded to the nearest or rounded down. This rounding of data would lead to a certain artificial hysteresis effect.

At this stage repeatability was approximately within  $\pm 0.1$ mm, and over a 28mm displacement range would therefore contribute an uncertainty of 0.36% either way in the observed output. The approximate incremental resolution of the sensor has already been set at 0.03mm and therefore the effect of repeatability in terms of resolution is to give an uncertainty of  $\pm 3$ LSB (0.1mm either way is approximately 3 times 0.03mm either way).

## 5.3.2.5 Results from the fifth test series

By this stage a reliable sensor had been developed whose characteristics were approaching those of conventional devices. The fifth test series made use of realtime online temperature compensation based on functions developed in earlier test series. The effectiveness of the temperature correction functions are demonstrated in the steady state test results. In addition to real-time temperature compensation the non-linearity of the displacement response of the prototype sensor was characterised for the first time with a view to providing a conversion function to linearity.

5.3.2.5.1 Steady state test result (temperature compensation)

This test was primarily concerned with proving the effectiveness of the temperature correction facility developed off-line in previous test series.



A temperature correction function as described earlier was applied in real-time during a steady state test conducted over a four day period. Figure 5.34 indicates how the raw prototype sensor signal drifted freely during this test.



Figure 5.35 Real performance result of on-line temperature correction

Figure 5.35 shows the corresponding temperature compensated signal produced by the real-time correction function. By comparison of figure 5.34 and 5.35 it can be seen that drift has been entirely eliminated. The Y axis of both figures is on the same scale in that an increment on each has identical magnitude (however the correction function introduced an offset from 7000 range to 550 range) such that the prior drift of 28 incremental steps (figure 5.34) has been reduced to 3 incremental steps after correction, an improvement of approximately 9 times. The effect on overall error will become apparent as displacement range is discussed.

#### 5.3.2.5.2 Range test

The chart in figure 5.36 shows the nonlinearised displacement response of the prototype sensor.





It is clear that the response is not linear as noted previously. The chart also demonstrates that the prototype output range was 840 discrete incremental steps. With reference to temperature correction the real-time compensation function reduced drift error to less than  $\pm 0.2\%$  of the full scale.

The data shown in figure 5.36 was used as a reference set upon which to base a transfer function which would provide output proportional to displacement input. The method used in this instance was application of a polynomial fit. In order that the polynomial transfer function could be implicitly resolved to give linearised output based on non-linear input, it was necessary to design the function such that

the solution was numerical and not estimated (i.e., y should be the output and x the input of a polynomial function). To achieve this the axes of the chart in figure 5.36 were transposed and a polynomial fit performed on the resulting trace. This process is indicated in figure 5.37.





A curve fit was performed on the data set shown in figure 5.37. Figure 5.38 shows the function which provided the best fit.

Figure 5.38 Application of curve fit to data set



On production of the curve fit,  $6^{th}$  order polynomial, as indicated in figure 5.38 it was possible to apply this function to the data set used in its formation. This would give a practical impression of the performance of the function. The graph in figure 5.39 indicates how the  $6^{th}$  order polynomial would linearise the raw data.



On first inspection it can be seen that there is a clear improvement in linearity of response over the raw data set. However, as shown in figure 5.40, there remains an element of non-linearity which accounts for an error range of approximately 10 incremental output steps when compared to the ideal function.

Figure 5.40 Error remaining after polynomial correction



The solution to the remaining error is either to increase the polynomial order, or to consider other appropriate functions which might provide a better fit. These will be considered further in chapter 6. It was found that the best interim compromise was to employ an  $8^{th}$  order polynomial fit over the complete displacement range, this afforded a maximum non-linearity error of just ±3LSB as indicated in figure 5.41. The standard deviation of the error from the ideal straight line was 1.67.

Now non-linearity error sums with any drift error hence it was anticipated that total maximum error should be in the order of  $\pm 4.5$ LSB which equates to an uncertainty of 0.27mm.





5.3.2.5.3 Application of linearity correction in real-time

The 8<sup>th</sup> order polynomial was applied to the real-time data processing function within the host PC. The sensor input measurand was then cycled through its full range. The chart in figure 5.42 shows how the sensor response became visibly linearised.



Figure 5.42..Visible linearity of sensor response

The chart in figure 5.43 shows the remaining error as the processed data deviates from the ideal function. Figure 5.43 indicates that a significant error remains on the linearised output. This can be quantified as within  $\pm 3LSB$  or  $\pm 0.35\%$  full scale. It was important to know the stability of this system over time as various inputs such as temperature changed. One day later and then several days later the same experiment was performed.





The chart in figure 5.44 shows the three raw response curves (no temperature correction or linearisation) on the same graph. It may be observed that the shape of the response curves is similar, and that the small offset in y is due to temperature.

It may be seen from figure 5.44 that there is a small offset and slight curve variation between the first day's response and the two subsequent responses. The two subsequent responses were nearly identical. The oscillator temperature during these respective experiments was 28.4°C, 27.25°C, and 23.5°C, which would account for offset but not for shape variance (because the temperature effect has been observed as linear). The chart in figure 5.45 shows how the response curves compared after temperature correction.



The information in figure 5.45 is important because it shows how the linear temperature correction function is effective over the range of temperatures experienced during these tests.



Figure 5.45Temperature corrected response curves over several days

The chart in figure 5.46 shows how the same polynomial function was applied to each response curve (in real-time) and the resulting linearisation achieved. It is not entirely clear from figure 5.46 how the system has performed in terms of accuracy. To establish performance of the linearisation function a comparison was made to the chart shown in figure 5.43. The charts shown in figures 5.47 and 5.48 show how the linearised response curve deviated from the ideal function for the next day test and the test conducted several days later respectively.



Figure 5.46 Variance in linearising performance over a period of days

From figures 5.47 and 5.48 it is clear that devaince has increased since the polynomial coefficients were first generated, this has tended to introduce a nonlinearity as well as a slight gradient in response. The effect of this increasing deviance was to increase remaining non-linearity error to a typical  $\pm$ 4LSB. Whilst this error may be tolerable in terms of full scale displacement range, it is clear that the function combinations were not performing reliably and it was reasonable to suppose that given time (without recalibration) the error would continue to grow to intolerable levels.



Figure 5.47 Error after linearisation one day after function generation

Figure 5.48 Error after linearisation several days after function generation



It was found that the coefficient generation process is most sensitive and in order that a reliable functional description could be produced it was necessary to ensure that all the data used was valid (i.e., there should be no spikes within the sample data used for curve fitting). There was also an issue concerning the precision of the coefficients generated. It was found in general that low order polynomials did not require more than perhaps 4 decimal places of precision in their coefficients to work effectively (although low order polynomials did not give a good fit anyway). However, larger polynomials were found to be sensitive to 20 decimal place precision, and these were not favoured. The significance of mantissa precision is well illustrated in the series of charts shown in figure 5.49. From figure 5.49 it is clear that beyond 4 decimal places of precision there was little improvement in the linearised response. It is also clear that in this instance a  $6^{th}$  order polynomial curve fit would lead to considerable errors in the order of ±6LSB.



Figure 5.49 Series to show effect of numeric precision in polynomial curve fit

With reference to figure 5.47 and 5.49, it was found that the observed deviant trends were due to a technicality with temperature correction. It will be recalled that earlier experimentation prior to signal linearisation, involved recording of temperature corrected sensor output in an incremental manner. It was stated at this time that rounding of the sensor output to its nearest increment would introduce an error of  $\pm \frac{1}{2}$ LSB. Removal of this incrementation process would allow more accurate temperature corrected data to be presented to the linearisation function. The result of this correction is presented in the series of charts in figure 5.50.



Figure 5.50 series to show effective operation transfer functions

5.3.2.5.4 Application of metric calibration and tolerance setting

From figure 5.50 it is possible to observe the characteristic oscillation of a high order polynomial curve fit. It is also possible to observe that time based gradients have been eliminated from the linearised response. However, although standard deviation remains consistently at approximately 1, the deviance from average is seen to peak at approximately  $\pm 3LSB$  (accounting for an overall tendency to offset at -1). From figure 5.51 it can be observed that the incremental range of the fully condition (temperature corrected and linearised) output signal (prior to metric calibration) is 655.5. The input displacement range to achieve this output range was measured as 23.5mm (less than modelled due to the 4mm thickness of the modulator's coil). Therefore the displacement uncertainty introduced by 6

increments would be 0.22mm. As such it was possible to predict that the performance of the sensor would allow positional accuracy within  $\pm 0.11$ mm.



Having established a linear relationship between input measurand and output signal, the final step in setting up the prototype sensor for 'real' measurement performance testing was to apply a metric calibration. The calibration included both a functional conversion to displacement in millimetres and a definition of tolerance. The calibration to millimetres was achieved through a linear transfer of the form y=mx+c based on the known input displacement range and the known output signal range. From examination of the remaining deviance from the ideal function after a polynomial fit has been applied it was known that a maximum error range of 6LSB or 0.22mm existed. This being the case it was expected that the actual displacment of the input measurand could be as much as 0.11mm either side of the calibrated reading (in a worst case). Therefore it would be inappropriate to set a tolerance within 0.11mm and it was decided in the first instance to define the tolerance of repeatability and positional accuracy as  $\pm 0.125$ mm maximum. The calibrated output signal was rounded to the nearest 0.25mm accordingly. The chart in figure 5.52 shows the calibrated output signal (rounded to 0.25mm) plotted against the reference LVDT. The LVDT cannot give information on the accuracy of the prototype sensor but it does confirm the linearity. The observed quantisation in this figure is due to the rounding of the output signal to the nearest 0.25mm.


Figure 5.52 Response profile of calibrated prototype sensor

To determine the positional accuracy of the prototype sensor over its range it was necessary to make reference to the vernier scale mounted on the measurand drive. The positional accuracy test was conducted as repeatability testing. For each test a successful result occurred if the position achieved was within  $\pm 0.125$ mm of that measured by the prototype. The prototype was used to guide the input measurand drive to stop at absolute displacements of 15mm, 20mm, 25mm, and 30mm (the displacement range on the absolute scale was 9.7mm to 33.2mm). The tables in figure 5.53 indicate the success, in every case of the prototype to guide the measurand drive back to within  $\pm 0.125$ mm of the requested displacement. These results show how the sensor was able to reliably measure displacement within a tolerance of 0.25mm. The limiting factor appeared to be the oscillation of the output signal caused by the polynomial curve fit.

Ingui C Sist		Um campi	aicu repca	LADIMLY COL			
displacement 15mm		]		displaceme	ent 20mm		
actual		direction	success	ł	actual	direction	success
1	15.1	forward	lyes	1 1	19.9	reverse	lves
2	15	reverse	yes	1 2	20.1	forward	yes
3	15.1	forward	yes	] 3	19.9	reverse	yes
4	15	reverse	yes	] 4	20.1	forward	yes
5	15.1	forward	yes	5	19.9	reverse	yes
6	15	reverse	yes	6	20.1	forward	yes
7	15.1	forward	yes	7	19.9	reverse	yes
8	15	reverse	yes	8	20.1	forward	yes
9	15.1	forward	yes	9	19.9	reverse	yes
10	15	reverse	yes	10	20.1	forward	lyes

Figure 5.53 results from calibrated repeatibility test

displacement 25mm		J		displaceme	ent 30mm	<b></b>	
	actual	direction	success		actual	direction	success
1	25	reverse	yes_	1	30	reverse	yes
2	25	forward	yes	2	30	forward	yes
3	25	reverse	yes	3		reverse	yes
4	25	forward	yes	4	30	forward	yes
5	25	reverse	yes	5	30	reverse	yes
6	25	forward	yes	6	30	forward	yes
7	25	reverse	yes	7	30	reverse	yes
8	25	forward	yes	8	30	forward	yes
9	25	reverse	yes	9	30	reverse	yes
10	25	forward	yes	10	30	forward	yes

# 5.4 Performance Under Bending

Bending of the sensor during operation is considered as an 'input' liable to cause cross-talk. The objective of testing under bending conditions was to ascertain how the output signal was affected if all other inputs were stable or accounted for. The main difficulty in designing a suitable test method was both philosophical and practical. On the one-hand it was important to ensure the principal input measurand (i.e., core displacement) was held constant during bending. This was important to truly attribute any change of output to the cause of bending (this being a practical problem because bending of the device is liable to cause displacement of coil/core in its manifestation). On the other-hand it was realised that in 'real-world' implementation bending and axial displacement would inherently occur in unison so that testing under such controlled conditions would not reflect 'real-world' performance. It was supposed that a bending jig could be used to hold the sensor whilst its coil/core were variably displaced. However, under these conditions it would not be possible to use any conventional measuring devices such as an LVDT as a reference to observe cross-talk. At this stage it was suggested that rigorous analysis under bending conditions would only come from practical implementation at a later stage. As a compromise and to allow theoretical testing of the sensor under controlled conditions the core was held in fixed axial displacement relative to the coil whilst the whole device was manually bent in various directions. The output of the sensor was monitored during this aggressive test. Unfortunately it was not possible to quantify the applied loads and magnitude of bending at this time because equipment necessary was not

available (however chapter 6 describes an experimental process which was able to collect useful data on bending performance).

Figure 5.54 shows how the calibrated output signal of the prototype sensor did not appear to respond to mechanical bending, although there was a regular flip between two output states. Figure 5.55 is a plot of the raw sensor output in terms of oscillator increments within the sampling microcontroller. Here the effect of bending is much clearer, the signal can be made to 'drift' by 14 incremental steps in this case.



Figure 5.54 Bending had little effect on calibrated output signal

Figure 5.55 Bending had noticeable effect on raw output signal



Figure 5.55 shows how the sensor was aggressively manipulated for approximately 40 seconds. If the calibrated output as shown in figure 5.44 were set to a higher tolerance of perhaps 0.1mm range then the effect of bending might be more obvious. However, as already discussed the point of this experiment was academic and it remains to be seen how the sensor really performs in practical application.

### 5.5 Tests conducted

The tests conducted fall into two distinct groups. There was one group concerned with the stability of the sensor system over time. There was a second group concerned with performance in active use. These two groups are expanded in the sub-sections below.

### 5.5.1 Testing sensor system stability

It was critical to understand the consistency of performance over time, particularly with a view to identifying systematic variations distinct from un-attributable noise. Tests were designed to hold the system in a steady state so that variables were reliably fixed or else reliably monitored. The test was conducted so that regular samples were taken over a long period of time, in some cases up to 3 weeks. Chapter 5 details how systematic variations were identified and compensated for. The stability tests were not exhaustively conducted with tests for every variable combination because time would not allow this. A range of random tests were conducted, together with several tests at specific variable settings (for example with the minimum and maximum measurand settings).

