TIME DOMAIN WORK ON
BRASS INSTRUMENTS

Anne Margaret DEANE
(née DUFFIELD)

A thesis submitted to the Faculty of
Mathematical and Physical Sciences
of the University of Surrey
in fulfilment of the requirements
for the degree of Doctor of Philosophy

October, 1986
This work investigates brass instruments in the time domain, rather than the traditional frequency domain, and considers first, impulse measurements and secondly, their analysis.

An existing apparatus for measuring the response to an acoustic impulse at the input of a brass instrument has been refined. Problems of impulse inconsistency, ambient temperature variation and source reflections have been resolved.

Developments of the above equipment are used to test the quality of brass instruments on a factory production line. A prototype and a test instrument are compared by taking the arithmetical difference of their impulse responses. The equipment has detected small faults missed by normal inspection methods. The usefulness of this technique to brass instrument manufacturers is discussed.

Links between the instrument’s measured transient response and its bore geometry have been developed. The stages involved are deconvolution and bore reconstruction. Various deconvolution methods have been studied systematically by applying them to simulated noiseless and noisy data. Noise introduces errors, particularly at high frequencies, so deconvolution of real measured data is distorted. Techniques to reduce the effects of noise have been investigated. Attempts to employ the Gerchberg restoration algorithm to restore high frequency information proved unsuccessful.

A new inverse method, based on an iterative z-transform procedure, of reconstructing an instrument’s bore shape and damping profile from its transient response has been developed. It produces perfect results for noiseless model data, but even the smallest amount of noise renders the method unstable. Regularisation is therefore required. The corresponding direct process of predicting the transient response from bore and damping data is stable and
produces results which compare well with measured responses.

The work strengthens relationships between an instrument's shape and its musical quality, and will enhance the design of better instruments. Further research on the link between transient response and subjective quality is recommended.
ACKNOWLEDGEMENTS

I would like to express my gratitude to Dr. John Bowsher for his patient supervision during the course of this project. The support of Dr. Mike Jones is also gratefully acknowledged.

My Industrial Supervisor, Dr. Richard Smith has provided useful insights into the brass instrument industry and made many helpful contributions to the work. I am grateful to Boosey and Hawkes (MI) Ltd. for the loan of a variety of brass instruments. Thanks is also due to Dr. Dick Bacon for lending his trumpet, and to Miss Hilary Bowsher for lending her cornet.

Members of staff and fellow students within the Physics Department, University of Surrey, have provided much practical and moral support throughout; particular thanks is given to Mr. Dave Munro for his unparalleled computing expertise.

My thanks goes to the SERC, Boosey and Hawkes, and my parents, Mr. and Mrs. John Duffield for financial support received.

Finally, I would like to thank my husband, Mr. Ron Deane, who has been a continual source of encouragement.
# TABLE OF CONTENTS

1. INTRODUCTION

2. REVIEW OF THE FIELD OF RESEARCH
   2.1 Brass Instrument Research in the Frequency Domain
      2.1.1 Wave Motion in Brass Instruments
      2.1.2 Physical Assessment of Instruments Using Input Impedance
         2.1.2.1 Introduction
         2.1.2.2 Measurements with Constant Input Volume Velocity
         2.1.2.3 Direct Measurement of Velocity Using Anemometer
         2.1.2.4 Comparison of Computed and Measured Input Impedance Values
         2.1.2.5 Relationship between Input Impedance and Playing Frequencies
      2.1.3 Subjective Assessment of Instruments
      2.1.4 Summary of Section 2.1
   2.2 Time Domain
      2.2.1 Introduction
      2.2.2 Time Domain Work on Wind Instruments
      2.2.3 Impulse Response
      2.2.4 Summary of Section 2.2
   2.3 Deconvolution Methods
      2.3.1 Time Domain Deconvolution
      2.3.2 Frequency Domain Deconvolution
         2.3.2.1 Discrete Fourier Transform and Fast Fourier Transform
2.3.2.2 Windowing

2.3.2.3 Problems with Frequency Domain Deconvolution

2.3.3 Constrained Iterative Methods

2.3.4 Other Methods of Deconvolution

2.3.4.1 Cepstral Deconvolution (Quefrency Domain Deconvolution)

2.3.4.2 Cross Correlation

2.3.4.3 Inverse Filtering

2.3.5 Restoration Algorithms

2.3.5.1 Gerchberg-Papoulis Algorithm

2.3.5.2 Method of Projections onto Convex Sets (POCS)

2.3.5.3 Compensation for Slow Convergence

2.3.6 Measurements of Input and Output Signals

2.3.6.1 Constraints on Size Resulting from Convolution Properties

2.3.6.2 Method of Independent Measurements of Input and Response

2.3.7 Summary of Section 2.3

2.4 Reconstruction of Internal Shape from Transient Response

2.4.1 Introduction

2.4.2 Sondhi et al

2.4.3 Descout, Tousignant, Lecours and Lefèvre

2.4.4 Stansfield and Bogner

2.4.5 Discussion

2.4.6 Summary of Section 2.4

2.5 Overall Summary of Chapter 2
3. IMPULSE MEASUREMENTS

3.1 Apparatus

3.1.1 Overall Description

3.1.2 Choice of Source

3.1.2.1 Considerations
3.1.2.2 Spark Source
3.1.2.3 Practical Notes on the Use of the Spark Source

3.1.3 Coupler

3.1.3.1 Introduction
3.1.3.2 Mouthpiece Coupler
3.1.3.3 Source Tube
3.1.3.4 Flexible Source Tube
3.1.3.5 Use of Source Tube to Measure Attenuation

3.1.4 Choice of Microphone

3.1.4.1 Probe Microphone
3.1.4.2 1/8" Microphone
3.1.4.3 Electret Microphone

3.1.5 Insulation

3.1.6 Summary of Section 3.1

3.2 Software

3.2.1 Main Features of Data Acquisition Software

3.2.2 Interpolation

3.2.2.1 Method of Interpolation
3.2.2.2 Detailed Study of Interpolation

3.2.3 Impulse Inconsistency Investigations

3.2.3.1 Sources of Inconsistency and Remedies
3.2.3.2 Method of Monitoring Inconsistencies
3.2.3.3 Results of Inconsistency Investigations

3.2.4 Compensation for Ambient Temperature Drifts
3.2.4.1 Why Temperature Compensation is Necessary
3.2.4.2 Compensation by Changing Sampling Rate
3.2.4.3 Relationship between Temperature and Sampling Rate

3.2.5 Signal Averaging

3.2.6 Summary of Section 3.2

3.3 Measurements for Analysis or Comparison
3.3.1 Introduction
3.3.2 Summary of Practical Details
3.3.3 Standard Post-Measurement Processing
3.3.4 Additional Measurements Required for Analysis
3.3.5 Examples of Measurements of Brass Instruments

3.4 Overall Summary of Chapter 3

4. TECHNIQUE FOR INSTRUMENT COMPARISON AND INDUSTRIAL APPLICATION

4.1 Introduction

4.2 Method of Instrument Comparison
4.2.1 Methods Considered
4.2.2 Arithmetical Differencing Technique
4.2.3 Analysis of Arithmetical Difference Data
4.2.4 Investigations of the Limit Of Sensitivity of the Arithmetical Differencing Technique
4.2.4.1 Method
4.2.4.2 Results and Discussion
4.2.4.3 Temperature
4.2.4.4 Other Factors
4.2.4.5 Summary of Section 4.2.4
4.2.5 Summary of Section 4.2

4.3 Preliminary Investigations

4.3.1 Leak Detection (Trombone)

4.3.2 Detection of Small Constrictions (Cornet)

4.3.3 Detection of Changes in the Bell

4.3.4 Detection of Changes in Length

4.3.5 Differences Introduced by Bends

4.3.6 Summary of Section 4.3

4.4 Relaxation of the Experimental Arrangement

4.4.1 Shorter Source Tube

4.4.2 Single Standard Source Tube with Different End Couplers

4.5 Application of Instrument Comparison Technique in Industry

4.5.1 Quality Control

4.5.1.1 Limitations of Current Testing Procedures

4.5.1.2 Example of Simulated Production Line Conditions

4.5.2 Valve Misalignment

4.5.2.1 Detection of Simulated Valve Misalignment

4.5.2.2 Example of Actual Valve Group Fault

4.5.3 Tuba Variation

4.5.3.1 Examples of Tuba Variation

4.5.4 Common Manufacturing Faults

4.5.5 Detection and Diagnosis of Faults which Develop

4.5.5.1 Impulse Apparatus as a Diagnostic Tool

4.5.5.2 Example of Trumpet Fault Diagnosis

4.5.6 Summary of Section 4.5

4.6 Equipment Required for Factory System
4.6.1 Introduction
4.6.2 Data Acquisition Options
4.6.3 Microcomputer System

4.7 Overall Summary of Chapter 4

5. ANALYSIS OF RESULTS

5.1 Introduction

5.2 Deconvolution
  5.2.1 Method of Assessment of Different Deconvolution Techniques
  5.2.2 Time Domain Deconvolution
    5.2.2.1 Regularisation
    5.2.2.2 Input Data Shift
  5.2.3 Frequency Domain Deconvolution
    5.2.3.1 Software
    5.2.3.2 Results
    5.2.3.3 Spectral Windowing
    5.2.3.4 Input Data Shift
    5.2.3.5 Smoothing
    5.2.3.6 Discussion
  5.2.4 Constrained Iterative Deconvolution
    5.2.4.1 Method and Software
    5.2.4.2 Results and Discussion
  5.2.5 Final Assessment of Deconvolution Methods

5.3 The Gerchberg Algorithm
  5.3.1 Results for Noiseless Model Data
  5.3.2 Results for Model Data with Added Noise
  5.3.3 Results for Deconvolved Real Data
  5.3.4 Discussion
5.4 Bore Reconstruction

5.4.1 Theory
   5.4.1.1 Essence of the Method
   5.4.1.2 Reflectance Synthesis
   5.4.1.3 Bore Reconstruction

5.4.2 Performance of Bore Reconstruction Algorithm
   5.4.2.1 Test Data
   5.4.2.2 Results
   5.4.2.3 Discussion

5.4.3 Reflectance Synthesis Algorithm
   5.4.3.1 Performance
   5.4.3.2 Velocity Variation
   5.4.3.3 Discussion

5.5 Overall Summary of Chapter 5

6. DISCUSSION

6.1 Introduction

6.2 Summary of Achievements of Research

6.3 Existing Development Methods

6.4 How Impulse Measurements Enhance Design and Development Procedures
   6.4.1 Precise Copies Ensured
   6.4.2 Comparison of Old and New Models
   6.4.3 Comparison with Competitors' Models
   6.4.4 Further Potential of the Apparatus as a Research Tool

6.5 Relationship between Bore Shape and Musical Quality
   6.5.1 Indirect Relationship via Input Impedance
   6.5.2 Indirect Relationship via Transient Response
6.5.3 Link between Bore Shape and Transient Response
6.5.4 Links between Transient Response and Subjective Quality
6.5.5 Discussion of Future Subjective Experiments
6.5.6 Potential Use of Indirect Transient Response Relationship in Instrument Design

6.6 Summary of Suggestions for Future Work
6.7 Overall Summary of Chapter 6
TABLE OF FIGURES

Chapter 2
2.1 Schematic Representation of the Gerchberg Algorithm.
2.2 Experimental Arrangement Used by Sondhi and Resnick (1983) to Measure the Transient Response of the Human Vocal Tract.
2.3 Sketches to Illustrate the Measurements Obtained by Sondhi and Resnick (1983).
2.4 Schematic Representation of the Reflection and Transmission of Forward ($p_i$) and Backward ($b_i$) Travelling Waves in Successive Cylindrical Elements.

Chapter 3
3.1 Block Diagram of Impulse Measurement Apparatus.
3.2 Response when Loudspeaker Source is Coupled to the Source Tube via a Conical Coupler.
3.3 Car Spark Plug with Outer Electrode Removed.
3.4 Illustration of where Sparks Can Form.
3.5 Original Damped Mouthpiece Coupler.
3.6 Response of Trombone with Damped Mouthpiece Coupler.
3.8 Illustration of how Secondary Spikes Arise with Mouthpiece Coupler.
3.9 Response of Damped Mouthpiece Alone Measured Using 1/8" Microphone.
3.10 Source Tube. (a) Basic Source Tube Design, (b) Division of Original Source Tube into Sections.
3.11 Demonstration that the Source Tube Makes the Source Non-Reflecting.
3.12 Experimental Arrangement for Investigation of Flexible Source Tube.
3.13 Comparison of Impulses Measured at Three Different Positions Along the Source Tube.
3.14 Plot of Impulse Amplitude Versus Distance From Source.
3.15 Comparison of Amplitude Spectra of Impulses Measured at Three Different Positions Along Source Tube.
3.16 Comparison of the Performances of (a) Probe, (b) 1/8", (c) Knowles electret Microphones.
3.17 Photographs of the Impulse Measurement Equipment.
3.18 Flow Chart of General Data Acquisition Programme "IMPULSE5".
3.19 Comparison of Polynomials of Different Orders Fitted Through the Impulse Peak.
3.20 Distribution of Impulse Amplitudes in Volts for 1475 Successive Impulses.
3.21 Run-by-Run Variation of FPSN for 200 Consecutive Spark Impulses During a Temperature Rise of 0.3 °C.
3.22 Variation of Mean-FPSN-Over-Previous-10-Impulses For Each Impulse in Turn.
3.23 Variation of Running Average of Spark Impulse Amplitude with Increasing Number of Averages for 1475 Successive Impulses.
3.24 Run-by-Run Variation of Spark Impulse Amplitude for 200 Successive Impulses.
3.25 Variation of Running Average of Impulse Amplitude with Increasing Number of Averages for 200 Successive Spark Impulses.
3.26 Overall FPSN Distribution for 110 Consecutive Spark Impulses.
3.27 Run-by-Run Variation of FPSN for 200 Consecutive Loudspeaker Impulses During a Temperature Drop.
3.28 Arithmetical Difference Between Two Transient Responses of the Same Trombone Measured at Different Temperatures.
3.29 Flow Chart to Illustrate the Steps Involved in Finding the Correct Sampling Rate.
3.30 Comparison of Arithmetical Difference Plots with and without the Use of Interpolation During the Measurements.
3.31 Relationship Between Sampling Rate and Temperature for Consistent FPSN.
3.32 Variation of Mean RMS Level of N Sequences of a Computer-Generated Random Numbers with Successive Values of N up to 200.
3.33 As for Figure 3.32, but with 0.03 Added to Each Generated Random Number.
3.34 As for Figure 3.32, but Using 200 Sequences of 200 Samples of Measured Background Noise (Before the Arrival of the Acoustic Impulse).
3.35 Pressure Variation for Three Different Source Tube Terminations (Rigid Cap, Semi-Infinite Tube, and Instrument).
3.36 True Incident Impulse and Instrument Response Obtained by Subtraction of Responses in Figure 3.35.

Chapter 4
4.1 Flowchart Showing Proposed New Instrument Testing Scheme.
4.2 Arithmetical Difference Plots for the Euphonium.
4.3 Plots of Difference Level Versus (a) Time Interval (Hours), (b) Change in % Humidity, (c) Change in Temperature (°C), (d) Change in Ambient Pressure (dB).
4.4 Plot of Difference Level Against Temperature Interval in °C.
4.5 Sketch Plot of Feature Displacement Against Time.
4.6 Effect of Opening Trombone Waterkey.
4.7 Effect of Small Constriction in the Bell of a Cornet.
4.8 Effect of Placing a Duster in the Bell of a Cornet.
4.9 Effect of Placing a Mute in the Bell of a Trumpet.
4.10 Effect of Pulling Trombone Playing Slide Out by 7cm.
4.11 Effect of Pulling Trombone Tuning Slide Out by 4cm.
4.12 Comparison of Responses of a Normal Trumpet (Continuous) and the Equivalent Straight Trumpet (Dashed).
4.13 Demonstration of How, with a Shorter Source Tube, a Single Bore Constriction May Produce Two Arithmetical Difference Spikes.


4.15 Detection of a Sleeve Missing From a Cornet.

4.16 Diagram to Illustrate Missing Sleeve of Figure 4.15.

4.17 Diagram to Illustrate Valve Misalignment.

4.18 Illustration of How Large Valve Misalignments Change the Character of the Euphonium Response.

4.19 Detection of 1mm Valve Misalignment.

4.20 Relationship Between Size of Arithmetical Difference Spike and Extent of Valve Misalignment for all Euphonium Valves.

4.21 Valves Up Comparison of Two B-Flat Euphoniums.

4.22 First Part of Valves Down Comparison of the Same Two Euphoniums as in Figure 4.21.

4.23 Photographs Illustrating Detected Euphonium Valve Group Constriction.

4.24 Expanded Comparison of the First Part of the "Valves Up" Responses of the Three Tubas Superimposed.

4.25 Detection of Constriction in the Bell Bend of Trumpet.

4.26 Photograph to Illustrate the Bell Constriction of Figure 4.25.

4.27 Block Diagram to Illustrate the Peripheral Interfaces of the Microcomputer.

Chapter 5

5.1 Noiseless Model Data (f, h and g).

5.2 Real Experimental Data. (a) Incident Pulse, f, (b) Response of Straight Trumpet.

5.3 Comparison of (a) True Model h with (b) h Calculated by Time Domain Deconvolution, f having Been Shifted by 25 Samples to Start at its Peak. (The Resulting h has Been Shifted back by 25 Samples for Ease of
5.4 Comparison of h's for Real Data Calculated By Time Domain Deconvolution, f having Been Shifted by (a) 2 Samples (b) 3 Samples (c) 6 Samples.

5.5 Comparison of (a) True Model h with (b) h Calculated by Frequency Domain Deconvolution of Data Set (B), i.e. Model Data with 1% Gaussian Noise Added to g.

5.6 Difference Between the Spectra Corresponding to Figure 5.5.

5.7 Comparison of (a) Best Approximation of True h of Straight Trumpet with (b) h Computed by Frequency Domain Deconvolution.

5.8 Spectrum Corresponding to Figure 5.7(b).

5.9 Spectra Corresponding to Figure 5.2(a) and (b): f Dashed, g Continuous.

5.10 Results of Applying Rectangular Windows to Spectrum of Figure 5.8.

5.11 Expanded Upper Frequency Region of f-Spectrum of Figure 5.9, Shifted by 0 (Continuous), 1 (Dashed) and 2 (Dotted) Samples.

5.12 Comparison of (a) True g of Straight Trumpet with g's Obtained by Convolution of f with the Following Frequency Domain Deconvolved h's: (b) h Spectrally Windowed After Frequency Domain Deconvolution, and (c) h Obtained when f was Shifted by 2 Samples Beforehand.

5.13 Flow Chart of Constrained Iterative Deconvolution Programme ("DECONVIT2").

5.14 Comparison of (a) Best Approximation of True h of Straight Trumpet with Results of Constrained Iterative Deconvolution (b) After 10 Iterations, (c) After 50 Iterations.

5.15 Results of Application of Gerchberg Algorithm to Noiseless Model Data with q=240, k=127.

5.16 Inverse Fourier Transforms Corresponding to Figure 5.15.

5.17 Demonstration of How Gerchberg Algorithm Diverges when Applied to Real Deconvolved Data with q=1303, k=372.
5.18 Spectra Corresponding to Figure 5.17.

5.19 Schematic Representation of the Reflection and Damping of Forward and 
Backward Waves in the Region \( k\Delta < x < (k+1)\Delta \).

5.20 Model Bore Profile (Radius in cm Versus Sample Number) Used to Test Bore 
Reconstruction Algorithm.

5.21 Results of Analysis on Synthesised \( \{h(k)\} \) with 1% Gaussian Noise Added. 
(a) Original Synthesised \( \{p_k\} \), (b) \( \{p_k\} \) After Analysis.

5.22 Results of Analysis on Synthesised \( \{h(k)\} \) with 1% Gaussian Noise Added. 
(a) Original Synthesised \( \{q_k\} \), (b) \( \{q_k\} \) After Analysis.

5.23 Results of Reflectance Synthesis. (a) Measured Radius Data of Straight 
Trumpet, (b) Reflectance Synthesised From (a), (c) Best Approximation to 
True Reflectance Obtained Using Time Domain Deconvolution.

5.24 Comparison of Synthesised and Measured Responses.

5.25 Computed Approximation of Velocity Variation Along the Bore of the 
Straight Trumpet (Shown in Figure 5.23(a)).

Chapter 6

6.1 Flow Chart Showing Traditional Design Process.

6.2 Valves Up Comparison of Two B-Flat Euphoniums.

6.3 Complete Valves Up Comparison of Two Makes of B-Flat Euphonium.

6.4 Valves Down Comparison of the Same Euphoniums as in Figure 6.3.

6.5 Comparison of Two Makes of Trumpet. (a) Boosey and Hawkes "Symphony", (b) 
Vincent Bach, (c) Arithmetical Difference (Expanded Vertical Scale).

6.6 Comparison of Two Makes of Cornet. (a) Boosey and Hawkes, (b) Unknown 
Make, (c) Arithmetical Difference.

6.7 Flow Chart Showing Modified Design Process Suggested by Pratt.

6.8 Flow Chart Showing Modified Design Process Suggested by Goodwin.

6.9 Input Impedance of Straight Trumpet Computed From its Synthesised Reflectance.
6.10 Comparison in the Time Domain of (a) Synthesised Reflectance and (b) Input Impedance Computed From it.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
<th>Section in which Symbol First Appears</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a$</td>
<td>Duct Radius (in m)</td>
<td>5.4.3.2</td>
</tr>
<tr>
<td>$a_k$</td>
<td>Radius of $k$ th Element (in m)</td>
<td>5.4.2.1</td>
</tr>
<tr>
<td>$a(\omega)$</td>
<td>Real Part of Spectrum</td>
<td>2.3.4.1</td>
</tr>
<tr>
<td>$a(x)$</td>
<td>Pulse Amplitude (in Pa)</td>
<td>3.1.3.5</td>
</tr>
<tr>
<td>$b(\omega)$</td>
<td>Imaginary Part of Spectrum</td>
<td>2.3.4.1</td>
</tr>
<tr>
<td>$b_i(t)$</td>
<td>Backward Travelling Wave Pressure Amplitude in Element $i$ at Time $t$ (in Pa)</td>
<td>2.4.3</td>
</tr>
<tr>
<td>$b(x,y)$</td>
<td>Blurred Image</td>
<td>2.3.4.3</td>
</tr>
<tr>
<td>$c$</td>
<td>Arbitrary Data Zero</td>
<td>3.1.3.5</td>
</tr>
<tr>
<td>$c$</td>
<td>Velocity of Sound in Free Air (in m s$^{-1}$)</td>
<td>5.4.3.2</td>
</tr>
<tr>
<td>$d(i)$</td>
<td>Representation of Differences between Responses (No Units)</td>
<td>4.2.1</td>
</tr>
<tr>
<td>$d_k$</td>
<td>Damping Factor in $k$ th Element</td>
<td>5.4.1.1</td>
</tr>
<tr>
<td>$e$</td>
<td>Arbitrary Regularisation Parameter</td>
<td>5.2.2.1</td>
</tr>
<tr>
<td>$f$</td>
<td>Frequency (in Hz)</td>
<td>2.1.2.1</td>
</tr>
<tr>
<td>$f_s$</td>
<td>Sampling Frequency (in kHz)</td>
<td>3.2.2</td>
</tr>
<tr>
<td>$f_{s1}, f_{s2}$</td>
<td>Sampling Frequencies (in kHz)</td>
<td>3.2.4.2</td>
</tr>
<tr>
<td>$f(t)$</td>
<td>Input Signal (Arbitrary Units)</td>
<td>2.2.3</td>
</tr>
<tr>
<td>$f$</td>
<td>Unknown Input Signal</td>
<td>2.3.3</td>
</tr>
<tr>
<td>$f^{(k)}, f^{(k+1)}$</td>
<td>Successive Approximations to Unknown Input Signal $f$</td>
<td>2.3.3</td>
</tr>
</tbody>
</table>
\( f(i) \)  
Finite Discrete Input Sequence  
2.2.3

\( f(n) \)  
Finite Discrete Input Sequence  
2.3.3

\( f(j) \)  
Finite Discrete Input Sequence  
2.4.3

\( f \)  
General Discrete Input Sequence  
5.2.1

\( f(0) \)  
First Element of General Discrete Input Sequence  
5.2.2

\( f_p \)  
Peak Value of General Discrete Input Sequence  
5.2.4.1

\( f \)  
Discrete Time Sequence Expressed as a Column Vector  
2.3.6.3

\( f_0, f_1, f_2 \)  
Constituent Column Vectors which Together Make Up  
Column Vector \( f \)

\( f(i) \)  
Discrete Periodic Input Sequence  
2.3.6.1

\( f_{F,B}^k(t) \)  
Forward or Backward Incident Wave Pressure Amplitudes in \( k \) th Element at Time \( t \) (in Pa)  
5.4.1.1

\( f_{F,B}^k[m] \)  
Discrete Equivalent of \( f_{F,B}^k(t) \) (in Pa)  
5.4.1.1

\( f(x,y) \)  
Image to be Restored  
2.3.4.3

\( f(x,y) \)  
Approximation to Original Image \( f(x,y) \)  
2.3.4.3

\( g(t) \)  
Response Signal  
2.2.3

\( g \)  
Measured Response Signal  
2.3.3

\( g(n) \)  
Discrete Response Sequence  
2.2.3

\( g(n) \)  
Predicted Discrete Response Sequence (Using Convolution)  
2.3.3

\( g(i) \)  
Finite Discrete Response Sequence  
2.3.6.1

\( g(j) \)  
Measured Discrete Response Sequence  
2.4.3

\( g_{a}(j) \)  
Computed Discrete Response for Assumed Area Variation  
2.4.3
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>$g_1(i)$, $g_2(i)$</td>
<td>Measured Discrete Responses</td>
<td>4.2.1</td>
</tr>
<tr>
<td>$g$</td>
<td>General Discrete Response Sequence</td>
<td>5.2.1</td>
</tr>
<tr>
<td>$g_p$</td>
<td>Value of General Discrete Response $g$ at which $f_p$ Occurs</td>
<td>5.2.4.1</td>
</tr>
<tr>
<td>$g$</td>
<td>Latter Part of Discrete Response Expressed as a Vector</td>
<td>5.2.4.1</td>
</tr>
<tr>
<td>$g_p(i)$</td>
<td>Discrete Periodic Response Sequence</td>
<td>2.3.6.1</td>
</tr>
<tr>
<td>$g^{F,B}_{k}(t)$</td>
<td>Forward or Backward Outgoing Wave Pressure Amplitudes in $k$ th Element at Time $t$ (in Pa)</td>
<td>5.4.1.1</td>
</tr>
<tr>
<td>$g^{F,B}_{k}[m]$</td>
<td>Discrete Equivalent of $g^{F,B}_{k}(t)$ (in Pa)</td>
<td>5.4.1.1</td>
</tr>
<tr>
<td>$g(x,y)$</td>
<td>Distorted Image</td>
<td>2.3.4.3</td>
</tr>
<tr>
<td>$h(t)$</td>
<td>Impulse Response (No Units)</td>
<td>2.2.3</td>
</tr>
<tr>
<td>$h$</td>
<td>General Impulse Response</td>
<td>2.3.3</td>
</tr>
<tr>
<td>$h(k), h(k+1)$</td>
<td>Successive Approximations to Impulse response $h$</td>
<td>2.3.3</td>
</tr>
<tr>
<td>$h(i)$</td>
<td>Finite Discrete Impulse Response</td>
<td>2.2.3</td>
</tr>
<tr>
<td>$h^{(m)}(n)$</td>
<td>$m$ th Estimate of Discrete Impulse Response $h(n)$</td>
<td>2.3.3</td>
</tr>
<tr>
<td>$h^{(i)}$</td>
<td>Discrete Reflectance (No units) (See Also $r(t)$)</td>
<td>2.4.3</td>
</tr>
<tr>
<td>$h$</td>
<td>General Discrete Reflectance</td>
<td>5.2.1</td>
</tr>
<tr>
<td>$h$</td>
<td>Discrete Impulse Response Expressed as a Vector</td>
<td>5.2.4.1</td>
</tr>
<tr>
<td>$h^{(k+1)}$</td>
<td>$(k+1)$ th Estimate of $h$</td>
<td>5.2.4.1</td>
</tr>
<tr>
<td>$h_{p}(i)$</td>
<td>Discrete Periodic Impulse Response</td>
<td>2.3.6.1</td>
</tr>
<tr>
<td>$h(x,y)$</td>
<td>Point Spread Function</td>
<td>2.3.4.3</td>
</tr>
<tr>
<td>$i$</td>
<td>Integer Index</td>
<td>2.2.3</td>
</tr>
<tr>
<td>$i$</td>
<td>$\sqrt{-1}$</td>
<td>2.3.4.1</td>
</tr>
<tr>
<td>$i$</td>
<td>Cylindrical Element Index</td>
<td>2.4.3</td>
</tr>
</tbody>
</table>
\begin{itemize}
    \item \(j\) Integer Index 2.4.3
    \item \(k\) Integer Index 2.3.3
    \item \(k\) Defines Size of Column Vector, \(F_1\) to be \((2k+1)\) 2.3.5.3
    \item \(l\) Length of Cylinder (in m) 2.1.2.4
    \item \(l\) Number of Non-Zero Elements in \(h(i)\) 2.3.3
    \item \(l\) Integer Index 3.2.2.2
    \item \(m\) Integer Index 2.3.3
    \item \(m\) Number of Non-Zero Elements in \(f(i)\) 2.3.3
    \item \(m\) Number of Samples of Random Noise in a Sequence 3.2.5
    \item \(n\) Integer Index 2.2.3
    \item \(n\) Number of Non-Zero Elements in \(g(i)\) 2.3.3
    \item \(n\) Period of Discrete Periodic Response Sequence \(g_p(i)\) 2.3.6.1
    \item \(n(x,y)\) Noise on Image 2.3.4.3
    \item \(p(t)\) Pressure Variation with Time (in Pa) 2.2.3
    \item \(p(x,t)\) Pressure Variation with Distance and Time (in Pa) 2.4.2
    \item \(p_i(t)\) Forward Travelling Wave Pressure Amplitude in \(i\) th Element at Time \(t\) (in Pa) 2.4.3
    \item \(p(0), \ldots, p_{n-k}\) Coefficients of Polynomial \(P_k(w)\) 5.4.1.1
    \item \(q\) Dimension of Column Vector \(f_1\) 2.3.5.3
    \item \(q(0), \ldots, q_{n-k}\) Coefficients of Polynomial \(Q_k(w)\) 5.4.1.1
    \item \(r(t)\) Reflectance as a Function of Time (No Units) 2.4.2
    \item \(r_d(t)\) Direct Reflection from Farthest Junction (in Pa) 2.4.4
\end{itemize}
$r_m(t)$  Multiple Reflections from Intermediate Junctions (in Pa)

$r_i$  Reflection Coefficient at $i$th Element  2.4.3

$r_k$  Reflection Coefficient at $k$th Element  5.4.1.1

$r_n$  Reflection Coefficient at the End of the Duct  5.4.1.1

$s$  Dimension of Column Vector $s^f_\theta$

$s(k)$, $s(k)$  Coefficients of Polynomial $S_k(w)$  5.4.1.3

$\theta$  Time (in s)

$t$  Time Delay (in s)