### 5.5.2 Testing active performance of sensor system

A variety of tests were designed to optimise data obtained for characterisation of specific performance criteria. First and simplest was the determination of output range for a given input range. This test was conducted for different frequency sample times as discussed in section 4.4.2.5.1. The test involved a simple sweep of the input measurand (via the measurand drive) over its full range, to record both the reference LVDT reading and the prototype sensor reading. From this information it was possible to interpret information on displacement resolution including maximum resolution average resolution and instantaneous resolution.

Further to this it was possible to infer hysteresis between a rising input measurand and a falling input measurand. The second test was designed to expose error due to repeatability. A high level function of the user interface was designed specifically to allow repeatability testing. The process involved driving the input measurand to some random displacement within the valid range. The drive was then required to return to some specified displacement based only on the prototype sensor's series of readings. The difference between position achieved and position requested gave rise to a reading of repeatability. The absolute positions were measured both by the reference LVDT and by the operator using the vernier scale. It was possible to infer other characteristics such as overall error and accuracy from a considered combination of all test data obtained. The results of these tests are presented in chapter 5.

## 5.6 Conclusions

This chapter has presented the results of testing the prototype displacement sensor, which was developed in chapter 3. The scope of testing was good in that a good impression of performance in comparison to conventional devices can be made in terms of conventional characteristics in conventional circumstances. It is clear that the device has a high sensitivity coupled with low hysteresis, characteristics which allow reliable measurement to within one quarter of a millimetre. There is clearly potential to increase the performance of this device in terms of response profile and sensitivity, such that reliability of output approaches the micron level. The signal conditioning processes appear to be stable once correctly defined, although more time is required to characterise performance over months as opposed to days and weeks. With regard to performance under specialised conditions such as in this case mechanical bending, it was found that problems lay in the manner of reliable and meaningful testing rather than in the ability of the sensor to continue operating as desired. It was seen that bending does introduce a limited amount of output response although it was argued that this might be entirely expected and compatible with the real axial displacement of coil/core in practical application.

In summary, the prototype sensor was shown to have a working displacement range of approximately 23.5mm. Its output response to input measurand was found to be linear subject to oscillations introduced by a polynomial fit function and quantisation noise from the frequency capture method. The combined effect of these error contributors was to introduce an uncertainty of  $\pm 0.11$ mm about an ideal linear relationship between calibrated output and calibrated input. It was shown that the sensor reliably measures to the nearest 0.25mm. Amongst other things chapter 6 will discuss some alternatives to the use of a polynomial transfer function.

# 6 Refinements and Applications

#### 6.1 Introduction

Chapter 6 addresses some of the concerns raised during the course of previous chapters with regard to unexpected results. It then goes on to suggest how the prototype and the principle may be applied to both the specific context of this research and also generically. Some detail is given as to how a self-contained device might be developed to allow the realistic possibility of commercial use.

#### 6.2 Further modelling of coil inductance

It was observed in chapter 5 that the range of frequency output of the Colpitts oscillator within the prototype sensor was not as was predicted by the basic model presented in chapters 3 and 4. Although this observation did not greatly affect the ability to use the principle in displacement measurement, there may be scope to increase output range and hence increase overall sensitivity of the prototype sensor.

## 6.2.1 Refined parameter quantification

In the original model development, many parameters were as specified by manufacturers data sheets without regard for tolerance. It was supposed that any divergence from ideal to actual values of electrical parts might affect the Colpitts oscillator in particular. Further to this it was assumed that the powdered ferrite core used in the manufacture of a flexible tapered ferrite core might have a relative permeability of about 1000. It was additionally estimated that the relative permeability of the flexible ferrite core might be approximately 100. A permeability factor of one hundred would suggest a difference in extremes of coil inductance of one hundred as the relationship is proportional.

The coil/core inductance was measured using a specific instrument designed for this purpose. When the core was fully extracted such that the coil was considered to have an air core the inductance was measured at  $51.1\mu$ H. When the core was positioned such that the coil enclosed a maximum volume of ferrite material the inductance was measured at  $66.8\mu$ H. This compares with the originally modelled values of  $66.1\mu$ H minimum and 3.4mH maximum. The difference in range

magnitude was therefore an expected 51 times compared to an observed 1.3 times. This result suggests that it was not a good estimate to value relative magnetic permeability of the doped silicon rubber at 100. In experimentation the relative permeability of this compound material was inferred as approximately 1.3 (this value cannot be accurately inferred because there was always a quantity of air enclosed within the coil along with the doped silicone rubber owing to the tapered form).

In addition to refinement of inductance within the Colpitts oscillator it was determined useful to know the accurate values of capacitance used to complete the oscillator tank. Again a specific instrument was employed to measure component values as they were positioned within the sensor circuitry. It was found that the value of capacitance used during modelling  $(4.845e^{-9}F)$  was more accurately quantified as  $6.08e^{-9}F$ .

It was decided that in the first instance the inductance variable and the capacitance parameter were the two critical influences on modelling accuracy.

## 6.2.2 Equation refinement

It was stated in chapter 3 that the inductance approximation equation used for basic modelling was perhaps too approximate and would not provide a real enough comparison with experimental results. This has been found to be the case and as such some time was spent researching more accurate formulae for the determination of coil inductance.

As a good starting point Murgatroyd [88] provides an insight into the complexity of calculations for just one type of coil geometry. Points raised which the basic approximation equation has no capacity to account for are, multi-layer windings, high frequency effects, significance of coil insulator volume, and cross-sectional shape of the windings (i.e., do the windings form a perfect rectangular section, or is it more of a dome shape). Grover [89] provides a definitive compilation of many techniques applying to the specific case and to more general cases. From this text it was clear that much effort had been made to simplify the procedure for inductor design in terms of tabulating common combinations. In each case an air core coil was considered such that it would be possible to account for insertion of ferrite cores after the minimum inductance had been derived. The effect of tabulation and simplification into specific cases allowed the selection of a particular formula from Curtis [90] which applies to 'thick' coils (where the coil diameter is greater than the coil length):

$$L = 0.001 N^2 a P'$$
 (1)

where P' is some factor based on geometric properties of the coil, a is the geometric mean radius of multi-layer windings, and N is the number of turns. This equation is simple enough in the form presented as equation 1. However, it is not simple to compute the value of P'. Grover explains that P' is the product of two geometric factors P and F. P is a function of winding depth/diameter and F takes into account the loss of inductance due to winding separation, which is forced by conductor shape and insulator thickness. The computational method of arriving at values for these factors is as has been stated involved and case specific. Curtis [90] provides tables to simply look-up correct values for P and F given the presented geometric data.

With reference to the geometric definition given in chapter 3 the value of P was found to be 35.058, and the value of F was 0.55085. Therefore the working equation for inductance of the coil in question (with core fully extracted) was:

## $L = 0.001 N^2 a \times 19.311699$

Hence the minimum inductance of the coil was calculated to be  $24.1\mu$ H, and the maximum, with the tapered core fully inserted,  $31.3\mu$ H. This result remains small in comparison to the actual measured inductances by a factor of two. The numerical method was verified by employment of the companion equation for 'thin' coils as formed by Nagaoka [91]:

$$L = 0.019739 \left(\frac{2a}{b}\right) N^2 a K'$$
<sup>(2)</sup>

Where a is again the radius to mean centre of coil windings, and b is the thickness of the coil, K' is a geometric factor determined in a similar manner as for P' in (1).

It was subsequently found that the coil used in the prototype had erroneously been constructed with 75 turns as opposed to the specified 50 turns, this was due to a misunderstanding of the winding machine used. When the calculations were revised to this effect it was found that the expected range of inductance was 54.3 $\mu$ H minimum and 70.6 $\mu$ H maximum, this result being much closer to the actual recorded inductance range. The remaining difference between model and reality could be attributed to the slight non-uniformity in windings and the tendency for a dome towards the middle of the coil.

#### 6.2.3 Re-calculation of frequency range of sensor output

It was possible to reconsider the output range of the prototype sensor based on the revisions described in the preceding sub-sections. It will be recalled from chapter 3 that the frequency of oscillation for a Colpitts oscillator was given as:

$$f = \frac{1}{2\pi\sqrt{LC}}$$
(3)

We have revised values for both L and C such that the new modelled minimum output frequency should be 243KHz, and the new modelled maximum output frequency should be 278KHz. When compared to experimental results of 260KHz and 292KHz respectively it may be seen that an offset remains although the range is correct. It is perhaps sufficient to say at this time that remaining error might be due to for example capacitance within the 1m analogue cable connecting the coil to the remainder of the oscillator circuit.

## 6.3 Improving linearity

It was found during experimentation that the effect of a high order polynomial fit to achieve a linear output of the prototype sensor was introducing a small but limiting oscillation. It is proposed that this situation might be improved in either of two ways. The first involves alternative signal conditioning to the polynomial fit, the second involves modification to the sensing mechanism to reduce the requirement for signal processing.

## 6.3.1 Improved signal conditioning

It was suggested in chapter 3 that conventional software methods of achieving linearisation correction of a response profile were not very appropriate to microcontroller applications. For example a detailed look-up table would not easily fit into the limited memory resources of a microcontroller. For this reason it was necessary to attempt an efficient curve fit using a mathematical description of the raw data set. However, it has been observed that the application of a  $9^{th}$ order polynomial fit neither provides an efficient solution in terms of the quality of fit or in terms of the number and precision of coefficients which must be stored, or indeed, the large number of computation cycles which would be necessary to complete the transform. As such it would be beneficial to find a better mathematical description than a single polynomial fit over the entire displacement range. It was proposed that there might be feasible approaches, the first to nit multiple functions in series to provide successive descriptions of portions of the response profile, the second approach is to attempt a custom defined function which provides a better description than the generic polynomial fit.

#### 6.3.1.1 Multiple function description

It was noted that many regions of the raw response profile were smooth in gradient. It was suggested that a low order polynomial function might be able to give a good description of a particular region of the response profile, rather than using a single but high order polynomial to describe the entire profile. As such it was proposed that regions exhibiting particular gradient features could be assigned their own local function description. In this way the entire response profile would be covered by a sequential series of simple functions. Function selection would be based simply on the current operating output of the raw response signal (after temperature calibration). The general method of applying multiple functions is known as splining, and a commonly used form of splining is cubic spline interpolation [92]. Cubic spline interpolation forms cubic polynomial function descriptions between known data points. The cubic polynomial, being third order, is more efficient in coefficient and calculation terms than the higher order polynomials worked with in chapter 5. It was found however, that unless high order polynomials were employed no significant improvement over the previous method of curve fitting could be made. This is largely due to the level of uncertainty in the reference LVDT signal which at this stage of development in the project had comparable accuracy with the prototype sensor (i.e., it was not significantly more stable than the prototype sensor). Subsequent development of the project may consider an improved reference measurement system so that more refinement can be made in the area of curve fitting.

## 6.3.1.2 Custom function description

It was noted that the raw response curve was somewhat similar in shape to a tangent plot, therefore it was supposed that a function could be designed around this. Various tools were employed to investigate the suitability of forms of the tangent plot, these include Microsoft Excel, National Instruments Lab View, and Matlab. This work is ongoing and it cannot be determined as yet whether a more efficient single function is available to replace the current polynomial function. Current impressions are that the refining coefficients required to make a good fit are as many as when using a polynomial function.

## 6.3.2 Mechanical alteration

It was considered that there would be an advantage to reducing the requirement for software based signal processing if it were possible to improve the linearity of the raw output signal. The principle was to apply the functional correction to the shape of a subsequently manufactured tapered ferrite core.

### Figure 6.1 Use of non-linear tapered core to improve response profile

Straight taper produces non-linear response curve



Variable taper produces more-linear response curve



In this manner a straight taper would be characterised as in chapter 5, and the resulting curve fit (high order polynomial) would be used to define the taper shape of the next tapered core to be manufactured. Figure 6.1 illustrates this proposal. Application and testing of this principle are currently ongoing and it is too early to state whether or no significant improvements have been made.

### 6.4 Context specific application

It will be recalled from chapters 1 and 2 that the application in question was to profile the surface form of a worm-like underground vehicle. Having developed a suitable type of sensing device to assist in this requirement it was necessary to consider the exact manner in which multiple sensors would be arranged and controlled.

It was considered that there exist two feasible methods of employing networks of sensors such as that developed during this research. Each of the two methods would allow operation in a slightly different manner and each would have its benefits and its drawbacks. The following subsections attempt to clarify these two options with a view to demonstrating how the system, if implemented, might be expected to perform.

## 6.4.1 Simplistic surface form measurement by interpolation

As the title of this sub-section suggests the method presented here is considered simple, both in terms of implementation and in terms of the information which can be obtained. The method involves arrangement of displacement sensors within or on the surface material in circumferential and longitudinal layouts, as described in chapter 2 (figure 2.28). Assuming that one node of intersection between an end circumferential displacement sensor and a selected longitudinal displacement sensor is defined in 3D space as an origin (this point may be defined by position tracking systems as described in chapter 2), the 'lattice' of sensors then extends away from this node in three directions, along the longitude and in both directions around the circumference. If a single sensor were employed to measure the longitude at defined intervals around the circumference, then a difference in length between any of these sensor readings would imply bending. However, the limit would have no ability to differentiate between first and higher order bending

(i.e.,  $1^{st}$  order is a single bend in one direction,  $2^{nd}$  order is a double bend or 's' bend, see figure 6.2), and there would be a possibility of missing  $2^{nd}$  order and higher bending in this configuration.





If two or more longitudinal sensors were employed in series along the longitude at each interval around the circumference, then it would be possible to detect 1<sup>st</sup> and higher order bending, the relationship in general terms being that one additional bend can be detected for every sensor in series. The limitation on this approach is that only one bend per sensor may be detected. This means that if two longitudinal sensors in series were exposed to no bending whilst a third sensor was exposed to a double bend, then the system would be limited to detecting just a 1<sup>st</sup> order bend in the 3<sup>rd</sup> segment. The issue becomes one of bend resolution proportional to density or length of displacement sensor. As such it would be necessary to determine the minimum bend which is of significance. Therefore resolution of bend detection is ultimately determined by displacement sensor size rather than by the number of sensors present.

A change in length of one or more series longitudinal displacement sensors can imply one or both of two things, extension, and/or bend in the surface. In order that this information can be determined (including the direction of a bend) it is necessary to consider displacement data in comparison with data from other series longitudinal displacement sensors around the circumference. For example, suppose that longitudinal sensors were positioned around the circumference such that each had a twin exactly 180° around the axis. In simplistic circumstances it can be imagined that any load or condition effecting a bend in one sensor will also effect a corresponding bend in its twin on the other side of the body (this is simplistic in that we imagine a uniform bend over the whole cross-section rather than a dent in one side). If the condition had been a mere extension or contraction of the segment then the sensor pair would be expected to record displacement changes of equal proportion. If as suggested a uniform bend is applied then one sensor would be expected to increase its displacement measurement relative to its pair. The sensor whose measurement was lower would be closest to the origin of the bend and so direction of bend could be inferred. Further to this, for the simple case, additional longitudinal sensors around the circumference could be expected to record a gradual transition from the two extremes recorded by the sensor pair in-plane with the bend. As such it would be possible to determine the plane of the bend around the longitudinal axis of the body.