$u(t)$  Volume Velocity Variation with Time (in $m^3 s^{-1}$)  2.2.3

$u(x,t)$  Volume Velocity Variation with Distance and Time (in $m^3 s^{-1}$)

$v$  Velocity of Impulse in Duct (in $ms^{-1}$)  5.4.1.1

$v$  Time Delay (in s)  2.3.4.2

$v_1, v_2$  Two Random Variables  3.2.5

$w$  $z^{-2}$, where $z$ is the z-Transform Index  5.4.1.1

$x$  Arbitrary Function  2.3.3

$x$  Cartesian Coordinate  2.3.4.3

$x$  Distance along Axis (in m)  2.4.2

$x_1$  Measured FPSN  3.2.4.2

$x(i)$  Arbitrary Sequence  4.3.4

$y$  Cartesian Coordinates  2.3.4.3

$y(i)$  Arbitrary Sequence  4.3.4

$z(t)$  Input Impedance in the Time Domain  2.2.3

$z$  z-Transform Index  5.4.1.1

$A$  Cross-sectional Area (in $m^2$)  3.1.3.5

$A_i$  Cross-Sectional Area of $i$th Element (in $m^2$)  2.4.3
$A(x)$

Cross-Sectional Area as a Function of Axial Distance, $x$ (in $m^2$)

$A_1(f), A_2(f)$

Amplitude Spectra

$A(\omega)$

Spectral Amplitude

$B(\omega, \sigma)$

Fourier Transform of $b(x,y)$

$C$

Constraint Operator

$C_1, C_2$

Arbitrary Constants

$C_{ff}(t)$

Autocorrelation Function of $f(t)$

$C_{fg}(t)$

Cross Correlation Function of $f(t)$ and $g(t)$

$C_{xy}(k)$

Sample Cross Correlation Function of $x(i)$ and $y(i)$

$D$

Arbitrary Operator

$F^{-1}$

Inverse Fourier Transformation

$F(\omega)$

Fourier Transform of $f(t)$

$F$

DFT Expressed as a Column Vector

$F_0, F_1, F_2$

Column Vectors which Together Make up Column Vector $F$

$L_1$

Latter Part of Discrete Response in Matrix Form

$F_p(k)$

DFT of $f_p(i)$

$z^{-1}$

$z$-Transform of $F_{k}^{F,B}(m)$

$F(\omega, \sigma)$

Fourier Transform of $f(x,y)$

$G$

Constant Related to Attenuation in a Cylinder

$G(\omega)$

Fourier Transform of $g(t)$

$G_p(k)$

DFT of $g_p(i)$

$z^{-1}$

$z$-Transform of $G_{k}^{F,B}(m)$

$G(\omega, \sigma)$

Fourier Transform of $g(x,y)$

$H$

Known Distortion Operator
\( \text{H}(\omega) \) Fourier Transform of \( h(t) \) 2.2.3

\( \text{H}_0(z) \) z-Transform of Reflectance of Whole System 5.4.1.1

\( \text{H}_k(z) \) z-Transform of Reflectance of Latter Part of Duct 5.4.1.1

\( \text{H}_p(k) \) DFT of \( h_p(i) \) 2.3.6.1

\( \text{H}(\omega,\sigma) \) Fourier Transform of \( h(x,y) \) 2.3.4.3

\( \text{H}_0 \) Harris Spectral Weighting Function 2.3.4.3

\( \text{I} \) Identity Operator 2.3.3

\( \text{I} \) Identity Matrix 5.2.4.1

\( L \) Number of Elements in Arbitrary Sequence \( y(i) \) 4.3.4

\( M \) Subscript Index within Discrete Convolution Matrix 5.2.4.1

\( N \) Number of Points to be Fourier Transformed 2.3.2.1

\( N \) Dimension of Column Vectors \( f \) and \( F \) 2.3.5.3

\( N \) Dimension of DFT Square Matrix, \( \alpha \) 2.3.5.3

\( N \) Number of Successive Sequences of \( m \) Random Numbers 3.2.5

\( N \) Number of Elements in Region of Interest (ROI) 4.2.3

\( N \) Number of Elements in Arbitrary Sequence, \( x(i) \) 4.3.4

\( N(\omega,\sigma) \) Fourier Transform of \( n(x,y) \) 2.3.4.3

\( P \) Arbitrary Operator 2.3.3

\( P(f) \) Complex Pressure as a Function of Frequency (in Pa?) 2.1.2.1

\( P_k(w) \) Numerator Polynomial when \( \text{H}_k(z) \) is Expressed as a Rational Function, \( \text{H}_k(z) = \frac{P_k(w)}{Q_k(w)} \) 5.4.1.1

\( Q_k(w) \) Denominator Polynomial (as above) 5.4.1.1

\( R(\omega) \) Fourier Transform of Reflectance, \( r(t) \) 6.5.3

\( R(f) \) Reflection Coefficient at Input as a Function of Frequency 2.4.4

\( S_k(w) \) A Polynomial \( (d_k^2 P_k(w)) \) 5.4.1.3
Sampling Period (in s)  2.4.3
Temperature (in °C)  3.2.4.3
Temperatures (in °C)  4.2.4.3
Complex Volume Velocity as a Function of Frequency
(in m s⁻¹)  2.1.2.1
Input Impedance (in Ω, i.e. kgm⁻⁴ s⁻¹)  2.1.2.4
Complex Input Impedance (in Ω)  2.1.2.1
Input Impedance in the Frequency Domain (in Ω)  6.5.3
Characteristic Impedance (in Ω)  2.1.2.4
Load Impedance (in Ω)  2.1.2.4
Source Impedance (in Ω)  2.4.4
Attenuation Coefficient (in m⁻¹)  3.1.3.5
δ - h  2.3.3
Propagation Coefficient (in m⁻¹)  2.1.2.4
Unit Impulse  2.3.3
Delta Function  2.3.4.2
Error Squared  2.3.3
Feedback Coefficient  2.3.3
A Constant  3.1.3.5
Wavelength (in m)  2.4.3
Eigenvalue of Matrix F  5.2.4.1
Parameter to Control Rate of Convergence  2.3.3
A Constant  5.2.4.1
Spatial Frequency  2.3.4.3
Delay in Time (in s)  2.2.3
Phase of Spectrum (in rad)  2.3.4.1
Angular Frequency (in rad s⁻¹)  2.2.3
$\omega$  Spatial Frequency  2.3.4.3

$\Gamma$  A Constant  2.3.4.2

$\Delta$  Length of Cylindrical Elements (in m)  5.4.1

$\Pi$  Product  2.4.4

$\Sigma$  Summation  2.2.3

$\Omega$  DFT Matrix, Square Matrix of Dimension N  2.3.5.3

$\Omega_{11}, \Omega_{12},$ etc.  Partition Matrices of $\Omega$  2.3.5.3

$^\dagger$  Complex Conjugation  2.3.5.3

$^*$  Convolution  2.2.3
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADC</td>
<td>Analogue to Digital Converter</td>
</tr>
<tr>
<td>B and K</td>
<td>Brøel and Kjaer</td>
</tr>
<tr>
<td>BCD</td>
<td>Binary Coded Decimal</td>
</tr>
<tr>
<td>DAC</td>
<td>Digital to Analogue Converter</td>
</tr>
<tr>
<td>DFT</td>
<td>Discrete Fourier Transform</td>
</tr>
<tr>
<td>FFT</td>
<td>Fast Fourier Transform</td>
</tr>
<tr>
<td>FPSN</td>
<td>Fractional Peak Sample Number</td>
</tr>
<tr>
<td>G-P Algorithm</td>
<td>Gerchberg-Papoulis Algorithm</td>
</tr>
<tr>
<td>ISP</td>
<td>Inter-Sample Points</td>
</tr>
<tr>
<td>PDCS</td>
<td>Projections onto Convex Sets</td>
</tr>
<tr>
<td>PSN</td>
<td>Peak Sample Number</td>
</tr>
<tr>
<td>QA</td>
<td>Quality Assurance</td>
</tr>
<tr>
<td>rms</td>
<td>Root Mean Square</td>
</tr>
<tr>
<td>ROI</td>
<td>Region of Interest</td>
</tr>
<tr>
<td>S-N Ratio</td>
<td>Signal-to-Noise Ratio</td>
</tr>
<tr>
<td>SVD</td>
<td>Singular Value Decomposition</td>
</tr>
<tr>
<td>VG</td>
<td>Valve Group</td>
</tr>
</tbody>
</table>
Dedicated to my parents
With trumpets and the sound of the horn
make a joyful noise before the King, the Lord!

Psalm 98, Verse 6
CHAPTER 1
INTRODUCTION

A better instrument is the goal of both the musical acoustician and the instrument manufacturer, not to mention the player and the listener. In this thesis ways in which progress towards this goal may be made for brass wind instruments are considered.

At the outset it is necessary to consider what is meant by a "better instrument". Many different factors determine how good an instrument is. Edwards (1978) asked over 120 professional trombone players which features they considered to be of critical importance for a high quality instrument. For convenience, these are divided into two categories - first, musical features and secondly, non-musical features.

Musically, the timbre is considered most important. It should be pleasing and the instrument should be such that the player can vary the timbre to suit different types of music if he so desires. Good "responsiveness" or "ease of blowing" is also necessary so that the player can start and change a note easily and swiftly. The intonation should be good, although good players can compensate for intonation deficiencies. Note production should be consistent, and there should not be certain notes ("duff notes") which do not sound properly. Investigation of these musical qualities is the concern of the acoustician.

Important non-musical features include smoothness of slide and valve action, balance, comfort, weight, appearance and price. These qualities are the concern of the manufacturer.

For the purposes of this thesis, a better instrument will be considered to be one in which one or more of the above musical qualities has improved. To determine whether an improvement has occurred, some method of subjective assessment must be used. Pratt (1978) has carried out pioneering work in this field. The manufacturer has an additional concern. He needs to be able to
assure customers of the consistency of given instrument models. At present this cannot be guaranteed.

The evolution of a better instrument will be facilitated by rationally-based methods of designing and developing new instruments. Traditional design procedures have been largely empirical rather than scientific, in that modifications have been made using "craftsmen's intuition" instead of an accurate understanding of how the acoustics will be affected.

The major influence on the acoustics of an instrument is the shape of its internal bore. An understanding of the direct relationship between the bore geometry and the acoustics would greatly enhance development procedures, but such a relationship is not easy to establish. Instead, attempts have been made to forge links by way of an intermediate, physically-measurable quantity.

Early attempts to assess brass instruments physically are reviewed in Chapter 2. Initially these involved the measurement of resonance frequencies (Webster, 1947). Later Kent (1956) built an impedance measuring device, and input impedance then became the focus of attention for many years. Many workers have been able to compute numerically the input impedance of a given bore shape. The link between input impedance and intonation is now also well understood (Wogram, 1972). However, the relationship between input impedance and other subjective qualities is not so clear (Pratt, 1978), and so the link between the bore and the acoustics remains incomplete.

Returning to the manufacturer's requirement of guaranteed consistent high quality of instruments, the method of physical assessment of instruments should be such that small inconsistencies between instruments can be detected. There are indications that input impedance is not helpful in this respect (Pratt, 1978, Section 4.9.4). Therefore, a more suitable method is needed.

In this thesis, the use of transient response as an alternative (or complement) to input impedance is proposed. The aims of the research may be outlined as follows:
(1) to devise a means of measuring accurately an approximation of the transient response (Chapter 3);

(2) to develop a method whereby measured transient responses are used to compare instruments and detect inconsistencies between them (Chapter 4);

(3) to process the measured transient response in order to obtain the true transient response, that is, the response to a Dirac delta function (this process is termed deconvolution) (Chapter 5);

(4) to investigate the direct and inverse relationships between transient response and bore geometry (Chapter 5);

(5) to examine links between transient response and subjective quality (Chapter 6).
CHAPTER 2
REVIEW OF THE FIELD OF RESEARCH

This chapter consists of a review of the relevant literature. Section 2.1 contains a general summary of brass instrument research, concentrating mainly on frequency response.

In Section 2.2, transient measurements are discussed. These must be processed in order to extract the true transient response, and various methods for doing this are explained in Section 2.3.

Finally in Section 2.4, ways of calculating the profile of a duct from its transient response are examined; such a relationship provides a valuable link between the bore of an instrument and its acoustical properties.

2.1. Brass Instrument Research in the Frequency Domain

Most investigations on brass instruments to date have involved the frequency domain. Therefore, although the bulk of this thesis is devoted to time domain work, the frequency domain research in the field must first be reviewed.

2.1.1. Wave Motion in Brass Instruments

A brass instrument consists of an air column of varying cross-sectional area, confined by tubing. This tubing has distinct sections, starting with the mouthpiece, then the main bore and finally the bell section (or flaring part). The mouthpiece acts as a coupler between musician and instrument. Brass instruments are sometimes referred to as cup-mouthpiece instruments.

Oscillations are set up in the instrument when the player blows into it and his lips vibrate within the mouthpiece cup. The frequency of these oscillations is controlled by the tension and inertia of the player's lips. Most of the resulting pressure wave is reflected back when it encounters the marked change in wave impedance at the instrument bell; frequencies above a certain limit are radiated out of the bell. At certain frequencies, called resonance
frequencies, the returning and incoming waves reinforce each other and a standing wave is set up; here, large pressure variations in the mouthpiece actually aid the lip vibration, causing "self-sustained oscillations" (a term coined by Benade (1973)). At other frequencies, the reflected wave is completely out of phase with the incident wave, producing a minimum response. In a later section (2.1.2.5), the frequencies at which resonance occurs [*] are examined more closely. This model of instrument behaviour is termed the pressure-controlled lip-valve model.

For convenience, wave propagation in brass instruments is normally assumed to be planar. A.G. Webster (1919) made this simplifying assumption (and various others) when he derived the well-known Horn Equation. It is known that waves travelling in a cylinder are plane and those travelling in a cone are spherical. However, the precise shape of the wavefront in a rapidly flaring tube cannot be calculated analytically, because of the non-separability of the Wave Equation. Benade and Jansson (1974) undertook a close study of this problem. They reduced Webster's Horn Equation so that it became identical in form to Schrödinger's Time Independent Equation, and postulated the existence of a "horn function barrier" in the flaring part of the instrument. They used this barrier to estimate various horn properties; the predicted horn behaviour varied according to which wavefront shape was assumed - plane or spherical. After comparing predicted and measured instrument behaviour, Benade concluded that the true wavefront shape at the bell lay somewhere between plane and spherical because of "mode conversion" in the bell region. Further investigations of the wavefront have been attempted at the University of Surrey by measuring the isophase contour just outside the instrument bell (Downes, 1980; Lewis, 1981).

[*] It should be noted that the intervals between resonance frequencies are too large (up to five tones) for musical purposes; the notes between them are produced by lengthening the instrument using valves or slides.
The accuracy of some of Benade's theoretically predicted results could be questioned. He used Euler's Method for a step-by-step numerical investigation; yet Euler's Method involves only a first order approximation and can have a sizable cumulative error. More accurate methods exist, eg. Euler's Improved Method or the Runge-Kutta Method (see, for example, Peckham, 1971).

Ranade and Sondhi (unpublished work, referred to in Sondhi and Resnick, 1983) have also looked at the effect of non-planar wave motion. They used a finite-element technique to solve a two-dimensional wave equation for an axially symmetric tube of variable cross-sectional area.

2.1.2. Physical Assessment of Instruments using Input Impedance

2.1.2.1. Introduction

A means of quantifying the acoustical behaviour of brass instruments is essential if ways of improving instruments are to be found. The existence of resonance frequencies suggests that measurement of the frequency response of an instrument will be useful. Early resonance measurements involved excitation of the instrument at the mouthpiece and measurement of pressure variation at the bell end, eg. J.C. Webster (1947), Carmichael (1968). R.A. Smith and Daniell (1976) measured resonance frequencies by placing a microphone in the mouthpiece and comparing the phases of the incident and reflected waves.

Measuring input impedance as a function of frequency was subsequently discovered to be a more satisfactory way of assessing instrument response. Input impedance, \( Z(f) \) in \( \text{kg m}^{-4} \text{s}^{-1} \) is the complex ratio of pressure, \( P(f) \) to volume velocity, \( U(f) \).

\[
Z(f) = \frac{P(f)}{U(f)} \tag{2.1}
\]

When measuring input impedance, both excitation and measurement take place at the mouthpiece end. The development of methods of measuring input impedance will now be examined.
2.1.2.2. Measurements with Constant Input Volume Velocity

When volume velocity is kept constant in the mouthpiece, a measurement of pressure response will be proportional to the magnitude of the input impedance. One way of maintaining constant volume velocity is to excite the instrument by feeding a sinusoidal loudspeaker signal of controlled constant amplitude through a fine cylindrical capillary into the mouthpiece cup. A probe microphone within the cup measures the amplitude of the resulting pressure variations. As the frequency of the loudspeaker signal is varied, a chart recorder plots the amplitude of the pressure response against frequency. The first such system was built by Kent (1956) at Conn Ltd. Since then there have been various developments of the system. Coltman (1968) and Merhaut (1969) maintained constant volume velocity by monitoring the motion of a diaphragm which was directly connected to the loudspeaker driver. Wogram (1972) and Benade (1973) have built systems similar to Kent's. Backus (1976) used a system identical to that of Kent except that he used a specially devised annular capillary to connect the loudspeaker to the instrument. Earlier, Backus and Hundley (1971) had used an impedance tube method. More recently, Caussé et al (1984), Krüger (1984), Agulló and Badrinas (1985), and Kergomard and Caussé (1986) have used methods similar to that of Backus.

The inventors of the STL ionophone (Fransson and Jansson, 1978) claim that it is a constant volume velocity source, and Dekan (1974) has used it for brass instrument resonance measurements. However, Gant (1983) found that it is not, in fact, a constant velocity source.

All the above methods are limited by the fact that volume velocity is not measured directly.

2.1.2.3. Direct Measurement of Velocity using Anemometer

A major step forward was the establishment of an acoustic impedance measurement system at the University of Surrey by Pratt (1978), with later developments by Elliott (1979). In this system, acoustic pressure and particle
velocity are measured simultaneously using a probe microphone and hot-wire anemometer respectively, allowing calibrated data for the acoustic impedance, both modulus and phase, to be obtained. The plane of measurement is the throat of the mouthpiece. A minicomputer controls the apparatus, records the pressure and velocity data, and calculates the input impedance from it. Careful calibration procedures are carried out (Pratt, Elliott and Bowsher, 1977).

The Surrey system has several advantages over other systems. One is the well-defined plane of measurement. Other investigators have used the mouthpiece rim as the plane of measurement. Here the volume of air displaced by the players lips was neglected. Usually the microphone was just beyond the rim, leading to an effective increase in mouthpiece cup volume. Having the plane of measurement at the throat of the mouthpiece eliminates the effect of the mouthpiece cup, but the measured impedance can be modified to include it by modelling the mouthpiece cup as a lumped element (Elliott, 1979, Section 2.4.7).

The microphone and anemometer are in the same plane, meaning that measured impedance is independent of what is happening upstream (Elliott, Bowsher and Watkinson, 1982).

Keefe and Benade (1981) discuss problems of source and microphone proximity effects. The source produces a local disturbance caused by evanescent modes, and the microphone must be carefully placed to avoid these. Such problems do not arise in the Surrey system.

2.1.2.4. Comparison of Computed and Measured Input Impedance Values

The ability to predict the input impedance of a given bore configuration would be useful, particularly in the area of improving the design of brass instruments. (Instrument design is discussed further in Chapter 6.) The problem is not straightforward because of the complicated shapes of brass instruments.
The input impedance of a simple straight-sided cylinder can be predicted by treating it as a 2-port network and using standard transmission line theory. Here, one must assume a linear system and plane wave propagation. For a lossy cylinder of length $l$,}

$$Z = Z_c \left[ \frac{Z_1 + Z_c \tanh(\gamma l)}{Z_c + Z_1 \tanh(\gamma l)} \right]$$

(2.2)

Here, $Z$ is input impedance, $Z_c$ is characteristic impedance, $\gamma$ is propagation coefficient, and $Z_1$ is load impedance. Pratt (1978, Section 2.15) has compared measured and theoretical input impedances for a simple cylinder. In his case, measured values were higher, but he had neglected to take account of the presence of the mouthpiece in his theoretical formulation. Watkinson (1981) also attempted such a comparison; an interesting consequence of his work was the development of a technique for determining acoustic losses at tube walls.

To derive the input impedance of a brass instrument, one has to approximate it as a series of simple geometrical shapes. The air column of the instrument is linear, and any linear system can be represented as a 2-port network. The whole 2-port network can be broken down into a succession of simpler 2-port networks. One starts at the load end (bell) of the instrument; the input impedance of each element is calculated and used to load the next element until the mouthpiece is reached.

Much work has been done in this area. The fundamental elements have taken on various shapes. Young (1960, 1966) used cylindrical elements when calculating the input impedance of smooth horns with no side holes or losses. Plitnik and Strong (1979), when finding the input impedance of an oboe, used cylindrical elements to approximate the conical bore, and treated each as a lossy transmission line. Goodwin (1981) termed this process the "Look Back Method". His theoretical results were comparable with measured values, the greatest source of error being temperature variation - a factor not normally referred to in the literature. Wogram (1972, Section 3.1) has made measurements which
confirm that ambient temperature variations have a significant effect on instrument intonation.

Some recent researchers have used truncated cones as the fundamental elements - in particular, Caussé et al (1984). Wogram (1972) constructed an experimental instrument consisting of a cylindrical tube and an exponential horn, and compared measured and computed input impedance.

All these techniques have served to predict input impedance to a reasonable degree of accuracy. In this way the acoustics of a given bore shape can be anticipated to a certain extent before any acoustical measurements are made.

2.1.2.5. Relationship between Input Impedance and Playing Frequencies

To examine this relationship, it is necessary to look more closely at how the player interacts with the instrument. Helmholtz (1895) suggested a linear interaction whereby the lip-valve collaborated with a single impedance maximum to maintain a stable oscillation, and each vibrational mode operated independently of other modes. One would therefore expect that the easiest frequencies to play would correspond to the tallest of the impedance maxima. This does not necessarily happen.

Benade and Gans (1968) pointed out that the equation governing the flow through the lip orifice is non-linear. This leads to a non-linear interaction between the lip-valve and the air column, whereby several impedance maxima collaborate with the lip-valve mechanism. Benade calls this a "regime of oscillation", and the resulting stable oscillation contains several harmonically-related components. Therefore, values of impedance at the harmonics of the fundamental frequency are also important.

Worman (1971) analysed this non-linear interaction for clarinet-like systems. He found that as the dynamic level of a note increased, the harmonics made up an increasing proportion of the sound, and that the note produced was more stable. For quiet playing levels, the linear theory of Helmholtz holds
true.

Fletcher (1979) and Elliott (1979) have both re-analysed the regeneration process. Elliott made some simplifying assumptions, and his model was based on a model of speech production. He found that the relative magnitudes of the "average" resistance of the lip opening and the input impedance of the instrument played a large part in determining the harmonic structure of the note produced.

It is worth mentioning at this point the similarities which exist between the vocal tract and brass instruments. Measurements of the input impedance of the vocal tract at the lips have been made - see, for example, Schroeder (1967) - and the relevance of the work on vocal tract area function recovery to the present work on brass instruments will be seen later (Section 2.4).

Wogram (1972) introduced the now well-known "Sum Function", calculated by summing the real parts of the impedance at the fundamental frequency and its integral multiples (up to a certain maximum frequency). He postulated that frequencies of Sum Function maxima, that is, frequencies of maximum energy transfer from the instrument to the surroundings, corresponded to playing frequencies. His measurements confirmed this quite well. However, Pratt (1978) found Wogram's theory deficient as it was not satisfied by cylindrical tubes, and suggested a refinement whereby the dynamic level of the played note was taken into account.

2.1.3. Subjective Assessment of Instruments

In addition to physical measurements, another equally important way of assessing brass instruments is to find out what the listener or player himself thinks. If relationships between subjective qualities and bore shape or some physically measurable acoustical properties could be established, this could well pave the way towards the design of better instruments [*].

[*] For explanation of "better", see Chapter 1.
This aspect has received little attention. The only systematic subjective studies carried out on brass instruments have been at the University of Surrey (Pratt, 1978; Pratt and Bowsher, 1978; Pratt and Bowsher, 1979), and in Czechoslovakia (Melka, 1981).

At Surrey, a way of quantitatively evaluating subjective quality has been sought. A survey (Edwards, 1978) identified the most important factors governing instrument quality as timbre and responsiveness, and Pratt (1978) developed Semantic Differential Scaling (SDS) and Multidimensional Scaling (MDS) techniques for assessing these qualities. R. A. Smith (1984) describes how Pratt's techniques have been used to assist greatly in the development of a new Boosey and Hawkes cornet. Their usefulness was also demonstrated during an investigation into an intonation fault in a bass trombone (Pratt, Bowsher and Smith, 1978).

Pratt also sought to relate subjective quality to some physically measurable attribute, in this case input impedance. However, no clear relationship was found; Bowsher noticed that in certain cases the gradients of lines joining the peaks of the input impedance curves can be used to make a fair prediction of the responsiveness and richness of tone of trombones (see Pratt, 1978, Section 4.7); but the physical basis of such a relationship is not apparent. Further discussion of the relationship between physical and subjective characteristics will be given in Section 6.5.

In recent years, it has come to light that the player himself can exert a large influence on the behaviour of an instrument. Therefore, in the latest work by Bowsher (1983, 1984) and Simpson (1983) the emphasis has changed from making accurate physical measurements of the instrument alone to studying the ways in which the sound production can be controlled by the player. A complicated model of the control structure governing the player/instrument interface has been formulated by Bowsher (1983). However, this thesis has concentrated solely on the instrument, rather than the player/instrument combination.
2.1.4. Summary of Section 2.1

In this section, the research done on brass instruments has been briefly reviewed. The non-linear valve mechanism and the existence of resonance frequencies has been described. The problem of the wavefront in flaring regions has been examined. The useful quantity of input impedance has been discussed. Many attempts have been made to measure it and calculate it. The influence of the non-linear regeneration mechanism on the relationship between input impedance and playing frequencies has been seen. The subjective quality of brass instruments and its relationship with physical characteristics is an important area which has received little attention to date.

2.2. Time Domain

2.2.1. Introduction

Time domain measurements of system response have become increasingly popular in recent years. The measurements involved are simpler and less time-consuming than those in the frequency domain. In addition, they may be cheaper than steady-state swept-sine measurements, since anechoic facilities are not necessary. Fast data acquisition, storage and analysis have been made possible by the availability of fast, low-cost digital computers.

In a recent review of transient measurements by Downes (1985), the three main types of transient measurements were identified as impulse measurements, tone burst measurements, and time delay spectrometry. Here the emphasis is specifically on impulse measurements, that is, measurements in which the test signal is an impulse. Before discussing the quantity of "impulse response" in more detail, some examples will be given of areas in acoustics where time domain techniques have been used.

Impulses have been used in acoustic measurements for well over two centuries. Some of the earliest experiments involving impulses were carried out by members of the Paris Academy in 1738; canons were fired, and the retardation
of the reports at different distances was used to determine the velocity of sound in free air (see Barton, 1908). In 1855, Thomas Young was reported to have used primitive impulses (explosions and stamping of feet) to investigate the acoustics of passages and rooms (Wood, 1940). Barton (1908) noted the phenomenon of the "musical echo"; a single impulse incident on a number of equidistant parallel reflecting surfaces may be converted into a succession of reflected pulses, which is perceived as a short musical note. This can be observed when walking past an overlapping wooden fence, or clapping hands in front of a flight of steps. Richardson (1927) reported that in 1921, Langevin introduced ultrasonic echo-sounding as a means of determining the depth and nature of the sea-bed.

In architectural acoustics, impulsive signals have been used to find the impulse responses of auditoria. Other useful quantities such as the frequency response and reverberation time can be derived from the impulse response (Bar­ron, 1984). Barron discusses suitable forms of impulse signal. Impulse sources have included pistols, balloon explosions, spark sources and loudspeakers driven by a pulse generator.

Before low-cost digital computers were so readily available, Louden (1971) was using a simple single-pulse technique to measure transmission properties of various acoustic systems. His approach was to photograph oscilloscope traces of the input and response signals, and compare their calculated Fourier components.

To overcome problems of inadequate signal-to-noise (S-N) ratio, some investigators, for example, Berkhout et al (1980) proposed that the source be a loudspeaker driven by a long-duration computer-generated signal containing all frequencies of interest. The technique of using the computer to generate a desired impulse shape has been tried for various applications. Aoshima (1981) has used an extension of this technique to measure speaker response, sound reflection and sound attenuation. The impulse shape can be synthesised in such
a way that the frequency response of the acoustic transducer is taken into account (Davies et al, 1981; Salikuddin et al, 1984).

Berman and Fincham (1977) have obtained the frequency response of loudspeakers by direct measurement of impulse response. Downes and Elliott (1985) have developed a similar technique to evaluate responses of microphones. Merhaut (1986) has used impulse measurements to determine the frequency response and electrical impedance of a loudspeaker driver loaded with an "ideal" infinite horn.

Impulse measurements have been carried out on the vocal tract to determine its area function (Sondhi and Resnick, 1983; Descout et al, 1976). These will be discussed in more detail in Section 2.4.

2.2.2. Time Domain Work on Wind Instruments

Some of the first published work on impulse measurements on brass instruments appears to be that of Krüger (1979). His aim was to diagnose response deficiencies of various instruments in terms of the reflections from within the instrument, measured at the input. Impulses were generated using an STL ionophone (Fransson and Jansson, 1978) placed in the mouthpiece. The reflected signal was measured using a 1/4" microphone placed in the mouthpiece. In a series of simple measurements, Krüger showed how small modifications to the instrument affected the reflection pattern, for instance, a hole in the bore of less than one millimetre introduced an extra reflection of large amplitude. The strength of the reflection from the bell end was shown to vary depending on the shape of the bell. Players found that with instruments whose bells produced weak bell reflections, it was difficult to maintain a steady tone. More recently, Krüger (1984) has returned to assessing instrument response using frequency domain measurements.

At the University of Surrey, Pratt (1978) and Elliott (1979) first approximated the impulse response by inverse Fourier transforming their frequency domain measurements of input impedance and transfer function. Crude
initial measurements of impulse response were made by slapping the mouthpiece of the attached instrument with the palm of the hand, measuring pressure variation in the mouthpiece, and displaying the result on an oscilloscope. Wogram (1972, Section 2.2.1.3) had used a similar technique, but measured the response at the bell. Later, more serious attempts to measure the impulse response were made by Goodwin (1981), who designed and built a simple computer-controlled apparatus (Goodwin and Bowsher, 1980). Here, a spark source introduced an impulse into the instrument mouthpiece and a probe microphone in the mouthpiece detected the reflected signal. The amplified microphone signal was sampled and stored in the computer in digitised form. Different features from within the instrument bore could clearly be identified from Goodwin's data. His apparatus has since undergone several refinements which will be discussed in Chapter 3. Goodwin proposed the use of various signal processing techniques on the resulting measurements. He also attempted to use his data records to reconstruct the instrument bore shape using techniques previously used for vocal tract area function calculation. The extent to which his attempts at reconstruction were successful will be discussed later in Section 2.4.5. Chapter 5 of this thesis is devoted to a thorough analysis of the same subject.