The circumferential displacement sensors would be evenly distributed along the length of the body. In the simplest case one sensor would be employed to measure the entire circumference at the given point along the longitudinal length of the body. The information returned from the sensor would allow only to infer sectional perimeter length and not additional information such as uniformity of sectional form.

In this configuration it would be possible to determine the body form of the vehicle in terms of segment length, diameter, and bend position, size, and orientation. The degree of resolution would be directly proportional to the number of longitudinal displacement sensors connected in series along the length of the body.

## 6.4.2 Complex surface form measurement by 3D vector calculation

The complex alternative to simple displacement measurement aims to treat the body surface as a thin sheet, therefore affording it a  $3^{rd}$  dimension of thickness. A geometric arrangement of theoretical nodes are placed on each side of the surface. Displacement sensors are connected between nodes to form a triangulated structure as briefly discussed in chapter 3. The principle of operation is that from a single reference plane and origin at one end of the body it is possible to

determine the position in 3D space of every other node by 3D trigonometry as long as the distance between nodes is known from the displacement sensors.

The benefit of this complex system is that, depending on the density of nodes more complex surface information can be provided such as localised deformation (dents), and complex sectional forms. The drawbacks are: the additional complexity of implementation, the requirement for many more sensors than in the simple network arrangement, and the summation of errors as distance increases from the origin. It is also not yet known whether or not such detailed surface information would be required by the project described in chapter 2.

# 6.5 Further tests and results on bending

A relatively simple bend test was designed about a cantilever sliding bed. The mechanical implementation involved an add-on modification to the original test rig as described in chapter 4. The basic design was as indicated in the figure 6.3, figure 6.4 show photographs of the rig in operation. The sliding bed was driven by a rigid member linked to the drive of the main test rig. Varying degrees of bend were applied to the modified bed by adjusting the deflection load on the end of the cantilever.



#### Figure 6.3 Schematic representation of bending test rig

Figure 6.4 photographs of the bending rig in operation



It was accepted that a major limitation of this rig was that bending was not along the axis of the prototype sensor, it was along the central axis of the cantilever. This geometry meant that a bend in the cantilever member resulted in a bend and extension in the prototype sensor's axis. Thus the effect of bending in the prototype sensor output was always accompanied by the effect of an extension. It would in principle be possible to filter the extension effect of bending when the system is explained in two dimensional geometry as in figure 6.5.

#### Figure 6.5 Relevant geometry of bending rig



Figure 6.5 demonstrates that the cantilever and mounted sensor are taken to be uniform arcs such that arc length l is defined by the product of radius r and angle subtended  $\theta$ . Now arc length l is known to be relatively constant (ignoring temperature effects etc) and either angle or radius are measured/estimated to give a working value of angle subtended. The arc on which the prototype sensor axis lies has radius d greater than r and d is known and constant. Therefore the length of this arc pl may be calculated in the same way as l. The extension factor in the axis pl is simply pl/l.

In practise it was observed that the cantilever bending profile was not a perfect arc and it was not easy to estimate the angle subtended during bending. These difficulties could be overcome by design consideration in future work. However, it was noted that it was not necessary to compensate for extension offset as the data obtained would hold most useful information in the level of linearity retained in the working output range of the prototype sensor. This is demonstrated by figure 6.7 which shows a plot of displacement response linearity under zero bending and then under an unquantified but constant bending. Figure 6.6 shows the sensor output during the test period from which the data in figure 6.7 is taken.



Figure 6.6 Sensor output during bending tests

Figure 6.7 is an early indication that little crosstalk is introduced when the prototype sensor is bent through an estimated angle 10° from the horizontal. It may be seen that some deviation occurs under bending at the extremes of range although this might be an effect of the experimental design in that the ratio of prototype sensor displacement to LVDT displacement increases as bending increases (because of the extension symptom under bending). This phenomenon most notably introduces the offset observed in figure 6.7 and 6.6, however, towards the extremes of range where the raw response begins to loose sensitivity

(and hence becomes increasingly non-linear) the linearisation function (9<sup>th</sup> order polynomial in this instance) is very sensitive to deviation from the response curve it was designed to work for. Hence the non-linearity is observed.





Further work is required both in the design of a suitable experimental apparatus and in the design of an appropriate experimental procedure. It must be recognised that laboratory experimentation can only provide so much information about the properties of a device or system. It would be most advantageous to rapidly place this prototype sensor in a system for which it was developed so that its limitations can be measured in context.

## 6.6 Development of a commercial device

This section does not aim to produce detailed designs for a commercial form of the prototype, which has been developed during this thesis. The intention is to clarify how it is feasible to produce such a device in a reasonably simple manner. It will be shown how it may contain all mechanical and electrical/electronic parts within one body. This feature enhances simplicity in use, offering to the external environment two electrical connections for power and a further electrical connection or connections for digital communication. Connections for communications would depend on the standard used, e.g., I<sup>2</sup>C and CAN Bus would require just one data line, RS454 and RS232 would require 2 data lines.

## 6.6.1 Mechanical design

There is scope for many variations on the mechanical form of the device. This is particularly due to the proposed production method of casting rubber/rubber-like compounds, and the tolerant nature of the 'active' parts embedded within. It would of-course be necessary for some further consideration of suitable materials as this work has not focused on the properties of the encapsulating medium (in this case catalyst curing silicone rubber compound). However, apart from materials it is only necessary to consider the housing of electrical parts and the displacement of the core/coil interface. As such the current mechanical design as employed for production of the prototype sensor would be a suitable basis.

### 6.6.2 Electrical design

During development of the prototype, the electrical components of the sensor (taken to include: coil, Colpitts oscillator, frequency capture components of microcontroller, signal processing components of microcontroller, communications components of microcontroller, and signal processing components of host PC), were distributed both electrically and physically. The main challenge of commercial development will be to integrate all these functions, both mechanical and electrical, and software, into a confined space within the body of the sensor.

There are various options open to the developer in addressing what amounts to a hybrid technology problem (hybrid is due to the analogue and digital aspect of the electrical system). It may be possible for mass-production to make use of a custom ASIC which would encompass both Colpitts oscillator and microprocessor in a single package, although it is unlikely that this functionality currently exists. It is more reasonable to suppose that the Colpitts oscillator will remain a system constructed from discrete surface mount components and that a separate microcontroller will be employed in a similar manner to that of the prototype development, except that it will be wholly devoted to servicing the device it is mounted within. The microcontroller would be required to capture frequency data from the Colpitts oscillator, apply all required signal conditioning (this would include a temperature sample and compensation), store captured data,

communicate with a host device to upload data and download configuration instructions.

There are microcontroller devices in production with small footprints and high functionality. An unusual requirement in this instance is to measure temperature directly. One device was found, from Arizona Microchip, which has capability to sample its own temperature to 16 bit resolution. This device also includes a communications port and capture facility. Reasonable programme and data memory is available to allow signal processing. A complete electrical system based on this design might look as show in figure 6.8.



· • • • •

OSC:	Crystal oscillator module
COLPITTS OSC:	Circuitry for Colpitts Oscillator
V+:	Input terminal for supply voltage
GND:	Input terminal for supply ground
DAT:	Input terminal for serial data communications
PROG:	Programme memory
REG:	Data registers
CAP:	Capture register and logic
TMR:	Counter timer register and logic
TMP:	Temperature measurement peripheral
SPI:	Serial port interface peripheral
PROC:	Main processor

Figure 6.9 Mechanical forms of integrated sensor



It is supposed that it would be possible to mount this circuit entirely on a two sided surface mount printed circuit board of area no greater than  $2cm^2$ . This estimation is based on the dimensions of the largest components of the system:

- Microcontroller P14C000: 10 x 8 x 2 (mm)
- Op-Amp LM1624: 4 x 4 x 2 (mm)

The printed circuit board, along with the modulator coil, is the limiting factor on mechanical flexibility for a wholly integrated system. It would be for the ingenuity of the designer to make best use of these limitations in specific circumstances (i.e., to lie rigid surface objects normal to the plane of greatest distortion). Some options for mechanical arrangement of electrical components are shown in the figure 6.9.

# 6.7 Other applications

It is expected that a technology such as that developed through this thesis might find alternative applications, due to its novel characteristics. The principle can be used to measure the perimeter length of irregular bodies and curved surfaces. The principle can also be applied to both incremental and absolute encoding of continuous systems such as belt conveyors.

# 6.8 Conclusion

Chapter 6 has been used to explain and resolve some of the discrepancies found between results expected from models and those achieved during experimentation. Some discussion has been made on the potential applications of this technology and how a tangible and robust product might be formed. An experimental procedure to measure performance under synthesised bending conditions is described and favourable first results are presented. Chapter 7, the final chapter, is mostly concerned with on-going and future work in the development of this flexible displacement sensor. However, there is also a summary of the state of the tunnelling robot concept and suggestions as to which areas of this research would be ideal for further investigation.

# 7 Conclusions

## 7.1 Summary of thesis

The initial investigation conducted in this research was concerned with the efficiency of accessing buried infrastructure. A very conceptual and technically challenging solution was proposed and it was realised that it would be most appropriate to target one specific area for in-depth research. Sensing was selected as an area of greatest interest and a novel approach was taken to the measurement of complex surfaces. A prototype sensor was developed and tested.

#### 7.2 Achievements

The following is a list of accomplishments which contribute both to the author's experience and knowledge, and to the general body of knowledge.

- A novel displacement sensor has been developed from first principles. In academic terms this development has benefited general knowledge as a demonstration of technological integration. Integration is specific to the use of embedded microcontroller technology to enhance the performance of simple and otherwise unsuitable sensing mechanisms.
- In terms of practical application to the class of sensors, a new category of device can be established within the field of displacement measurement unique to mechanical flexibility and elasticity of both the sensor body and the relationship of the points measured.
- This thesis contains the starting point for subsequent development and research, both academic and industrial in the fields of minimal intrusion repair, and contact measurement of complex surfaces. The prototype developed for testing in this work is crude when compared to the degree of engineering applied to the development of a commercial displacement sensor. However, despite these limitations, performance over a 30mm input displacement range resulted in a device having incremental resolution to 30 microns (worst case), a guaranteed repeatability of ±0.25mm, a zero temperature coefficient, and overall accuracy error within ±1% of the full displacement range. The potential performance increase might be a factor of 10 or greater after further refinement.

• The author has designed and manufactured in person, every aspect of the project components. This includes:

The prototype sensor, the mould tools to manufacture the prototype sensor, the test rig, the systems to drive the test rig in a reliable and safe manner, the software to control the mechanical automation of the test rig, and the electronic hardware and software to provide real-time control over sensor sampling and axis drive, the various interfaces between different elements of the system the PC based user interface, the microcontroller based real-time controller, output to the drive amplifier, input from the reference LVDT, input from the various temperature sensors, and input from the prototype sensor.

• Content from this thesis has been published in Mechatronics Journal (see appendix D). It is expected that further publications will arise directly from the work presented in this thesis.

# 7.3 Suggested further work

This thesis has probably presented many more problems than it has solved and there would be starting points for engineers of many varied disciplines within the early chapters. The technical challenges are many, some of the core areas of development for a prospective researcher might include:

- Tunneling vehicle topography and locomotion
- Underground navigation and location
- Ground cutting and spoil handling
- Automated/teleoperated repair of infrastructure

Each of these core areas might have its own network of sub topics like for example, 'vehicle topography and locomotion' includes the area of form sensing, which has been partially addressed in this thesis.

A prospective researcher may choose to follow up the more specific work of this thesis. There is considerable engineering and philosophy required in the next stage of sensor development – to begin practical application of a small sensor mesh on an elastic surface. There is the issue of how to control and extract data efficiently from multiple sensors, and there is the issue of refined mechanical form of the sensor and its constituent materials.

It may also be the ambition of a researcher to apply the sensing concepts presented here to some other field of science or engineering. It is possible for the author to only imagine a few potential applications of the flexible displacement sensor, these might be: measurement of shell structures in civil engineering as an industrial example, and measurement of test surfaces in a wind tunnel as a scientific example. However, the technology has been shown to be very generic and adaptable to different forms, and applicable to displacement measurement over a wide range of input. These attributes readily lend themselves to a system which is not just mechanically flexible but application flexible too.

# 8 References

[1] Byatt I. C. R., Annual Report on Leakage 1997. OFWAT publication

[2] Heim P. M., Conducting a Leak Detection Search. Distribution Systems – Actions and Innovations. 1980.

[3] Kingston W. L., A do-it-yourself Leak Survey Benefit-Cost Study. Distribution Systems – Actions and Innovations. 1980.

[4] General Excavating, 6701 Cornhusker Hwy, Lincoln. NE 68507 – 3113. www.generalexcavating.com

[5] Slabach Enterprises 'VacStar', PO BOX 814, Kalona 1A 52247. www.vacstar.com

[6] Vacmasters of Bourne, 5879 West 58<sup>th</sup> Avenue, Arvada, Colorado 80002. www.vacmasters.com

[7] Iseley T., Trenchless Excavation Construction Methods – Classification
 And Evaluation. Journal of Construction Engineering and Management ASCE.
 1992.

[8] Evin C., et al., Planning the Rehabilitation of Water Distribution Systems -Principal Document. *Water Research Centre*. 1989.

[9] Akkerman, <u>www.akkerman.com</u>

[10] Cement Mortar Lining Company Inc, 5 Greenway Plazza, Suite 1775, Houston, Texas, 77046. <u>www.cementlining.com</u>

[11] W. Walsh Company Inc, 32 Walton Street Attleboro, MA 02705. www.wwalsh.com

[12] CASE Construction Equipment. www.jicase.com

[13] John Deere Commerical Worksite Products. <u>www.deere.com</u>

[14] Premiere Pipe Services. <u>www.premierepipe.com</u>

[15] Pipelining Products Inc, 251 West Chatham Street, Cary, NC 27511. www.pipelining.com

[16] Byatt I. C. R., Speech given at the 1997 Water Summit. OFWAT report.

[17] T. D. Williamson Inc. T. D. Williamson (UK) Ltd, Faraday Road, Dorcan Way, Swindon, Wiltshire SN3 5HF. <u>www.tdwilliamson.com</u>

[18] Gonen T., Electric Power Transmission System Engineering. 1988.

[19] Dudeney, W, L. "Robotic Tunnelling with Biological Inspiration" CCAM, 1997.

[20] Dudeney, W, L. "Mechanical Earthworm to facilitate the access and repair of buried infrastructure" Conference proceedings, Mechatronics '98, 1998.