Benade and J.H. Smith (1981) made pulse reflectometry measurements of tubas. One motivating factor was that Benade (1976) had previously proposed that premature bore reflections damaged tone onset, making the instrument hard to blow or "stuffy". Their experimental apparatus also incorporated a spark source at the mouthpiece rim, with an electret microphone to pick up the reflected signal. Low-pass filters reduced electrical noise and the signal was fed into a real time analyser. Measurements revealed a poor impulse shape caused by mouthpiece ringing and reflections within a small cavity between the impulse source and the mouthpiece. An FFT was applied to the time record of a tuba, and Benade and Smith assumed that this resulted in the input impedance.
Section 2.2.3 will show that this is not necessarily so. Benade and Smith found that the positions of the first few impedance peaks corresponded well with playing frequencies at quiet playing levels; this agrees with Helmholtz theory (Section 2.1.2.5). They compared their frequency responses with the tuba input impedance response measurements of Caussé et al (1984). The comparison was reasonable except for errors in Benade's results at higher frequencies caused by a defect in the synchronisation of the apparatus.

Benade later built a new version of his impulse apparatus using low-cost components (Ibisi and Benade, 1982). The spark source was replaced by a small piezoelectric disk bonded to the end of a 2-centimetre section of tube. The equipment was used to determine the input impedance of the vocal tract (Benade, 1985). The impulse was fed into the vocal tract and the resulting pressure response was Fourier transformed. Here, Benade was investigating the effect of the player's vocal tract on the playing properties of wind instruments (Benade and Hoekje, 1982).

Recently, Ishibashi and Idogawa (1986) have reported measurements of the input impulse response of bassoons.

McIntyre, Schumacher and Woodhouse (1983) have published an important paper in which they formulate a simple time domain model which describes the behaviour of non-linear musical oscillators. The model is versatile and can be extended to many musical instruments. Their investigations so far have involved the clarinet, violin and flute. Using a computer, mouthpiece pressure and volume flow rate can be simulated in the time domain, given the non-linear flow characteristic and the reflectance (defined in Section 2.2.3). The general model neglects reed dynamics, but these have been included in an earlier paper by Schumacher (1981), and would be important if the model were to be applied to brass instruments. McIntyre et al have suggested that accurate measurements of reflectance (and non-linear characteristics) are needed to extend the use of their model. In the light of this, Hoekje and Benade (1986)
have recently reported work on the determination of the reflectance of
cylinders and musical instruments.

This section has shown that through the work of several investigators,
considerable progress has been made in the realm of time domain work on wind
instruments - both experimental and theoretical. The precise meaning of
impulse response is discussed in the next section.

2.2.3. Impulse Response

For linear, causal, time-invariant systems, as defined in Rabiner and
Gold (1975, Sections 2.4 and 2.5), the impulse response is generally defined
as the response when the input is the Dirac delta function or ideal impulse.
The impulse response, h(t) provides a complete description of the response of
the system in that it can be used to determine the response, g(t) for any
given input, f(t). The following convolutional relationship exists.

\[ g(t) = f(t) * h(t) \tag{2.3a} \]

that is [\*]

\[ g(t) = \int f(t-\tau) h(\tau) \, d\tau \tag{2.3b} \]

where \( \tau \) denotes time delay. For finite discrete signals, this integral can be
approximated by the summation,

\[ g(n) = \sum_{i=1}^{n} f(n-i+1) h(i) \tag{2.4} \]

By the Convolution Theory, convolution in the time domain is equivalent
to multiplication in the frequency domain. Therefore, if \( F(\omega) \), \( G(\omega) \) and \( H(\omega) \)
are the Fourier transforms of \( f(t) \), \( g(t) \) and \( h(t) \) respectively, then

\[ G(\omega) = F(\omega) H(\omega) \tag{2.5} \]

Any definition of impulse response for a brass instrument must be made
with care. For clarity, it is necessary to specify input signal, input

[\*] Here, "\*" denotes convolution
condition, and position at which the response is measured. Three descriptions of impulse response have been given.

(1) **Inverse Fourier transform of transfer function.**

Defined in the frequency domain, the transfer function is the complex ratio of pressure seen at the output of the instrument to pressure at the input. For measurements made at the University of Surrey, the output position is taken to be one bell radius from the end of the instrument, and the input position is the throat of the mouthpiece. Elliott (1979) defined the impulse response as the pressure seen one bell radius from the end of the instrument when a perfect delta-function of pressure is generated at the throat and pressure maintained at zero thereafter, i.e. the input is terminated with zero impedance (open end). Elliott measured the transfer function in the frequency domain, windowed it and inverse Fourier transformed to obtain the impulse response.

The other two descriptions involve response measurements at the input of the instrument rather than the output.

(2) **Inverse Fourier transform of the input impedance.**

From the definition of input impedance given in equation (2.1) (Section 2.1.2.1),

\[
Z(f) = \frac{P(f)}{U(f)}
\]

\[
P(f) = Z(f) U(f)
\]

and the Convolution Theorem, it may be seen that input pressure as a function of time can be expressed as the convolution of the inverse transform of input impedance and input volume velocity, i.e.

\[
p(t) = z(t) \ast u(t).
\]

\[z(t)\] must therefore be the pressure at the input as a function of time in response to a delta-function of volume velocity applied at the input. Various
workers have used an open circuit input condition, as opposed to the short circuit condition of the previous transfer function case [*]. Elliott (1979) and Goodwin (1981) both state that the input should be rigidly terminated with an infinite impedance, that is, any pressure will be totally reflected in phase. Goodwin states that "the input must be open-circuit or no signal will be seen". Schumacher (1981), in his time domain work on the clarinet, also defines a certain "impulse response function" as the Fourier transform of the input impedance, and seems to imply that for this function, the mouthpiece behaves as a closed end. Certainly, in Benade's time domain measurements of vocal tract input impedance (Section 2.2.2), the source condition was like that of an open circuit, because of the flat piezoelectric impulse source. However, the input condition of Sondhi and Gopinath (1971a) was different; they measured pressure and velocity at the lips using an impedance tube which effectively removed source reflections.

Most definitions of input impedance do not include a clear specification of input condition for either the frequency or time domains. Section 2.1.2.3 mentioned that when independent frequency domain measurements of pressure and volume velocity are made at the input plane, the input impedance is independent of conditions upstream of that plane; so in that case, the clarification of input conditions is less important. However, input impedance in the time domain cannot be independent of source conditions, because source reflections clearly influence the pressure measured at the input. Thus, a precise specification of input conditions is essential.

Even when source conditions are assumed to be consistently open-circuit for measurements of input impedance in both time and frequency domains, it is difficult to understand the precise relationship between them in quantitative terms. In frequency, independent measurements are made at each frequency with steady-state conditions in which there is a continual interaction between

[*] Here, "open circuit" means the opposite of "open end".
incoming and outgoing waves. For a given frequency, the result will be the same over an infinite time period. However, in time domain measurements, a single brief measurement provides information over the entire frequency range.

Despite the difference between the methods of measuring $Z(f)$ and $z(t)$, both quantities contain the same information and are simply related by Fourier transformation.

(3) Reflectance

Here there is no ambiguity about source conditions. The source must be a perfect absorber, that is, the source impedance must equal the characteristic impedance at the instrument input. Reflectance $[\ast]$, defined in the time domain, is the pressure at the origin in the reflected (or left-going) wave when the pressure at the origin in the incident (or right-going) wave was a delta-function. Resnick (1980, p.34) defines reflectance as that generalised function which, when convolved with the input pressure waveform, yields the reflected pressure waveform. Schumacher (1981) points out that reflectance should diminish to zero more rapidly than input impedance (in time) because of the absence of source reflections. In practice, the reflectance of a brass instrument takes a long time to decay because of multiple reflections within the instrument.

Which of the above is the correct definition of impulse response is not clear; however, all three are acceptable definitions of transient response. Reflectance is the quantity which receives most attention in this thesis.

To determine the transient response from independent measurements of input signal and response (or output) signal, it is necessary to perform the process of deconvolution or inverse convolution. Various forms of this process will be discussed in the next section.

[\ast] Reflectance has also been termed "reflectivity function".
2.2.4. Summary of Section 2.2

In this section, time domain measurements of instrument response have been considered and been seen to be an acceptable alternative to frequency domain measurements. The utilisation of transient methods in many branches of acoustics to date was discussed. Time domain investigations of wind instruments have been attempted by various investigators; the simple measurements of Kräger, Goodwin, and Benade et al were reviewed. The theoretical model of McIntyre et al was described; this time domain model can be used to predict the acoustic behaviour of various types of instruments, including brass instruments. Finally, the problem of finding a suitable definition of impulse response for a brass instrument was discussed at length.

2.3. Deconvolution Methods

Some deconvolution methods which can be applied to discrete finite causal signals for a linear time-invariant system are surveyed. An analysis of the degree of success of such methods will be given in Chapter 5.

2.3.1. Time Domain Deconvolution

This method is based on a simple rearrangement of the terms within the convolution sum,

\[ g(n) = \sum_{i=1}^{n} f(n-i+1) h(i), \]

so that \( h \) can be calculated when \( f \) and \( g \) are known. The general principle can be understood by looking at the first few steps of the process.

\[ g(1) = f(1) h(1) \quad \text{so} \quad h(1) = g(1) / f(1) \quad (2.8a) \]

\[ g(2) = f(1) h(2) + f(2) h(1) \quad \text{so} \quad h(2) = \left( g(2) - f(2) h(1) \right) / f(1) \quad (2.8b) \]

Calculation of \( h(2) \) is now possible since \( h(1) \) is known.

\[ g(3) = f(1) h(3) + f(2) h(2) + f(3) h(1), \quad \text{so} \quad (2.8c) \]
\[ h(3) = \left( g(3) - f(2) \cdot h(2) - f(3) \cdot h(1) \right) / f(1) \]

As \( h(1) \) and \( h(2) \) are known at this point, \( h(3) \) can be calculated, and so on. Each successive value of \( h \) can be calculated using the \( h \)-values already found.

This is an unstable method of deconvolution (see Section 5.2.2 for examples of results). When initial values of \( f \) and \( g \) are very close or equal to zero, the resulting small denominators cause significant errors in the first few values of \( h \); these affect all later values of \( h \), causing the final solution to be one which has oscillations of ever increasing amplitude. The problem is therefore "ill conditioned" or "ill posed". The presence of noise in the measured signals also introduces error.

Sondhi and Resnick (1983) have attempted to overcome these instability problems by introducing a method of regularisation based on singular value decomposition. Another form of regularisation is used in Section 5.2.2.1.

2.3.2. Frequency Domain Deconvolution

Section 2.2.3 stated that convolution in time becomes multiplication in frequency, such that

\[ G(\omega) = F(\omega) \cdot H(\omega). \quad (2.5) \]

The frequency response is therefore

\[ H(\omega) = G(\omega) / F(\omega), \quad (2.9) \]

and this is the basis of frequency domain deconvolution. The measured input and response signals, \( f(t) \) and \( g(t) \) are Fourier transformed into the frequency domain. The spectra are divided and the result is inverse Fourier transformed back into the time domain. The process of obtaining the spectrum of a time signal involves two steps - windowing and a discrete Fourier transform.

2.3.2.1. Discrete Fourier Transform and Fast Fourier Transform

The discrete Fourier transform (DFT) is the Fourier representation of a sequence of finite duration, for a linear shift-invariant [*] system (Lynn, [*] Shift-invariant is to discrete signals what time-invariant is to continuous signals.)
It is not normally evaluated directly because of the existence of the Fast Fourier transform (FFT) set of algorithms. The FFT greatly increases the efficiency of the DFT calculation, since the computation time is reduced from being proportional to \(N^2\) to becoming roughly proportional to \(N \log N\), where \(N\) is the number of points to be Fourier transformed.

One generally deals with purely real finite-length sequences, and for these, the DFT has the following symmetry properties: its real part and modulus are periodic even sequences, and its imaginary part and phase are periodic odd sequences. Therefore, if one half of the spectrum is known, the other half can be reconstructed from it. Consequently, it is only necessary to show one half of the spectrum for display purposes.

2.3.2.2. Windowing

Before the DFT of the sampled time signal is found, it is necessary to window it. Windowing is the process of multiplying the data sequence by a weighting sequence.

Windowing can be used to shorten a given sequence to a desired length, or to remove undesirable features from the end of the signal, for instance, reflections from the walls of a room (Downes, 1985). Beauchamp and Yuen (1979, pp.157-159) have found the theory of this method of windowing unsatisfactory for the following reasons. It does not conserve power. The weighting applied is different for different data points; yet each data element has the same variance and each should be given equal weight.

Despite these arguments, windowing is widely used. Its generally accepted purpose is to reduce leakage of spectral energy from one frequency into adjacent frequencies. The problem of leakage is inherent with any finite data record, because to make it finite a rectangular window must be applied,
and this results in convolution of the spectrum with the sinc function.

The DFT of the ideal window would be the delta function, which would have no effect on the spectrum, but this is impossible to achieve. The desired window DFT will have a tall narrow central lobe and negligible side lobes. Normally these two features cannot occur together. The Hanning window is generally regarded as the optimum compromise, and so has been the main one used in this work.

Discontinuities in the time sequence cause spectral ripple and should be avoided (F.J. Harris, 1978). The half Hanning window tapers the signal to zero, thus removing any end-discontinuity.

In this work, the main motivation for windowing was to shorten the data record to a pre-determined number of samples, say N, so that an N-point FFT could be performed. The maximum time for which the instrument response is required is the time taken for the impulse to travel from the input to the bell and back to the input just over twice (see Section 5.4.1.3). The actual response continues well beyond this time. Windowing was also occasionally used to remove source reflections.

2.3.2.3. Problems with Frequency Domain Deconvolution

The method of dividing the DFT of the response sequence by the DFT of the input sequence does not appear in most of the signal processing literature as a recognised method of deconvolution, although the equation from which it is derived (start of Section 2.2.3) is well known.

In addition, frequency domain deconvolution is a technique which does not appear to have been widely used. Some investigators of vocal tract impulse response have investigated it; Descout et al (1976) found that it gave "meaningful" results but did not pursue it because in their case the frequency resolution was not sufficient. Resnick (1980) attempted a method involving the division of Laplace transforms; he encountered problems with noise and eventually opted for the time domain deconvolution technique mentioned earlier.
Silverman and Pearson (1973) reason that frequency domain deconvolution will only produce a direct solution in idealised cases in which $G(\omega)$ and $F(\omega)$ contain all pertinent frequencies.

Noise constitutes the main problem when attempting frequency domain deconvolution. When noise is present and spectral values at high frequencies are low, then division of spectra leads to distorted values at high frequencies. This distortion in frequency results in a distorted deconvolved time record. Investigations of these issues are discussed more fully in Section 5.2.3.

One possible way of removing the undesirable effects of noise is to use some form of digital filtering to the data sequence so that the noise can be removed before deconvolution proceeds. One problem with this approach is that the spectral distribution of the noise is not known. Noise does not occur in a well-defined frequency band. It occurs throughout the spectrum and therefore in the same part of the spectrum as the signal itself. Thus, filtering out noise will inevitably lead to loss of some important high frequency components of the signal.

Later the way in which signal averaging can improve the S-N ratio is discussed; it is not, however, sufficient to remove the noise problem altogether.

A possible solution to the noise problem in frequency domain deconvolution is the use of restoration algorithms, which are discussed in Section 2.3.5.

### 2.3.3. Constrained Iterative Methods

Constrained iterative methods are used to improve the quality of a distorted function, eg. a signal or image, progressively. A priori knowledge of properties of the function is combined with the distorted approximation to the function. Iterations result in convergence towards an improved solution. In Section 2.3.5, such techniques are used in signal restoration. In this
section, they are used to carry out deconvolution.

First, a generalised iterative procedure for the recovery of a distorted signal is described, and then the way in which it can be applied to deconvolution is shown. The general situation can be represented mathematically by

\[ g = Hf \]  \hspace{1cm} (2.10)

where \( f \) is the unknown input signal to be recovered, \( g \) is the measured response signal, and \( H \) is a known distortion operator (Schafer et al, 1981). When it is not possible to find the inverse operator \( H^{-1} \), one alternative approach is the method of successive approximations to \( f \). This involves the use of an iteration equation,

\[ f^{(k+1)} = Df^{(k)} \]  \hspace{1cm} (2.11)

Here, \( D \) is an operator which is dependent on \( H \) and may incorporate constraints based on a priori known properties of \( f \). These properties can be expressed in terms of a constraint operator \( C \), so that

\[ f = Cf, \]  \hspace{1cm} (2.12)

that is, any given \( f \) will be modified as necessary by operator \( C \) so that it satisfies the required constraints. Therefore,

\[ g = H Cf \]  \hspace{1cm} (2.13)

From these expressions, the following basic functional equation can be formed

\[ f = Cf + \mu(g - H Cf), \]  \hspace{1cm} (2.14)

from which the following iteration equation is obtained,

\[ f^{(k+1)} = Df^{(k)} = \mu g + P f^{(k)}, \]  \hspace{1cm} \text{where } P = (I - \mu H)C \hspace{1cm} (2.15)

\( \mu \) is a parameter which can be used to control the rate of convergence and \( I \) is the identity operator. Schafer et al (1981) use the "contraction mapping theorem" to prove that if \( D \) (or \( P \)) are "contraction mappings", then the procedure converges to a unique solution for \( f \).
For a linear shift-invariant system with known impulse response $h$,

$$g = h * f$$  \hspace{1cm} (2.16)$$
and the problem becomes deconvolution of input signal $f$. The following iteration for constrained iterative deconvolution can be used:

$$f_{(k+1)} = \mu g + \beta * C f_{(k)}$$  \hspace{1cm} (2.17)$$
where $\beta = \delta - \mu h$, and $\delta$ is the unit impulse.

Schafer et al (1981) have investigated the constraints of finite duration and positivity, and the conditions necessary for a convergent solution in each case.

So far, the discussion in this section has involved the recovery of system input from measured response and impulse response, and perhaps some a priori knowledge of the input. Constrained iterative deconvolution can also be used to recover the impulse response from the measured response and system input. Descout et al (1976) used such a technique to recover the impulse response of a vocal tract. Their approach was to minimise the square of the error between the real response of the system $g(n)$, and the response $\hat{g}(n)$ predicted by convolution of the input signal $f(n)$ with the estimated impulse response $h_{(m)}(n)$.

$$\hat{g}(n) = f(n) * h_{(m)}(n)$$  \hspace{1cm} (2.18)$$

$$= \sum_{i=1}^{n} f(n-i+1) h_{(m)}(i) \epsilon(n)$$

$$= (\hat{g}(n) - g(n))^2$$

where $\epsilon(n)$ is the error squared. The following iteration equation, based on the "gradient (optimisation) method" was used to obtain the next ($m+1$th) estimate of $h(n)$.

$$h_{(m+1)}(n) = h_{(m)}(i) + \kappa \epsilon(n) f(n-i+1)$$  \hspace{1cm} (2.19)$$
Here, $k$ is a carefully chosen feedback coefficient. Descout et al found that about 100 iterations were necessary to produce a good 20-point impulse response.

At Surrey a simple constrained iterative deconvolution algorithm has been implemented, which makes use of the following iteration equation,

$$h_{(k+1)} = h_{(k)} + \frac{1}{\mu} D\{g - f \ast h_{(k)}\}$$  \hspace{1cm} (2.20)

Here $g$, $f$, and $h$ are non-zero in the ranges $0$ to $(n-1)$, $0$ to $(m-1)$, and $0$ to $(l-1)$ respectively, such that $n$, $m$, and $l$ satisfy the size constraints discussed later (Section 2.3.6.1). The operator $D$ is such that $Dx$ moves to the point in $x$ at which $f$ is a maximum and counts $l$ elements forward thereafter. $\mu$ is a parameter which is chosen to ensure convergence and may also control the rate of convergence. The proposed iteration equation above does not correspond exactly with the generalised iteration equation of Schafer et al (1981) which would be

$$h_{(k+1)} = C h_{(k)} + \mu(g - f \ast C h_{(k)})$$  \hspace{1cm} (2.21)

where $C$ is the constraint operator. Further discussion of this algorithm, together with results of its application will be presented in Section 5.2.4.

2.3.4. Other Methods of Deconvolution

2.3.4.1. Cepstral Deconvolution (Quefrency Domain Deconvolution)

The cepstrum can be defined in two ways. The power cepstrum, the original quantity defined by Bogert et al (1963), is the inverse Fourier transform of the logarithmic power spectrum. However, here the complex cepstrum is discussed, which is defined as the inverse Fourier transform of the complex logarithm of the complex spectrum. The complex cepstrum retains phase information which is lost in the power cepstrum, and in this respect is more useful.

To find the complex cepstrum of a given time signal $f(t)$, the following steps are carried out:
(1) Fourier transform to obtain the complex spectrum,

\[ F(\omega) = a(\omega) + ib(\omega) = A(\omega) \exp(i\phi(\omega)) \]  \hspace{1cm} (2.22)

where \( a \) and \( b \) are the real and imaginary components respectively, and \( A \) and \( \phi \) are the amplitude and phase.

(2) Take the complex natural logarithm of \( F(\omega) \)

\[ \ln F(\omega) = \ln A(\omega) + i\phi(\omega) \] \hspace{1cm} (2.23)

Here, the real and imaginary components of the log spectrum are \( \ln A(\omega) \) and \( \phi(\omega) \) respectively.

(3) Inverse Fourier transform into the quefrency domain to form the cepstrum. Before the inverse Fourier transformation, it is necessary to carry out phase unwrapping and removal of the linear phase component, because the phase function must be continuous rather than modulo \( 2\pi \) when calculating the complex cepstrum. Phase unwrapping can sometimes present problems.

Cepstral deconvolution involves the following three steps:

(1) transforming the signal into its cepstrum,

(2) performing some editing operation in the cepstrum, eg. subtraction, "liftering" (the cepstral equivalent of spectral filtering).

(3) transforming back into the time domain; this inverse operation involves a forward Fourier transform, restoration of any linear phase component that was removed when the cepstrum was being formed, complex exponentiation, and an inverse Fourier transform.

The potential use of cepstral deconvolution to recover the impulse response can be understood if the measured response \( g(t) \) is considered as the convolution of the input \( f(t) \) and the impulse response \( h(t) \). By the Convolution Theorem this transforms to multiplication in the frequency domain. In the log spectrum, the multiplication becomes a sum

\[ \ln G(\omega) = \ln F(\omega) + \ln H(\omega) \] \hspace{1cm} (2.24)

This addition is maintained in the complex cepstrum because of the linearity
of the Fourier transform. If the cepstrum of either \( f(t) \) or \( h(t) \) is known, and is subtracted from the cepstrum of \( g(t) \), this constitutes deconvolution. For investigation of brass instrument impulse response, an accurate knowledge of \( f(t) \) and \( g(t) \) should enable \( h(t) \) to be recovered by cepstral deconvolution. However, the process tends to be unstable in practice. The normal way of proceeding is to remove the features caused by the system (e.g. echoes) from the cepstrum of \( g(t) \), perform an inverse cepstrum, and thus obtain a good estimate of \( f(t) \).

This technique has been applied in many fields such as seismology, speech processing and echo removal (Brödel and Kjaer, 1981; Oppenheim and Schafer, 1975), but has not yet been used in brass instrument investigations.

### 2.3.4.2. Cross Correlation

The cross correlation of two signals \( f(t) \) and \( g(t) \) is a measure of the dependence between them, and is defined as

\[
C_{fg}(\tau) = \int f(t-\tau) g(t) \, dt
\]

\[
= \int f(t) g(t+\tau) \, dt
\]  

The autocorrelation of a wideband noise signal is approximately a delta function, and this suggests an alternative way of finding the impulse response of a system. When the input signal is wideband noise, then the cross correlation of this input with the response will be proportional to the impulse response. This is a result of the Wiener-Lee relation which will now be derived.

The convolutional relationship between input \( f(t) \), impulse response \( h(t) \) and response \( g(t) \) can be written as
\[ g(t) = \int h(v) f(t-v) \, dv \]  
\[ g(t+\tau) = \int h(v) f(t+\tau-v) \, dv \]

Cross correlation of \( g(t) \) with \( f(t) \) gives

\[ C_{fg}(\tau) = \int f(t) h(v) f(t+\tau-v) \, dv \, dt \]

\[ = \int h(v) C_{ff}(\tau-v) \, dv \]  
\[ = \Gamma h(\tau), \]

This is the Wiener-Lee relation and states that the cross correlation between the input and response is the convolution of the impulse response with the autocorrelation of the input. For a wideband noise input, the autocorrelation is a delta function, so

\[ C_{fg}(\tau) = \Gamma \int h(v) \delta(\tau-v) \, dv \]  
\[ = \Gamma h(\tau), \]

and the impulse response is produced (Beauchamp and Yuen, 1979, p.205). (Here, \( \Gamma \) is a constant.) Schroeder (1979) has used this method to determine the impulse response of a concert hall. However, the method is inappropriate for finding the impulse response of brass instruments because then the response is measured at the input. The input signal is continuous, and so it will be impossible to make a measurement of \( g(t) \) which is independent of \( f(t) \).

2.3.4.3. Inverse Filtering

So far attention has been concentrated on the case in which the response and input signals are known, and deconvolution is used to find the impulse response. However, sometimes the impulse response and response signal are known and deconvolution is required to calculate the input signal. One such technique which is commonly used is inverse filtering; this is essentially the same as frequency domain deconvolution discussed in Section 2.3.2, with the same inherent problems. In this section some of the methods used to overcome these problems are discussed; these methods may also prove useful in frequency
domain deconvolution.

An example of a field in which inverse filtering has been applied is image restoration. Here, the response is a two-dimensional image which has been degraded by various distortions and noise in the optical system. The distortion phenomena can be represented by the "point spread function", which is the two-dimensional equivalent of the impulse response function (the response of the system when the input is a point source of light considered as the Dirac delta function). The degraded image is the convolution of the original undistorted image with the point spread function, with noise added. Let \( f(x,y) \) be the original image, \( g(x,y) \) be the distorted image, \( h(x,y) \) be the point spread function, and \( n(x,y) \) be the noise. Then

\[
g(x,y) = b(x,y) + n(x,y) \quad \text{where} \tag{2.30a}
\]

\[
b(x,y) = h(x,y) * f(x,y) \tag{2.30b}
\]

Deconvolution is necessary to determine the original image given the distorted image and the point spread function.

Inverse filtering attempts a perfect restoration [\*]. Fourier transformation of the above two equations gives

\[
B(\omega,\sigma) = B(\omega,\sigma) + N(\omega,\sigma) \tag{2.31a}
\]

\[
B(\omega,\sigma) = F(\omega,\sigma) \cdot H(\omega,\sigma) \tag{2.31b}
\]

where \( \omega \) and \( \sigma \) represent the \( x \) and \( y \) spatial frequencies. The approximation, \( f(x,y) \) to the original image is obtained by division and inverse Fourier transformation

\[
f(x,y) = F^{-1}(B/H) = f(x,y) + F^{-1}(N/H) \tag{2.32}
\]

for \( H(\omega,\sigma) \neq 0 \). There are two main problems with this technique. The first is

[\*] In contrast, the iterative restoration algorithms, to be discussed in the next section, attempt a gradual improvement of image quality.
that at high frequencies, \( H \) normally decreases rapidly, so noise at high frequencies will be enhanced. Secondly, \( H \) often has zeroes, and division by zero causes the method to break down. J.L. Harris (1966) has suggested a way of overcoming these drawbacks, namely, multiplying equation (2.32) by a function \( H_0 \) before dividing by \( H \).

\[
\hat{f}(x,y) = F^{-1} \left( \frac{(BH_0 + NH_0)}{H} \right) \tag{2.33}
\]

\( H_0 \) is carefully chosen so that it is zero when \( B \) is dominated by \( N \), to reduce high frequency noise, and so that \( H_0/H \) is finite wherever \( H \) is zero. Sondhi (1972), in his review of image restoration techniques, finds that Harris's method is not entirely satisfactory, as \( H_0 \) introduces further distortion.

Schafer et al (1981) simply set the quotient of the transforms to zero for frequencies higher than a certain level where the Fourier transform \( H \) becomes too small. They noted that such abrupt bandlimiting can cause additional unwanted large oscillations in the spectrum. In Section 5.2.3.3, a spectral windowing technique analogous to that of Schafer et al was used to help overcome frequency domain deconvolution instability problems.

2.3.5. Restoration Algorithms

Restoration can be considered as the process of improving the quality of a signal (or image) by combining the incomplete measured data with a priori known properties of the signal (or image). In Section 2.3.3 it was shown how such algorithms can form a means of deconvolution. To date most restoration algorithms have been used to improve image quality. However, a restoration algorithm could also help solve the problem of distortion introduced by noise in frequency domain deconvolution. It could restore the high frequency section of the deconvolved spectrum without the distortion. Some of the existing restoration algorithms will now be examined.
2.3.5.1. Gerchberg-Papoulis Algorithm

Gerchberg (1974) proposed a method whereby the resolution of an optical image could be improved by extending the spectrum beyond its known limits. If the object function is finite in extent, then a given section of complex spectrum can be extended beyond its known limits. This is because a finite function has an analytic spectrum; a known finite segment of an analytic function can, in principle, be used to determine the whole of the rest of the function.

In practice, a perfect reconstruction of the spectrum to a new upper limit cannot be achieved using imperfect data. Gerchberg instead attempted to converge iteratively on an extended spectrum of progressively improving quality. His algorithm assumes prior knowledge of the following two constraints:

(1) a finite segment of spectrum,
(2) the finite limits of the corresponding object.

The algorithm is summarised in Figure 2.1. The known portion of spectrum, with spectral values outside its range initially set to zero, is inverse Fourier transformed. The resulting object is set to zero outside its known finite limits — this is the application of the second constraint — and forward Fourier transformed. The first constraint is then applied by giving the correct values to the portion of the generated spectrum in which the true spectrum is known. This procedure is repeated until convergence has been achieved. Convergence is tested for by ensuring that the estimated and known spectra over the range of the known spectrum are sufficiently close. In addition, the object estimate outside the known object limits should be negligible.

Gerchberg explains that convergence must occur by defining an "error energy" as either

(1) the object energy outside the known object extent, or
(2) the difference function energy between the estimated and true spectrum over the range of known spectrum.

He shows how this error energy is reduced at each stage of the iteration. The
Figure 2.1
Schematic Representation of the Gerchberg Algorithm

Start

Start with known frequency-limited spectrum; outside this frequency range, spectral values are initially set to zero.

Sufficient iterations?

Yes

Forward FFT to obtain new spectrum $F(\omega)$.

No

Inverse FFT to obtain new time record, $f(t)$.

Modify $F(\omega)$ by replacing the appropriate part with the original frequency-limited spectrum.

Modify $f(t)$ by setting it to zero outside its known finite limits.

Reduction is initially rapid, but becomes more slow as the limit is approached. This convergence can only occur with perfect data. Noisy and distorted data can severely upset the procedure, and can cause the algorithm to diverge! Gerchberg has devised a scheme whereby the ill effects of distortion can be substantially reduced.

Papoulis (1975) provided a more mathematically rigorous treatment of the Gerchberg algorithm, and proposed a new use for it, namely, the extrapolation of a bandlimited function $f(t)$ beyond its original limits, and computation of its spectrum $F(\omega)$.

Image restorations of the Gerchberg-Papoulis (G-P) type have generated considerable interest, and various important potential applications have been suggested.
2.3.5.2. Method of Projections onto Convex Sets (POCS)

Youla (1978) developed a generalisation of the G-P algorithm. He showed that it has a natural geometrical interpretation, and can be formulated in terms of linear orthogonal projections onto linear subspaces. This picture served to clarify some of the limitations of the G-P algorithm (Youla and Webb, 1982). The G-P algorithm, being exclusively linear, cannot incorporate nonlinear constraints, and must therefore, in general, operate with incomplete information. This can lead to an excessive number of required iterations, and the algorithm may prove ineffective in the presence of noise. Youla and Webb proposed that an algorithm based on the method of projections onto convex sets (POCS) would overcome some of the limitations of the G-P algorithm; the advantages of the POCS method are that it enables many a priori nonlinear constraints to be incorporated, and ensures convergence.

Sezan and Stark (1982) have applied the method of POCS to various simulated images. They have tended to use the following four constraints on the signal $f$ to be restored:

1. **Space constraints**: $f$ must vanish outside a prescribed region,
2. **Fourier transform constraints**: the Fourier transform of $f$ must assume a prescribed set of values within a certain region of the Fourier plane,
3. **Energy constraints**: the energy contained in $f$ must not exceed a certain limit,
4. **Level constraints**: the amplitude of $f$ must lie within a prescribed interval.

The POCS algorithm is a general algorithm; the G-P algorithm is a special case of POCS which involves the first two of the four aforementioned constraints.

Sezan and Stark (1983) have found that the performance of the POCS algorithm is superior to that of the G-P algorithm in the presence of noise.

The problem of the onset of "saturation", that is, the point at which.
restoration becomes very slow indeed, seems to be unavoidable in all cases. A possible way of overcoming this is proposed in the next section.

### 2.3.5.3. Compensation for Slow Convergence

Jones (1986) has formulated a discrete version of the Gerchberg algorithm. The discrete time function and its discrete Fourier transform are represented as column vectors

\[
\mathbf{f} = \begin{pmatrix}
\mathbf{f}_0 \\
\vdots \\
\mathbf{f}_{N-1}
\end{pmatrix}
\]  

(2.34)

and

\[
\mathbf{F} = \begin{pmatrix}
\mathbf{F}_0 \\
\vdots \\
\mathbf{F}_{N-1}
\end{pmatrix}
\]  

(2.35)

Discrete Fourier transformation becomes premultiplication by a square matrix $\Omega$ of dimension $N$. The two column vectors are each divided into three partitions as follows:

\[
\mathbf{f} = \begin{pmatrix}
\mathbf{f}_0 \\
\mathbf{f}_1 \\
\mathbf{f}_2
\end{pmatrix}
\]  

(2.36)

with $\mathbf{f}_0$ of dimension $s$, $\mathbf{f}_1$ of dimension $q$, and $\mathbf{f}_2$ of dimension $(N-s-q)$;

\[
\mathbf{F} = \begin{pmatrix}
\mathbf{F}_0 \\
\mathbf{F}_1 \\
\mathbf{F}_2
\end{pmatrix}
\]  

(2.37)

with $\mathbf{F}_0$ of dimension $(N-k)/2$, $\mathbf{F}_1$ of dimension $(2k+1)$ and $\mathbf{F}_2$ of dimension $(N-k-1)/2$. These partitions induce a partitioning of $\Omega$ into nine partition matrices. If $\mathbf{f}_0$ and $\mathbf{f}_2$ are fixed, and $\mathbf{F}_0$ and $\mathbf{F}_2$ are known, then the discrete Gerchberg algorithm can be used to determine $\mathbf{F}_1$ and $\mathbf{f}_1$. Jones proves that convergence to a unique result is guaranteed provided
in other words, there must be at least as many fixed F's as unfixed f's.

However, the rate of convergence is dependent on the largest eigenvalue of the matrix $1/N \Omega_{11} \Omega^t_{11}$, [*] where $\Omega_{11}$ is the central partition matrix of $\Omega$. Eigenvalues may lie between 0 and 1. Eigenvalues close to 1 lead to a very slow rate of convergence, and the phenomenon of "saturation" mentioned earlier.

Jones proposes an acceleration procedure for cases where there are only a few eigenvalues close to 1. These eigenvalues may be identified and approximated from results of earlier iterations. Corresponding eigenvectors may then also be calculated. These eigenvectors and eigenvalues are used to form a correction term, which, when added to the results of the Gerchberg iteration, produce a better approximation to the correct result. For precise details of this acceleration procedure, see Jones (1986).

In Section 5.3 some results of the above algorithm are presented.

To the best of the author's knowledge, the present work constitutes the first instance in which restoration algorithms have been used in acoustics.

2.3.6. Measurement of Input and Output Signals

2.3.6.1. Constraints on Size Resulting from Convolution Properties

The measured input and response signals are finite duration sequences, and thus involve linear convolution. It can be shown that circular convolution of two periodic sequences $f_p(i)$ and $h_p(i)$ of period $n$ can be carried out by multiplication of their DFT's followed by inverse Fourier transformation (Rabiner and Gold, 1975, p.59).

\[
\begin{align*}
\hat{f}_p(i) \ast \hat{h}_p(i) &= \hat{g}_p(i) \quad (2.39a) \\
\hat{F}_p(k) \cdot \hat{H}_p(k) &= \hat{G}_p(k) \quad (2.39b)
\end{align*}
\]

[*] " $^t$ " denotes "complex conjugation".
A given periodic sequence and its related finite duration sequence (corresponding to one period) share the same DFT. This suggests a way of performing linear convolution of finite sequences. The response sequence $g_p(i)$ of the above circular convolution has period $n$, the same as the two inputs. This means that to obtain a finite duration response sequence $g(i)$ of period $n$, the two sequences $f(i)$ and $h(i)$ must also be of length $n$. Zeroes must be appended to $f(i)$ and $h(i)$ to satisfy this condition for the following reason. If the duration of $f(i)$ is $m$ samples, i.e. $f(i)$ is non-zero in the interval $1 \leq n \leq m$, and the duration of $h(i)$ is 1 sample, i.e. $h(i)$ is non-zero in the interval $1 \leq n \leq 1$, then direct linear convolution will yield $g(i)$ of duration $n = m + 1 - 1$ samples. To make $f(i)$ and $h(i)$ up to $(m+1-1)$-point sequences, the appropriate number of zeroes must be appended. If this is not done, wraparound difficulties occur.

To summarise, the following constraints must be applied during linear convolution of finite duration sequences:

1. $n = m + 1 - 1$,
2. all three sequences must be of duration $n$, which involves appending the appropriate number of zeroes to $f(i)$ and $h(i)$.

When measuring the response of brass instruments, these constraints mean that the combined length of $f(t)$ and $h(t)$ must not exceed the length of $g(t)$. Ideally, the measurement of $g(t)$ should proceed until the signal has died out sufficiently to become negligible. This takes a considerable time with brass instruments because of multiple reflections travelling back and forth within the instrument. This timescale is too long for practical purposes, so a shorter measurement must be made.

In addition, the measurement of $f(t)$ should be as short as possible. In practice, the input pulse has a negative overshoot which takes a long time to decay to negligible proportions. To satisfy the constraints, the process of windowing (see Section 2.3.2.2) is used to cut down the time records to the correct lengths, but inevitably this will introduce error into the deconvolu-
2.3.6.2. Method of Independent Measurements of Input and Response

Independent measurements of input and response are necessary for deconvolution. In the brass instrument case, the input and response are both measured at the input, and their separation is not entirely straightforward.

Sonoki and Resnick (1983) encountered the same problem when measuring the transient response of the human vocal tract at the lips. Their apparatus is illustrated in Figure 2.2. Three independent measurements were made:

(A) with a rigid termination at point Y,
(B) with an extension tube longer than the vocal tract at point Y,
(C) with the vocal tract at point Y.

Measurement A provides the incident signal. It contains a forward-going and backward-going pulse relatively close to each other (Figure 2.3(a)). The incident signal is taken to be the latter, i.e. that reflected from the rigid termination. This is to compensate for the fact that the microphone is slightly displaced from the actual input (Y) of the vocal tract. Correction must be made for the extra transfer function, attenuation and delay introduced when the pulse travels from the microphone to the lips, that is, along the section of tube XY.

Measurement B is an independent measure of the forward-going pulse (Figure 2.3(b)). When it is subtracted from measurement A, the result is the best estimate of the actual incident signal at point Y.

Measurement B is also subtracted from measurement C (Figure 2.3(c)), resulting in the best estimate of the vocal tract response at point Y.

The resulting incident wave (A - B) and reflected wave (C - B) both differ from the corresponding waves at the lips by the same transfer function, that is, the transfer function between points Y and X, and this will cancel out during deconvolution.
Figure 2.2 Experimental Arrangement Used by Sondhi and Resnick (1983) to Measure the Transient Response of the Human Vocal Tract.
Figure 2.3 Sketches to Illustrate the Measurements Obtained by Sondhi and Resnick (1983). (a) Rigid Termination, (b) Semi-Infinite Termination, (c) Vocal Tract.
Sondhi and Resnick's approach is the one adopted for the measurements in this work.

2.3.7. **Summary of Section 2.3**

Various approaches to the problem of deconvolution (or inverse convolution) have been examined. Time domain deconvolution was shown to be ill-posed unless some form of regularisation was used. Frequency domain deconvolution is susceptible to the effects of noise. A possible solution to this problem is the use of restoration algorithms of the Gerchberg-Papoulis type which may be a means of correcting the distorted regions of the deconvolved spectrum. Constrained iterative deconvolution was also explained. All the aforementioned techniques have been systematically investigated, and the results will be compared in Chapter 5. For completeness of the survey, other deconvolution methods including cepstral deconvolution, cross correlation and inverse filtering were described.

Finally, the necessary constraints on the sizes of the measured signals were considered, and a method whereby independent measurements of input and response may be made was described.

2.4. **Reconstruction of Internal Shape from Transient Response**

2.4.1. **Introduction**

After the transient response at the input of the brass instrument has been obtained by some appropriate method of deconvolution, it can potentially be used to reconstruct the shape of the instrument. Such a technique was first applied to the vocal tract.

Early attempts to determine vocal tract shape involved the use of X-rays. A later approach was the inverse problem of estimating vocal tract area from acoustical measurements at the lips. Schroeder (1967) and Mermelstein (1967) estimated the area from frequency domain measurements of input impedance at the lips. Sondhi and Gopinath (1971a) were the first to propose the use of
time domain measurements in this context. In this section, some of the developments in this field will be reviewed, and then the feasibility of applying the same techniques to brass instruments will be discussed.

2.4.2. Sondhi et al

Sondhi and Gopinath (1971a) showed that the knowledge of the reflections coming out of the vocal tract in response to an acoustic impulse presented at the lips, for a duration corresponding to the time required for the impulse to travel to the end of the vocal tract and back, is necessary and sufficient to determine its area function [\#]. The starting point of their derivation was two well-known linear time-invariant partial differential equations - Newton's Law and the Continuity Equation respectively. These lead to Webster's Horn Equation.

\[
\frac{\partial}{\partial x} p(x,t) = \frac{-1}{A(x) \partial t} u(x,t) \quad (2.40a)
\]
\[
\frac{\partial}{\partial x} u(x,t) = -A(x) \frac{\partial}{\partial t} p(x,t) \quad (2.40b)
\]

Here, \(A(x)\) is cross-sectional area, \(p(x,t)\) is pressure and \(u(x,t)\) is volume velocity. The form of these equations has been simplified because of the choice of units. Implicit in these equations are the following assumptions.

(1) Wave propagation is linear and planar.
(2) No losses are caused by viscosity, i.e. the air filling the tract is assumed to be a perfect gas.
(3) No losses are caused by wall vibration, i.e. the tract walls are assumed to be rigid.

Assumptions (2) and (3) can be relaxed, leading to a modified form of the above equations, and thus an area recovery procedure which takes losses into account (Sondhi and Gopinath, 1971b; Sondhi, 1974). Ranade and Sondhi (unpublished) have also looked into the possibility of relaxing assumption (1) (see [\#] Area function is cross-sectional area as a function of distance along the axis.}
The above equations are used to derive \( A(x) \) from the transient response of the vocal tract. There are three different transient responses from which \( A(x) \) can be derived (Sondhi and Resnick, 1983) - input impedance \( z(t) \), reflectance \( r(t) \), and step reflectance (the integral of reflectance).

Deconvolution is necessary to obtain the true transient response from the measured response. Sondhi's method of deconvolution has been referred to in Section 2.3.1 (time domain deconvolution with regularisation procedure). Thereafter, an inversion algorithm is required to find \( A(x) \). Sondhi and Resnick (1983) have developed algorithms for the input impedance and step reflectance cases.

Their experimental arrangement was shown earlier in Figure 2.2, Section 2.3.6.2. It consists of an impedance tube with an impulse source at one end and the vocal tract at the other. The impulse source is an electrostatic transducer which produces impulses containing frequencies up to 8 kHz. To satisfy the assumption of planar wave motion, transverse modes must be avoided, i.e. the wavelengths contained in the signal must be greater than the vocal tract diameter. The maximum radius of the vocal tract during normal speech is about 2.3 centimetres; therefore, transverse modes would be expected to occur at frequencies greater than about 5.4 kHz (Stevens et al, 1953); so Sondhi and Resnick's signals probably contain transverse modes.

When measuring reflectance, the source must be a perfect absorber, i.e. non-reflecting (see Section 2.2.3). This can effectively be brought about by making the distance between the microphone and the impulse source greater than the distance between the microphone and the glottis. This condition has not been satisfied in Sondhi and Resnick's experimental arrangement. By implication, the former distance is 40 centimetres (16 inches) and the latter distance is 61 centimetres (24 inches) plus the length of the vocal tract (nominally 17 centimetres). Sondhi and Resnick, however, state that the exact
dimensions of the impedance tube are not critical to the success of the experiments.

Independent measurements of incident and reflected signals were made using the technique described in Section 2.3.6.2, with data being stored in digital form.

Sondhi and Resnick's technique gives quite good reconstructions of area function for relatively open vocal tract shapes. However, when the tract contains a narrow constriction, the quality of the reconstruction suffers. Three possible reasons were investigated:

1. bandlimiting of reflectance,
2. non-planar wave motion,
3. viscosity.

The conclusion was that bandlimiting was the primary cause of quality loss. Bandlimiting was shown to smooth the reconstructed area function and reduce the magnitude of the computed area beyond the constriction.

2.4.3. Descout, Toussignant, Lecours and Lefavre

This group of French workers have used the experimental technique proposed by Sondhi et al. They also measure the response of the vocal tract at the lips, but their analysis of the results is done in a different way.

The vocal tract is modelled as a succession of elementary cylindrical tubes of equal length, each treated as a transmission line propagating plane waves. Each junction between adjacent cylinders has a reflection coefficient associated with it which is assumed to be real and independent of frequency. The length of each element must be small compared with the wavelength \( \lambda \) of the highest frequency contained in the signal. For plane wave propagation, \( \lambda \) must also be greater than the vocal tract diameter, as mentioned in the previous section. Incident and reflected waves are assumed to propagate independently of each other.
The iterative procedure for recovering the vocal tract area function can be summarised in the following six steps.

1. The input impulse $f(t)$ and the reflected response $g(t)$ are independently measured by sampling.

2. By inspection of $g(j)$, the point at which the greatest area variation occurs is located. The reflection coefficient at this point can be estimated by calculating the ratio of the value of the response at that point to the amplitude of the input impulse.

3. The area variation in this region is estimated. The bandlimited nature of the signals means that the area variation must be smoothed using a transition profile (Tousignant et al, 1979; Lefèvre et al, 1981).

4. The reflection coefficients corresponding to the above area variation are calculated using the following expression derived from transmission line theory (Skudrzyk, 1971)

$$r_i = \frac{A_i - A_{i+1}}{A_i + A_{i+1}}$$  \hspace{1cm} (2.41)

The discrete impulse response $h(j)$, i.e. reflectance as defined in Section 2.2.3, is computed from the reflection coefficients by following the progress of the unit impulse from the lips through the vocal tract and back to the lips using the following simple transmission model (Descout et al, 1976; Lefèvre et al, 1981). At each junction between elements, part of the wave is reflected and the other part is transmitted. For a pulse travelling towards the throat, at junction $i$ it will be reflected by $r_i$ and transmitted by $(1+r_i)$. For a pulse travelling backwards towards the lips, at junction $i$ it will be reflected by $-r_i$ and transmitted by $(1-r_i)$. The forward- (p) and backward- (b) travelling waves in each element $i$ at time $t$ must be considered. Figure 2.4 gives a schematic representation of the situation at time $t$. After one further sample, the forward- and backward-travelling waves in element $i$ are given by the following
Figure 2.4

Schematic Representation of the Reflection and Transmission of Forward ($p_i$) and Backward ($b_i$) Travelling Waves in Successive Cylindrical Elements.

By recursive application of these equations, the value of the backward-travelling wave at the input, i.e. in the first element, can be found, and this gives $h(i)$. Losses may be included by multiplying the pressure by the attenuation factor of an element each time it is traversed. (For a slightly different approach to the problem of recursively computing reflectance from bore shape, see Section 5.4.1.2.)

[*] NOTE: The French authors used a different notation.
(5) \( f(j) \) is convolved with the above approximation of \( h(j) \) to obtain the response \( g_a(j) \) which would have been obtained from the vocal tract whose area varies in the manner identified so far.

(6) The calculated response \( g_a(j) \) is subtracted from the measured response \( g(j) \) to obtain the residual response that can be analysed further by repeating the above procedure from step (2).

The area function is thus progressively built up, each area constriction appearing in order of importance.

Lefèvre et al (1983), using the above method have obtained vocal tract shapes with an additional "residual tract" at the back of the throat. This error was thought to be caused by the presence of noise, limited computational precision and shortcomings of the model used. The quality of reconstruction in the region of the throat was poor in some cases. The accuracy was also poor behind an almost total constriction. (Sondhi also found this - Section 2.4.2.)

2.4.4. Stansfield and Bogner

Unlike the investigators mentioned so far, Stansfield and Bogner (1973) do not make impulse measurements of the vocal tract. Instead they require the transfer impedance in the frequency domain. The transfer impedance is the ratio of the output pressure at the load to the input volume velocity at the source. Here the glottis is regarded as the input and the lips as the output. Stansfield and Bogner measure the output pressure, but the input volume velocity can only be approximated.

The vocal tract model is the same as the one used by the French workers. An additional assumption is the presence of infinitely long source and load tubes at the glottis and lips respectively, that is, source and load are both non-reflecting.

The area function of the vocal tract is calculated from the transfer impedance in the following way. First, the input impedance \( Z(f) \) is calculated
from transfer impedance. Then the reflection coefficient \( R(f) \) at the input (glottis) as a function of frequency is obtained using the following well-known relation:

\[
R(f) = \frac{Z(f) - Z_s}{Z(f) + Z_s}
\]  

(2.43)

where \( Z_s \) represents the source impedance. Using a Fourier transform relationship, the reflectance \( r(t) \) is obtained from \( R(f) \). The reflectance is the response at the glottis to an initial unit impulse of pressure. It is interesting to note that Sondhi et al. and Leffèvre et al. measured the reflectance at the lips, whereas Stansfield and Bogner have computed it at the glottis.

The next stage in the procedure is to calculate the reflection coefficients \( r_i \) for all junctions in succession, starting at the source. This involves the same model of transmission as was used by the French workers, but it is used in the inverse sense (to calculate reflection coefficients from reflectance rather than vice versa). The reflected signal seen at the input is considered to consist of two parts: the direct reflection \( r_d(t) \) from the farthest (\( k \) th) junction reached at that time, and the multiple reflections \( r_m(t) \) from all intermediate junctions, i.e.

\[
r(2kT) = r_d(2kT) + r_m(2kT).
\]

(2.44)

From the transmission model, it is clear that

\[
r_d(2kT) = r_k \prod_{i=1}^{(k-1)} (1 + r_i)(1 - r_i)
\]

(2.45)

\[
r_m(2kT) = r_k \prod_{i=1}^{(k-1)} (1 - r_i)^2
\]

(2.46)

\( r_m(2kT) \) is found by considering the forward- and backward-travelling waves in each element, as explained in the previous section (Section 2.4.3). By rearrangement, an expression for the reflection coefficient at the \( k \) th junction is obtained:
\[ r_k = \frac{r(2kT) - r_m(2kT)}{k-1} \prod_{i=1}^{k-1} (1 - r_i^2) \] (2.46)

From this all the reflection coefficients can be calculated recursively, and from them all the cross-sectional areas can be evaluated using the following recursive relation (related to equation 2.42):

\[ A_{i+1} = A_i \left( \frac{1 - r_i}{1 + r_i} \right) \] (2.47)

Stansfield and Bogner obtained area functions which were good approximations to smoothed versions of the actual area functions. They suggested the smoothing was caused by bandlimiting of the signal. Sondhi et al also found this.

2.4.5. Discussion

The applicability of a radius recovery procedure to brass instruments will now be discussed. All the aforementioned methods made the assumption of plane wave propagation. This assumption is satisfied in the vocal tract provided that the signal frequency content does not exceed about 4 kHz. In a cylinder, waves are plane until transverse modes appear.

Brass instruments contain conical and flaring sections as well as cylindrical sections, so wave propagation cannot be planar throughout. In conical sections the waves are spherical. The precise nature of the wavefront in a region of rapid flare is subject to debate, as discussed in Section 2.1.1. Clearly the bore diameter in the bell region will greatly exceed the vocal tract diameter. For a cylinder this would necessitate a lower maximum frequency than 4 kHz to avoid transverse modes, which would lead to an undesirable loss of resolution. However, Benade and Jansson (1974) would argue that higher order modes in the bell region lead to a "flattening" of assumed spherical wavefronts; if this is true, then it would be better to keep high frequencies in the signal. The precise choice of signal bandwidth will be discussed later (Section 3.1.1).
Errors resulting from losses in the vocal tract become significant at frequencies below about 1 kHz (Sondhi, 1974). Sondhi has attempted to compensate for visco-thermal losses and losses caused by yielding walls by use of an idealised model of wall impedance. Such compensation has not been applied in the cases where the tract was modelled as a cascade of uniform sections. In such cases, "it is difficult to introduce factors varying with frequency" (Descout et al, 1976). For elements treated as transmission lines, reflection coefficients must be real and independent of frequency. This is unrealistic as losses at cylinder walls are known to vary in proportion to the square root of the frequency (Kinsler and Frey, 1950, Section 9.7, p. 240). (See also Section 3.1.3.5.) The French workers used a frequency-independent attenuation coefficient for each individual section (Stevens et al, 1953). When a cylindrical-element type model was used for brass instruments, compensation for losses also had to be made in the time domain (see Section 5.4.1).

The assumption of a non-reflecting source has not been emphasised by the vocal tract workers. The source used by Sondhi et al was not strictly non-reflecting (Section 2.4.2). The French workers did not mention the requirement of a non-reflecting source. It is possible that the source reflections which did occur in the above cases were of sufficiently small amplitude as to be negligible.

Goodwin (1981) was the first person to seek to recover the radius function of a brass instrument using impulse techniques. His model was closely based on the Stansfield-Bogner model. He also proposed a simplified version of the model in which the multiple reflection contribution was considered negligible. Goodwin's input impulse supposedly contained frequencies well in excess of 100 kHz, thus making non-planar wave propagation unavoidable. His treatment of losses was simply to multiply a measured time record by a linear gain function so that the bell reflection became as large as the initial impulse (Goodwin, 1981, p.34). This cannot be justified theoretically. The amplitude
of the impulse would be expected to decay exponentially with time, and so an exponential gain function would be more realistic. In addition, the amplitude of the bell reflection must be less than the amplitude of the initial impulse because of the energy lost by reflection at the junctions along the impulse path. Goodwin's source was not non-reflecting. Details of how his source was modified to make it non-reflecting will be given in the next chapter (Section 3.1.3). Goodwin noted the need for deconvolution of the measured instrument response before the radius could be accurately recovered, but he did not actually use deconvolution. Thus, any results he obtained were based on non-deconvolved responses. However, one apparently promising result is displayed in Section 2.5 of his thesis (Goodwin, 1981).

The problem of recovering the radius of a brass instrument will clearly be different in some ways from the corresponding problem for a vocal tract. A brass instrument is much larger. A trombone is about 2.7 metres long, whereas a vocal tract is 17 centimetres long. This difference in length has several consequences. The brass instrument measurement will take longer. The time required for a single trombone measurement is about 45 milliseconds; for the vocal tract, the corresponding time is about 1 millisecond. For a vocal tract, the speed of the measurement is important because the tract must be kept effectively stationary during the measurement. In addition, many measurements of the tract are desired per second so that the manner in which the tract shape varies can be followed. The difficulty of not being able to ensure that the tract remains stationary means that signal averaging cannot be applied. No such problem exists with brass instruments.

To have a non-reflecting source, a source tube must necessarily be extremely long for brass instruments, as will be shown later (Section 3.1.3.3). The greater length of the brass instrument will also make the effects of losses more noticeable.

The greater diameter of the flaring parts of brass instruments necessi-
tates a smaller signal bandwidth so that the planar wave propagation assumption be satisfied. The original motivation for the application of the radius recovery procedure to brass instruments was the detection of small differences between instruments. Such differences would become undetectable if the signal bandwidth were too narrow. The conical and flaring sections of brass instruments clearly violate the plane wave assumption and thus bring into question the applicability of the radius recovery procedure to brass instruments.

Another possible source of error with brass instruments is the presence of bends. They cause additional reflections and scattering (Nederveen, 1969), when in real terms there is no significant area change.

Jones (1983) has recently proposed an original method in which reflection and damping coefficients may be recovered from a knowledge of the impulse response (see also Duffield and Jones, 1985). The brass instrument is again modelled as a series of cylindrical elements propagating plane waves (as in the models of the French workers and Stansfield and Bogner). Here, the reflection and damping coefficients are real and necessarily independent of frequency. The details of this method will be given in Section 5.4, together with some results of its applications.

2.4.6. Summary of Section 2.4

The possibility of reconstructing the shape of a brass instrument from its transient response measured at the input has been discussed. This technique was first proposed by Sondhi and Gopinath as a means of determining the area function of the vocal tract. The pioneering theoretical and experimental work of Sondhi et al was reviewed. They modelled the tract using a version of the Webster Equation. Subsequent work by Descout et al involved similar experimental measurements, but a different analysis technique based on a piecewise-constant transmission-line model of the vocal tract. Stansfield and Bogner used a similar analysis method, but their experimental arrangement was different; they obtained the time-domain quantity of reflectance indirectly
via frequency domain measurements. The applicability of the vocal tract technique to brass instruments was considered carefully. Potential problems exist in some areas, particularly in the assumption of plane waves. Goodwin has already attempted to reconstruct brass instrument radius functions using a model of the Stansfield-Bogner type, but with limited success. Jones has proposed an extension of this model such that losses within the instrument as well as its area function can be recovered. In Section 5.4, a detailed investigation of his method will be presented.

2.5. Overall Summary of Chapter 2

The large amount of research carried out on brass wind instruments has been reviewed; most investigations to date have involved the frequency domain. Input impedance has been shown to be a particularly useful quantity that can be related to instrument intonation. Methods of subjective as well as physical assessment of instruments have been developed, but clear links between the results of these have not yet been established. The current increased interest in transient acoustics measurements has led to the development of impulse measurements on brass instruments. To obtain the response of the instrument to a delta function, deconvolution must be carried out, and with imperfect experimental data this can be a difficult task. Various deconvolution methods have been discussed and their relative successes will be reported in Chapter 5. The transient response of a vocal tract, measured at the input, has been used to reconstruct its shape. The viability of using such a technique to reconstruct the radius function of a brass instrument will be investigated in Chapter 5.
CHAPTER 3

IMPULSE MEASUREMENTS

This chapter contains a description of the apparatus used to measure the response of brass instruments to an impulsive acoustic excitation. The original basic experimental arrangement was designed and built by Goodwin (1981), but since then several refinements have been made, particularly in the software which controls the apparatus.

3.1. Apparatus

3.1.1. Overall Description

Here a general description of the measurement system and its components is given. Detailed discussion of some of the individual components and experimental refinements will appear in later sections.

Impulse production and data acquisition are controlled by a Data General Nova 4 16-bit minicomputer. A block diagram showing the experimental arrangement is given in Figure 3.1. The essential working of the system is described in Goodwin and Bowsher (1980) and Goodwin (1981, Sections 3.4 and 4.3). A brief resumé will be given here.

The impulse source is a modified car spark plug. The source is discussed in more detail in the next section. The spark source is driven by a specially-built EHT pulse generator which transforms the TTL pulse from the DAC (Micro Consultants Type DW 3504) into a high amplitude short duration (10 microseconds) voltage impulse.

Pressure variation at the input of the brass instrument is picked up by a microphone, amplified, filtered, and recorded by sampling. The sampling rate is variable and is set using a programmable oscillator (Adret Codasyn 201) via an 8-decade BCD interface. The sampling rate can be set to an accuracy of .1 Hertz. It will be seen later (Section 3.2.4.2) that such precision is necessary in order to compensate for changes in ambient temperature. The sinusoidal
Figure 3.1
Block Diagram of Impulse Measurement Apparatus

output of the oscillator is converted to pulses using an electronic pulse shaper box. The pulses then clock the ADC. The ADC (Micro Consultants Type DW 3258, 14-bit, +10 volt range) accepts data at rates of up to 150 kHz. Data is written away on disk in binary form. In this work, it has been observed that a small DC voltage tends to build up on the ADC as it is used. Others have noted this as well (Fincham, 1982). Any error invoked by this is removed during post-measurement processing by detecting and subtracting any DC level present (Sections 3.3.3 and 4.2.2).

The anti-aliasing filter ensures that all frequency components above the Nyquist frequency are removed. Various filters have been tried. Goodwin used
a B and K Type 2121 filter/amplifier unit which has a cut-off frequency variable up to 20 kHz, but its fall-off was found to be insufficiently sharp. A Chebyshev filter, made by Julian Downes, with a cut-off at 75 kHz was extremely useful at a sampling rate of 150 kHz, but could not be used at lower rates. Eventually a programmable filter (Kemo Type VBF/22) was used. The cut-off frequency may vary between .1 Hz and 100 kHz, and may be set either manually or by the computer via a BCD interface. The filter is elliptic with a fall-off of 135 dB/octave. As the sampling rate normally used for measurements is about 46 kHz, for reasons to be discussed later (Section 3.1.3.3), the cut-off frequency is normally set to 20 kHz.

Pulse production and the start of data acquisition occur simultaneously. Such synchronisation is necessary in order that signal averaging may be carried out to improve the S-N ratio. This synchronisation is not as perfect as was at first assumed, as will be discussed later (Section 3.2.3.1).

Section 3.1.3 describes a major practical development - a new means of coupling the impulse source to the instrument input. The different microphones used are discussed in Section 3.1.4. Section 3.2 concentrates on software developments, and the different types of measurements are explained in Section 3.3.

3.1.2. Choice of Source

3.1.2.1. Considerations

Impulse measurements are the simplest means of obtaining the impulse response of a system, but there are some difficulties associated with the technique which will now be clarified. Barron (1984) has summarised three general requirements for an impulse test signal:

(1) short duration,
(2) adequate energy,
(3) reasonably flat spectrum in the measurement range.