[21] Nichols, D. "The Oxford Book Of Invertebrates" 1971.

[22] Hegner, W. "Invertebrate Zoology" 1968.

[23] Herrenknecht Tunnelling Systems (Microtunnelling laser alignment). Schlehenweg 2 D-77963 Schwanan. <u>www.herrenknecht.com</u>

[24] Geo Services International (UK) Ltd. Holden House 26 Bridge Streey Witney Oxon OX8 6HY. <u>www.geophysics.co.uk</u>

[25] Katchalsky. Experientia, Vol 5 pp319-320, 1949.

[26] Khun. Experientia, Vol 5 pp318-319, 1949.

[27] Solari, M. "Evaluation of the Mechanical Properties of a Hydrogel Fibre in the Development of a Polymeric Actuator" Journal of Intelligent Material Systems and Structures, Vol 5, 1994.

[28] Flory, P, J. "Principles of Polymer Chemistry"1967.

[29] Shahinpoor, M. "Ionic Polymeric Gels as Artificial Muscles for Robotic and Medical Applications" Iranian Journal of Science and Technology, Vol 20 No 1 1996.

[30] Brock, D. "A Dynamic Model of a Linear Actuator Based on Polymeric Hydrogel" Journal of Intelligent Material Systems and Structures, Vol 5 1994.

[31] Otero, T, F. "Electrochemomechanical properties from a bilayer: polypyrrole/non-conducting and flexible material – artificial" Journal of Electroanalytical Chemistry 1992.

[32] Kanteto, K. "Artificial muscle: Electromechanical actuators using polyaniline films" Synthetic Metals 1995.

[33] Otero, T, F. "Artificial Muscles from Bilayer Structures" Synthetic Muscles 1993.

[34] Baughman, R, H. "Conducting Polymer artificial Muscles" Synthetic Metals 1996.

[35] Hill G. "Chemistry in Context" 1989.

[36] Neagu, C, R. "An Electrochemical Microactuator: Principle and First Results" Journal of Microelectromechanical Systems, Vol 5 No 1 1996.

[37] Robson, M. "The Nature of Electrolyte Solutions" 1988.

[38] Janocha, H. "Neue Aktoren" Actuator Proceedings 1988.

[39] Kempe, W. "Electrochemical Actuators" Actuator Proceedings 1990.

[40] NAFION is a E.I. du Pont de Nemours & Co Inc registered trademark, catalogue no 27 pp470-474.

[41] Friend, C. M., Morgan, N. B. Shape-memory alloys in medicine – from new prosthetic devices to 'cuddly' instruments. "*Medical Applications for Shape-Memory Alloys (SMA)*". IMechE Seminar Publication1999.

[42] Waram. "Design Principles For Ni-Ti Actuators", Engineering aspects of SMA,1990.

[43] Tautzenberger. "Thermal Actuators: A comparison of SMA with Thermostatic Bimetals and Wax Actuators", Engineering aspects of SMA, 1990.

[44] Osada, Y. "Biologically Inspired Robots" 1993.

[45] Brindley, K., Sensors and Transducers. 1988.

[46] Morris, A. S., Principles of Measurement and Instrumentation. 1993.

[47] Boyle, H. B., Transducer Handbook. 1992.

[48] Smith, E., Principles of Industrial Measurement for Control Applications.1984.

[49] Norton, H. N., Handbook of Transducers for Electronic Measuring Systems. 1969.

[50] Ruocco, S. R., Robot Sensors and Transducers. 1987.

[51] Stadler, W., Analytical Robotics and Mechatronics. 1995.

[52] Tse, F. S., Measurement and Instrumentation in Engineering. 1989.

[53] Hansman, J. R., Characteristics of Instrumentation. "The Measurement, Instrumentation, and Sensors Handbook." 1999.

[54] Elgar, P., Sensors for Measurement and Control. 1998.

[55] Turner, J. D., Instrumentation for Engineers. 1988.

[56] Ristic, L., Sensor Technology and Devices. 1994.

[57] Usher, M. J., Sensors and Transducers. 1996.

[58] Harris, R., Progress in Radar Remote Sensing of Vegetation and Hydrology. 1984.

[59] Munday, T. J., A Review with Particular Emphasis on Applications in the Geological Sciences and Geomorphology. 1984.

[60] Jay Zwallag, H., Surface Elevation Contours of Greenland and Antarctic Ice Sheets. *Journal of Geophysics Research* 88(c3) 1983.

[61] Allan T. D., An eye on Ocean Waves and Currents. *Physics Bulletin* No35 1984.

[62] Gracey, W., Measurement of Aircraft Speed and Altitude. *NASA reference publication 1046* 1980.

[63] Watton, W. L., A direct-Capacitance Aircraft Altimeter. *Proceedings of the Institute of Electrical Engineers* V96 pt3 1949.

[64] Renishaw Group Profile Brochure. 1997. <u>www.renishaw.co.uk</u>

[65] Ossana, P. H., Form Measurements a Bridge Between Production Metrology and Biomechanics. *International Journal of Machine Tools and Manufacture* v35 n2 1995.

[66] Ip, W. L. R., Range Image Measurement and Shaded Image Measurement: two Non-contact Coordinate Measurement Approaches Designed for Special Tasks. *Proceedings of the Institute of Mechanical Engineers* v213 Part B 1999.

[67] Besl, P. J., Active Optical Range Imaging Sensors. *Machine Vision and Applications* n1 1988.

[68] Ip, W. L. R., Shaded Image Machining, a reverse engineering approach using machine vision techniques and analytical methods to reproduce sculptured surfaces. *International Journal of Production Research* 34(7) 1996.

[69] Ip, W. L. R., Enhanced Procedures to Model, Machine and Measure Free Form Surfaces Using Analytical and Artificial Intelligence Techniques. PhD dissertation, University of Birmingham, 1993.

[70] Stout, K. J., Application of 3-D Topography to Bio-Engineering. International Journal of Machine Tools and Manufacturing. V35 N2 1995.

[71] Chen, F., Overview of three-dimensional shape measurement using optical methods, *Opt. Eng.* 39(1) 10-22 2000.

[72] Moring, I., Active 3D vision system for automatic model-based shape inspection, *Opt. Lasers Eng.* 10, 3-4 1989.

[73] Ji, Z., Design of triangulation devices, *Opt. Lasers Tech.* 21(5), 335-338 1991.

[74] Takasaki, H., Moiré topography, Appl. Opt. 9, 1467-1472 1970.

[75] Wang, J.Y., Imaging Laser Radar – an overview, Proc. 9<sup>th</sup> intl. Conf. Laser
'86, pp. 19-29 1986.

[76] Haines, K. A., Contour generation by wavefront construction, *Phys. Lett.*19, 10-11 1965.

203

[77] Wester-Ebbinghaus, W., Analytics in non-topographic photogrametry, *ISPRS Cong., Com. V., Kyoto* 27(B11), 380-390 1988.

[78] Schertenleib, W., Measurement of Structures (surfaces) utilising the SMART 310 laser tracking system, *optical 3-D measurement techniques III*, Wichmann 1995.

[79] Coggrave, C. R., Optimization of a shape measurement system based on spatial light modulators, *Opt. Eng.* 39(1) 91-98 2000

[80] Gasson, P. C., Geometry of Spatial Forms. 1983.

[81] Nelkon, M., Advanced Level Physics. 1985.

[82] Nelkon, M., Advanced Level Magnetism and Electricity. 1966.

[83] Jiles, D., Introduction to Magnetism and Magnetic Materials. 1990.

[84] Wobschall, D., Circuit design for Instrumentation. 1979.

[85] Cirovic, M. M., Electronic Devices, Circuits, and Systems. 1987.

[86] Hicks, C. R. Fundamental Concepts in the Design of Experiments. 1973.

[87] Kempthorne, O. Design and Analysis of Experiments. 1960.

[88] Murgatroyd, P. N., The Brooks Inductor: a study of optimal solenoid cross-sections. IEE Proceedings Vol.133 Pt.B No 5 1986.

[89] Grover, F. W., Inductance Calculations. 1946.

[90] Curtis, H. L., Research Paper 685. B. of S. Journal of Research 12, 665 1934.

[91] Nagaoka, Science Papers of Tokyo Institute. Vol.II p1 p75 1927

[92] Etter, D. M., Fortran 77 With Numerical Methods. 1992.

#### Appendix A

This appendix details mechanics and electrics specif to the manufacture of the prototype sensor. The first part is a series of engineering drawings to produce the moulds for forming the silicone rubber based prototype. The second part indicates the Colpitts circuit's used to produce TTL compatible pulse outputs



A2



A3



A4



. . . . . . . . .
#### Working drawings



A6



A7





A9



#### Electrical diagrams



## Op-Amp based version of Colpit ts Oscillator



Configuration C1 > C 2  $1 < \frac{C1}{C2} < 2$   $\frac{C1}{C2} < \frac{Rf}{R1}$   $\frac{Rf}{R1} \le 2$   $C = \frac{C1.C 2}{C1 + C2}$   $f_{osc} = \frac{1}{2\pi \sqrt{LC}}$   $@ f_{osc} : Cc \Rightarrow 0Z$ 

## Appendix B

This appendix includes details of the test rig as developed through chapter 4. The main mechanical elements of the rig include the measurand drive and the motor gear box. Also included is the real-time code used in the microcontroller

### Drawings to produce motor gearbox.















# Un-annotated drawings of measurand drive.





.







Microcontroller code

/\* Data acquisition module \*/ /\* Version 1 Initial Code W L Dudeney 8/1/1999 \*/ /\* W L Dudeney 26/1/1999 Version 1.1 Major debug \*/ Version 2 /\* W L Dudeney 29/1/1999 Serviceable code \*/ /\* W L Dudeney 10/2/1999 version 2.1 port from emulator to ic \*/ /\* W L Dudeney 17/2/1999 version 2.11 minor alterations \*/ /\* W L Dudeney 15/3/1999 version 2.5 i2c temp sensing \*/ /\* W L Dudeney 30/3/1999 version 2.51 higher resolution temp sensing \*/ 15/6/1999 /\* W L Dudeney version 2.6 Binarisation of data comms \*/ 15/6/1999 version 2.7 /\* W L Dudeney ergonomic/interface debug \*/ // This code links the P17c756 micrcontroller to a host PC LabView // interface via an RS232 serial interface. Operation if full duplex // asynchronous. // Useable version. Modification sample\_rate becomes 16bit to allow // sample periods longer than 4 seconds - new max period approx 18minutes #include <P17C756.H> #include <STDDEF.H> #include <USART16.H> #include <delays.h> #include <timers16.h> #include <adc16.h> #include <stdlib.h> #include <captur16.h> #include <string.h> #include <i2c16.h> /\* constants \*/ #define no inst 6 #define temp device 1 0x99 #define temp device 2 0x9B #define temp\_device\_3\_write 0x96 #define temp\_device\_3\_read 0x97 #define temp device 4 write 0x9E #define temp\_device\_4\_read 0x9F #define temp\_device\_5\_write  $0 \times 92$ #define temp device-5 read 0x93 /\* global variables \*/ static char rec\_buff[16]; static char \*rec buff ptr;

```
static char header chk, total exp, already relieved=1,
limit switches=0;
static char header refs[no inst][3] =
{"chk", "stp", "mot", "lvd", "dat", "sam"};
static char desired direction, desired energise;
static char direction, energise, slowing, timing, reversing=0;
static unsigned short int TMR0_preset, speed, desired_speed store,
desired speed;
static char *sbr_ptr, *sbw_ptr, *sbreset_ptr;
static char snd_buff[5][20], message_count = 0,current_read_message=0;
static unsigned char LVDT request=0, sensor capture=0,
current write message=0;
static unsigned int sensor reading=0, LVDT reading = 0;
static unsigned char there has been trouble, image of PORTB,
bint count=0;
static unsigned char PIE1 store, PIE2 store, INTSTA store,
done already=0;
static unsigned char WREG store, FSR0 store, FSR1 store, PRODL store,
PRODH store;
static unsigned int sample_rate, T1_overflow;
static int temp sensor 1, temp sensor 2, temp sensor 3;
static unsigned char temp_byte_high, temp_byte_low,
sampling complete=0, temp, local counter;
static unsigned char capturecounter, freqcapstrt;
/* function prototypes */
void initialise(void);
void chk data in(void);
void chk header(void);
void process rec buff(void);
void reply to chk(void);
void reply to stp(void);
void reply_as_bad(void);
void update_motor(void);
void chk motor speed(void);
void get_LVDT_sig(void);
void config sample rate(void);
void send_data_to_snd_buff(void);
void send_trouble_to_snd_buff(void);
void check message box(void);
void relieve microswitch(void);
void stop_relieving_microswitches(void);
void get_temperatures_only_for_LM75_chip(void);
void get temperatures(void);
void configure_temp_device(unsigned char);
void get device temp(unsigned char);
void frequency_capture(void);
//Interrupt handlers go here
void default handler(void)
Ł
      return;
}
void peripheral ints(void)
```

// this function handles all peripheral interrupts. // a case statement is executed to determine which interrupt occurred. // first house keeping - store critical registers: WREG store=WREG; FSR0\_store=FSR0; FSR1\_store=FSR1; PRODL store=PRODL; PRODH store=PRODH; if ((PIR2bits.CA3IF==1)&&(PIE2bits.CA3IE)) // Timer 3 (capture)? ſ WriteTimer3(60); sensor reading=ReadCapture3(); capturecounter++; PIE2bits.CA3IE=0; PIR2bits.CA3IF=0; } if ((PIR2bits.CA3IF==1)&&(PIE2bits.CA3IE)) // Timer 3 (capture)? 11 11 £ 11 if (capturecounter++==1) //2 caps but drop first 11 Ł 11 sensor reading=ReadCapture3(); //transfer reading sensor capture=0; //update flag to say process is 17 complete //then disable interrupt til next time 11 11 PIE2bits.CA3IE=0; 11 PIR2bits.CA3IF=0; 11 PIE1bits.RC1IE=1; 11 PIE1bits.TMR1IE=1; 11 } 11 //reset capture for accurate cycle else 11 Ł 11 WriteTimer3(4); //offset to compensate for service delay 11 PIR2bits.CA3IF=0; 11 } 11 ł 11 if ((PIR2bits.CA3IF==1)&&(PIE2bits.CA3IE)) // Timer 3 (capture)? 11 { PORTCbits.RC0=!PORTCbits.RC0; //DEVELOPOMENT FLAG 11 11 WriteTimer3(0); // when a capture occurs it is critical to service it asap 11 // the capture input source must be disabled to ensure the next // capture is accurate. 11 //PORTDbits.RD7=!PORTDbits.RD7; //disable the capture input source //CloseTimer3(); 111 //CloseCapture3();

{

```
11
            sensor reading=ReadCapture3(); //transfer reading
11
            PIR2bits.CA3IF=0; // manual clear of capture int
11
            sensor capture=0; //update flag to say process is complete
11
            // disable capture interrupt until next sample period
11
            PIE2bits.CA3IE=0;
            // restore interrupt enable status's
11
            PIE1bits.RC1IE=1;
11
            PIE1bits.TMR1IE=1;
            //PIE1=PIE1_store;
//PIE2=PIE2_store;
11
            //PIE2bits.RC2IE=1;
            //INTSTA=INTSTA||INTSTA store;
            //done already=0;
11
     }
      // check for emergency stop request
      if(PIR1bits.RBIF==1)
      Ł
            PORTCbits.RC1=!PORTCbits.RC1; //DEVELOPOMENT FLAG
            //if(PORTB!=0xF8)
            if((PORTB|0b11111000)!=0xF8)
            {
                  PORTDbits.RD0=0; // motor off
                  energise=0;
                  desired energise=0;
                  desired_speed=0;
                  speed=0;
                  CloseTimerO(); // motor won't restart automatically
            }
                  CloseTimer1();
            11
            11
                  CloseTimer3();
            11
                  CloseCapture3(); // stop sampling as well
            image of PORTB=PORTB;
            PORTB=image of PORTB;
                                   //write PORTB to clear int
            there has been trouble=1;
            PIR1bits.RBIF=!PIR1bits.RBIF; // clear interrupt request
            bint_count=bint_count+1;
            PIE1bits.RBIE=0; // must disable to stop effect of bounce
      if(PORTBbits.RB1||PORTBbits.RB2) {limit_switches=1;reversing=0;}
            else limit switches=0;
      }
      if(PIR1bits.RC1IF==1)
      Ł
            PORTCbits.RC2=!PORTCbits.RC2; //DEVELOPOMENT FLAG
            // if bytes have arrived at USART1
            *rec buff_ptr=getcUSART1();
```

```
rec buff ptr=rec buff ptr+1;
            PIRIbits.RC1IF=0;
      }
      if (PIR1bits.TMR1IF==1) // Timer 1?
      {
            // on request for LVDT measurement T1 interrupt initiates
an ADC
            // conversion and flags the main programme to poll for
completion
            // the interrupt does not wait for the conversion to
complete
            PORTCbits.RC3=!PORTCbits.RC3; //DEVELOPOMENT FLAG
            T1 overflow=T1 overflow+1;
            if (T1 overflow==sample rate)
                  //initiate both sensor capture and LVDT conversion
                  //sensor capture:
                  PORTCbits.RC4=!PORTCbits.RC4; //DEVELOPOMENT FLAG
                  PIR1bits.TMR1IF=0; //manual clear of interrupt
request
            11
                  OpenCapture3(CAPTURE INT_ON&C3_EVERY 16 RISE EDGE);
            11
                  PR3=0xFFFF;
                  OpenTimer3(TIMER_INT_OFF&T3_SOURCE_INT);
            11
            11
                  PORTDbits.RD7=!PORTDbits.RD7; // enables capture
input clock source
                  sensor capture=1; // set capture flag
            11
                  T1 overflow=0; // reset the counter
                  //LVDT conversion: ADC was configured at startup
                  ConvertADC();
                  LVDT request=1;
            // allow next capture to interrupt
            11
                  PIE2=0;
                  PIR2bits.CA3IF=0;
            11
            11
                  PIE2bits.CA3IE=1;
            11
                  capturecounter=0;
            // try disabling usart for capture accuracy
            11
                  PIE1=0;
            // flag sample start
                  freqcapstrt=1;
            }
            else
                  PIR1bits.TMR1IF=0;
      }
```

// other peripheral interrupts are handled here

```
// update house keeping before return
      WREG=WREG_store;
      FSR0=FSR0_store;
      FSR1=FSR1 store;
      PRODL=PRODL store;
      PRODH=PRODH store;
} // end of peripheral interrupt handle
void Motor_timer()
ł
      //interrupt handler for timer 0.
      // first house keeping
     WREG store=WREG;
      FSR0_store=FSR0;
      FSR1_store=FSR1;
      PRODL_store=PRODL;
      PRODH_store=PRODH;
      //stop TMR0
      CloseTimer0();
      // when servicing a limit switch
      if(limit switches)
      ł
            PORTDbits.RD1 = 1; // begin duty cycle
            Delay10TCYx(2); //wait 100 cycles >> 25usec
            PORTDbits.RD1 = 0;
            OpenTimer0(TIMER_INT_ON&T0_SOURCE_INT&T0_PS_1_32);
            WriteTimer0(speed);
            INTSTAbits.TOIF=0;
            already relieved=1;
            if (energise)
                              PORTDbits.RD0 = 1;
            else
                        PORTDbits.RD0 = 0;
            if (direction)
                              PORTDbits.RD2 = 0;
            else PORTDbits.RD2 = 1;
            return;
      }
      //storing speed settings when changing direction or stopping
      if(((direction!=desired direction)||((energise)&&((energise!=
            desired_energise))))&&!slowing)
      {
            slowing=1; //slowing is set when slowing due to dir or en
            desired speed store=desired speed;
```

```
//incase of slowing for a direction change
if((direction!=desired direction) && (speed==0))
ſ
      direction=desired direction;
      desired speed=desired speed store;
      slowing=0;
}
else
      if ((direction!=desired direction) && (speed!=0))
            desired speed=0;
//incase of slowing to stop
if((energise) &&(!desired_energise) &&(speed==0))
Ł
      energise=0;
      slowing=0;
      CloseTimer0(); //disable TMR0 interrupt
}
else
      if((energise==1)&&(desired_energise==0)&&(speed!=0))
            desired speed=0;
//update motor driver
if (energise) PORTDbits.RD0 = 1;
                  PORTDbits.RD0 = 0;
      else
if (direction)
                  PORTDbits.RD2 = 0;
      else PORTDbits.RD2 = 1;
// modified clock generator:
PORTDbits.RD1 = 1; // begin duty cycle
Delay10TCYx(2); //wait 100 cycles >> 25usec
PORTDbits.RD1 = 0;
OpenTimer0(TIMER INT_ON&TO SOURCE_INT&T0_PS 1 32);
WriteTimer0(speed);
INTSTAbits.TOIF=0;
// finaly update house keeping before return
WREG=WREG store;
FSR0=FSR0 store;
FSR1=FSR1 store;
PRODL=PRODL store;
PRODH=PRODH_store;
```

} // end of TMR0 interrupt handler

}

```
void main(void)
{
      /* initialise code */
      initialise();
      while(1) /* main programme loop */
      {
      /* first check to see if new data has arrived from host PC */
      chk_data_in();
      timing=timing+1;
      if ((timing==2)&&(energise))chk_motor_speed(); //motor speed ramp
      if (there_has_been_trouble)send_trouble_to_snd_buff(); //Port B
ints
      if (limit_switches&&!reversing)relieve_microswitch();
      //Limit switches
      if (reversing&&!limit switches)stop relieving microswitches();
//Limit switches
      if
((LVDT request & !ADCONObits.GO) & !sensor_capture)get temperatures();
//sample temperatures
      if (sampling complete) send data to snd buff(); //conversion done
and capture done
      if (freqcapstrt) frequency_capture();
      } /* end of main while loop */
  /* end of main function */
ł
/* function definitions */
void frequency capture()
{
                  //PIE2=0;
                  WriteTimer3(0);
                  capturecounter=0;
                  PIR2bits.CA3IF=0;
                  PIE2bits.CA3IE=1;
                  while(!capturecounter);
                  while(!PIR2bits.CA3IF);
                  sensor reading=ReadCapture3();
                  sensor_capture=0;
                  PIR2bits.CA3IF=0;
                  PIE1bits.RC1IE=1;
                  PIE1bits.TMR1IE=1;
                  freqcapstrt=0;
}
void initialise ()
{
      /* install TMR0 vector function */
```

```
Install INT(default handler);
Install TMR0 (Motor timer);
Install PIV (peripheral ints);
Install TOCKI(default_handler);
/* initiate the motor control registers */
timing = 0;
speed = desired speed = 0;
direction = desired direction = 0;
energise = desired energise = 0;
slowing=0;
DDRD = 0;
PORTDbits.RD1=PORTDbits.RD0=0;
// disable capture source input
PORTDbits.RD7=1;
// set up development indicator port c
DDRC=0;
PORTC=0;
//set up i2c port
OpenI2C(MASTER, SLEW OFF); //configure as master
SSPADD=39; //set baud rate to 100KHz @16MHz
configure temp device(temp device 3 write);
configure temp device(temp device 4 write);
configure_temp_device(temp_device_5_write);
      /* configure analogue port */
OpenADC (ADC INT OFF&ADC FOSC 64&ADC RIGHT JUST&
      ADC VREF INT&ADC GANA GDIG, ADC CH10);
/* configure RS232 port */
OpenUSART1 (USART TX INT_OFF & USART RX INT_ON & USART ASYNCH MODE
& USART EIGHT BIT & USART CONT RX, 12);
/* setup buffer pointers */
rec buff ptr = &rec buff[0];
/* setup flags */
header_chk = total_exp = 0;
/* setup the message buffer to host machine */
sbr ptr = sbw ptr = sbreset ptr = &snd buff[0][0];
/* set TMR3 to 0 */
```

```
WriteTimer3(0);
      CA3L=CA3H=0;
      /* set TMR1 2 period registers */
      PR1=PR2=255;
      /* configure PORTB for emergency stop hardware */
      DDRB=0x07; // PORTB bits 0..2 are inputs
      PORTAbits.NOT RBPU=0; // turn on PORTB pull-ups;
      PORTB=0xF8; //set PORTB latch to expected input levels
      PORTB=PORTB;
      /* check emergency stop system for trouble */
      if(PORTB!=0xF8)there has been trouble=1;
      PIE1bits.RBIE=1; // enable PORTB interrupt
      freqcapstrt=0;
      capturecounter=0;
      /* finally enable unmasked interrupts */
      CPUSTAbits.GLINTD=0;
} /* end of initialise function */
void chk_data_in(void)
{
      if ((header chk==0)\&\&((rec buff ptr - 3) == \&rec buff[0]))
      /*
         if 3 unchecked bytes are available then check them */
            chk header(); /* then identify header */
      else
            if (((rec_buff_ptr - total_exp) ==
&rec_buff[0]) & (header_chk==1))
            /* if total bytes received are what's expected */
                  process_rec_buff(); /* then process buffer */
} /* end chk data in function */
void chk header(void)
//Check out incomming data by analysing its header
ł
      static char counter = 0;
      counter = 0;
     PIE1bits.RC1IE=0;
     while (counter<=no_inst)
```

```
{
            if ((rec buff[0]==header refs[counter] [0] )&&
                (rec buff[1]==header refs[counter] [1] )&&
                (rec buff[2] == header refs[counter] [2] ))
            ł
                  switch(counter)
                  {
                         case 0:
                               {total exp=3; /* no data associated */
                               rec_buff_ptr = &rec_buff[0];
                               PIE1bits.RC1IE=1;
                               reply_to_chk();
                               counter = no inst;
                               break;}
                         case 1: //request to stop acquisition.
                                                                  No
additional data associated
                               ł
                               rec_buff_ptr = &rec_buff[0];
                               PIE1bits.RC1IE=1;
                               reply_to_stp();
                               total exp=3;
                               counter = no inst;
                               break;}
                         case 2: // motor control expects more bytes
                               {PIElbits.RC1IE=1;
                               total_exp=11;
                               header chk=1;
                               counter=no inst;
                               break;}
                         case 3: // read LVDT input send to host PC
                               ł
                               rec buff ptr = &rec buff[0];
                               PIE1bits.RC1IE=1;
                               total exp=3;
                               get_LVDT_sig();
                               counter=no inst;
                               break; }
                         case 4: // checking for stored messages
                               ł
                               rec_buff_ptr= &rec_buff[0];
                               PIE1bits.RC1IE=1;
                               total_exp=3;
                               check message box();
                               counter=no inst;
                               break; }
                         case 5: // configuring sample rate
                               {PIE1bits.RC1IE=1;
                               total exp=7;
                               header chk=1;
                               counter=no inst;
                               break; }
                  } /* end switch */
            } /* end if true */
```

```
else
            Ł
                  counter=counter+1;
                   if(counter==no_inst)
                   {
                         PIE1bits.RC1IE=1;
                         rec_buff_ptr=&rec_buff[0];
                         reply_as_bad();
                   }
            }
      } /* end main while loop */
} /*end header check function */
void process rec buff(void)
{
// first determine which message has been sent then branch to handle
it.
      static char count = 0;
      count = 0;
      PIElbits.RC1IE=0;
      while (count <= no inst)
      ł
            if ((rec buff[0]==header refs[count] [0] )&&
                 (rec buff[1] == header refs[count] [1] )&&
                 (rec_buff[2]==header_refs[count] [2] ))
            {
                 switch(count)
                   Ł
                         case 0: //shouldn't happen
                               ł
                               break;}
                         case 1: //shouldn't happen
                               ł
                               break;}
                         case 2: // motor control message
                               rec_buff_ptr = &rec buff[0];
                               PIElbits.RC1IE=1; //enable int
                               update_motor();
                               count=no inst;
                               header chk=0;
                               reply to chk(); //respond ok
                               break;}
                         case 3: // shouldn't happen
                               {break;}
                         case 4: // shouldn't happen
                               {break;}
                         case 5: // getting sample rate data
                               //rec buff ptr = &rec buff[3];
```

```
PIE1bits.RC11E=0; //disable int
                               config_sample_rate();
                               count=no inst;
                               header_chk=0;
                               rec buff ptr = &rec buff[0];
                               PIElbits.RC1IE=1; //enable int
                               reply_to_chk(); //respond ok
break;}
                         default:
                               ł
                               if (count==no inst)
                               £
                                     reply as_bad();
                                     rec buff ptr=&rec buff[0];
                               }
                               break; }
                   } /* end switch */
            } /* end if true */
            else
                   count++; // end if else
      } /* end while structure */
} /*end process receive buffer */
void reply to chk(void)
      static char ok[]="ok";
ſ
      while(BusyUSART1()); /* ensure ready to transmit */
      putcUSART1(ok[0]);
      while(BusyUSART1());
      putcUSART1(ok[1]);
} /* end of reply to chk command */
void reply_to_stp(void)
ł
      static char *stp;
      static char stopped[]="stopped";
// here we would place any critical control routines
// e.g., stop that stepper motor right now!
      CloseTimer1();
      CloseCapture3();
      PORTDbits.RD7=1;
      sensor_capture=0;
// then we respond to the host PC to let it know what's happening
      stp = &stopped[0];
```