As the impulse response is the response produced when the input signal is the delta function, it would seem that the best test signal should be a good approximation to the delta function, i.e., it should be as narrow and as tall as possible. Narrowness has the additional advantage of giving greater spatial resolution in the brass instrument response. However, when the duration of the impulse is short, it is difficult to deliver enough energy to the system to overcome the noise that is present, and the resulting S-N ratio may be inadequate. Hence the need for large impulse amplitudes; but these are limited by the range of linearity of the system. Signal averaging may be necessary to improve the S-N ratio, in which case, impulse production must be repeatable. A quiet environment is also helpful. The width of the impulse is limited by the range of frequencies allowed by the system. The limiting criterion here is the need to eliminate aliasing; since the sampling rate is 46 kHz, the upper frequency limit is 20 kHz. An electrostatic actuator source was not chosen, as its frequency range was not broad enough. (See, for example, Sondhi and Resnick, 1983, where the bandwidth was just 8 kHz.)

It is difficult to ensure that energy is equally distributed over all frequencies of interest. In this work, it was considered that it was not crucial that this condition be satisfied because in theory, deconvolution could be used to obtain the response that would have been measured had the energy distribution been flat.

In cases where an approximation to the impulse response is required, without the use of deconvolution, a good impulse shape is desired, i.e., a sharp rise and clean decay with no secondary spikes or ringing.

In the light of the above criteria and the limitations of the experimental arrangement, a suitable source must be chosen. Goodwin (1981, Section 3.4.1) considered various sources before opting for the spark source. He noted that pulses produced by an exploding wire or shock tube were not repeatable.
The STL ionophone (Fransson and Jansson, 1978) produces good pulses, but was disregarded on the grounds of expense and complexity. The spark source is cheap, simple and robust.

A pulsed loudspeaker has since been considered as it produces more consistent pulses than a spark source (Barron, 1984). In Section 3.2.3 the relative consistencies of loudspeaker and spark sources are compared, and the superior consistency of the loudspeaker is confirmed. A further advantage of the loudspeaker is that it can produce both negative and positive pulses. If, during data acquisition, pulses are alternately positive and negative, and the negative records are multiplied by -1, then any bias in the background noise distribution will cancel out during signal averaging. The problem of DC build-up on the ADC's, mentioned earlier, (Section 3.1.1) will also be removed automatically. This technique has been applied successfully by Fincham (1982).

The rim of the smallest loudspeaker available was too large to be coupled directly to the brass instrument input, so a conical brass coupler had to be made. This coupler caused severe secondary reflections which spoiled the impulse shape (see Figure 3.2). Descout et al (1976) also found that with tweeters the ringing was too long. One could argue that deconvolution might remove these effects. However, with the spark source a good approximation to the impulse response can be obtained immediately, without the need for deconvolution first. With the loudspeaker this advantage does not exist. The results in Chapter 5 show that deconvolution of real measurements does not work as well as one might hope; therefore the best approximation of the true impulse response will be obtained using the spark source.

3.1.2.2. Spark Source

Spark generators have been used as impulse sources by other investigators. Salikuddin et al (1980) used a high voltage spark discharge to measure the transmission properties of a duct-nozzle system. They chose a spark source "because of its flatter, more uniform spectrum, pulse uniformity and ease of
Figure 3.2 Response when Loudspeaker Source is Coupled to the Source Tube via a Conical Coupler.
Measurements show that the spectrum of the spark impulse is not completely flat, but smooth up to frequencies of about 15 kHz (see Figure 3.15, Section 3.1.3.5). Here, the spectrum has been modified by the passage of the impulse along the source-to-instrument coupler (see Section 3.1.3.5).

Salikuddin et al experienced many problems with equipment caused by EM radiation interference each time the spark discharged. This led to a DC shift in the microphone potential that resulted in saturation of the microphone signal. EM radiation also caused cable ringing. In addition, discharge currents were so strong that changes were induced in the local ground potential. Careful earthing and shielding were necessary to minimise these problems. Such drawbacks have been present at various stages in the investigations (see Sections 3.1.3.2 and 3.1.4.3). Chinoy (1982), in his recommendation of the spark generator as an impulsive signal source, has also noted the problem of EM radiation interference with ancillary equipment.

Salikuddin et al. (1984) later noted that the main problems with spark impulses are

1. low S-N ratio, and
2. non-linear acoustic propagation of high intensity spark impulses.

In this work, impulse amplitude is sufficiently low that equations of linear acoustics are valid.

Barron (1984) used spark sources for impulse testing of auditoria, and noted that signals from spark sources tend to be deficient in low frequency energy, leading to signal-to-noise problems at low frequencies. In this work, this type of problem has been more prominent at high frequencies.

An inherent feature of spark sources is that the impulses produced by them are somewhat inconsistent. The discharge intensity varies from one spark to another. When signal averaging is to be carried out, this presents a problem, since signal averaging assumes that each generated pulse waveform is
identical. The most obvious inconsistency is in amplitude, but variations in timing and shape also occur. Much time has been spent investigating these inconsistencies and formulating a way of overcoming them, as will be discussed in Section 3.2.3. The variations appear to be random, following no predictable pattern. The physical reasons for them are not well understood. Spark formation occurs when the voltage across a gap between two electrodes reaches a breakdown value, resulting in ionisation of the air in a narrow cylindrical region between the electrodes (Chinoy, 1982). "De-ionisation" or "recovery" requires a certain amount of time, about a fraction of a second (Rubchinskii, 1964). In the Nova-based spark source, the time between successive firings is at least one second. An investigation into the effects of increasing this delay to 5 seconds revealed no change in impulse consistency; therefore impulse variability is not caused by insufficient time for full recovery of the spark gap.

In this work, the spark source is a car spark plug with its curved outer electrode removed to increase the spark gap length and thence the power of the impulse (see Figure 3.3). In consequence, the spark can form at any position around the outer electrode ring (see Figure 3.4). It was felt that this might be contributing to impulse inconsistency, and that a more precisely defined gap would lead to more consistent impulse characteristics. For this reason, tests were carried out on a spark plug with its curved and central electrodes each filed to a point. In the latter case the impulses were found to be even less consistent, even though the impulses formed at the same position each time; so we conclude that defining the spark gap more precisely does not enhance consistency. After each spark discharge, small structural changes occur at the tip of each electrode; these may be partly responsible for impulse inconsistency.

Barron and Chinoy (1979) were aware of the inconsistency problem. For improved reproducibility they designed a triggered source using a car induc-
Figure 3.3 Car Spark Plug with Outer Electrode Removed.

Figure 3.4 Illustration of where Sparks Can Form.
tion coil, or a third electrode (Chinoy, 1982) for the trigger pulse and tungsten electrodes sharpened to a point for the main impulse.

3.1.2.3. Practical Notes on the Use of the Spark Source

The EHT pulse generator was designed with the aim that the impulse amplitude and width could be varied independently (Goodwin, 1981, Section 3.4.2). Amplitude can be increased by increasing the voltage across the capacitance across the ignition coil. Width variation is achieved by varying the capacitance but is accompanied by amplitude variation. The higher the capacitance, the higher the pulse amplitude and the greater the pulse width. The general strategy adopted was to use the maximum capacitance and voltage in order to obtain the maximum pulse height.

The performance of the spark plug has been found to vary with time. The impulse amplitude tends to decrease as the plug is used. Cleaning the region around the spark gap with a small stiff brush (e.g., a toothbrush) usually serves to raise the amplitude. The unpredictability of impulse amplitude has led to the practice of normalising each averaged measurement (Section 3.3.3), so that the impulse amplitude is unity and comparison of measurements is more appropriate.

Each spark plug has a limited lifetime. As time goes on, its performance deteriorates and the spark gap region becomes corroded because of the sparks. Eventually a crack may form in the ceramic insulator, leading to leakage and an unsuitable impulse shape. At this point it would be time to replace the plug.

3.1.3. Coupler

3.1.3.1. Introduction

In Section 2.4, it was shown that to reconstruct the bore profile from the reflectance, it is necessary to have a non-reflecting source when the impulse measurement is made. In Goodwin's original apparatus, the source was
coupled to the instrument using a mouthpiece packed with absorbing material (see Figure 3.5). Here the microphone was inserted in the throat of the mouthpiece. To do this, the mouthpiece is cut in two at its throat and normal to its axis, and a paxolin washer containing a hole for the microphone is inserted. The absorbing material was for reducing resonances in the mouthpiece cup and multiple reflections from the source (Goodwin, 1981, pp. 69, 70). However, complete removal of these features was not possible with Goodwin's arrangement, as will be seen in the next section (3.1.3.2).

A uniform semi-infinite tube with the same cross-section as that at the source is a perfect absorber. This fact was used to replace Goodwin's coupler with a source tube which resulted in a non-reflecting source and several other advantages. These points will be discussed in Section 3.1.3.3.

3.1.3.2. Mouthpiece Coupler

Early measurements of instrument response using a damped mouthpiece coupler clearly show multiple spikes on the decay of the impulse (see Figure 3.6). To carry out deconvolution, an independent measurement of the input impulse must be made. This measured impulse should be the signal which enters the instrument and travels along it, that is, all its features should be travelling in the forward direction. Some of the multiple spikes were suspected to be travelling in the reverse direction, and for this reason they were closely investigated. Preliminary ideas that they were caused by ringing of the electrical LCR circuit were quickly dismissed when it was seen that changing the capacitance across the coil had no effect on the timing of the spikes.

Eventually the cause of the spikes became clear. Figure 3.6 reveals three positive-going secondary spikes on the decay of the main impulse. Let us call the main impulse the first spike, followed by the second, third and fourth spikes. When the instrument is removed, leaving just the mouthpiece, there are two positive-going spikes followed by two negative-going spikes (see Figure 3.7). The two negative-going third and fourth spikes are the reflections of
**Figure 3.5** Original Damped Mouthpiece Coupler.
Figure 3.6 Response of Trombone with Damped Mouthpiece Coupler.

Pressure in Pascals

Time in milliseconds
Figure 3.7 Response of Damped Mouthpiece Alone (No Instrument) Measured Using Probe Microphone.
the first two positive spikes from the open end of the mouthpiece backbore. The cause of the second spike is illustrated in Figure 3.8. The first spike travels directly from the source to the microphone. Some of this direct energy is reflected back at the sudden increase in acoustic impedance (decrease in cross-sectional area) at the mouthpiece throat; it hits the spark plug and is reflected back towards the microphone as the second spike.

Given that the above reasoning is correct, a measurement of input impulse made with a mouthpiece coupler will be unsatisfactory because the third and fourth spikes are not travelling in the forward direction. Therefore, an alternative measurement arrangement is required.

Further disadvantages are associated with the use of a mouthpiece coupler. Figure 3.9 shows a measurement made with an 1/8-inch microphone (B and K Type 4138) using a sampling rate of 150 kHz and a 75 kHz anti-aliasing filter. Much high frequency energy is present; the impulse cannot be distinguished clearly, and there is a large amount of post-impulse ringing and disturbance. A Fourier transform reveals that frequencies greater than 75 kHz are still present in the signal, despite the presence of the filter. This high frequency energy is unwanted and can only be destructive.

In the early stages of the work, the diaphragm of the 1/8-inch microphone used to break. This could well have been caused by the high acoustic pressures produced by the spark; the actual pressures produced could be far greater than those measured by the computer-controlled apparatus. To investigate this, free field measurements of the sound emitted by the spark were made at various distances using the 1/8-inch microphone, a pulse generator and an oscilloscope, so that the pressure generated at the mouthpiece measurement plane could be calculated by extrapolation. The calculated value was of the order of several hundred pascals, yet the microphone is supposed to be able to withstand pressures of 8000 pascals. Another possible reason for the breakages was that sparks were hitting the diaphragm.
Figure 3.8 Illustration of how Secondary Spikes Arise with Mouthpiece Coupler.
Figure 3.9 Response of Damped Mouthpiece Alone Measured Using 1/8" Microphone.
In any case, increasing the distance between the source and the microphone appeared to be highly advantageous. Another reason for doing this was the occurrence of a "pre-pulse" before the actual acoustic pulse. This was produced by electrical pick-up of the EM spark radiation (as mentioned in Section 3.1.2.2). When the source and microphone are close, the pre-pulse can overlap with the acoustic pulse (see, for example, Figure 3.9), which can lead to confusion.

When the mouthpiece is replaced by a source tube, all the aforementioned disadvantages are overcome neatly.

3.1.3.3. Source Tube

The basic concept of the source tube is this: source reflections are effectively eliminated when they are made to occur after the desired section of impulse response. In other words, the impulse must travel from the start of the instrument to the bell and back to the microphone more quickly than from the start of the instrument to the source and back to the microphone. Therefore, the length of the source tube between the source and microphone must exceed the length of the instrument.

The basic source tube design is shown in Figure 3.10(a). The interior of the spark plug end of the source tube is threaded so that the spark plug can screw in. The other end is externally tapered to the shape of the backbore of the appropriate mouthpiece so that the instrument will fit on easily and tightly. The microphone is inserted into the tube via a paxolin washer, with appropriate holes, which fits between the two sections of the tube.

The large-bore trombone (0.547 inches input diameter) is taken to be the "standard" instrument in the investigations. Its length is 2.7 metres, so, in Figure 3.10 length SX should be comfortably in excess of that to allow for extension of the trombone using slides. SX was chosen to be 3.58 metres. The choice of distance XY between the microphone and the open end of the source tube was 0.48 metres. This distance was introduced so that the input impulse
Figure 3.10 Source Tube. (a) Basic Source Tube Design, (b) Division of Original Source Tube into Sections.
had time to decay to a suitably low level before the first reflection at the tube-instrument join. The diameter of this standard source tube is 12.7 millimetres (1/2 inch), the diameter at the input of the large-bore trombone. For a cornet or trumpet, the tube diameter will be 9.5 millimetres. These instruments are only half the length of a trombone, so normally a half-length source tube would suffice. A full-length narrower source tube was maintained because some of the theory to be presented in Section 5.4 requires a double-length impulse record. A medium-bore trombone would require a source tube of intermediate diameter. Ideally, a set of source tubes of differing diameters should exist to suit each instrument in order that reflections before the start of the instrument be eliminated. However, for some measurements the standard source tube is adequate, regardless of the instrument, with an appropriate tapering coupler between the microphone and the start of the instrument. This will be demonstrated later in Section 4.4.2. In addition, there are occasions when the length criterion for the source tube can be relaxed so that, for instance, the standard source tube can be used when measuring a tuba, even though it is double the length of a trombone (see Sections 4.4.1 and 4.5.3.1).

The improvement in pulse shape is immediately apparent when Figure 3.11 is compared with Figure 3.6. With the source tube, the impulse is clean with no secondary spikes. Superimposed on Figure 3.11 is a plot of a measurement taken with no instrument attached to the source tube. This clearly shows that the first source reflection occurs well after the bell reflection, demonstrating that the source is now effectively non-reflecting.

With the source tube arrangement, it is now particularly straightforward to make an independent measurement of input impulse. The method used is Sondhi et al.'s, referred to earlier (Section 2.3.6.2); its application will be discussed later in Section 3.3.4.

The source tube has the effect of attenuating much of the unwanted high frequency content of the impulse, so that the excessive energy which may have
Figure 3.11 Demonstration that the Source Tube Makes the Source Non-Reflecting. Response of Trombone with Source Tube (Continuous) and Response with No Instrument on Source Tube (Dashed) to Show the Location of Source Reflections.
been causing microphone breakage is removed. The anti-aliasing filter would have had to remove this energy in any case, but would not have provided the necessary protection for the microphone.

The electrical pre-pulse is now separated from the acoustic impulse by about 10 milliseconds. This time delay, during which the impulse travels from the source to the microphone, provides an extremely useful way of monitoring changes in ambient temperature caused by the changing velocity of sound in the source tube. Using this information, changes in temperature can be compensated for, as will be explained in Section 3.2.4.2. This has been another advantage of using the source tube coupler. After a measurement is completed, the unwanted 10 millisecond delay may be removed by windowing.

Using the source tube has led us to choose a standard sampling rate of 46 kHz in association with an anti-aliasing filter cut-off frequency of 20 kHz (as mentioned in Section 3.1.1). The reason for this choice of sampling rate is as follows. For frequency domain deconvolution (Section 5.2.3), an FFT of the complete measured instrument response record, from the instrument input to just after the bell, was required. Originally, all numerical processing was carried out on the Nova minicomputer, whose memory capacity was such that the largest double precision FFT which could be performed was 1024 points. The time taken for the impulse to travel from the microphone to just beyond the bell of the standard length instrument (trombone) and back is about 22 milliseconds. The ratio of the maximum number of array elements (1024) to the time taken is about 46 kHz. The total number of data points written away is 2048, giving 44.5 milliseconds worth of data; so even with the initial 10 millisecond delay, the instrument record fits in easily. The "essential" 1024-point record is extracted by windowing.

A typical input pulse using the standard source tube has an amplitude of about 60 pascals and a half amplitude pulse width of about 110 microseconds. Goodwin's data for the damped mouthpiece coupler were pulse amplitude 57.4
pascals and pulse width 20 microseconds (Goodwin, 1981, Section 3.4.3). The increased pulse width is clearly caused by the presence of the source tube; high frequencies are attenuated, and different frequencies travelling at different velocities causes spreading. Pulse amplitude is higher for narrower source tubes. Pulse characteristics are also dependent on the spark plug being used.

The obvious disadvantage of the source tube is the impracticability and inconvenience of having to use a 4.06-metre straight tube. In addition, much swapping around of tubes is necessary, as different diameter tubes are required for different instruments. In these respects a mouthpiece coupler is far superior. The possibility of using a flexible source tube is investigated in the next section.

3.1.3.4. Flexible Source Tube

A flexible source tube has the advantage that it can be coiled so that the apparatus takes up less space.

A 3-metre length of re-enforced rubber hose was attached between two short sections of brass tube, as illustrated in Figure 3.12.

When the hose was coiled tightly, with two or three loops, the measured impulse contained a double peak. When the curvature was gentler with just one large loop, the impulse had a single smooth peak. However, the pulses were far more inconsistent than with the straight brass tube. Pulse amplitude was also lower because of the greater attenuation in the hose material than in the brass.

For these reasons the work on a flexible source tube was not pursued.

3.1.3.5. Use of Source Tube to Measure Attenuation

The original source tube was made in sections, as shown in Figure 3.10(b), so that there were three different positions at which the microphone could be placed. This made it possible to investigate the way the impulse
Figure 3.12 Experimental Arrangement for Investigation of Flexible Source Tube.
changed as it travelled along the tube.

Measurements were made at each of the three positions using a sampling rate of 150 kHz, a broadband (100 kHz) microphone (Section 3.1.4.2), and the 75 kHz cut-off low-pass anti-aliasing filter. The impulses at the three positions are compared in Figure 3.13. It appears that bringing the microphone closer to the source has the following effects on the impulse:

(1) it has higher amplitude,
(2) it is narrower,
(3) it has a more pronounced double-peak,
(4) its decay is less smooth,
(5) the amplitude of the post-spark noise is greater.

A plot of pulse amplitude \( a \) in pascals against distance \( x \) from source in centimetres is given in Figure 3.14. (Three of the data points were obtained from a different source tube of the same diameter.) When \( \ln(a) \) was plotted against \( x \), and a least squares fit applied to find the best straight line, the following exponential relationship resulted

\[
a(x) = 221 e^{-0.004x}
\]

(3.1)

This relationship can only be considered approximate, as there were few data points and pulse amplitude varies with other factors besides distance, eg. age of spark plug (Section 3.1.2.3). However, an exponential relationship has also been proposed by Fant (1970), namely

\[
\text{Attenuation factor} = e^{-\alpha x}
\]

with \( \alpha = \kappa (0.007 \sqrt{\pi/A}) \) 

(3.2)

where \( \alpha \) is the attenuation coefficient, \( A \) is cross-sectional area and \( \kappa \) is an appropriate constant.

The double-peak and post-spark noise near the source probably result from a combination of reflections, scattering from walls, and transverse modes, which would begin to occur at about 13 kHz in a 12.7-millimetre diameter tube. Figure 3.15 shows the spectra of the impulses at the different positions,
Figure 3.13 Comparison of Impulses Measured at Three Different Positions Along the Source Tube. (a) Position 3 (Closest to Source), (b) Position 2, (c) Position 1 (Furthest From Source).
Figure 3.14 Plot of Impulse Amplitude Versus Distance From Source.
Figure 3.15: Comparison of Amplitude Spectra of Impulses Measured at Three Different Positions: Position 3 (Continuous), Position 2 (Dotted), and Position 1 (Dashed).
calculated by applying a half Hanning window to the impulse and using a 1024-point FFT. Clearly, the closer the impulse is to the source, the greater its high frequency content. Thus, the source tube is shown to attenuate the high frequency content.

By comparing the three spectra, the dependence of attenuation on frequency was investigated. Only the smooth first 15 kHz of the three amplitude spectra were used. Spectral division resulted in three deconvolved spectra

\[ \frac{A_1(f)}{A_2(f)} \]

where \( A_2(f) \) is the first, higher amplitude spectrum and \( A_1(f) \) is the second, lower amplitude spectrum. The following equation expresses the spectral attenuation,

\[ A_1 = A_2 e^{-\alpha x} \]

The deconvolved spectra were subjected to three separate methods of analysis.

**Method 1**

According to Kirchoff’s theory on visco-thermal losses in a cylindrical tube (Kinsler and Frey, 1950, p.241),

\[ \alpha = G \sqrt{f/A} \]

Here, \( f \) is frequency, \( A \) is cross-sectional area in square metres and \( G \) is a constant which should take the value \( 2.76 \times 10^{-5} \).

To verify that the power of \( f \) is 0.5, \( \ln [-\ln(A_1/A_2)] \) is plotted against \( \ln f \). The gradient should be 0.5; the data zero, \( c \) is used to determine \( G \), where

\[ e^c = 8x/A \]

**Method 2**

The second method uses the same theory as the first, but it assumes that the power of \( f \) is 0.5, and proceeds to calculate \( G \). Here, \( \ln(A_1/A_2) \) is plotted
against $\sqrt{f}$; the result should be a straight line through the origin having gradient $-Gx/A$.

**Method 3**

Fay (1940) proposed that $\alpha$ had an additional $f$-dependent term related to sound absorption in air, as well as the $\sqrt{f}$-dependent term, ie.

$$\alpha = C_1f + C_2\sqrt{f}$$

This gives

$$-\ln\left(\frac{A_1}{A_2}\right) = C_1 \sqrt{f} + C_2$$

So a plot of $-\ln\left(\frac{A_1}{A_2}\right) / \sqrt{f}$ versus $\sqrt{f}$ has a slope of $C_1$ and data zero $C_2$, from which $G$ is calculated. Fay's value for $G$ is $2.92 \times 10^{-5}$.

The three methods of analysis were applied to several different sets of measurements, using different diameter source tubes, different sampling rates and different spark plugs. A selection of results for tubes with diameters of 12.7 and 9.5 millimetres is summarised in Table 3.1.

**TABLE 3.1**

Results of Source Tube Attenuation Analyses

<table>
<thead>
<tr>
<th>Deconvolved spectrum file</th>
<th>Power of $f$ predicted by Method 1 (Should be 0.5)</th>
<th>$G$ using Method 1 ($\times 10^{-5}$) (Should be 2.76)</th>
<th>$G$ using Method 2 ($\times 10^{-5}$) (Should be 2.76)</th>
<th>$G$ using Method 3 ($\times 10^{-5}$) (Should be 2.92)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) STANDARD 12.7mm diameter</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FDC1WITH2</td>
<td>0.448</td>
<td>2.41</td>
<td>2.17</td>
<td>2.44</td>
</tr>
<tr>
<td>FDC1WITH3</td>
<td>0.471</td>
<td>2.14</td>
<td>2.03</td>
<td>2.18</td>
</tr>
<tr>
<td>FDC2WITH3</td>
<td>0.496</td>
<td>1.94</td>
<td>1.91</td>
<td>1.96</td>
</tr>
<tr>
<td>(2) NARROW 9.5mm diameter</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FDC1B1W2D</td>
<td>0.450</td>
<td>1.82</td>
<td>1.63</td>
<td>1.87</td>
</tr>
<tr>
<td>FDC1B1W3D</td>
<td>0.432</td>
<td>1.86</td>
<td>1.59</td>
<td>1.93</td>
</tr>
<tr>
<td>FDC1B2W3D</td>
<td>0.415</td>
<td>1.90</td>
<td>1.56</td>
<td>1.99</td>
</tr>
</tbody>
</table>

For Method 1, the exponents of $f$ were close to .5, but the corresponding
G's were too low. In the second method, when the f-exponent was forced to .5, the G's were lower still. The highest G-values were obtained using the third analysis method. Results were better for the 12.7 millimetre tube, although theoretically G should be the same for all smooth cylindrical tubes.

These results were pleasing, considering the possible sources of error in the measurements. The measurements at the three positions were not of the same pulse; improved results might be obtained if data at the three positions could be recorded simultaneously, using three microphones. In addition, some error was introduced as a result of windowing (Section 2.3.2.2).

The above method could also be used to investigate the attenuation of pulses in flaring tubes, and thus provide useful information about losses in brass instruments.

### 3.1.4. Choice of Microphone

Three microphones have been used.

#### 3.1.4.1. Probe Microphone

The horn-coupled probe microphone (B and K Type 4170) is the most expensive of the three, retailing at £1456 (1986 price; the microphone has its own built-in preamplifier).

With a 10 kHz cut-off frequency, it removes the need for the anti-aliasing filter with 20 kHz cut-off frequency, but also reduces the high frequency detail, much of which has already been attenuated by the source tube. The low frequency response is poor, as indicated by the dip after the impulse (see Figure 3.16(a)).

#### 3.1.4.2. 1/8" Microphone

The 1/8" microphone (B and K Type 4138) is less expensive, costing about £1107 (1986 price, which includes extra for preamplifier and adapter), but has a broader response of 100 kHz. Such a broad response is not necessary when an anti-aliasing filter is removing the frequency content above 20 kHz, but is
Figure 3.16 Comparison of the Performances of (a) Probe, (b) 1/8", (c) Knowles Electret Microphones.
useful for investigations such as those in Section 3.1.3.5.

The post-impulse dip shown in the probe microphone response was not present, indicating a superior low-frequency response (see Figure 3.16(b)).

In view of the good impulse shape and optimum high frequency detail, we have preferred to use the 1/8" microphone for instrument measurements.

3.1.4.3. Electret Microphone

Two Knowles electret microphones (CA2832 and BT1759), having reasonably flat responses over the 20 kHz audio frequency range, were both affected by electrical pick-up from the spark source. This caused an additional decaying voltage on the microphone output signal (see Figure 3.16(c)). Attempts to remove this by electrical screening were unsuccessful. Placing a resistor and capacitor across the output high-pass filtered the signal so that the electrical pick-up level decayed to zero more quickly; but this spoiled the low frequency response of the remaining acoustic signal, as evidenced by the characteristic post-impulse dip.

The Knowles microphones are inexpensive and produce acoustic measurements which are detailed enough to be used successfully in the instrument comparison technique (Chapter 4). However, the performance of the 1/8" microphone is superior.

3.1.5. Insulation

Transient acoustic measurements will be affected by ambient conditions. To reduce these effects, thermal and acoustic insulation have been added to the impulse apparatus.

The source tube is surrounded with domestic pipe lagging. The brass instrument is enclosed within a polystyrene box. Ambient background noise enters the bell of the instrument and is efficiently coupled back into the microphone. Completely enclosing the instrument reduces the effect of this to some extent.
When the inside wall of the insulating box is close to the instrument bell, the shape of the bell reflection is altered. This does not matter for simple instrument comparison measurements (Chapter 4), but should be avoided when accurate measurements of transient response are required for analytical purposes (Chapter 5).

3.1.6. Summary of Section 3.1

In this section, the equipment for measuring the transient response of brass instruments has been described. A minicomputer controls impulse production and data acquisition by sampling.

A spark source was chosen because it produced impulses with a good (delta-function-like) shape, even before deconvolution. Good repeatability enabled signal averaging to be carried out. The main problems with the source were inconsistency and EM interference.

The source was rendered non-reflecting by the introduction of a long source tube. This had many other advantages. It removed unwanted high frequency energy; it made independent measurements of input signal straightforward; it provided a way of compensating for ambient temperature changes; it has been used to investigate the attenuation of pulses in cylindrical tubes.

Three types of microphone have been used. The performance of the electret microphones was spoiled by pick-up of EM radiation from the spark source. The 1/8" microphone was preferred to the probe microphone because of its broader frequency response, and better response at low frequencies.

Finally, the use of thermal and acoustic insulation was discussed.

Photographs of the apparatus (minus insulation) are shown in Figures 3.17 (a), (b) and (c).
Figure 3.17 Photographs of the Impulse Measurement Equipment. (a) Complete Apparatus, (b) Close-Up of Source End, (c) Close-Up of Instrument End
3.2. Software

3.2.1. Main Features of Data Acquisition Software

A flowchart illustrating the main points of the data acquisition programme is given in Figure 3.18.

Before actual data acquisition, a pre-measurement trial run of the equipment is carried out. During this run, amplitude (in volts) and timing (in samples) of each successive impulse is listed on the VDU. Amplitude is set to the desired level (normally about 8 volts) by adjusting the B and K amplifier (Type 2608) gain settings.

After this adjustment, the microphone is calibrated using a pistonphone and special microphone calibration software. The pistonphone generates a sinusoidal pressure signal of known rms value (31 pascals). The sinusoidal microphone output is read into memory via the ADC; its rms voltage is determined, and so the calibration constant in pascals/volt is calculated. The calibration procedure incorporates a means of compensating for changes in ambient pressure. It should be noted that accurate calibration is not always necessary, eg. when runs are normalised for comparison purposes (Section 3.3.3). However, the true impulse amplitude is always a useful indicator of the state of the spark plug. If the plug is producing impulses of lower amplitude than normal, it could need cleaning, or even replacing.

After calibration, the desired sampling rate is determined using the method to be described in Section 3.2.4.

Data acquisition begins immediately after the spark source has been triggered. Careful synchronisation between pulse production and data acquisition is necessary in order that signal averaging can be carried out. Small errors in this synchronisation are the subject of further discussion in Section 3.2.3. Signal averaging is considered further in Section 3.2.5.

The programme was designed to overcome problems of impulse inconsistency. Each incoming pulse is carefully monitored. If its amplitude or timing do not satisfy certain pre-determined criteria, the run will not be included in the
Figure 3.18
Flow Chart of General Data Acquisition Programme "IMPULSES"

Beforehand,
(1) ensure amplifier settings are correct,
(2) calibrate microphone,
(3) determine correct sampling rate and desired location of impulse peak.

- Input number of averages to be included, impulse amplitude acceptance window, and desired impulse peak location. Set programmable oscillator to desired sampling frequency.

- Send trigger pulse to DAC (and spark source). Collect 2048 data samples via the ADC and store in integer array. Convert from integer to real. Locate impulse peak and hence determine impulse amplitude and timing.

- Pulse amplitude within acceptance window? No
  - Correct location? No
    - Include in running average.
  - Correct location? Yes

- Enough running averages? No
  - Calibrate by multiplying all data points by calibration constant. Write away 2048 elements of averaged data

Yes
final averaged version. Further discussion of the pulse-scanning procedure and sources of inconsistency will be given in the following sections. The amplitude acceptance window additionally ensures that any runs with ADC overflow will be detected and rejected.

3.2.2. Interpolation

The original pulse-scanning procedure consisted of searching for the sample at which the signal amplitude was highest. This sample, corresponding to the peak of the input impulse, is termed the "peak sample number" or PSN. The amplitude at the PSN is the impulse amplitude. If the PSN did not correspond with the pre-determined desired PSN, the run was rejected.