B21

```
while(BusyUSART1());
     putsUSART1(stp);
} /* end of reply to stop command */
void reply as bad(void)
ſ
      char bad[]="bad";
      while(BusyUSART1()); /* ensure ready to transmit */
      putcUSART1(bad[0]);
      while(BusyUSART1());
      putcUSART1(bad[1]);
      while (BusyUSART1()); /* ensure ready to transmit */
      putcUSART1(bad[2]);
} /* end of reply as bad statement */
void update motor()
      static char *tmp ptr;
ł
      // first update control registers:
      tmp ptr=&rec buff[5];
      while (*tmp_ptr==0x20) tmp_ptr++; // remove padding spaces
      desired_speed=atoui(tmp_ptr);
      desired direction=rec_buff[3];
      desired energise=rec buff[4];
      // then plan change in motor status
      //if motor is currently denergised but requested energised
      //then initiate timer0 at lowest speed, and in direction required
      if((desired energise==1)&&(!energise)&&((PORTB==0xF8)||(reversing
))) // ensure no emergencies
      ł
            /* prescale set to 20bit overflow */
            energise=desired energise;
            direction=desired direction;
            OpenTimer0(TIMER_INT_ON&T0_SOURCE_INT&T0_PS_1 32);
      }
      //if the motor is already energised and we get here then either
      //direction change or speed change is required. These changes
are
      //handled in the TMRO interrupt routine.
      //if the motor is energised and is requested to stop, it must
slow
      //before stopping from high speeds so cannot be immediately
stopped.
      //TMR0 handles gradual stopping.
}
```

```
void chk_motor_speed()
```

```
{
    //incase of changing speed
    timing=0;
    if(speed>desired_speed)
        speed--;
    else
        if(speed<desired_speed)
            speed++;
    }
}</pre>
```

```
}
```

void get\_LVDT\_sig()

```
{
```

```
static char *LVDTstr;
static char sendstring[7]="okl2";
static int LVDTresult;
```

//this routine handles the reading and sending of LVDT
displacement info

}

void check\_message\_box(void)

```
{ static char default_str[4]="ok0";
 static char *default_str_ptr;
```

//this function is called when the host tests for stored messages //if there are stored messages the oldest one is sent to the host

```
if (message_count!=0)
```

{

//two cases to consider: is it a binary message or ascii?
//binary messages begin okl

//do some common setup:

sbr ptr=&snd buff[current\_read\_message][0];

//test for binary message:

```
if (*(sbr_ptr+2)=='1')
                   local counter=0;
                   while(local counter++<12)
                   Ł
                         while(BusyUSART1());
                         putcUSART1(*(sbr ptr++));
                   }
            }
            else //it's not a binary message so:
            £
                   while(BusyUSART1());
                  putsUSART1(sbr ptr);
            }
            //more common stuff
            current read message++;
            if (current_read_message==5) current_read_message=0;
            //pointer wraps round
            message_count--;
      }
      else
      ł
            // no messages waiting so send default reply to host
            default_str_ptr=&default_str[0];
            while(BusyUSART1());
            putsUSART1(default str_ptr);
      }
void config sample rate(void)
                   char *src rate ptr;
      static
      // this function reads the desired sample rate and configures the
      // appropriate timer.
      src_rate_ptr=&rec_buff[3];
      while(*src_rate_ptr==0x20)src_rate_ptr++; // remove padding
spaces
      sample_rate=atoui(src_rate_ptr); // holds overflow target
      T1 overflow=0; // reset overflow
      OpenTimer1(TIMER INT ON&T1 SOURCE INT&T1 T2 16BIT);
      //PIR1bits.TMR1IF=0; //manual clear of interrupt request
      OpenCapture3(CAPTURE_INT ON&C3_EVERY_16_RISE_EDGE);
      PR3=0xFFFF;
      OpenTimer3(TIMER_INT_OFF&T3_SOURCE_INT);
      PORTDbits.RD7=0; // enables capture input clock source
sensor_capture=1; // set capture flag
      T1_overflow=0; // reset the counter
```

}

ł

```
void send trouble to_snd_buff(void)
{
      static char *trouble pointer;
      static char trouble_message[]="ok2....";
      // this function is called when there has been trouble (emergency
stop)
      // the cause of the stop is interpreted and sent to the host PC
      trouble pointer=&trouble_message[3];
      ubtoa(image of PORTB, trouble pointer);
      // assume buffer pointer not to be pre-set:
      sbw ptr=&snd buff[current write message][0];
      trouble pointer=&trouble message[0];
      strcpy(sbw ptr,trouble pointer); //write to buffer
      current_write_message++;
      if (current write message==5) current write message=0; //pointer
wraps round
      // update message counter:
      message count++;
      //reset other flags
      there_has_been_trouble=0;
      PIE1bits.RBIE=1; //re-enable RB int after bounce settled
}
void send data to snd buff(void)
            char *message_pointer;
{static
 static
            char data message[20];
      // construct header
      data_message[0]='o';
      data_message[1]='k';
      data message[2]='1';
      // this function is called when sensor data is ready to be sent
      //CPUSTAbits.GLINTD=1;
      // first get the LVDT reading
      LVDT_reading=ReadADC();
      // build message string:
      message_pointer=&data_message[3];
      //move bytes from LVDT registers as sequential chars to buffer
```

```
*(message pointer++)=(unsigned char)LVDT reading;
```

```
*(message pointer++)=(unsigned char)(LVDT reading>>8);
      //and then for the test sensor
      *(message pointer++)=(unsigned char)sensor reading;
      * (message pointer++) = (unsigned char) (sensor reading>>8);
      //and then for the temp sensor 1
      *(message_pointer++)=(unsigned char)temp_sensor_1;
      *(message pointer++)=(unsigned char)(temp sensor 1>>8);
      //and then for the temp sensor 2 \
      *(message pointer++)=(unsigned char)temp sensor 2;
      *message pointer=(unsigned char)(temp sensor 2>>8);
      sbw ptr=&snd buff[current write message][0];
      message_pointer=&data_message[0];
      //manual copy of strings
      local counter=11;
      while(local counter-->0)*(sbw ptr++)=*(message pointer++);
      //memcpy(sbw ptr,message pointer,11); //write to buffer
      current write message++;
      if (current write message==5) current write message=0; //pointer
wraps round
      // update message counter:
     message count++;
      //reset other flags
      sampling_complete = 0;
      LVDT request = sensor capture = 0;
      PORTCbits.RC5=!PORTCbits.RC5; //DEVELOPOMENT FLAG
      //CPUSTAbits.GLINTD=0;
}
void relieve microswitch(void)
{
      // this function overrides motor control to bring the
      // bed back from a limit hit.
      // first ensure this routine isn't repetitively called
      if (reversing) return;
      // reverse the motor (whilst still off)
      if (desired direction)
            desired direction=0;
      else
            desired direction=1;
      direction=desired direction;
```

```
// energise motor
      desired energise=1;
      energise=1;
      reversing=1; //flag to ensure this routine happens once
      OpenTimer0(TIMER_INT_ON&T0_SOURCE_INT&T0_PS_1_32);
      WriteTimer0(speed);
}
void stop relieving microswitches (void)
ł
      if (desired direction)
            desired_direction=0;
      else
            desired direction=1;
      desired_speed=0;
      desired energise=0;
      reversing=0;
      OpenTimerO(TIMER INT ON&TO SOURCE INT&TO PS 1 32);
      WriteTimer0(speed);
}
void get_temperatures_only_for_LM75_chip(void)
Ł
      // this function accesses two temperature sensing devices on the
      // i2c bus, one provides temperature in proximity of transducer
      // the second provides temperature in proximity of sensing
      // circuitry.
      // the i2c interface is configured in the initialise routine
      // clear buffers
      temp sensor 1=temp sensor 2=0;
      // first check bus is idle
      IdleI2C();
      //initiate comms
      StartI2C();
      IdleI2C();
      temp=WriteI2C(temp_device_1); //write address byte
      //wait for byte to return then read it
      IdleI2C();
      temp byte high=ReadI2C();
      IdleI2C();
      //acknowledge reception
      AckI2C();
      IdleI2C();
      //wait for low byte to return then read it
      temp byte low=ReadI2C();
      IdleI2C();
      NotAckI2C();
      IdleI2C();
      //stop sequence
      StopI2C();
```

```
//now transfer data to 16 bit int store
```

```
temp sensor 1=temp byte high;
      temp sensor 1<<=1;
      temp byte low<<=1;
      if (ALUSTAbits.C==1) temp sensor 1++;
      //repeat this sequence for the second temperature sensor
      IdleI2C();
      //initiate comms
      StartI2C();
      IdleI2C();
      temp=WriteI2C(temp device 2); //write address byte
      //wait for byte to return then read it
      IdleI2C();
      temp byte high=ReadI2C();
      IdleI2C();
      //acknowledge reception
     AckI2C();
      IdleI2C();
      //wait for low byte to return then read it
      temp byte low=ReadI2C();
      IdleI2C();
      NotAckI2C();
      IdleI2C();
      //stop sequence
      StopI2C();
      //now transfer data to 16 bit int store
      temp sensor 2=temp byte high;
      temp sensor 2<<=1;
      temp byte low<<=1;</pre>
      if (ALUSTAbits.C==1) temp sensor 2++;
      // set finished flag
      sampling complete=1;
      PORTCbits.RC6=! PORTCbits.RC6; //DEVELOPOMENT FLAG
                              //only for DS1624 chip
void get temperatures(void)
      // clear buffers
      temp_sensor_1=temp_sensor_2=temp_sensor_3=0;
      get_device_temp(temp_device_3_write);
      //now transfer data to 16 bit int store
      temp sensor 1=temp_byte_high;
      temp_sensor_1<<=8;
      temp sensor 1+=temp byte low;
      temp_sensor_1>>=3;
      get device temp(temp device 4 write);
      //now transfer data to 16 bit int store
```

ł

```
temp sensor_2=temp_byte_high;
      temp_sensor 2<<=8;</pre>
      temp sensor_2+=temp_byte_low;
      temp sensor 2>>=3;
      get device_temp(temp_device_5_write);
      //now transfer data to 16 bit int store
      temp_sensor_3=temp_byte_high;
      temp sensor 3<<=8;
      temp_sensor_3+=temp_byte_low;
      temp_sensor_3>>=3;
      // set finished flag
      sampling complete=1;
      PORTCbits.RC6=!PORTCbits.RC6; //DEVELOPOMENT FLAG
void get device temp(unsigned char device base no)
      //CPUSTAbits.GLINTD=1;
      // first check bus is idle
      IdleI2C();
      //initiate comms
      StartI2C();
      IdleI2C();
      temp=WriteI2C(device base no); //write address byte
      //wait for acknowledge
      IdleI2C();
      //then write instruction to i2c bus
      temp=WriteI2C(0xAA);
      //wait for acknowledge
      IdleI2C();
      //then issue restart
      RestartI2C();
      IdleI2C();
      //then address chip to read data
      temp=WriteI2C(device base no+1);
      IdleI2C();
      //read back high byte
      temp_byte_high=ReadI2C();
      IdleI2C();
      //acknowledge reception
      AckI2C();
      IdleI2C();
      //wait for low byte to return then read it
      temp byte low=ReadI2C();
      IdleI2C();
      NotAckI2C();
      IdleI2C();
      //stop sequence
      StopI2C();
      //CPUSTAbits.GLINTD=0;
```

}

ł

```
void configure temp device (unsigned char device_base_no)
{
//initiate ds1624,
      // first check bus is idle
      IdleI2C();
      //configure ds1624
      StartI2C();
      IdleI2C();
      temp=WriteI2C(device_base_no); //
      IdleI2C();
      temp=WriteI2C(0xAC);
      IdleI2C();
      temp=WriteI2C(0);
      IdleI2C();
      StopI2C();
      //wait 10ms for eeprom write
      Delay10KTCYx(4);
      // first check bus is idle
      IdleI2C();
      //initiate comms
      StartI2C();
      IdleI2C();
      temp=WriteI2C(device_base_no); //write address byte
      //wait for acknowledge
      IdleI2C();
      //then write instruction to i2c bus
      temp=WriteI2C(0xEE);
      //wait for acknowledge
      IdleI2C();
      //then issue start
      StopI2C();
```
# Appendix C

This appendix includes data sheet first pages for the electronic components of greatest significance to the test rig and the prototype sensor.



# DS1624 Digital Thermometer and Memory

#### www.dalsemi.com

# **FEATURES**

- Temperature measurements require no external components
- Measures temperatures from -55°C to +125°C in 0.03125°C increments. Fahrenheit equivalent is -67°F to +257°F in 0.05625°F increments
- Temperature is read as a 13-bit value (two byte transfer)
- Converts temperature to digital word in 1 second (max)
- 256 bytes of E<sup>2</sup> memory on board for storing information such as frequency compensation coefficients
- Data is read from/written via a 2-wire serial interface (open drain I/O lines)
- Applications include temperaturecompensated crystal oscillators for test equipment and radio systems
- 8-pin DIP or SOIC packages

# **PIN ASSIGNMENT**



DS1624S 8-PIN SOIC (208 MIL) See Mech Drawings Section



DS1624 8-PIN DIP (300 MIL) See Mech Drawings Section

# **PIN DESCRIPTION**

- SDA 2-Wire Serial Data Input/Output
- SCL 2-Wire Serial Clock
- GND Ground
- A0 Chip Address Input
- A1 Chip Address Input
- A2 Chip Address Input
- $V_{DD}$  Digital Power Supply (+3V +5V)
- NC No Connection

# DESCRIPTION

The DS1624 consists of a digital thermometer and 256 bytes of  $E^2$  memory. The thermometer provides 13-bit temperature readings which indicate the temperature of the device. The  $E^2$  memory allows a user to store frequency compensation coefficients for digital correction of crystal frequency due to temperature. Any other type of information may also reside in this user space.

August 2000



# LM6142 and LM6144 17 MHz Rail-to-Rail Input-Output Operational Amplifiers

# **General Description**

Using patent pending new circuit topologies, the LM6142/44 provides new levels of performance in applications where low voltage supplies or power limitations previously made compromise necessary. Operating on supplies of 1.8V to over 24V, the LM6142/44 is an excellent choice for battery operated systems, portable instrumentation and others.

The greater than rail-to-rail input voltage range eliminates concern over exceeding the common-mode voltage range. The rail-to-rail output swing provides the maximum possible dynamic range at the output. This is particularly important when operating on low supply voltages.

High gain-bandwidth with 650 µA/Amplifier supply current opens new battery powered applications where previous higher power consumption reduced battery life to unacceptable levels. The ability to drive large capacitive loads without oscillating functionally removes this common problem.

#### Features

At  $V_8 = 5V$ . Typ unless noted.

- Rail-to-rail input CMVR -0.25V to 5.25V
- Rail-to-rail output swing 0.005V to 4.995V
- Wide gain-bandwidth: 17 MHz at 50 kHz (tvp)
  - Slew rate: Small signal, 5V/µs
  - Large signal, 30V/µs
- Low supply current 650 µA/Amplifier
- Wide supply range 1.8V to 24V
- CMRR 107 dB
- Gain 108 dB with R<sub>L</sub> = 10k
- PSRR 87 dB

#### Applications

- Battery operated instrumentation
- Depth sounders/fish finders
- Barcode scanners
- Wireless communications
- Rail-to-rail in-out instrumentation amps



© 2009 National Semiconductor Corporation DS012057 www.national.com



# PIC17C7XX

# High-Performance 8-bit CMOS EPROM Microcontrollers with 10-bit A/D

#### **Microcontroller Core Features:**

- · Only 58 single word instructions to learn
- All single cycle instructions (121 ns), except for program branches and table reads/writes which are two-cycle
- · Operating speed:
- DC 33 MHz clock input
- DC 121 ns instruction cycle
- · 8 x 8 Single-Cycle Hardware Multiplier
- Interrupt capability
- 16 level deep hardware stack
- · Direct, indirect, and relative addressing modes
- Internal/external program memory execution, capable of addressing 64 K x 16 program memory space

Device	Memory	
	Program (x16)	Data (x8)
PIC17C752	8K	678
PIC17C756A	16 K	902
PIC17C762	8K	678
PIC17C766	16 K	902

#### **Peripheral Features:**

- Up to 66 I/O pins with individual direction control
- 10-bit, multi-channel Analog-to-Digital converter
- · High current sink/source for direct LED drive
- · Four capture input pins
- Captures are 16-bit, max resolution 121 ns
- Three PWM outputs (resolution is 1 to 10-bits)
- TMR0: 16-bit timer/counter with 8-bit programmable prescaler
- TMR1: 8-bit timer/counter
- TMR2: 8-bit timer/counter
- TMR3: 16-bit timer/counter
- Two Universal Synchronous Asynchronous Receiver Transmitters (USART/SCI) with independent baud rate generators
- Synchronous Serial Port (SSP) with SPI™ and I<sup>2</sup>C™ modes (including I<sup>2</sup>C Master mode)



#### **Special Microcontroller Features:**

- Power-on Reset (POR), Power-up Timer (PWRT) and Oscillator Start-up Timer (OST)
- Watchdog Timer (WDT) with its own on-chip RC oscillator for reliable operation
- Brown-out Reset
- Code protection
- · Power saving SLEEP mode
- · Selectable oscillator options

#### **CMOS Technology:**

- Low power, high speed CMOS EPROM technology
- · Fully static design
- Wide operating voltage range (3.0V to 5.5V)
- · Commercial and industrial temperature ranges
- · Low power consumption
- < 5 mA @ 5V, 4 MHz
- 100 µA typical @ 4.5V, 32 kHz
- < 1 µA typical standby current @ 5V</li>

© 2000 Microchip Technology Inc.

#### Appendix D

This appendix includes relevant publications as a result of this research work.

Proceedings of Mechatronics 98.
Mechatronics Journal 2001

#### Mechanical Earthworm to facilitate the access and repair of buried infrastructure.

W. L. Dudeney and M. R. Jackson

Mechatronics Research Group, Department of Mechanical Engineering, Loughborough University, Loughborough, LE11 3TU, U.K. Tel: +44 (0)1509 223231 Fax: +44 (0)1509 223934 e-mail: w.l.dudeney@lboro.ac.uk m.r.jackson@lboro.ac.uk

This paper introduces an access problem faced by industries owning buried infrastructure. Current solutions to the problem are assessed and their disadvantages are highlighted. A novel concept solution is proposed which will address current disadvantages, so providing an efficient solution with generic applicability. The proposed solution follows a keyhole surgery principle with an in-hole propulsion mechanism. The system is briefly described with additional detail given to biomechanical analogies used in the propulsion mechanism.

#### **1. INTRODUCTION**

of buried There are four main owners distributors, electricity infrastructure: gas distributors, water distributors/sewage collectors, and communications providers. Infrastructure types are pipe and cable (see figure 1). This infrastructure is buried, in most instances for the convenience of the public and to provide a mechanically and thermally stable bedding.

infrastructure. However various problems do arise for different reasons, often created by third party intrusion such as subsidence caused by heavy road vehicles mounting pavements, and sometimes by operational wear such as gradual degradation caused by encrustation of a cast iron water main. When problems occur it is often necessary to access the infrastructure in question, the challenge is to provide this access whilst minimizing urban disruption.

In normal circumstances it is unnecessary for the operator to physically access this type of

#### Figure 1. Two types of buried infrastructure.



Cable is either electrical or optical and used for power communications.

Insulation and mechanical protection are a major constituent of cable

Buried electric cables occasionally operate at up to 275,000 volts, although 11,000 volts down to 240 volts is more typical.

#### 2. STRATEGIES TO ACCESS BURIED INFRASTRUCTURE

Historically, access to any buried object has involved some form of excavation. Modern techniques still rely heavily on excavation of overlying and surrounding material, but in specific circumstances it is possible to gain access nonintrusively if the infrastructure concerned allows internal access (e.g., pipes with sufficient internal diameter).

rehabilitate strategies to under-Current performing buried infrastructure are many and varied, but the more novel technologies can be shown to be [4,5] specific rather than widely applicable. A core decision which will determine the nature of any access required is whether to replace a section or refurbish a section as the processes may be very different. Buried distribution and collection systems often run beneath other urban services such as road ways and pavements, it is for this reason that access is difficult and likely to cause urban disruption.

#### 2.1. Repair of buried cable

Power distribution cables are usually bedded directly in the ground, whilst modern optical communications networks are ducted in 15 centimetre diameter plastic pipe. Cable affords no internal access so excavation is the only method of providing access. Cable faults are failing insulator or failing conductor, or both. There are established techniques [1] to accurately locate points of failure to within 4 metres in typical circumstances. It is normal to replace a 2 metre section of cable around any point of failure. This is skilled work and is currently a job requiring hands-on labour (implying an excavation large enough to allow human access). Excavation of this nature is a common sight in urban U.K.

#### 2.2. Repair of buried pipe

Pipe is bedded directly in the ground. There are many different types of pipe in service, from fired clay, concrete, and cast iron, through to steel and plastic. Typical problems with pipe are blockage, corrosion, and failing joints and valves. There are well developed sonic techniques for leak detection [2,3], but the accuracy of pinpointing leak detection is variable dependent on factors such as: leak velocity, pipe material, mechanical qualities of surrounding ground. Pipe offers the possibility of internal access. There are systems in use to allow inspection, refurbishment and replacement of pipe. Briefly, these include in-pipe robotics for inspection, lining mechanisms, and various systems to replace the pipe with similar, smaller, or even larger diameter pipe. All these systems operate within the pipe, requiring at most an entrance and an exit point. These systems are detailed elsewhere [4,5].

However, Many pipes are too small to allow internal access, and often the nature of the problem may not suit any of the internal repair systems (for example leaking joints and valves). In these instances excavation is required to allow skilled manual labour (implying a man-in-hole sized excavation).

#### 2.3. Evaluation of current strategies

Although no data is available (inquiries were made at Department of Environment, Department of Transport, and many owners of buried infrastructure, as well as their regulators), it is easy to accept that the majority of buried infrastructure failings are accessed through open excavation. There are economic, social, and environmental disadvantages to open excavation in urban areas:

- Urban disruption includes road works generated to allow access to services beneath the road.
- Spoil becomes distributed.
- Excavation produces noise.
- Excavation often leads to permanent and progressive damage to roads and pavements.
- Excavation often requires plant hire.
- Third parties are required to reinstate surface structures.
- Excavation often damages other buried infrastructure.

Despite these disadvantages, excavation continues to be popular. This might be simply for the reason that repair operations are technical and require experience and hands-on skill to perform.

#### Figure 2. Concept of keyhole surgery applied to buried infrastructure.



However, with developments in teleoperation [6] it should now be possible to remove the need to physically place the skilled labourer at the work site and so reduce the size of any excavation.

#### 3. THE MINIMAL INTRUSION CONCEPT

The ideal way to minimize the disadvantageous effects of excavation is to adopt a key hole surgery principle. The key hole concept meets the desire to avoid disturbance of critical factors around the work site. It would be possible to access infrastructure below a road way without affecting that road in any way (see figure 2).

The device would be 'inserted' into the ground at a convenient location, away from critical services. From this point the device tunnels its way towards the desired location under the control of a human operator. On arrival at the desired location the damage is inspected and repaired by the operator, using a selection of on-board tools and materials. Following a successful remote repair the device is withdrawn along the tunnel it has created, backfilling as it goes. The only surface evidence of the operation will be a small circle where the device was inserted. At all times waste materials and spoil are managed. Noise will be reduced because cutting operations will be underground.

#### 4. THE MECHANICAL EARTHWORM

The construction of a feasible plan for the key hole surgery principle requires ambitious thinking. The desire is to create a self sufficient excavation and repair system with wide applicability and the ability to address many if not all the disadvantageous factors in traditional excavation.

From this the concept can be detailed as a system including: a tunnelling vehicle, a surface power and spoil/materials handling unit, a human control interface, an umbilical connection between the surface unit and the vehicle, and a set of interchangeable tools for tunnelling and repair/inspection. A graphical representation of this system is given in figure 3.

It is accepted that there are technological challenges in many areas of this system, some of those identified are:

• in-tunnel propulsion,

- underground vision,
- cutting through urban ground,
- directional tunnelling,
- obstacle identification,
- teleoperation procedures/human interfacing,
- interchangeable teleoperation tool design,
- vehicle extraction and ground reinstatement.

Brief consideration has been given to a possible topography for the tunnelling vehicle, this is best described by figure 4. Other than this the current



Figure 3. Schematic representation of the conceptual key-hole surgery system's main components.

#### Figure 4. Visual of possible topography of tunnelling vehicle.



research focus is on in-tunnel propulsion and the conclusion of this paper will concentrate on a novel method.

#### 5. BIOLOGICALLY INSPIRED IN-TUNNEL PROPULSION

Biological inspiration has been sought to address the issue of providing propulsion. The locomotion mechanics of the earthworm is a suitable system on which to base an artificial in-tunnel propulsion mechanism. Basic earthworm locomotion is described in figure 5, and is described in greater detail elsewhere [7,8]. The earthworm is analogous to an elastic sack containing a fixed volume of fluid. The elastic sack includes embedded antagonistic muscle groups which operate to pump fluid between different segments of the sack. Because the volume of the sack is constant any swell in one location must be taken up by either shortening of that location, or by constriction of another location. Alternatively the muscle groups may be used to elongate one section and maintain local volume by constricting the section. Articulation may be achieved by antagonistic operation of muscle groups which run the length of the sack (see figure 6).

#### 5.1. Mechanical implementation of earthworm



Figure 5. Description of basic earthworm locomotion.

Figure 6. Worm segment sketch to show muscle groups used in locomotion.



Circumferential muscles: Control local diameter. Longitudinal muscles: Control elongation and articulation.

#### locomotion principle

Three priority issues are recognized in the mechanical implementation of an earthworm locomotion system:

- •Power supply type
- Actuator requirements
- •Closed loop control

Current work is focused on an embedded/surface mounted positional feedback system to allow a control system to know the deformation shape of an elastic surface. Initially it is sought to provide 1 and 2 dimensional information and subsequently through wrapping the surface into a cylindrical form to interpolate 3 dimensional information.

#### REFERENCES

- 1. T. Gonen, Electric Power Transmission System Engineering. 1998.
- E. S. Cole, Methods of Leak Detection: An Overview. Distribution Systems - Actions and Innovations, AWWA, 1980.
- G. L. Laverty, Leak Detection: Modern Methods, Costs, and Benefits. Distribution Systems - Actions and Innovations, AWWA, 1980.
- M. Najafi, State-of-the-art Review Trenchless Pipelind Rehabilitation Systems. 2<sup>nd</sup> International Conference on Advances in Underground Pipeline Engineering, 1995.

- T. Takana, The Application of Pipe Repair, Rehabilitation Techniques and their Recent developments. Distribution/Transmission Conference of the Operatins Section of the American Gas Association, 1986.
- 3. R. Kalawski, The Science of Virtual Reality and Virtual Environments, 1993.
- 4. D. Nichols, The Oxford Book of Invertebrates, 1971.
- 5. T. Newton, Muscling in on Nature. EPSRC Newsline, Winter 1997.

Some information presented in this paper has been sourced directly from relevant industries and regulators in the U.K. These organisations include:

Scottish Power, The York Waterworks, Yorkshire Water, Transco, London Electricity, Bristol Water, OFWAT, OFTEL, OFFER, South Western Electricity, Thames Water, Southern Water, London Cable.

# AN EXAMPLE OF DIGITAL PROCESSING TO ENHANCE SENSOR PERFORMANCE

# W. L. Dudeney and M. R. Jackson

Mechatronics Research Group, Department of Mechanical Engineering, Loughborough University, Loughborough, LE11 3TU. Tel +44 (0)1509 263171, e-mail w.l.dudeney@lboro.ac.uk

Received 8 March 2000; accepted 3 August 2000

Abstract—This paper discusses techniques used to gain high performance from sensor devices. A suggestion is made that costly traditional techniques may now be superceded through use of embedded processing power with the advantage of attaining good performance and adaptability for less expense. An experimental project is presented to demonstrate the capability of embedded systems in transducer design.

#### **1. INTRODUCTION**

There are several important criteria to be considered when determining the performance of a sensing system. These criteria affect the overall critical factor, which is to know and trust the relationship between sensor input and sensor output. Traditional texts [1-2] have generic terms for the holistic performance of sensor systems, these are often expressed as guaranteed percentage accuracy or maximum percentage error. These factors are determined from the product of sub-criteria and cannot as such be directly 'tweaked' (i.e., no sensor system has an accuracy trimmer!). It is necessary to consider classic factors such as: proportionality of input to output, stability of output, hysteresis of output, environmental effects on output, range of output and sensitivity of output to input. Whichever of these factors introduces the greatest error, or whichever combination of these factors introduces the greatest error would determine for example the maximum percentage error.

It can be seen that each factor specific to a particular system must attempt to minimise its contribution to the overall error score. Traditional approaches fall into two categories:

- Refine components of the system to a point where their error contribution is within tolerance. This approach can be expensive because of the need to manufacture parts to tight specifications, and the need to trim or tune individual systems for best performance.
- Employ the principle of common mode rejection. This approach requires that a sensor system be comprised of at least two identical transducers. In many cases one transducer is coupled to the system being measured, whilst the other is held close-by without being coupled (e.g., foil-leaf strain gauge on a wein-bridge arrangement) see [3] for more details. The two transducer outputs are differentially summed to eliminate any common signal, so that in principle, the effect of say temperature is

common to both and removed, whilst a change of input to the coupled transducer which does not occur at the input of the non-coupled transducer will produce a signal proportional to that change of input. In special cases, such as the linear variable differential transformer [4] (LVDT) two transducer elements are mechanically configured in a differential sense such that common signals are eliminated, whilst the effect of valid input is doubled. When transducers are combined in this manner it is important that they are well matched such that their responses to undesirable influences are identical and opposite.