The time taken for the impulse to travel from the source to the microphone is given by

\[ t = \text{PSN} / f_s \]  \hspace{1cm} (3.6)

where \( t \) is time in milliseconds and \( f_s \) is sampling rate in kHz. It soon became clear that it was necessary to determine \( t \) more precisely. The two main reasons for the need for greater accuracy were:

1. so that smaller impulse timing inconsistencies could be detected and removed, thus effectively ensuring improved consistency in impulse timing (Section 3.2.3),

2. so that the correct sampling rate could be determined with greater precision (Section 3.2.4).

Two methods of improving the accuracy of \( t \) were considered. The first possibility was using a higher sampling rate. When 150 kHz (the maximum rate possible on the present apparatus) is used instead of 46 kHz (the rate normally used) there is a good improvement in the accuracy of \( t \) (see below).

The alternative is interpolation. Here one calculates data values at inter-sample points and finds the "fractional peak sample number" or FPSN at which the impulse peak occurs.
Interpolation led to greater accuracy, as the following figures show.

Time interval between two samples at 46 kHz = .0217 milliseconds

" " " " " " 150 kHz = .0067 milliseconds

(31% of that at 46 kHz without interpolation)

" " in .01 of a sample at 46 kHz = .0002 milliseconds

(1% of that at 46 kHz without interpolation)

Interpolation can be used with any sampling rate. When the sampling rate is increased, the essential parts of the instrument response may no longer fit into a 2048-element array; a larger array may be required.

For these reasons, interpolation was the preferred method.

3.2.2.1. Method of Interpolation

A polynomial is fitted through an odd number of sampled data points around the impulse peak. The number of points is odd so that the sample at which the peak impulse amplitude occurs (the PSN) is central, with an equal number of sample points on either side.

The polynomial fit is exact. This means that for n data points, the fitted polynomial is of order (n-1). The n polynomial equations are summarised in one matrix equation, and the polynomial coefficients are calculated using Gaussian elimination, a form of triangular decomposition (see, for example, Wilkinson, 1965). Using the resulting polynomial, pressure is calculated at inter-sample points (ISP) close to the impulse peak. The ISP at which the maximum pressure occurs is taken to be the FPSN. ISPs are usually .01 of a sample apart.

The subroutine "Interpeak" performs the interpolation.

3.2.2.2. Detailed Study of Interpolation

The above calculation involves many steps, meaning that cumulative computer-arithmetic errors may affect the accuracy of the final result. The higher the order of the polynomial, the more steps there are, and thus the
greater the likelihood of error. For polynomials of order 14 and below, this error was found to be negligible.

The n data points also had to be shifted along the "sample number axis" towards the origin, so that the first point occurred at sample 1, the n th at sample n, etc. This is because normally the n data points occur at high sample numbers; for instance, at 46 kHz, the impulse normally occurs at sample 483. If large sample numbers are used in the interpolation, errors are introduced. Thus the points are shifted to lower sample numbers during the calculations, and afterwards shifted back to their original positions.

The most suitable order of polynomial had to be chosen. For the impulse, it was found that the higher the order of the polynomial, the lower the calculated FPSN. See the first three columns of Table 3.2(a) for an example.

To decide which order of polynomial was best, plots showing the original data points and comparing the fitted polynomials were made. The highest order gave the smoothest fit through the data points around the peak. Lower order polynomials tended to produce loops nearer the peak. Figure 3.19 illustrates this; it compares polynomials of order 4, 8 and 14.

However, the higher the order of the polynomial, the longer the polynomial fitting procedure takes. In the factory, speed of data acquisition is an important consideration. If n is increased from 5 to 15, the time taken to gather and average 50 data runs increases by about 60 % (from 1 minute to 1 minute 36 seconds).

Another source of variation of calculated FPSN is the position of the PSN in relation to the n points to be used in the polynomial fit. Slightly displacing the PSN from the central position was thought to be a possible way of compensating for the fact that the impulse is asymmetrical. Column 4 of Table 3.2(a) shows that when points (1-1) to (m-1) are used instead of points 1 to m, the FPSN increases. When points (1+1) to (m+1) are used, the FPSN decreases, as shown in the fifth column. This latter decrease has a greater
TABLE 3.2
RESULTS OF DETAILED STUDIES OF INTERPOLATION AROUND THE IMPULSE PEAK

TABLE 3.2(a) Actual FPSN ≈ 483.0. Sampling rate ≈ 46kHz.

<table>
<thead>
<tr>
<th>n</th>
<th>ORDER</th>
<th>FPSN when points 1 to m are used, ie. FPSN is central</th>
<th>FPSN when points (1-1) to (m-1) are used instead of points 1 to m</th>
<th>FPSN when points (1+1) to (m+1) are used instead of points 1 to m</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>2</td>
<td>483.07</td>
<td>483.30</td>
<td>482.01</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>483.00</td>
<td>483.02</td>
<td>482.78</td>
</tr>
<tr>
<td>7</td>
<td>6</td>
<td>482.96</td>
<td>482.97</td>
<td>482.86</td>
</tr>
<tr>
<td>9</td>
<td>8</td>
<td>482.94</td>
<td>482.94</td>
<td>482.88</td>
</tr>
<tr>
<td>11</td>
<td>10</td>
<td>482.93</td>
<td>482.93</td>
<td>482.88</td>
</tr>
<tr>
<td>13</td>
<td>12</td>
<td>482.92</td>
<td>482.92</td>
<td>482.89</td>
</tr>
<tr>
<td>15</td>
<td>14</td>
<td>482.91</td>
<td>482.92</td>
<td>482.89</td>
</tr>
</tbody>
</table>

TABLE 3.2(b) Actual FPSN ≈ 483.50. Sampling rate ≈ 46kHz.

<table>
<thead>
<tr>
<th>n</th>
<th>ORDER</th>
<th>FPSN when points 1 to m are used, ie. FPSN is central</th>
<th>FPSN when points (1-1) to (m-1) are used instead of points 1 to m</th>
<th>FPSN when points (1+1) to (m+1) are used instead of points 1 to m</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>2</td>
<td>483.56</td>
<td>483.52</td>
<td>483.01</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>483.49</td>
<td>483.50</td>
<td>483.46</td>
</tr>
<tr>
<td>7</td>
<td>6</td>
<td>483.48</td>
<td>483.49</td>
<td>483.48</td>
</tr>
<tr>
<td>9</td>
<td>8</td>
<td>483.48</td>
<td>483.48</td>
<td>483.48</td>
</tr>
<tr>
<td>11</td>
<td>10</td>
<td>483.48</td>
<td>483.48</td>
<td>483.48</td>
</tr>
<tr>
<td>13</td>
<td>12</td>
<td>483.48</td>
<td>483.48</td>
<td>483.48</td>
</tr>
<tr>
<td>15</td>
<td>14</td>
<td>483.48</td>
<td>483.48</td>
<td>483.48</td>
</tr>
</tbody>
</table>

TABLE 3.2(c)

<table>
<thead>
<tr>
<th>Actual FPSN</th>
<th>Difference between FPSN’s calculated for n=5 and n=15 polynomials</th>
<th>Change in FPSN when points (1+1) to (m+1) are used instead of points 1 to m</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>for n=5 case</td>
<td>for n=15 case</td>
</tr>
<tr>
<td>482.5</td>
<td>.02</td>
<td>.04</td>
</tr>
<tr>
<td>482.6</td>
<td>.05</td>
<td>.08</td>
</tr>
<tr>
<td>482.7</td>
<td>.07</td>
<td>.12</td>
</tr>
<tr>
<td>482.8</td>
<td>.09</td>
<td>.17</td>
</tr>
<tr>
<td>482.9</td>
<td>.10</td>
<td>.21</td>
</tr>
<tr>
<td>483.0</td>
<td>.09</td>
<td>.22</td>
</tr>
<tr>
<td>483.1</td>
<td>.07</td>
<td>.23</td>
</tr>
<tr>
<td>483.2</td>
<td>.05</td>
<td>.22</td>
</tr>
<tr>
<td>483.3</td>
<td>.03</td>
<td>.15</td>
</tr>
<tr>
<td>483.4</td>
<td>.02</td>
<td>.07</td>
</tr>
<tr>
<td>483.5</td>
<td>.01</td>
<td>.03</td>
</tr>
</tbody>
</table>

magnitude than the former increase because the onset of the impulse is steeper than its decay. Both effects can be seen to decrease as the order of the poly-
Figure 3.19 Comparison of Polynomials of Different Orders Fitted Through the Impulse Peak: Order 4 (Dashed), Order 8 (Continuous), and Order 14 (Dot-Dash). The Higher the Order, the Less "Loopy" the Curve Near the Peak.
nominal increases.

A systematic survey was carried out to discover how the position of the actual FPSN with respect to its two neighbouring sample points influences the extent of the aforementioned variations. The variations are found to be less when the FPSN is half way between the two samples, rather than closer to one or the other. This can be seen by comparing Tables 3.2(a) and (b). In Table 3.2(a), the interpolated FPSN is virtually at the same position as sample 483. In Table 3.2(b), the FPSN is half way between two samples, at 483.5. Table 3.2(b) shows a smaller FPSN decrease with increasing polynomial order, and less FPSN variation with PSN position. In Table 3.2(b) the results are identical for polynomials of order 8, 10, 12 and 14; so in this case, results obtained when using a polynomial of order 8 are just as accurate as those obtained when using a polynomial of order 14, and they are obtained more quickly.

Table 3.2(c) summarises what happens as the interpolated FPSN moves further away from the mid-point of the two samples. Column 2 shows an increasing difference in FPSN between the n=5 and n=15 polynomials. Columns 3 and 4 show an increasing effect on the FPSN when the n points are shifted, and that this effect is much less when n=15 than when n=5.

From the above detailed investigations of the interpolation procedure, the "standard" FPSN was chosen to be 483.50 (rather than 483.00) during data runs using a sampling rate of about 46 kHz. Nine points rather than 15 are used for the polynomial fit, to give optimum accuracy and speed. The PSN is placed at the centre of the nine points, although shifting it to the right or left would make no difference to the result here.

3.2.3. Impulse Inconsistency Investigations

3.2.3.1. Sources of Inconsistency and Remedies

Ideally, when signal averaging is carried out, each impulse should be
identical. For this reason, much time has been spent considering the problem of impulse inconsistency and how to solve it.

The two most convenient indicators of impulse inconsistency are pulse timing and pulse amplitude.

Sources of inconsistency can be considered in three categories. The first is the spark source itself. In Section 3.1.2.2, spark sources were seen to be inherently inconsistent. Variation in amplitude is the most obvious result. Small random variations in shape will also produce small changes in the precise timing of the impulse peak.

A second source of impulse inconsistency is ambient conditions, in particular ambient temperature. A change in temperature produces a change in impulse velocity in the source tube, and hence in the timing of the impulse. The point is considered in Section 3.2.4.

An additional source of timing inconsistency is an inherent fault in the computer sampling process. The precise instant at which the computer sampling process begins after the pulse has been triggered may vary by one or two sample periods, because the CPU clock and the sampling (external) clock are not synchronised. This would result in the whole data record being shifted by one or two sample periods. The data acquisition programme now contains software which detects when this occurs and shifts the whole record back to its "correct" position. If this correction were not made, signal averaging would lead to blurring.

With interpolation, very small changes in pulse timing can be detected. Experience has shown that most random changes in timing occur within a range of .06 of a sample at 46 kHz (ie. 1.3 microseconds). Data runs whose FPSN lies outside this .06 "acceptability window", typically from FPSN's 483.47 to 483.53 inclusive, are now rejected. This test is therefore now much more stringent than it was before interpolation was introduced.

Similarly, when the impulse amplitude does not lie within certain limits,
the run is rejected. The ADC voltage range is from -10 to +10 volts. The amplitude window is usually 2 volts and placed near the upper end of the ADC range, although placing it at any position within the 10-volt range will give adequate resolution.

If all impulses were identical in shape an amplitude window would not be necessary (apart from the need of an upper limit to prevent ADC overflow). Tests have shown that the general impulse shape is the same but small variations may occur. When larger variations in shape occur, they are accompanied by a greater change in timing or amplitude; as each impulse is tested for its acceptability with respect to timing and amplitude, any pulses with particularly adverse shape will automatically be detected and excluded from the averaged version.

3.2.3.2. Method of Monitoring Inconsistencies

To investigate the problem of impulse inconsistency more carefully, computer software has been written which stores the amplitude and timing of each successive impulse. Additional parameters, for instance, the root mean square (rms) background noise level and the time interval between the impulse and later features in the record, may also be written away. This information can be analysed in four different ways.

First, for many impulses in succession, overall distributions of each parameter can be plotted, and various statistical values, eg. mean, mode and standard deviation, can be calculated. An example is given in Figure 3.20 which shows the distribution of impulse amplitude in volts for 1475 successive impulses. (This distribution was neither Gaussian nor Binomial.) The tail at lower amplitudes suggests cases in which the spark discharged before the voltage across the spark gap had reached its maximum value.

A plot of overall distribution of FPSN (timing) will normally only be useful if the temperature stayed constant throughout the readings. A temperature change will broaden the distribution. The overall distribution gives no
Figure 3.20 Distribution of Impulse Amplitudes in Volts for 1475 Successive Impulses.
information about how the parameter varies as time progresses. To look at this, one would plot the value of the parameter for each impulse in turn. Figure 3.21 shows the run-by-run variation of FPSN over 200 runs during a temperature rise of .3°C. This plot reveals a gradual decrease in FPSN, although many additional random variations can be seen. These random fluctuations could sometimes obscure a gradual increase or decrease of a parameter. Calculation of the mean and standard deviation over the previous ten runs for each impulse in turn can help to provide a clearer picture - see Figure 3.22. Care should be taken when interpreting these results; an apparent peak or trough may be caused by one spurious impulse.

The fourth way of obtaining useful information from the parameter variation data is to determine how the running average changes with increasing number of averages. With impulse amplitude, the variability of the running average was usually large for the first few impulses, but became steadier as the number of averages increased (see Figure 3.23). This illustrates why statistical variations are more likely to be reduced when more averages are used.

It has also been interesting to investigate the running average of background noise level; this work is presented in Section 3.2.5.

In the next section further results of the above methods of inconsistency investigation will be given.

3.2.3.3. Results of Inconsistency Investigations

Investigations into various aspects of impulse inconsistency have been carried out.

First, an investigation was carried out into whether there was a discernible pattern in the way that impulse amplitude varied as the number of successive firings increased. The run-by-run variation of amplitude over 200 successive runs revealed no clear trend (Figure 3.24). Calculating the mean over every previous ten runs was also unhelpful. Variation of running average sometimes showed a decrease (Figure 3.25), suggesting that impulse amplitude
Figure 3.21 Run-by-Run Variation of FPSN for 200 Consecutive Spark Impulses During a Temperature Rise of 0.3 °C.
Figure 3.22 Variation of Mean-FPSN-Over-Previous-10-Impulses For Each Impulse in Turn.
Figure 3.23 Variation of Running Average of Spark Impulse Amplitude with Increasing Number of Averages for 1475 Successive Impulses.
Figure 3.24 Run-by-Run Variation of Spark Impulse Amplitude for 200 Successive Impulses.
Figure 3.25 Variation of Running Average of Impulse Amplitude with Increasing Number of Averages for 200 Successive Spark Impulses.
decreased with time while the plug was continually firing. However, this did not always happen — see, for instance, Figure 3.23. From four independent investigations in this area no clear trend could be deduced.

The random variation of impulse timing can be seen in Figure 3.26, a plot of overall FPSN distribution for 110 consecutive impulses. Here the temperature stayed constant throughout. 83% of runs lay within a timing window of .06 samples (at 46 kHz). This percentage varies; on two other occasions at which the temperature stayed constant it was 85% and 53%.

Experience has shown that impulse consistency varies from one day to the next. Amplitude variation is worse on some days than on other days. For instance, on a certain day, the standard deviation of an amplitude distribution was .60 volts, and four days later it was .49 volts. The extent of the variation was originally thought to be dependent on ambient conditions such as calmness of weather. The measurement laboratory was susceptible to weather conditions, being situated on a hill. However, impulse readings were sometimes more consistent during stormy weather than during calm weather. The cause of variation of consistency between days still remains obscure.

An experiment was carried out to determine whether consistency was related to the extent to which the apparatus had warmed up after switching on. The first 50 runs immediately after switching on had an amplitude standard deviation of .50 volts. After 25 minutes, a further 50 consecutive impulses had a slightly lower standard deviation of .43 volts. Therefore the time for which the apparatus has been switched on may have a small effect on impulse consistency, but there is obviously a more important factor involved.

Comparison of loudspeaker and spark sources revealed that the loudspeaker was far less susceptible to random timing variations. Figure 3.27 shows run-by-run variation of FPSN for 200 loudspeaker impulses during a temperature drop. This is seen to be far smoother than for the spark source (see, for example, Figure 3.25). Surprisingly, the consistency of the loudspeaker
Figure 3.26 Overall FPSN Distribution for 110 Consecutive Spark Impulses.
Figure 3.27 Run-by-Run Variation of FPSN for 200 Consecutive Loudspeaker Impulses During a Temperature Drop.
impulse amplitudes was slightly worse with a greater amplitude range and standard deviation. The reason for this is not clear. One would have expected greater amplitude consistency with the loudspeaker.

Finally, the effect of a localised temperature change has been closely monitored. While the spark source was firing, a hand was placed on the open-ended source tube for one minute to produce a localised temperature rise. The hand was removed and the impulses continued during the ensuing temperature drop. For each impulse, the FPSN of the forward impulse and its reflection from the open end were monitored. The experiment was carried out twice, first with the hand placed between the spark source and microphone, and secondly with the hand between the microphone and the open end. In the first experiment the mean FPSN's of both the incident and reflected impulses dropped by about 0.07 samples, then gradually increased as the temperature decreased. In the second experiment, the FPSN of the reflected impulse was initially reduced by .29 samples with the incident impulse FPSN staying about the same. This greater reduction is because the impulse travels twice through the warmer region before reaching the microphone. These experiments illustrate that handling a brass instrument before measurements could result in small shifts in the timings of its transient response features. Therefore, thermal gloves should be worn when handling instruments. Instruments should be handled for as short a time as possible, and instruments should, if possible, be left on the apparatus for a while before they are measured, to give time for the warmer areas to cool.

3.2.4. Compensation for Ambient Temperature Drifts

3.2.4.1. Why Temperature Compensation is Necessary

From the Gas Laws, the velocity of sound in free air is proportional to the square root of the absolute temperature. The velocity of sound in a cylindrical tube varies in a similar way with temperature, except that it will
be lower because of visco-thermal losses (Weston, 1957). It also varies with frequency. Accordingly, an increase (or decrease) in ambient temperature will result in an increase (or decrease) in impulse velocity, and hence a decrease (or increase) in impulse FPSN; so all features of the measured transient response will occur slightly too early (or late), i.e. the response is "squashed" (or "stretched"). This creates problems when a direct comparison of two transient responses measured at different temperatures is required.

Transient responses are normally compared by finding their arithmetical difference (more detail on this is found in Section 4.2). A large arithmetical difference denotes a significant physical difference between instruments. Shifted transient response features will lead to additional spikes in the arithmetical difference plot which should not be there. See, for example, Figure 3.28, which shows the arithmetical difference between two transient responses of the same trombone, but measured at different temperatures, one at 20 °C and one at 23 °C.

3.2.4.2. Compromise by Changing Sampling Rate

The solution is simply to modify the sampling rate as necessary so that the impulse always occurs at the same fixed FPSN. Subsequent plots and processing are carried out as if the sampling rates were not changed.

The fixed FPSN is usually 483.50 (as explained in Section 3.2.2.2). The correct sampling rate is chosen on the basis of the velocity of sound in the source tube between the spark source and the microphone. The initial choice of sampling rate, \( f_{s1} \), is close to 46 kHz (and can be calculated from the measured ambient temperature using the relationship to be given in Section 3.2.4.3). This produces an FPSN, \( x_1 \). The correct sampling rate, \( f_{s2} \), results in an FPSN of 483.50. As the time taken for the impulse to reach the microphone is the same in both cases, the following equations can be written.

\[
\frac{x_1}{f_{s1}} = \frac{483.50}{f_{s2}} \quad (3.7a)
\]
Figure 3.28 Arithmetical Difference Between Two Transient Responses of the Same Trombone Measured at Different Temperatures (20 and 23 °C) Using a Sampling Rate of 46 kHz.
In practice, \( x \) is the mean of ten preliminary FPSN's because of the statistical variations in impulse timing. The software for obtaining the correct sampling rate is as follows (a flowchart is given in Figure 3.29).

1. The impulse source fires ten times. Data is read in at sampling rate \( f_{s1} \) and all FPSN's are noted. Some of these FPSN's will be one or two samples higher than the rest because of the computer error (mentioned in Section 3.2.3.1). The latest FPSN, termed "FPSNMAX", is noted.

2. The impulse source fires a further ten times and data is again collected at rate \( f_{s1} \). If an FPSN is more than .6 samples lower than FPSNMAX, the whole record is shifted on one sample to compensate for the computer error. The mean \( (x_1) \) of the ten FPSN's is calculated and used to determine the new sampling rate \( f_{s2} \). FPSNMAX is modified by multiplying it by \( (f_{s1}/f_{s2}) \).

3. A further ten runs are performed using the new sampling rate \( f_{s2} \), and the mean FPSN is calculated. If the mean is sufficiently close to 483.50, further runs are carried out using rate \( f_{s2} \). If not, a new \( f_{s2} \) is calculated based on the new mean, and the test is repeated until the correct rate is found.

With interpolation, the correct sampling rate can be determined far more accurately. To demonstrate this, a trombone was measured twice - first in its standard state, and secondly with its waterkey open. The arithmetical difference was found. The two measurements were carried out three times. In each case, the temperature compensation procedure was used; in two cases interpolation was used but in the other case it was not. Figure 3.30 compares the three arithmetical difference plots. Without interpolation the plot is less smooth with spikes occurring before the waterkey position. (This shows that

\[
f_{s2} = 483.50 \frac{f_{s1}}{x_1} \quad (3.7b)
\]
Figure 3.29

Flow Chart to Illustrate the Steps Involved in Finding the Correct Sampling Rate

(1) Record 10 impulses using sampling rate $f_{s1}$.
   Note FPSNMAX for computer-error compensation on all future runs.

(2) Record a further 10 impulses using sampling rate $f_{s1}$.
   Calculate the mean FPSN of these 10.

   Calculate the new sampling rate $f_{s2}$
   and modify FPSNMAX as necessary.

(3) Record 10 impulses using sampling rate $f_{s2}$.
   Calculate the mean FPSN of these 10.

Is mean FPSN sufficiently close to 483.50?

   Yes
   Continue with data acquisition

   No
Figure 3.30 Comparison of Arithmetical Difference Plots with and without the Use of Interpolation During the Measurements. (a) No Interpolation, (b) Interpolation to an Accuracy of .1 Samples, (c) Interpolation to an Accuracy of .01 Samples.
corresponding features on the two instrument response curves are at slightly different positions because of a slightly inaccurate choice of sampling rate.)

To set the selected sampling rate as precisely as possible, the frequency resolution of the programmable oscillator should be good. The Adret oscillator is accurate to .1 Hz. An accuracy of at least six decimal places is required.

3.2.4.3. Relationship between Temperature and Sampling Rate

Such a relationship would be useful because then the sampling rate could be automatically predicted from measured ambient temperature. However, the precise relationship between temperature and velocity of sound in a tube is not yet known. If tube velocity is assumed to be close to free air velocity, then tube velocity, and hence sampling rate, will be roughly proportional to the square root of the absolute temperature.

To investigate whether this was so, all measurements carried out using the standard 12.7 millimetre internal diameter source tube over a period of seven months were reviewed, and a note was made of sampling rates used at different temperatures. The result is plotted in Figure 3.31. The relationship is seen to be linear. A linear regression programme produced the following relationship.

\[ f_s = 44.2627 + 0.079 T \]  (3.8)

where \( f_s \) is sampling rate in kHz and \( T \) is temperature in ° Celsius.

As the temperature range covered in the above data is small (15 to 25 °C), it is possible that the proposed square root curve in that region may be a good approximation to a straight line.

The above linear relationship has been used to make a first approximation to the correct sampling rate.

3.2.5. Signal Averaging

Signal averaging improves the S-N ratio when the noise is
Figure 3.31 Relationship Between Sampling Rate and Temperature for Consistent FPSN.
(1) random,
(2) uncorrelated, or linearly independent, i.e. the average of a product of two random variables is equal to the product of the averages, i.e.
\[
\langle v_1 v_2 \rangle = v_1 \cdot v_2
\]
(3.9)
(3) ergodic, i.e. time averages of almost all possible sample sequences are equal to the same constant and equal to the same ensemble average,
(4) stationary, i.e. all probability functions are independent of a shift of time origin.

(See Oppenheim and Schafer, 1975, ch.8.) To satisfy these criteria, the sampling rate must not be a multiple of the mains frequency, 50 Hz, i.e. it must not take on values of 45.95, 46.00, 46.05 kHz, etc.

The incident impulse should ideally be identical for each data record. Although spark impulses are inconsistent, as discussed earlier, their consistency has effectively been improved by rejecting impulses which lie outside stringent limits of acceptability in amplitude and timing.

In Sections 3.1.1 and 3.1.2.1, the problem of a small extra DC voltage appearing on the ADC was mentioned. A simple model of background noise variation was developed to investigate whether this was happening. The noise level is taken to be the rms of m samples of random noise. When a number, N of successive sequences of m samples are added, the rms noise level of the resulting sequence increases with \( \sqrt{N} \). To find the running average, the rms level of the resulting sequence is divided by N. Thus, when signal averaging is carried out, the rms background noise level decreases with \( 1/\sqrt{N} \), i.e. the S-N ratio increases with \( \sqrt{N} \).

A computer programme, IBGMODEL2, was written to verify this. It was structured as follows.

(1) Generate a sequence of m random numbers between -1 and +1; find the rms of these (N=1).
(2) Add further sequences \((N = 2, 3, \ldots)\) of random numbers to the first sequence, each time calculating the rms of the resulting sequence and dividing it by \(N\).

Figure 3.32 shows the \(1/\sqrt{N}\) variation of mean rms level over 200 runs, as expected.

To simulate the effect of the small additional voltage on the ADC, a small number, 0.03, was added to each generated random number. Figure 3.33 shows that the curve tends towards 0.03, instead of zero.

The rms noise variation over 200 runs was then calculated for real measured data using a selected range of about 200 samples before the arrival of the acoustic impulse. The result is shown in Figure 3.34. This curve is not as smooth as the theoretical curves. The noise level certainly decreases after the first few runs, but then fluctuates, and for the final few runs appears to increase. The fluctuation suggests that the noise distribution may not be purely random. No conclusion about the presence of an additional voltage on the ADC can be made from this plot.

3.2.6. Summary of Section 3.2

In this section, the essential features of the data acquisition software were described, including microphone calibration, source triggering, collection of 2048 samples of data and signal averaging of suitable data runs.

The programme now contains refined procedures for detecting inconsistencies in impulse amplitude and timing. Inconsistent runs are not included in the final averaged version, so in effect, the problem of spark source inconsistency is solved.

Interpolation, using a polynomial fit by Gaussian elimination, determines the precise location of the impulse peak to an accuracy of 0.01 of a sample. It has enabled careful investigations of various aspects of impulse inconsistency to be carried out.
Figure 3.32 Variation of Mean RMS Level of N Sequences of m Computer-Generated Random Numbers with Successive Values of N up to 200.
Figure 3.33 As for Figure 3.32, but with 0.03 Added to Each Generated Random Number.
Figure 3.34 As for Figure 3.32, but Using 200 Sequences of 200 Samples of Measured Background Noise (Before the Arrival of the Acoustic Impulse).
Variations of ambient temperature cause the impulse timing to change, which is inconvenient when comparisons of responses are desired. This problem has been overcome by modifying the sampling rate to compensate for the temperature change. Interpolation ensures that the required sampling rate is determined as accurately as possible.

Finally, the effectiveness of the signal averaging technique has been assessed by comparing the true variation of background noise level with the results predicted by a theoretical model.

3.3. Measurements for Analysis or Comparison

3.3.1. Introduction

Measurements of instruments can be divided into two categories.

(1) Those to be used for later theoretical analysis; analysis procedures will be dealt with in Chapter 5.

(2) Those involving the comparison of two instruments, normally of the same model; this is discussed in Chapter 4.

Although the essential experimental techniques, to be described in the next section, are the same for both categories, there are differences in the subsequent processing. Standard post-measurement processing will be explained and then details of the extra measurements required for the analysis will be given. Finally some examples of measurements that have been made on a variety of brass instruments will be given.

3.3.2. Summary of Practical Details

The apparatus is switched on and given time to settle to a steady state (at least ten minutes). During measurements laboratory windows and doors are kept closed and curtains are drawn to minimise the effects of weather (wind, sun) on ambient conditions. Before a set of measurements is made, the spark plug electrodes are cleaned by brushing to reduce the effects of corrosion (Section 3.1.2.3).
When handling instruments, thermal gloves are worn to minimise heat transfer from the hands to the instrument (Section 3.2.3.3). This also protects the surface of the instrument (sometimes silver plate) from finger marks.

The adjustment of amplifier gain settings and calibration have been discussed in Section 3.2.1. The method of setting up the correct sampling rate was explained in Section 3.2.4.2.

Before data acquisition begins, the number of averages, and the impulse timing and amplitude acceptance windows are chosen (see Figure 3.18). 50 averages are normally used. The normal choice of windows is given in Section 3.2.3.1, although sometimes, when pulses are particularly inconsistent, these are broadened to speed up the measurement.

3.3.3. Standard Post-Measurement Processing

All measurements are subjected to a standardisation procedure before further analysis. The procedure involves the following steps.

1. Subtraction of pre-impulse DC-level from all data samples within the record. This DC-level could be caused by low-frequency ambient noise which has not been removed by signal averaging. Additionally, if there has been a small voltage build-up on the ADC, it will be automatically removed at this stage. In practice, the data record is shifted so that it starts at the foot of the incident impulse, that is, immediately before the impulse attack commences, and the value at the foot is subtracted from all elements. Thus, at the foot the value becomes zero.

2. Normalisation. All elements are divided by the impulse amplitude, so that the amplitude of the input impulse becomes unity. This is necessary because of impulse amplitude inconsistency; even after many signal averages and stringent pulse acceptability criteria, the final impulse amplitude varies from one averaged data record to the next. Another reason for normalisation is the variation of the spark plug performance with time;
the impulse amplitude decreases as the electrodes become more corroded. In addition, different spark plugs produce acoustic impulses of differing amplitude.

(3) Windowing. The precise shape and size of the window is dependent on the type of processing to be carried out afterwards. The window starts at the position corresponding to the start of the instrument, i.e. the source-tube instrument junction. For subsequent instrument comparison, windows are normally rectangular and end at the position of the bell reflection (see Section 4.2.2). For the more detailed analyses of Chapter 5, more sophisticated window types are required which extend beyond the bell reflection.

3.3.4. Additional Measurements Required for Analysis

Two additional measurements are required when the response data is to be used for deconvolution followed by bore reconstruction. Independent measurements of incident and reflected signals are needed, and in fact, the following three separate measurements are carried out to achieve this (see Section 2.3.4.2). (NOTE: Figure 3.10 shows points "X" and "Y" on the source tube.)

(A) A detachable rigid brass termination is placed at point "Y". The resulting pressure variation is shown in Figure 3.35(a).