Alone or combined these approaches have proven effective in increasing the signal to noise ratio of sensor outputs [5]. However there are obvious disadvantages in the cost and bulk of multiple transducers per sensor and/or precision components, and [6] is an example of the use of digital technology to lower costs.

It has recently become increasingly possible to replace traditional approaches to noise rejection and performance enhancement through the use of embedded computational devices and methods of adaptability (see [7-9] for more information), giving rise to what are often described as 'intelligent' or 'smart' sensors [10]. The remainder of this paper gives a practical example of the successful use of an embedded processor to dramatically improve the performance of a novel flexible but primitive displacement sensor. It will be seen that the combination of embedded processor system and basic transducer can be easier to produce, to install and to maintain than traditional sensor systems whilst achieving comparable performance.

## 2. NOVELTY IN DESIGN AND APPLICATION

There was a specific need to develop a displacement sensor based on a single coil to allow novel mechanical functions. These mechanical functions are the subject of subsequent publication and are not described here. However, the novelty of the instrument described in this paper is the proven use of an unbalanced inductive *displacement* sensor which, even at this preliminary stage of development suggests performance previously only obtained from balanced instruments. This is not to say that single coil inductive sensors do not exist, rather that these are traditionally classed as *proximity* rather than *displacement* sensors.

The advantage, in general terms, of the single coil sensor over, for example the LVDT and its derivatives are two fold: displacement range is no longer a function of coil length (and hence limited by body size); a single coil device precludes the need for rigid and precise alignment of multiple coils. This being the case there is potential for a diversity of mechanical forms of the sensor both in terms of geometry/topography and perhaps in materials. A new diversity in forms allows new measuring applications where bulk precluded the use of conventional devices before.

# 3. DEVELOPMENT AND PERFORMANCE OF A SIMPLE DISPLACEMENT SENSOR

We produced a simple displacement sensor from readily available and low cost materials, and used deliberately basic processes during manufacture. The principle of function was as follows:

A Colpitts oscillator [11] circuit is dependent on an LC tank to define its output frequency, if either the capacitative or inductive element of the tank become variable with a

couple to a displacement input, then the frequency of oscillation is related to the value of displacement. The equation for the Colpitts oscillator is:

$$f_{osc} = \frac{1}{2\pi\sqrt{LC}}$$

Therefore the relationship could be expected to be inverse root, and certainly not linear. A single Colpitts oscillator circuit was constructed (fig.1).



Fig. 1. Description of the Colpitts Oscillator

The inductive element of the oscillator was constructed from a short fat coil of approximately 50 turns and with a tapered ferrous core (fig. 2). The tapered core was coupled to the displacement system and was constructed such that it would move freely along the axis of the coil. The length of the taper on the core was approximately 28mm from its base to its apex.



Fig. 2. Inductor geometry

To analyse performance of the sensor system two main tests were conducted:

- Static testing. The system was held in a steady state whilst allowing environmental factors to change (temperature). Both temperature and the output of the sensor system were recorded. The aim was to observe the stability of the system and the effects of external factors.
- Dynamic testing. The system was repeatedly cycled through its maximum displacement range over a short period of time. Both sensor output and an external reference displacement sensor output (LVDT) were recorded at regular intervals. The aim was to observe the response profile of the test sensor when measured against the reference LVDT. This test would allow determination of range, linearity, and hysteresis.

The test rig used during experimentation was developed specifically to analyse the prototype transducer. The system was implemented in such a way that testing and data acquisition was fully automated, and that high level operations of the rig were programmable through a controlling PC interface to allow optimisation for obtaining information on specific performance parameters, e.g., short-term drift, or repeatability etc. A schematic of the key components of the test rig are shown in fig. 3 below. The test system was designed around methods and philosophies described in [12-13].



Fig. 3. Schematic diagram of test rig used

## 3.1 Static testing.

The chart in fig. 4 shows how the test sensor typically drifts over a period of time.



Fig. 4. Steady state drift of sensor output

The drift test shown in fig. 4 was started about midday and run continuously, sampling every 10 seconds until the third morning. It is clear that there is a low frequency drift pattern, which has a period of approximately 24 hours, and expected quantization. There is a trend downwards in frequency during the test period. In conclusion there was evident drift which appears to have both long-term and periodic components. Temperature was not controlled but allowed to drift, it was however, continuously recorded during this test.



Fig. 5. Relationship between temperature and sensor output drift

The chart in fig. 5 shows the test sensor output frequency plotted against recorded temperature during the test. It can be seen that the periodic drift component was caused by the effects of temperature. Further to this it can be seen from the chart that the relationship between drift and temperature change is more or less linear.

It was necessary to know the operating frequency range of the sensor system to determine the significance of this drift. The chart in fig. 6 shows the sensor's response curve to a full cycle of its operating range:



Fig. 6. Sensor output response curve to linear displacement

As can be seen from the figure above the response is not linear to input displacement and sensitivity rapidly approaches zero at the extremes of range. The range in frequency output from the oscillator was approximately 528KHz to 595KHz, the range of frequency due to drift was approximately 3KHz, therefore the expected error due to drift was approximately 4.5%.

In conclusion static performance of the system is dependent on temperature. Other random noise introducing a high frequency component to the signal may be considered within tolerable limits of 1% full scale.

# 3.2 Dynamic testing

Range and resolution:

Fig. 6 demonstrates a typical output response curve over the full displacement range. The maximum range over which a displacement can be sensed is approximately 28mm. The resolution of displacements peaks towards the middle of the range and rapidly tapers out to zero at each extreme. The nature of the frequency measuring system employed dictates it is possible to quantify absolute resolution of the system.



Fig. 7. Frequency capture method

As shown in fig. 7 a 16 bit timer (TMR) incrementing at 4MHz is reset to zero every sixteenth rising edge which arrives on the capture pin. When a capture event occurs the capture register (CAP) is fixed with the 16 bit value attained by TMR. A simple conversion is performed to determine the time passed during 16 rising edges, hence determining the frequency. The fact that displacement related information is quantified into a 16 bit value is useful in so much as it allows for definite resolution information to be known. For example, whereas fig. 6 displays output information in terms of calibrated

frequency, fig. 8 displays the same output information as raw 16 bit values captured in the CAP register.



Fig. 8. Sensor response curve as uncalibrated data

In this sample the range of values for TMR at the point of capture is incremental from 6875 through to 7759, giving an output scale which may be resolved to 884 discrete steps. Therefore on average, a sensitivity of 32 microns per step has been achieved. In practical terms the sensitivity is very poor over a total of 5mm at the range extremities, and the sensitivity in the working range is closer to 20 microns per step. It may be noted that resolution is affected by operational frequency range; a higher resolution may be obtained by increasing the active sample period (e.g., divide the incoming pulse train by 2). The benefit of increased sensitivity is offset by the longer frequency acquisition period, which increases the likelihood of dynamic errors.

The error due to non-linearity depends on the reference linear function used to fit the sample. A best fit was determined and a plot of the error in mm between the function and the sample was plotted as shown in fig. 9 below:



Fig. 9. Effect of non-linearity of output.

It can be seen from fig. 9 that the sensor response to input displacement cannot be taken as linear, the average error across the full displacement range is 1.8mm or 6.4% of the full scale with a maximum error of 21%.

In conclusion it may be seen that resolution is good for a relatively crude device but the response is inherently non-linear to displacement. Fig.'s 6 and 8 demonstrate the lack of hysteresis between rising and falling traces as the system was driven through its operational range several times to acquire this data. This is important as it suggests that a particular TMR value corresponds to one absolute and discrete displacement position and that repeatability to a single step should be possible.

# 4. ENHANCEMENT OF SENSOR USING MICROCONTROLLER BASED DIGITAL SIGNAL PROCESSING

There were three signal processing objectives sited for the embedded microcontroller:

- Filter high frequency components,
- Characterise and compensate for temperature effects,
- Characterise and compensate for non-linearity.

The functional process is as shown below in fig. 10.



#### Fig. 10. A functional process used in 'real-time' to process sensor output.

#### 4.1 Compensation for the effects of temperature.

A digital temperature sensor was applied to the system to provide temperature data for the compensation function g(f,t). This function was a simple linear transfer based on information graphically presented in fig. 5. The effect of applying this function in real-time to sensor output as it is generated may be seen in fig. 11.



Fig. 11. The effect on steady state drift of temperature compensation

It is clearly observable that the temperature compensation method applied has been effective. Error due to drift was reduced from 4.5% to less than 1% of full scale. Fig. 11 shows the results of a steady state test conducted continuously over a period of 5 days.

#### 4.2 Linearisation of displacement response

Fig. 9 indicates the use of a transfer function  $h(f_t)$  to compensate for the effects of nonlinearity in output. It was possible that any of several linearisation processes could be applied to the signal. These options included:

- Look-up table (microcontroller based example in [14]),
- Multiple non-continuous fit functions,
- Continuous polynomial or spline fit function.

Look-up tables were rejected in this instance because of the requirement for embedded hardware resources. Multiple non-continuous functions were considered most appropriate because of their relatively simple processing requirement. These will be considered further in subsequent research. A continuous polynomial fit was selected because it was the easiest system to implement in the early stages of this research. A test run was conducted to collect a raw data sample. An off-line process was implemented to generate the appropriate polynomial description for the data sample. This polynomial was then embedded within the signal processing software to provide the transfer function  $\{h(f_t) \text{ in fig. 10}\}$ . Fig. 12 shows how the use of a polynomial function was effective in improving linearity of the sensor output.



Fig. 12. The effect on linearity of the polynomial transfer function

In quantifiable terms linearity of the output response has been improved from 6.4% error to 1.1% over the full displacement range. Sensitivity has been unified in software to give a consistent 32 microns per step over the entire displacement range. This figure is not directly comparable with sensitivity of an equivalent LVDT. It is traditional to quantify LVDT sensitivity in terms of output voltage swing (or for more basic devices, phase shift) per unit distance because of the device's analogue nature. It is subsequently the sensitivity of the device reading the output of the LVDT which will determine operational sensitivity. This assumes that there is no error in the analogue output of the LVDT, which would be a false assumption. An output ripple, non-linearity, and overall error are often quoted for an LVDT device and it is typical that overall error might be 1% of full scale. As such an equivalent LVDT device with a 28mm operating range might typically be reliable to 0.28mm whilst sensitive to a single micron displacement.

An undesired effect of polynomial transfers is a tendency to introduce an oscillation into the conditioned output, this may be observed in fig. 12. It is expected that these oscillations may be eliminated through the use of alternative curve fitting techniques. It is likely that the novel transducer presented in this paper would be directly comparable and perhaps superior to the conventional LVDT in terms of performance.

#### 4.3 Repeatability

The prototype sensor was used for positional feed-back for a single axis sliding bed. The bed incorporated a vernier scale graduated to 0.02mm for reference. The slide was driven to a random position and the returned to a specified point by referencing the prototype sensor. This action was repeated many times and for a range of specified points. Fig. 13 shows the range of absolute position recorded (using the vernier scale) for each prototype-guided return to one specific point.



Fig. 13. Single repeatability test result

These results show that the range of actual position for a given sensor increment was 0.17mm. Therefore, in this test the repeatability error was 0.6% of the full scale, and if the tolerance were 1% then it could be said that the prototype is repeatable to within 1% full scale. However, in terms of resolution, 0.17mm accounts for 6 sensor output increments and hence repeatability would be said to be within  $\pm 3LSB$ .

#### 5. SUBSEQUENT RESEARCH

Work is currently ongoing in the following areas to further refine digital processing of simple sensor outputs:

- Improved frequency capturing circuitry to allow higher operational frequency and reduced capture times. A higher frequency tank circuit would also have the added advantage of using smaller discrete components.
- Highly integrated sensor/transducer technology. The aim is to reduce size of processing and conditioning components such that they may be physically embedded within the transducer body. Communications and electrical connection outside the sensor may be simply a standard serial protocol or bus network.
- Enhanced conditioning algorithms. To allow efficient use of limited processing resources and to build in high level functions such as self calibration.

### 6. CONCLUSION

This paper has demonstrated how a relatively simple electro-mechanical transducer with poor performance has been improved through use of a standard microcontroller device. A single coil inductive sensor has been used for the first time to allow accurate displacement measurement. With increasing use of higher level integration and hybrid technology it has become possible to embed intelligent nodes within individual sensors and actuators to allow the benefits of distributed control and management of systems. The advantage of this combination will enable new forms of inductive displacement sensing devices and new applications not previously possible.

## REFERENCES

- 1. Brindley K., Sensors and Transducers, 1988, ch.1.
- 2. Boyle H. B., Transducer Handbook, 1992, pp5-10.
- 3. Ruocco S. R., Robot Sensors and Transducers, 1987, ch.4.
- 4. Stadler W., Analytical Robotics and Mechatronics, 1995, ch.5.
- 5. Tse F. S., Measurement and Instrumentation in Engineering, 1989, ch.3.
- 6. Marlán C. B. et al, A Low Cost Circuit with Direct Digital Output for Pressure Measurement. *IEEE Transactions on Istrumentation and Measurement*. Vol.48 No.4, 1999, pp817-819.
- 7. Ehrlich J., Zerrouki A., Coche G., Smart sensor research at the French Laboratoire Central des Ponts et Chaussées, *Sensor Review*, Vol.17 No.3 1997, pp240-247.
- 8. Urena J. et al, Correlation Detector based on FPGA for Ultrasonic Sensors. *Microprocessors and Microsystems*. Vol.23 No.1, 1999, pp25-33.
- 9. Rhind A. W., Anjanappa M., Microcontroller based sensor to measure serum column height. *Mechatronics*. Vol.9 No.4, 1999, pp329-348.
- 10. White N., Intelligent Sensors, Sensor Review, Vol.17 No.2 1997, pp97-98.
- 11. Cirovic M., Electronic Devices, Circuits, and Systems, 1987, pp286-288.
- 12. Maguire L. P., McGinnity T. M., McDaid L. J., Issues in the development of an integrated environment for embedded system design. *Microprocessors and Microsystems*, Vol.23 No.4 1999, pp199-206.
- 13. Evans D. G., Green P. N., Morris D., A systems approach to embedded systems development. *Embedded Microprocessor Systems*, 1996, pp371-382.
- 14. Flammini A., Marioli D., Taroni A., Application of an optimal look-up table to sensor data processing. *IEEE Transactions on Instrumentation and Measurement*, Vol.48 No.4 1999, pp813-816.

. .