(B) The 48 centimetre end-section, "XY", of the source tube is replaced by a "semi-infinite" extension tube of the same diameter, having length of approximately 3.58 metres. The resulting pressure variation is shown in Figure 3.35(b). The finite length of the extension tube means that in practice, the reflection from its open end must be windowed out.

(C) The instrument is placed at point "Y". The resulting pressure variation is shown in Figure 3.35(c).

The independent measurement of incident impulse is obtained by subtracting measurement B from measurement A (see Figure 3.36(a)). The instrument response is obtained by subtracting measurement B from measurement C (see Figure
Figure 3.35 Pressure Variation for Different Source Tube Terminations.
Figure 3.36 True Incident and Reflected Signals Obtained by Subtraction.  
(a) Independent Measurement of Incident Impulse (A - B), i.e. Figure 3.35(a) minus Figure 3.35(b).  
(b) Measurement of True Hornet Response (C - B), i.e. Figure 3.35(c) minus Figure 3.35(b).
When subtracting B from A or C, it was sometimes noted that the decays of the two input impulses were not identical. This was probably because of impulse inconsistencies, or possibly slight differences in the position of the microphone in the paxolin washer. Consequently, small spikes sometimes occurred at the impulse decay position on the arithmetical difference plot. These spikes were windowed out and so did not spoil the results, but the fact remains that the assumed incident impulse is sometimes slightly different from the actual incident impulse. In these instances, an imperfect deconvolution result might be expected.

When making measurements to be used for sophisticated analysis, there are some additional points to note, which are less important when the measurements are to be used for instrument comparison only.

(1) In Section 3.1.3.5, it was noted that the original source tube was made in sections. On close examination of results, it was discovered that small, almost unnoticeable reflections were occurring at the junctions between the sections. Therefore the source was not completely non-reflecting as required by the theoretical analysis techniques. The analysis algorithms are extremely sensitive to small errors in experimental data; even though these extra reflections are almost unnoticeable, they may not be negligible. For this reason the sectioned source tube was replaced by a continuous source tube.

(2) The extreme noise sensitivity of the analysis algorithms means that it is essential to reduce background noise to an absolute minimum. This may be achieved by using a higher number of averages. Care should be taken not to use too many averages while a temperature change is occurring. This would lead to a progressive shift in features of the response which would cause smearing of the final averaged result. This may turn out to have worse effects than the background noise.
(3) For instrument comparison measurements, the response is only required as far as the reflection from the instrument bell. For the analysis measurements, the response is required for more than double this amount. At present this has been achieved for cornets and trumpets; their source tubes are sufficiently long to make the source non-reflecting for a double-length record (see Section 3.1.3.3).

3.3.5. Examples of Measurements on Brass Instruments

Instruments measured so far include cornets, euphoniums, trombones, trumpets and tubas.

Euphoniums, large-bore trombones and tubas all fit onto the standard 12.7 millimetre diameter source tube. Cornets and trumpets fit onto the narrower source tube, but each have their own 48 cm end-coupler because of the difference in mouthpiece shape.

Examples of responses of the different instruments are found in plots throughout this thesis, and are summarised in Table 3.3 for convenient reference.

The record of the trombone with no valves is the simplest. The most complex is the tuba because of its length and the complex nature of its valves and tubing.

3.4. Overall Summary of Chapter 3

This chapter included a discussion of the careful refinements which have been made to the impulse apparatus, in particular the source tube, which has proved highly beneficial.

Data acquisition software has been extensively developed to overcome problems of spark source inconsistency and ambient temperature variation.

Finally, a distinction was drawn between the measurements to be used for detailed analysis, and those to be used simply for instrument comparisons. The latter type will now be discussed in Chapter 4.
TABLE 3.3
Summary of where Responses of Different Instruments are Found in this Volume.

<table>
<thead>
<tr>
<th>Instrument</th>
<th>Figure(s) where response is found</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) CORNETS</td>
<td></td>
</tr>
<tr>
<td>Unknown make</td>
<td>3.35, 3.36, 4.7, 4.8, 4.19</td>
</tr>
<tr>
<td>Medium-bore Besson</td>
<td>4.15</td>
</tr>
<tr>
<td>Comparison of above two</td>
<td>6.6</td>
</tr>
<tr>
<td>(2) EUPHONIUMS</td>
<td></td>
</tr>
<tr>
<td>B&amp;H Sovereign B flat</td>
<td>4.2, 4.18, 4.21, 4.22</td>
</tr>
<tr>
<td>Comparison of Sovereign &amp; Proto</td>
<td>6.2</td>
</tr>
<tr>
<td>Comparison of Yamaha &amp; Sovereign</td>
<td>6.3, 6.4</td>
</tr>
<tr>
<td>(3) TROMBONES</td>
<td></td>
</tr>
<tr>
<td>Large-bore Sovereign</td>
<td>3.6, 3.11, 4.6, 4.10, 4.11</td>
</tr>
<tr>
<td>Large-bore Bass Sovereign</td>
<td>4.13</td>
</tr>
<tr>
<td>Medium-bore Besson Concorde</td>
<td>4.14</td>
</tr>
<tr>
<td>Arithmetical Difference Plots</td>
<td>3.28, 3.30</td>
</tr>
<tr>
<td>(4) TRUMPETS</td>
<td></td>
</tr>
<tr>
<td>Trumpet with Mute</td>
<td>4.9</td>
</tr>
<tr>
<td>B&amp;H Symphony</td>
<td>4.12, 6.5</td>
</tr>
<tr>
<td>B&amp;H Sovereign</td>
<td>4.25</td>
</tr>
<tr>
<td>Straight Trumpet</td>
<td>4.12, 5.2, 5.24</td>
</tr>
<tr>
<td>(5) TUBAS</td>
<td></td>
</tr>
<tr>
<td>B&amp;H Sovereign EE flat</td>
<td>4.24</td>
</tr>
</tbody>
</table>
4.1. Introduction

This chapter contains a description of how the measurements of the previous chapter may be processed so that any differences which exist between two instruments which are supposed to be the same will be detected. Such a technique would be of great use to a brass instrument manufacturer. At present there is no way of ensuring that a set of assembled brass instruments at the end of a production line will be perfectly identical in structure. Instruments are subject to various manufacturing defects, the most common being valve misalignment. Examples of other faults will be given later. In addition, the tools used in instrument manufacture may wear down after several years' use, and thus change certain features of the instrument; for example, versions made ten years after the initial launch of a particular model may have a slightly different bell shape. These defects, which may produce undesirable effects in acoustic behaviour, are not easy to detect and locate by normal testing methods. Players are able to perceive differences in musical quality between one production instrument and another; professional players have criticised makers whose products have been seen to change when a factory moves, or when instruments are made or assembled in a different country, or when the firm is taken over by another company (Edwards, 1978). The new instrument comparison technique would ensure that such changes in product do not occur.

The new technique tests each fully-assembled instrument against a perfect prototype, and any differences are located and quantified, and consequently identified. Figure 4.1 illustrates the instrument testing scheme. In Section 4.2 the procedure for achieving this by comparison of transient response measurements is explained. A brass instrument manufacturer could use this technique to assure his customers of the consistent high quality of his product,
and could back up his claims with scientific data!

In Section 4.3 some preliminary results of the technique are presented. Section 4.4 describes some ways in which the experimental arrangement may be relaxed. A discussion of how the comparison method can be used in quality control and fault diagnosis in the factory is given in Section 4.5. Finally Section 4.6 considers the equipment which is necessary for building a new system for use in the factory.

4.2. Method of Instrument Comparison

4.2.1. Methods Considered

The aim of the instrument comparison technique is to detect the significant physical differences between two instruments, and represent them in a way which can be readily understood by Quality Assurance (QA) personnel.

Originally, three methods of evaluating the differences, \(d(i)\) between
two measured responses, \( g_1(i) \) and \( g_2(i) \) were considered:

1. **Calculation of arithmetical ratio at each point \( i \), ie.**
   \[
   d(i) = \frac{g_1(i)}{g_2(i)}
   \]

2. **Calculation of arithmetical difference at each point \( i \), ie.**
   \[
   d(i) = g_1(i) - g_2(i)
   \]

3. **Calculation of percentage difference at each point \( i \), ie.**
   \[
   d(i) = \frac{(g_1(i) - g_2(i)) \times 100}{g_1(i)} \%
   \]

It soon transpired that the second option was the best. The other two methods tended to produce unstable results at positions where response values were small, for instance, \( d(i) \) contained spikes at points where the response crossed the zero axis.

The use of pattern recognition algorithms was also considered. Implementation of these would be complicated and time consuming. The idea was not pursued because, as will be shown later, the simpler arithmetical differencing technique was found to be perfectly adequate.

### 4.2.2. Arithmetical Differencing Technique

Before the arithmetical difference between the measured responses is calculated, the two data files are converted to a standard form using the following steps. (The first two of these have already been explained in Section 3.3.3.)

1. **Shift the data so that the record begins at the foot of the initial impulse and subtract the value at the foot from all elements.**
2. **Normalise the whole record by dividing all elements by the impulse amplitude so that the impulse amplitude becomes unity.**
3. **For each instrument model, define a region of interest (ROI) on its response data. This will begin at the reflection from the instrument-source tube junction and end just before the minimum of the bell reflec-
tion. Set all values outside the ROI to zero.

(4) Calculate the mean value within the ROI and subtract it from all elements within the ROI. Thus, the mean value becomes zero.

For instrument comparisons, only the direct reflections from within the instrument are required. In step (3) the parts of the record which are not needed are removed. Arithmetical differences at the positions of the initial impulse and the bell are usually high because of the steepness of the record at these points. Smaller arithmetical differences resulting from significant physical differences between instruments could be overshadowed by them. Therefore, for clarity they are removed.

Step (4) ensures that the arithmetical differences are evenly distributed around a mean value of zero. When step (4) was not carried out, arithmetical difference plots sometimes showed a "DC" difference between responses, as well as the differences caused by physical variations between instruments. This DC level was probably related to low frequency background noise. As will be discussed later in this section, results are analysed by identifying the locations of the highest arithmetical differences. The presence of a high DC level could easily lead to misinterpretation; one could mistakenly suppose that the high arithmetical difference was due to a significant physical difference. Removal of the DC level provides the simple solution to the problem.

After these preparatory stages, the response of the test instrument is subtracted from the response of the normal instrument. (This is the convention chosen for this work; there is no reason why it should not be done the other way round.) Then the arithmetical difference data are analysed to see whether any differences are significant.

4.2.3. Analysis of Arithmetical Difference Data

The following three criteria are used in conjunction to assess the arithmetical difference data.
(1) Twenty largest arithmetical differences

These are located and listed on the VDU, first in order of decreasing amplitude, and secondly in the order in which they appear along the instrument.

(2) Significance level

It may well be that the largest arithmetical difference is not significant. One must therefore find a way of assessing whether a given arithmetical difference results from a significant physical difference. This has been done by setting a "significance level". Below this level, arithmetical differences denote non-significant physical differences; above it, they denote significant physical differences. The way in which the significance level is set is now considered.

One of the main determining factors is that two independent measurements of the same identical instrument may produce arithmetical differences, even though there are no physical differences. In this case the "difference level" is defined as the maximum arithmetical difference within the ROI. Ideally, the significance level will be higher than the difference level. In an early series of measurements on cornets (Section 4.5.1.2) the highest difference level was $1.42 \times 10^{-2}$; the significant physical differences result in arithmetical differences of $2 \times 10^{-2}$ or more, so the significance level was originally set at $2 \times 10^{-2}$.

However, it has since transpired that the significance level tends to vary, depending on the instrument and the circumstances, and it is not always easy to set. On two occasions the difference level has actually exceeded $2 \times 10^{-2}$. An investigation into possible reasons for the variation in difference level is reported in the next section.

For some physical differences, for instance, simulated valve misalignment of 1 millimetre (Section 4.5.2.1), the significant arithmetical differences are less than $2 \times 10^{-2}$. In these cases, the significance level is usually...
reset to $1 \times 10^{-2}$; but on one occasion a 1 millimetre misalignment resulted in an arithmetical difference of $8.7 \times 10^{-3}$. A small constriction in a cornet bell (Section 4.3.2) produced an even smaller arithmetical difference of $6.42 \times 10^{-3}$.

The following general guidelines may be drawn up regarding the significance of certain arithmetical differences:

- $\geq 3 \times 10^{-2}$: definitely a physical difference;
- $\geq 2 \times 10^{-2}$: almost certainly a physical difference;
- $\geq 1 \times 10^{-2}$: either a small physical difference, or a high difference level;
- $< 1 \times 10^{-2}$: probably no significant physical difference.

Once the best significance level has been found, other illustrative information can be extracted. The position at which the first significant arithmetical difference occurs can sometimes be more useful than the position of maximum arithmetical difference (see, for example, Section 4.3.1). For this reason it should always be noted.

The number of elements for which the arithmetical difference exceeds the significance level can be a useful indicator of the nature of the physical difference. A small number of significant differences probably results from a small constriction at a single position. A large number may be caused by a large difference such as a change in length or a hole in the bore.

(3) Root mean square difference

This can also be a useful general indicator of the degree of similarity between two instruments. Here, the rms difference is defined as

$$\sqrt{\frac{\sum_{i=1}^{N} (g_1(i) - g_2(i))^2}{N}}$$

where $N$ is the number of elements in the ROI. The rms difference is most useful for detecting instances in which the differences between the two response
records are great. An example would be when one instrument has a leak; this would result in the part of the response after the leak having a completely different character (see Section 4.3.1). A valve misaligned by a large amount can also affect the character of the rest of the response (e.g. 5 millimetres simulated euphonium valve misalignment, Section 4.5.2.1). Two completely different models of the same instrument will also register a large rms difference, for instance, the cornets of Section 6.4.3. The rms difference is not so useful when there are small faults at particular positions, e.g. a small constriction. Such a fault may cause significant arithmetical differences at one or two elements, but their effect will be lost when they are averaged out over the whole record. To date, the "significance level" of rms difference appears to be approximately $5 \times 10^{-3}$. If the rms difference lies above this, physical differences between the instruments can be expected to be major.

The three criteria discussed provide useful guidelines on the types of differences which may be occurring. However, to make the picture complete, visual inspection of superimposed plots of the two responses, and the arithmetical difference is necessary. Then one can see immediately where the significant differences lie. This relies on human judgement, meaning that at present the fault detection procedure is not completely automatic.

A plotting programme has been developed whereby the two responses are superimposed on an oscilloscope, and the two areas in which the arithmetical differences are greatest are made brighter. This aids speedy analysis.

4.2.4. Investigations of the Limit of Sensitivity of the Arithmetical Differencing Technique

In the last section it was mentioned that the sensitivity of the arithmetical differencing technique was limited by the fact that differences are detected between measurements of the same instrument. Investigations into possible physical causes of these differences have been carried out using a
4.2.4.1. Method

To ensure that the difference level was not brought about by a change in the position of the euphonium or microphone, the whole impulse apparatus and the euphonium stayed in the same position for the entire experiment (about 3.5 days). All parts of the apparatus were left switched on throughout the experiment to ensure that the equipment stayed in a steady state, since it was thought that some inconsistency might be introduced if measurements were taken before the apparatus had reached this condition.

Initially the following two reference files were measured within minutes of each other:

Reference file A: Euphonium with Valve 1 depressed by 1 millimetre (held in position using adhesive tape)

Reference file B: Euphonium in its normal state (adhesive tape removed)

Thereafter, measurements of the instrument, still in its normal state, were made after the following intervals: Case 1: 10 minutes; Case 2: 3 hours; Case 3: 24 hours; Case 4: 63 hours; Case 5: 87 hours. For each measurement, 100 runs were averaged and the sampling rate was adjusted as necessary to compensate for changes in ambient temperature (as explained in Section 3.2.4.2). Ambient conditions were closely monitored at the time of each measurement as follows:

1) Percentage humidity was measured using a whirling hygrometer.
2) Temperature was measured in °Celsius to an accuracy of one decimal place using a digital thermometer. In cases where the temperature changed during the measurement, the mean of the initial and final temperatures was calculated.
3) Ambient pressure in decibels was measured using a barometer.

The aim of the experiment was to see whether the difference level corresponded to particular changes in ambient conditions.
Measurements for each of the five cases were compared, first with Reference file A, to see what arithmetical differences were produced by a controlled physical change of the instrument, and secondly with Reference file B, to see how these compared with the difference levels.

4.2.4.2. Results and Discussion

The results of the above comparisons are given in Table 4.1.

<table>
<thead>
<tr>
<th>Case</th>
<th>Time Interval</th>
<th>% Humidity</th>
<th>Mean Temperature (°C)</th>
<th>Pressure (dB)</th>
<th>Reference file A comparison (Amplitude of arithmetical difference spike x 10^{-2})</th>
<th>Reference file B comparison (Difference level x 10^{-2})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ref. file A</td>
<td>0 mins</td>
<td>45</td>
<td>21.20</td>
<td>-.04</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Ref. file B</td>
<td>0 mins</td>
<td>46</td>
<td>21.55</td>
<td>-.04</td>
<td>1.73</td>
<td>N/A</td>
</tr>
<tr>
<td>1</td>
<td>10 mins</td>
<td>45</td>
<td>21.35</td>
<td>-.04</td>
<td>1.73</td>
<td>.0039</td>
</tr>
<tr>
<td>2</td>
<td>3 hrs</td>
<td>41</td>
<td>20.25</td>
<td>-.02</td>
<td>1.51</td>
<td>.1270</td>
</tr>
<tr>
<td>3</td>
<td>24 hrs</td>
<td>48</td>
<td>20.60</td>
<td>0</td>
<td>1.50</td>
<td>.4300</td>
</tr>
<tr>
<td>4</td>
<td>63 hrs</td>
<td>43</td>
<td>16.90</td>
<td>.12</td>
<td>1.43</td>
<td>.3420</td>
</tr>
<tr>
<td>5</td>
<td>87 hrs</td>
<td>45</td>
<td>16.90</td>
<td>.16</td>
<td>1.33</td>
<td>.6050</td>
</tr>
</tbody>
</table>

Here, a significance level of $1 \times 10^{-2}$ was suitable, because all the "misalignment spikes" had amplitude greater than this, and all the difference levels were lower. (The results are not always so well-behaved!) A typical result is illustrated in Figure 4.2, which shows the arithmetical difference plots for Case (5) versus Reference files A and B.

Plots of difference level against change in each physical quantity are shown in Figure 4.3, which also includes a plot of difference level against time interval. Although these plots do not suggest abundantly clear trends, one could make the general conclusion that for a greater change in ambient conditions (humidity, temperature or pressure) or a longer time interval, one would expect a higher difference level. This trend was less obvious with humi-
Figure 4.2 Arithmetical Difference Plots for the Euphonium. (a) ROI of Measured Euphonium Response, (b) Arithmetical Difference Between Reference File B and Case (5) Measurement, Showing the Difference Level, (c) Arithmetical Difference Between Reference File A and Case (5) Measurement, Showing the Arithmetical Difference Spike Due to Valve Misalignment. (NOTE: (b) and (c) have an Expanded Vertical Scale.)
Figure 4.3 Plots of Difference Level Versus (a) Time Interval (Hours), (b) Change in % Humidity, (c) Change in Temperature (°C), (d) Change in Ambient Pressure (dB).
Closer examination of results showed that during the experiment, temperature change and pressure change had varied in a similar way, so that the effects of each could not be separated. An earlier version of the same experiment proved helpful here. There the relationship between pressure change and difference level was not clear, but a plot of difference level against temperature interval (Figure 4.4) showed quite clearly that a larger temperature interval resulted in a larger difference level.

4.2.4.3. Temperature

The above observation suggested that the method of changing sampling rate to compensate for temperature change was not working sufficiently accurately. An error in the choice of sampling rate would lead to progressively larger displacements of plot features from their correct positions as time progressed, and so greater arithmetical differences would be likely.

This point was investigated carefully. For an incorrect choice of sampling rate, the error in position of features (hereafter called "feature displacement") increases linearly with time. This was verified by measuring an instrument and immediately afterwards remeasuring it using a different sampling rate; feature positions were accurately determined to ±0.005 of a sample using interpolation (see Section 3.2.2); a trombone was used as its response features are more distinct and evenly spread out than a euphonium's. Feature displacement was plotted against time. For large changes in sampling rate (>20 Hz), a straight line through the origin was obtained; here the feature displacement was large, e.g. for a rate change of 556 Hz, it varied between 0 and 16 samples. For smaller changes in sampling rate (<10 Hz), the feature displacement was only a fraction of a sample, and the straight line became distorted because of limitations in the accuracy of the interpolation process.

The effect was then investigated of remeasuring at a different sampling rate after a longer time interval of a day or more, but such that the initial
Figure 4.4 Plot of Difference Level Against Temperature Interval in °C.
impulse occurred at the same sample number. The result was not a straight line through the origin. Often the result had the form shown in Figure 4.5.

**Figure 4.5**

*Sketch Plot of Feature Displacement Versus Time*

Here a change in sampling rate did not produce the expected effect. A possible explanation was that the overall temperature distribution throughout the source tube and the instrument had changed between the first and second measurements. A localised change in temperature results in a change in velocity and therefore a feature displacement at that point. For clarification, the system is considered to consist of three sections:

1. **Source tube.** It was insulated and made of thicker metal than the instrument; so it reacted more slowly to changes in air temperature.

2. **Interior of instrument.** The trombone was not placed inside a polystyrene box for the above measurements; it reacted more quickly to ambient temperature changes because of its thinner walls and absence of insulation.

3. **Exterior of instrument.** The bell region will respond most quickly to changes in air temperature.

Examples follow of how the relationship between the temperatures of each section may change at different times of the day. The temperatures of each
section are $T_1$, $T_2$ and $T_3$ respectively. The apparatus may cool down to a steady state overnight so that the whole system is at the same temperature:

$$T_1 = T_2 = T_3$$

During the morning the temperature may rise; the different sections of the system react at different rates:

$$T_1 < T_2 < T_3$$

The equipment, having been evenly warmed throughout, may begin to cool down at different rates in the evening:

$$T_1 > T_2 > T_3$$

Localised temperature changes may also be brought about by other factors such as weather (draughts, sunlight, etc.). Handling the instrument causes temperature variations at particular positions (Section 3.2.3.3), and for this reason, gloves — preferably thermal — were worn. Central heating can produce particularly fast temperature rises when the radiators come on, and this will also change the temperature distribution in the apparatus.

The sampling rate was chosen on the basis of temperature $T_1$ only, so changes in the overall temperature distribution could not be taken into account. Changes in temperature distribution would certainly account for increased difference levels.

In the factory, a temperature controlled room in which the instruments could be stored for a few hours and then measured would be desirable. This also re-emphasizes the usefulness of the polystyrene box.

### 4.2.4.4. Other Factors

After three further repeats of the experiment (Section 4.2.4.1), difference levels were found to be reduced by the following:

1. having the instrument inside the polystyrene box instead of out in the open,
leaving the apparatus switched on between measurements, not switched off,
using a greater number of averages for each measurement.
Out of the above three factors, increasing the number of averages appeared to
produce the greatest reduction in difference level.

Difference level was affected by many other factors including background
noise and "wildness" (see Section 3.2.3.3). The state of the spark plug
underwent changes, which in turn affected the impulse shape, amplitude and
timing; a change in impulse shape could well affect the whole record and hence
the difference level. Variations in the position of the microphone in the
washer hole could also have an effect, so that whether the microphone was
removed between measurements could be a significant factor. The microphone was
normally removed whenever the instrument on the end of the source tube was
removed, in order to protect the microphone diaphragm.

4.2.4.5. Summary of Section 4.2.4

A "difference level" was found to exist between two responses even when
physical differences between the instruments were negligible or non-existent
(for repeated measurements of the same instrument). This limited the sensi­
tivity of the arithmetical differencing technique and called for the introduc­
tion of a "significance level"; arithmetical differences above or below this
level were assumed to be significant or non-significant respectively.

Careful investigation showed that changes in the overall temperature dis­
tribution of the source tube / instrument system, for instance, those intro­
duced by localised temperature changes, seemed to result in the most sig­ni­
cificant increases in difference level. Difference level could also be increased
by changes in ambient pressure, temperature and humidity. Other factors which
could increase it included random background noise variation, spark source
inconsistencies and change in microphone position.

Difference levels were sometimes so high that they obscured arithmetical
differences caused by actual physical differences between instruments. This
made the choice of significance level difficult. Obviously one would have liked the significance level to be as low as possible in order to maximise the sensitivity of the technique, but when difference levels were high this was not possible.

The difference level was greatly reduced by the ambient temperature compensation and interpolation procedures (Sections 3.2.4 and 3.2.2). An increased number of averages reduced random noise effects. When the equipment was left switched on between measurements, the difference level was less likely to be high. A temperature-controlled room would have helped overcome the problem of changing temperature distribution, but as this was expensive, insulation around the source tube and instrument was used instead.

4.2.5. Summary of Section 4.2

In this section the instrument comparison technique was explained. Before comparison, each response measurement is standardised. The test instrument response is subtracted from the normal instrument response. The resulting arithmetical difference data are analysed by identifying the greatest differences, by finding which differences lie above a certain "significance level", and by looking at the rms difference. The problem of the limit of sensitivity was thoroughly examined, and ways in which the sensitivity could be improved were suggested.

4.3. Preliminary Investigations

Early tests of the arithmetical differencing technique involved investigating whether it could be used to detect simple deliberate changes made to an instrument. For these tests, an instrument would be measured in its normal state; then it would be modified in some way and remeasured. The two measurements would be compared. Such tests have also been the means of identifying which features of a measured response correspond to the various physical features of the instrument.
Some examples follow which highlight important aspects of the arithmetical differencing technique. A variety of different instruments were used.

4.3.1. Leak Detection (Trombone)

This test was carried out on a large-bore trombone (Boosey and Hawkes "Sovereign"). A leak was simulated by opening the waterkey. The result is shown in Figure 4.6. The leak was easily detected using the three criteria discussed in Section 4.2.3 and visual inspection of plots.

The maximum arithmetical difference was .126 at sample 242; this is well above the highest significance level of $3 \times 10^{-2}$. The rms difference was $3.72 \times 10^{-2}$; this is also well above the rms significance level of $5 \times 10^{-3}$. The first arithmetical difference greater than $2 \times 10^{-2}$ occurs at sample 224. This corresponds to the position of the waterkey and comes before the position of maximum arithmetical difference. From this point onwards, the whole character of the response has been modified because of the leak; so large arithmetical differences occur throughout the latter part of the response. This creates many apparent physical differences when there is only one, illustrating why it is necessary to look not just at the positions at which the arithmetical difference is greatest, but also at the position at which the first significant difference occurs, as noted in Section 4.2.3.

If there had been any further physical differences beyond the waterkey, they could not have been detected because of the way the response had been altered. Thus, when physical differences between instruments are major, only the first can be detected reliably.

It should be noted that the leak here was large, and was used purely to illustrate certain points. Leaks of this size would not be expected to occur on high quality manufactured instruments!

4.3.2. Detection of Small Constrictions (Cornet)

A small ball of "BLU TACK" (approx 7 millimetres diameter) was placed 110
Figure 4.6 Effect of Opening Trombone Waterkey. (a) Response with Waterkey Open, (b) Normal Trombone Response, (c) Arithmetical Difference.
millimetres inside the bell of a cornet (make and model unknown). This corresponds to a constriction of approximately 5% of the bell cross-section. Figure 4.7 shows the result of comparison with the unmodified cornet. In the arithmetical difference plot, a single clear spike of amplitude $6.42 \times 10^{-3}$ occurs at the "BLU TACK" position. Although this falls below the significance level, the difference could easily be seen using graphical inspection. The rms difference here was only $1.5 \times 10^{-3}$, indicating no major widespread physical differences.

This result shows that a single constriction is easy to detect. (The above test worked just as well when repeated on a euphonium.) Thus, a lump of solder accidentally left in an instrument would be detected.

The size of the extra reflection produced by a given constriction depends on two factors:

1. **The bore diameter.** The larger the diameter, the smaller the reflection. This can be deduced from equation (2.42).

2. **The position in the instrument.** The further the incident pulse has travelled, the lower its amplitude because of reflections at discontinuities and attenuation, and so the smaller the reflection.

Here, the "BLU TACK" was detected in the bell - the widest part of the instrument and the furthest point from the input.

This ability to detect constrictions suggests a way in which a highly accurate form of instrument calibration may be carried out. Wires may be carefully inserted through the instrument wall into the bore at regular intervals, starting at the bell. For each wire, a measurement will show its position on the response. In this way one can build up a precise picture of which parts of the response correspond to particular sections of the instrument. (Unfortunately, the instrument is ruined in the process!)
Figure 4.7 Effect of Small Constriction in the Bell of a Cornet. (a) Normal Cornet Response, (b) Arithmetical Difference Between (a) and the Response of Cornet with a Small Constriction in the Bell (Expanded Vertical Scale).
4.3.3. Detection of Changes in the Bell

Further investigations of the sensitivity of the differencing technique to modifications in the bell region were carried out.

Figure 4.8 shows the result of placing a duster in the bell of a cornet. The character of the bell reflection changes completely. The maximum arithmetical difference of .146 is well above the significance level. The rms difference is $1.77 \times 10^{-2}$.

The presence of a mute in the bell of a trumpet also modifies the shape of the bell reflection; Figure 4.9 shows this. Here the rms difference is $1.19 \times 10^{-2}$, and the maximum arithmetical difference is $5.86 \times 10^{-2}$. Caussé and Sluchin (1982) have investigated the effect of mutes on the input impedance of brass instruments measured in the frequency domain. Comparison of the transient responses of instruments with and without mutes would complement their work.

4.3.4. Detection of Changes in Length

Length can be added using slides or valves.

The length of a trombone may be increased using either the playing slide (Figure 4.10) or the tuning slide (Figure 4.11). In the former case the playing slide was pulled out by 7 centimetres. This resulted in 7 centimetres of extra tubing being added at two different positions:

1. just before the playing slide bend,
2. just after the playing slide bend.

The first addition results in response features to the right of point "A" (Figure 4.10) being shifted 19 samples to the right. The second addition results in all features to the right of point "B" being shifted a further 19 samples to the right, making a total shift of 38 samples to the right.

In the latter case, the tuning slide in the bell section was pulled out by 4 centimetres, and the corresponding shifts were 10 samples after the point "C" and a further 10 samples after point "D".
Figure 4.8 Effect of Placing a Duster in the Bell of a Cornet. (a) Normal Cornet Response (Continuous) and Response when Duster Placed in Bell (Dashed), (b) Arithmetical Difference.
Figure 4.9 Effect of Placing a Mute in the Bell of a Trumpet. (a) Normal Trumpet Response (Continuous) and Response when Mute Placed in Bell (Dashed), (b) Arithmetical Difference.
Figure 4.10 Effect of Pulling Trombone Playing Slide Out by 7cm. (a) Normal Trombone Response, (b) Response with Slide Out, (c) Arithmetical Difference. (Points A and B Correspond to Points Just Before and Just After the Playing Slide Bend Respectively, for the Normal Trombone; Points A' and B' are the Same Points on the Extended Trombone.)
Figure 4.11 Effect of Pulling Trombone Tuning Slide Out by 4cm. (a) Normal Trombone Response, (b) Response with Slide Out, (c) Arithmetical Difference. (Points C and D Correspond to the Beginning and End of the Tuning Slide for the Normal Trombone; Points C' and D' are the Same Points on the Extended Trombone.)
Figures 4.10 and 4.11 also display the corresponding arithmetical differences. From these one can see that a major change has occurred; apart from that the information is not useful and may even be misleading, eg. one might mistakenly suppose that an arithmetical difference spike was caused by a constriction. However, if one knew a change in length had occurred, the position of the first significant arithmetical difference would help to pinpoint where the increase or decrease occurred.

Visual inspection of the two superimposed responses immediately shows that a lengthening has occurred. With suitable graphics facilities - cursor, etc. - the shifts at different points would be easy to locate and identify.

Detection of length changes without user interaction is not straightforward. One possibility is to use the powerful signal processing technique of cross correlation. The "Sample Cross Correlation Function",\( C_{xy}(k) \), for two sequences

\[
x(i), \quad i = 1, 2, \ldots N \quad \text{and} \quad y(i), \quad i = 1, 2, \ldots L \quad (L < N)
\]

where \( \{y(i)\} \) is shorter than \( \{x(i)\} \) is defined as follows:

\[
C_{xy}(k) = (1/N) \sum_{i=1}^{L} x(i+k-1) y(i) \quad (L < N - k + 1) \quad (4.4a)
\]

\[
C_{xy}(k) = (1/N) \sum_{i=1}^{N-k-1} x(i+k-1) y(i) \quad (L > N - k + 1) \quad (4.4b)
\]

The cross correlation of two identical sequences is the autocorrelation (see also Section 2.3.4.2). The autocorrelation and cross correlation functions both contain a peak. If the peaks are at different positions, it means that one sequence is shifted in relation to the other. In principle, this shift can be used to quantify any difference in overall length between two instruments. The positions of the bell reflections are compared. The bell reflection data of the prototype (or normal) instrument are placed in the shorter \( y \)-array. The data from the same region of the test instrument, only extended in case of lengthening or shortening, are placed in array \( x \). The technique was originally
tested on two measurements of a euphonium— one with all valves up, and one with the second valve depressed, which caused a lengthening of 17.5 centimetres. The actual change in position of the bell reflection was 50 samples. The cross correlation technique calculated it to be 51 samples. The accuracy of this result depended on the sizes of sequences x and y; the shorter the sequences, the less reliable the result. Although the above technique can detect overall length changes, it cannot show where the change occurred. Additional cross correlation of successive sections of the two measurements would be necessary, but if the sections are too small, an inaccurate result may be produced.

In reality, the types of length changes which occur on production line instruments are small shifts of a few samples at one or more points along the instrument (see, for example, the tubas in section 4.5.3.1). These may not necessarily produce a significant change in the position of the bell reflection, in which case they would be missed by the cross correlation technique.

Clearly, changes in length are not the most straightforward to detect. The arithmetical differencing technique will detect that a major difference has occurred, but may not be able to identify it as a change in length. It would be identified more easily by visual inspection.

For a thorough test of valved instruments, two comparisons are made—one with all valves up (minimum length) and one with all valves down (maximum length) to test the quality of the valve tubing. The detection of valve misalignment will be considered later in Section 4.5.2.

4.3.5. Differences Introduced by Bends

Boosey and Hawkes have made a special "one-off" "straight equivalent" of their "Symphony" trumpet. The responses of the normal and straight trumpets are compared in Figure 4.12. The straight trumpet has a much smoother response. Work on the effects of bends in wind instruments has been carried out by, for example, Nederveen (1969), Brindley (1973), and Keefe and Benade.
Figure 4.12 Comparison of Responses of a Boosey and Hawkes 'Symphony' Trumpet (Continuous) and the Equivalent Straight Trumpet (Dashed).
(1983). The use of transient measurements may provide a fresh and helpful approach to this problem.

Sondhi (1986) has recently reported an investigation of the effects of curvature on the resonance frequencies of the vocal tract.

4.3.6. Summary of Section 4.3

The use of the arithmetical differencing technique in the detection of leaks, constrictions and changes in the bell region has been seen. When a change in length occurs, the technique may detect that a major change has occurred, but visual inspection may also be necessary before a change in length can be diagnosed correctly.

Further results will be presented in later sections as illustrations and examples of how the instrument comparison technique can be applied in industry.

4.4. Relaxation of the Experimental Arrangement

Measurements made for the sole purpose of instrument comparison need not adhere to such stringent conditions as those to be used for theoretical analysis (see Section 3.1.3.3). The theory demands a non-reflecting source. For instrument comparison a non-reflecting source is not essential, but it makes interpretation of results easier. In this section it is shown that even when the source tube is shortened so that source reflections occur in the data record, differences between instruments can still be detected adequately. This is illustrated using measurements of a bass trombone.

The same standard source tube could be used for all instruments, together with a series of source-tube-to-instrument couplers. This removes the inconvenience of having to swap source tubes round each time a different type of instrument is measured. Measurements on a medium-bore trombone illustrate this.

In each of the following examples, the test applied to the relaxed exper-
imental arrangement was to determine whether a slight displacement of a rotary valve (as opposed to a piston valve) could be detected.

4.4.1. Shorter Source Tube

Four different lengths of source tube between the spark source and the microphone were investigated - 3.58 metres (full length), 2.3 metres, 2.08 metres, and .80 metres. (It was possible to vary the length in this way because the original source tube was made in sections - see Section 3.1.3.5.)

For each length, three measurements were made:
(1) Boosey and Hawkes large-bore "Sovereign" bass trombone,
(2) the same trombone with one of the two rotary valves displaced slightly,
(3) no instrument.

The third measurement was made in order to determine the positions of the source reflections. The shorter the source tube, the earlier and larger the source reflections; in addition, less of the high frequency content of the impulse is attenuated, so the record contains more high frequency noise and is less smooth.

For comparison of two identical instruments the presence of source reflections makes no difference, because on subtraction of the two responses they disappear. When one of the instruments has a misaligned valve or some other bore modification, source reflections can produce two corresponding spikes instead of one in the arithmetical difference plot. Figure 4.13 shows this for the .80 metre source tube. The valve misalignment produces an extra positive then negative double-spike at samples 457 and 462 respectively. This feature reappears just before the bell reflection at samples 672 and 676, the time interval corresponding to the time taken to travel from the microphone to the source and back to the microphone. On the arithmetical difference plot, spikes appear at the two corresponding positions. Therefore, one physical difference has produced two significant arithmetical difference spikes (amplitudes \(-3.06 \times 10^{-2}\) and \(3.26 \times 10^{-2}\), and \(-1.99 \times 10^{-2}\) and \(1.76 \times 10^{-2}\)).
Figure 4.13 Demonstration of How, with a Shorter Source Tube, a Single Bore Constriction May Produce Two Arithmetical Difference Spikes. (a) Normal Bass Trombone Response, (b) Response of Trombone with One Rotary Valve Displaced Slightly, (c) Arithmetical Difference.
One could argue that the presence of the second spike could lead to confusion. One might think it was produced by a second physical difference in the bell region. Therefore, when a short source tube is being used, one can take the following precautions to guard against misinterpretation. One should note only the largest arithmetical difference, and the position at which the first significant difference occurs. If there is a source reflection of the first arithmetical difference spike, it will have lower amplitude and will occur later, and will therefore be missed. If, however, the second spike is picked up, one can determine whether it is a source reflection by looking at the time interval between the two spikes. These precautions would tend to mean that if there is more than one physical difference between instruments, only the first would be detected reliably. Yet this restriction already exists for some differences, such as leaks or changes in length, when a full-length source tube is used (Sections 4.3.1 and 4.3.4).

4.4.2. Single Standard Source Tube with Different End Couplers

The .48 metre end-coupler of the standard source tube was replaced by a conical coupler whose internal diameter decreased linearly from large to medium bore size. (The coupler was made from the mouthpipe of a French horn.)

A medium-bore Besson "Concorde" trombone was measured with and without its rotary valve displaced slightly. The test was to see whether the misalignment was detectable. Figure 4.14 shows that it was. The tail after the initial impulse was no longer smooth because of the coupler, but this did not affect the sensitivity of the arithmetical differencing technique.

The coupler for a small-bore instrument would have a more rapid taper, but the above result suggests that it would also work adequately.

4.5. Application of Instrument Comparison Technique in Industry

In this section the potential uses of the instrument comparison technique in the brass instrument manufacturing industry are examined. Current instru-
Figure 4.14 Detection of Valve Displacement in a Medium-Bore Instrument Using a Standard Source Tube. (a) Complete Normal Medium-Bore Trombone Response (Including Incident Impulse), (b) Arithmetical Difference Between (a) and the Response when the Rotary Valve is Displaced Slightly (ROI only, Expanded Vertical Scale).
ment evaluation procedures are described, and limitations identified. Ways in which the new technique helps overcome these are explained. Some of the common manufacturing defects and instrument faults which develop are discussed. Examples of how the new technique has detected and diagnosed these are given.

4.5.1. Quality Control

4.5.1.1. Limitations of Current Testing Procedures

At present in the brass instrument industry, there is no rigorous way of testing the consistency of the completed brass instruments.

Before assembly, the physical quality of the instrument parts is tested using standard inspection gauges. After assembly, the decision about the level of quality is based on subjective human judgement rather than objective scientific measurements.

The musical quality of instruments at the end of the production line is tested by an experienced player. He plays each instrument for about 30 seconds, checking the quality of each note produced, and the ease of valve or slide operation, etc. Some manufacturers claim that a skilled tester can detect small inconsistencies in large batches of instruments. While this may be true, the consistency of such tests must be questioned. Sometimes, time constraints dictate that this final testing procedure must be omitted.

With the impulse testing equipment, each assembled instrument would be physically measured, and compared with its prototype. Small physical inconsistencies would be obvious immediately. The result of the next section (Section 4.5.1.2) shows that the equipment can detect faults which have been missed by normal inspection methods at the factory.

The impulse measurements take longer than the present quality tests. A run in which fifty measurements are averaged may take about two minutes. This time could easily be reduced to less than one minute if the number of averages were reduced, but generally a higher number of averages is preferred as then
statistical fluctuations are more likely to be ironed out. Clearly, the time taken to measure each instrument will be an important consideration in a factory.

4.5.1.2. Example of Simulated Production Line Conditions

This example shows how the new apparatus is a powerful quality control tool.

Six medium-bore Besson cornets of the same model (Model 723) were selected at random from wholesale stores, and labeled A to F inclusive. Instrument A was treated as the prototype. Comparison of measured responses showed that the cornets fell into two distinct groups:

Group 1: A, D, F
Group 2: B, C, E

Plots of arithmetical difference between "Group 2" cornets and the prototype contained a prominent spike at a point corresponding to the latter end of the tuning slide (see Figure 4.15). Such a spike was not present when instruments within "Group 1" were compared.

Careful inspection of instruments by Boosey and Hawkes QA personnel revealed that in Group 2 instruments a short thin-walled sleeve was missing from the inside part of the instrument immediately after the latter end of the tuning slide. This is illustrated in Figure 4.16. There should be a sleeve at both ends of the tuning slide to create a smoother transition between the end of the slide and the main body of the instrument. When the sleeve is missing, a premature reflection results, which is detected easily by the equipment. This fault had been missed completely by the normal factory testing methods.

A good player may have been able to detect a difference between the two classes of instrument, but would have been unable to pinpoint the fault.

4.5.2. Valve Misalignment

Valve misalignment is one of the most common manufacturing faults. When
Figure 4.15 Detection of a Sleeve Missing From a Medium-Bore Besson Cornet. (a) Cornet (A) with Sleeve, (b) Cornet (C) without Sleeve.
Figure 4.16 Diagram to Illustrate Missing Sleeve of Figure 4.15.
it occurs, the holes in the valve piston do not line up correctly with the holes in the valve wall. Figure 4.17 illustrates the problem. Both up and down valve misalignment are caused by the accumulation of the errors in several component parts. Tolerances are permitted in the manufacture of each part, and on occasion the combination of these may result in misalignment. The worst offenders are the piston buffers (felt or rubber) which are not made to exact tolerances, and may well be affected by environmental conditions.

Normally valve misalignment is inspected by the removal of some detachable part of the instrument, eg. a tuning slide. However, when the instrument is assembled, it is often impossible to see whether a valve is misaligned, even using lights and mirrors, eg. tuba family. A few of the valve positions may be visible for the small instruments only.

The measurements of transient response can be used to test the alignment of all valves, including the invisible positions (see next section for examples). To achieve this, the test instrument must be compared with an instrument in which the valves are known to be correctly lined up. Therefore, in the factory it would be essential to ensure that the prototype instrument did not have misaligned valves. With a fully assembled prototype tuba, such a check cannot be made after assembly; extra special care must therefore be taken when manufacturing the prototype.

Top quality players may be able to detect a difference in the quality of instruments with and without valve misalignment. The results of the next section show that the testing procedure can detect a misalignment of 1 millimetre. Thus, it is probably more sensitive than most players. The equipment can also pinpoint the position of misalignment, whereas a player could not. It would be valuable to run a series of subjective tests to see whether players could distinguish instruments with valves correctly and incorrectly aligned. The testing techniques developed by Pratt (1978) would be suitable. Of course, if the difference cannot be discerned by players, the value of the extreme
Figure 4.17 Schematic Sectional Diagram of Valve Line-Up Errors When Piston is Depressed (Cumulative Errors of at Least Seven Parts).
sensitivity of the equipment could be called into question. At present, however, some players are prepared to pay large sums of money to specialist "trumpet doctors" who guarantee that they will re-align valves correctly.

4.5.2.1. Detection of Simulated Valve Misalignment

In this section valve misalignment is shown to be detectable by the new instrument testing procedure. Valve misalignment was simulated by pressing a valve down by a measured amount (1, 3, 4 or 5 millimetres) and holding it in position using adhesive tape.

Systematic investigations were carried out on a B-flat euphonium (Boosey and Hawkes "Sovereign") and a cornet (make unknown). Each valve was investigated in turn. The aims were as follows:

(1) to find the minimum amount of misalignment which could be detected reliably,

(2) to find whether the extent of the misalignment bore any relation to the size of the resulting arithmetical difference spike.

Generally, misalignments of 4 or 5 millimetres greatly altered the character of the latter parts of the response. Figure 4.18 shows this for the 4th valve of the euphonium. Here the rms difference was $2.36 \times 10^{-2}$, indicating major differences between the two responses (refer to Section 4.2.3). Misalignments as small as 1 millimetre could be detected. Figure 4.19 illustrates this for the first valve of the cornet. The arithmetical difference at the position of misalignment was $-1.14 \times 10^{-2}$ which is above the significance level of $1 \times 10^{-2}$.

For each valve of the euphonium, the amplitude of the arithmetical differences spike(s) was plotted against the extent of the misalignment (in millimetres). The results are displayed in Figure 4.20. The relationship is reasonably linear for all valves, with gradients varying between 0.9 and 1.9. (For the cornet the linear relationships were not so apparent.)
Figure 4.18 Illustration of How Large Valve Misalignments Change the Character of the Response. (a) Normal Euphonium Response, (b) Same Euphonium with Valve 4 Misaligned by 5mm.
Figure 4.19 Detection of 1mm Valve Misalignment. (a) Normal Cornet Response, (b) Arithmetical Difference Between (a) and Response when First Valve is Misaligned by 1mm (Expanded Vertical Scale).
Figure 4.28: Relationship between size of arithmetical difference spike and extent of valve misalignment for all euphonium valves.
It has also been found that more than one misalignment can be detected at the same time, for instance, when all three cornet valves were misaligned by 3 millimetres, each misalignment could be seen individually in the arithmetical difference data.

4.5.2.2. Example of Actual Valve Group Fault

Two Boosey and Hawkes "Sovereign" B-flat euphoniums were compared. Figures 4.21 and 4.22 show the "valves up" and "valves down" comparisons. These plots showed that in both cases one instrument response had a much larger spike between valves 1 and 2 than the other. Originally it was supposed that this was caused by valve misalignment.

The euphonium with the suspected fault was examined by removing its mouthpipe. (This was necessary because valve alignment within a fully assembled euphonium cannot be inspected - see Section 4.5.2.) This showed that the alignment of the first valve was good. The extra spike was caused by a small piece of metal swarf lodged in the bow between the first and second valves. Figures 4.23 (a), (b) and (c) show photographs of the detached valve group, the blockage in the valve tubing where it was originally discovered, and the piece of swarf after careful removal from the instrument. This was a serious manufacturing fault, which, before the transient response measurements, had remained undiscovered.

4.5.3. Tuba Variation

The method of manufacturing tubas has changed little during the last century. The instrument is still not yet fully jigged [*]. Hence variations may arise.

The instrument is made in two sub-assemblies which are then joined. The valve group is made separately, and slides are assembled onto it. Little vari-

[*] A jig holds all the parts of the instrument in position prior to soldering.
Figure 4.21 Valves Up Comparison of Two B-Flat Euphoniums. (a) Silver-Plated Sovereign, (b) Clear-Lacquered Sovereign, (c) Arithmetical Difference.
Figure 4.22 First Part of Valves Down Comparison of the Same Two Euphoniums as in Figure 4.21. (a) Silver-Plated Sovereign, (b) Clear-Lacquered Sovereign, (c) Arithmetical Difference.
Figure 4.23 Photographs Illustrating Detected Euphonium Valve Group Constriction. (a) Detached Valve Group with Arrow Indicating Location of Constriction, (b) Constriction as Discovered in Tubing, (c) Constriction After Removal (Metal Swarf)
ation occurs between manufactured valve groups. Elsewhere, the main body of
the instrument - the bell, bottom end and branches - is formed, such that at
the innermost part of the instrument a gap is left for the valve group to fit
in. Because the instrument is not jigged, there may well be errors in the size
of this gap. If the gap is too large, an extra piece of tubing will be
required to reduce it; the mouthpipe might then be trimmed with shears so that
the instrument maintains the correct overall length. If the gap is too small,
it has to be widened by trimming off some more tubing; here, the mouthpipe
would be lengthened by adding extra tubing. The positioning of the valve group
may vary by 13 millimetres or so. Its position is determined using a metre
rule; this is rather crude.

An example of detection of tuba variation is presented in the next sec-
tion. Whilst the equipment cannot cure the above situation, it can be used to
assess the typical variations which occur between tubas.

4.5.3.1. Examples of Tuba Variation

Three EE-flat tubas (Boosey and Hawkes "Sovereign", Model 982) were
labelled A, B and C. Each was measured with all its valves up and again with
all its valves down. All slides were pressed in fully during measurements to
ensure that all instruments were in the same state.

As the tuba record is the longest and most complicated so far considered,
additional measurements had to be made in order to identify the different
regions of the record, eg. valves depressed slightly, waterkeys opened, pieces
of plasticine placed just inside tuning slides.

Arithmetical differences were well above the significance level of
$2 \times 10^{-2}$ for all possible "valves up" and "valves down" comparisons, as shown
in Table 4.2. These differences occur in the valve group region. The main
cause is small variations in length, which give particularly large arithmeti-
cal differences because of the many spikes in the valve group region of the
response record. Figure 4.24 compares the three tuba responses in the valve
**Figure 4.24** Expanded Comparison of the First Part of the "Valves Up" Responses of the Three Tubas Superimposed. Tuba A (Blue), Tuba B (Red), and Tuba C (Green).
TABLE 4.2

Maximum Arithmetical Differences for Tuba Comparisons

<table>
<thead>
<tr>
<th>Tubas compared</th>
<th>Maximum Arithmetical Difference (x10^-2)</th>
<th>Valves up Comparison</th>
<th>Valves down Comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>A, B</td>
<td>5.5 (End of VG)</td>
<td>4.7 (Valve 1)</td>
<td></td>
</tr>
<tr>
<td>A, C</td>
<td>2.9 (End of mouthpipe)</td>
<td>3.8 (Valve 1)</td>
<td></td>
</tr>
<tr>
<td>B, C</td>
<td>3.8 (End of VG)</td>
<td>3.0 (Valve 1)</td>
<td></td>
</tr>
</tbody>
</table>

The presence of length variations is obvious from visual inspection of this plot. Determination of the positions and amounts of the length variations was carried out by visual inspection of superimposed plots shifted relative to each other by different amounts.

In the mouthpiece region of Figure 4.24 the tuba A response is three samples later than the other two responses, indicating that the mouthpipe of tuba A is slightly longer than those of the other two tubas (about 10 millimetres longer). However, in the valve group region, tuba A was one sample ahead of tuba C, suggesting that at this position the length of tuba A was about 4 millimetres too short. It would appear that the mouthpipe of tuba A had been made 10 millimetres longer than standard to compensate for a loss of length elsewhere; possibly the tubing just before the valve group had had to be trimmed in order for the valve group to fit onto the main body of the instrument. The measurements would suggest that because instrument A is still slightly too short, its mouthpipe should have been made even longer.

Comparison of tubas B and C shows that B appears to "lose" 2 samples (about 7 millimetres) between the mouthpipe and the valve group, and a further 2 samples towards the end of the valve group.

The major difference suggested by the "valves down" comparison was possible misalignment of the first valve in tuba A. This was suspected because of a high amplitude negative-going spike in the valve 1 region of the response. This spike, however, was not present in the "valves up" comparison. (The valve alignment was not checked in this case, since the instrument actually belonged
to a professional player, and its disassembly may have proved inconvenient.) Other differences between responses appeared to be produced by minor structural differences within the valve group rather than variations in length.

These results show that variations in length of just a few millimetres can be detected by comparing the response measurements. Detection of these requires user interaction; it is not automatic as yet. Close study of the response was necessary before the above discoveries were made. As more experience in interpreting responses is gained, the diagnosis procedure will be speeded up.

4.5.4. Common Manufacturing Faults.

A detailed examination of how valves may become poorly aligned and how variations in length may arise has already been given. A variety of other manufacturing faults may occur.

(1) Constrictions. Tubing may become completely or partially blocked by something being left in the instrument unintentionally, eg. cleaning material, small pieces of metal; unwanted debris left after soldering. An example of an actual constriction was given in Section 4.5.2.2. A further example follows in Section 4.5.5.2.

(2) Structure. A join between sections or the ends of a slide may be left rough instead of smoothly chamfered; an example was given in Section 4.5.1.2. The wrong part may be fitted accidentally, eg. mouthpipe or bell.

(3) Faults. Small leaks may occur around the slides and valves. Small dents may occur. The process of "plugging", that is forcing ball bearings down cylindrical sections, is designed to smooth out small dents and other imperfections.
4.5.5. Detection and Diagnosis of Faults which Develop
4.5.5.1. Impulse Apparatus as a Diagnostic Tool

A powerful feature of the apparatus is its ability to diagnose faults. It is proposed that every manufactured instrument is measured, and that this measurement is stored on the computer. If the instrument develops a fault some years later, it can be remeasured, and the new measurement compared with the original measurement. Some examples of how this technique could be of great use follow.

Various faults may develop after the instrument is sold. It may be dropped or dented. Loose items in a player's case may find their way into the bell.

The present procedure used by the Boosey and Hawkes factory in these situations is to pay a professional musician (or a resident factory musician) to judge the faulty instrument by playing it. The musician is not always able to diagnose the fault.

In one such instance, a player returned a euphonium complaining that it was "stuffy". Eventually it was necessary to disassemble the whole instrument to discover that a cigarette lighter had become lodged well inside the instrument. Disassembly is expensive because after reassembly, the whole instrument has to be repolished, lacquered and plated. By comparing a new measurement of the euphonium with its original measurement or a measurement of the prototype euphonium, the equipment would have immediately detected the precise position and extent of a constriction at the cigarette lighter position.

In another case, a player complained of his instrument "buzzing". This was because a ticket had got stuck in a bend. Another player discovered that his handkerchief had found its way inside his instrument. In these instances, the apparatus would have been able to locate the fault.

Once the equipment has located a constriction, a rod can be inserted to that point and an attempt made to clear it (or at least identify it). Following this, another measurement would show whether the obstruction had been
removed successfully. In this way complete disassembly may be avoided.

A player may also return his instrument complaining that it has gone out of tune. The measurement technique would show whether any changes have occurred in the instrument since its original measurement, i.e., at the end of the production line before it was sold. Any detected changes could provide clues about why the instrument has become mistuned. A partial blockage could cause mistuning.

Clearly, with the new apparatus, fault diagnosis is quicker and simpler (and perhaps cheaper!) than at present.

In the next section, an example of how a trumpet fault was diagnosed is given.

4.5.5.2. Example of Trumpet Fault Diagnosis

During a series of subjective tests on a set of Boosey and Hawkes "Sovereign" trumpets, a particular trumpet was diagnosed as being "stuffy". (The term "stuffy" is explained in Section 6.5.4.) This trumpet and five others of the same type were measured and compared to see whether the stuffy trumpet could be identified, and whether the cause of the stuffiness could be discovered.

Comparison of the responses of all possible pairs (15) of instruments revealed that one trumpet consistently gave rise to a large arithmetical difference "bump" in the bell bend region of the response. This trumpet was correctly identified as the stuffy trumpet. The amplitude of the bump varied between $5.2 \times 10^{-2}$ and $6.7 \times 10^{-2}$. Figure 4.25 shows the bump clearly. This result led one to suppose that something lodged inside the bell bend was causing the stuffiness. No constriction could be seen by looking into the bell. When the instrument was cut open, it was discovered that some "Scotchbrite" (cleaning material) and flakes of silver plate had formed the constriction in the crook of the bell. A photograph of this is shown in Figure 4.26. It would not have been possible to view this constriction even if the bell had been
Figure 4.25 Detection of Constriction in the Bell Bend of Trumpet. (a) Trumpet without Constriction, (b) Trumpet with Constriction, (c) Arithmetical Difference.
Figure 4.26  Photograph to Illustrate the Bell Constriction of
Figure 4.25
removed from the trumpet. The equipment was able to locate the constriction easily.

Also in the above tests, the equipment was able to detect differences between the other five trumpets. One trumpet had the wrong mouthpipe fitted; this caused arithmetical differences of approx $2.1 \times 10^{-2}$ in the mouthpipe region. Another trumpet was a slightly different model; it had a discontinuity at the start of the tuning slide which was not present in the other five trumpets, and resulted in arithmetical differences which varied between $2.3 \times 10^{-2}$ and $3.7 \times 10^{-2}$.

4.5.6. Summary of Section 4.5

A scientific and reliable means of instrument quality control has not been available to the brass instrument industry to date. It has been shown that the arithmetical differencing technique provides this. It has been able to detect common manufacturing faults, such as valve misalignment and length variation. The technique has found serious faults which would not be detected by normal factory inspection methods. This helps to explain why players claim to be able to detect differences between mass produced instruments of the same type. The technique can also be used to diagnose the causes of complaints about returned instruments.

The potential benefits of the new equipment to the brass and woodwind instrument industries worldwide are clear.

4.6. Equipment Required for Factory System

4.6.1. Introduction

To date, all measurements on instruments have been made using the Nova 4-based equipment. The Nova 4 is now an old machine. A faster more up-to-date microcomputer would be required in the factory. The design of a new microcomputer-based apparatus will now be considered.
4.6.2. Data Acquisition Options

Two options exist. Data samples can be read directly from the ADC into computer memory. This would require a highly sophisticated microcomputer with a parallel interface which could cope with high data rates.

Alternatively, data can be read into a transient recorder from which it would be fed at a slower rate into the controlling microcomputer. In this case, a less sophisticated and therefore cheaper microcomputer would suffice.

The latter option is preferable as it is cheaper, simpler and more reliable.

The other hardware components necessary for data acquisition are the spark source, source tube, microphone, amplifier, filter, programmable oscillator and ADC.

Either a Knowles electret or a B & K microphone may be used (see Section 3.1.4). The Knowles is far cheaper; its performance is adequate, but spoilt by electrical pick-up. The filter would be a simple low-pass filter with a fixed cut-off frequency of 20 kHz.

The programmable oscillator (or frequency synthesiser) should have frequencies ranging to at least 70 kHz. The essential feature is that the sampling rate must be set to an accuracy of .1 Hz, ie. 6 digits, so that small changes in ambient temperature can be compensated for.

The ADC may already be built into the transient capture module. It should have either a built-in or additional sample-and-hold device. Accuracy should be at least 12 bits. It should be able to cope with data rates up to at least 70 kHz, ie. 70000 words per second.

4.6.3. Microcomputer System

The desired features of the microcomputer are as follows.

(1) Fast 16-bit operations.
(2) Back-up facilities in order that back-up copies of each measurement can be stored; this can be achieved with either an extra disk drive or tape drive.

(3) All the necessary interfaces to peripheral equipment. These are illustrated in Figure 4.27.

Figure 4.27
Block Diagram to Illustrate Required Peripheral Interfaces of Microcomputer

(4) High resolution colour graphics (at least 512 x 512).

(5) Adequate memory - at least 256 kilobytes.

(6) Adequate software.

High resolution screen colour graphics have not been available on the Nova. They would give flexibility and potential for further development. One could divide each response record into different sections to show the different regions of the instrument. Positions of maximum arithmetical difference could be pointed out automatically with flashing coloured arrows. In addition, with a cursor, the user could locate the precise positions at which
the differences appear to be most important.

For hard copy plots, a colour digital plotter is required.

A printer could be used to print some form of quality control certificate, including details of the instrument model and serial number, date of measurement, result of differencing test, etc. A high quality printer with an automatic sheet feeder could produce such certificates automatically.

4.7. Overall Summary of Chapter 4

In this chapter, the technique of computing the arithmetical difference between transient responses has been shown to be a valuable way of ensuring consistency of manufactured brass instruments. Common manufacturing faults, eg. leaks, constrictions, wrong mouthpipe, have been detected. Faults which were missed by existing quality control methods have been found.

The sensitivity of the new technique is limited by the fact that a measured response can be slightly affected by factors other than the instrument itself, eg. ambient conditions, but valve misalignments of 1 millimetre can still be detected.

People within the brass instrument industry have reacted with enthusiasm to the proposed system, recognising its potential benefits.
CHAPTER 5

ANALYSIS OF RESULTS

5.1. Introduction

In this chapter the measurements of instrument response are subjected to a more rigorous analysis than that described in the previous chapter. The aim of the analysis was to reconstruct the shape of the instrument bore accurately. The first stage in the process is deconvolution. In Section 5.2, different methods of deconvolution are systematically investigated. In Section 5.3, the application of the Gerchberg restoration technique to the deconvolution results is discussed.

The bore reconstruction method is explained and assessed in Section 5.4. The associated direct process, by which the response of an instrument can be calculated from its bore shape, is discussed also.

Computer programming was done on the Gould "Powernode 6000" minicomputer, which has 16 Megabytes of virtual memory (4 Megabytes of physical memory). To maintain the highest possible numerical accuracy throughout, all arithmetical operations were carried out in double precision; this required large amounts of memory which were unavailable on the Nova minicomputer (only 32 kilobytes of memory). Hence, data acquired on the Nova had to be transferred to the Gould via a serial link for later analysis.

5.2. Deconvolution

5.2.1. Method of Assessment of Different Deconvolution Techniques

In Chapter 2, various deconvolution methods were introduced. The three selected for systematic investigation were the following:

(1) time domain deconvolution,

(2) frequency domain deconvolution,
(3) constrained iterative deconvolution.

These were chosen as they represent completely different approaches to the problem, and so enable a thorough survey of the possible techniques. To compare the methods, each was applied to the same "model" data and also to real experimental data. The following four data sets, two model and two real, were used, each consisting of a response, g (n samples) and an input signal, f (m samples).

DATA SET (A) Noiseless Model Data

The reflectance, h (l samples) was generated from model bore data (Figure 5.20) using the method described later in Section 5.4.1.2. The input f was a synthetic impulse. The response g was obtained by convolution of f with h. Figure 5.1 shows g, h, and f. Here l=440, m=73 and n=512. This satisfies the criterion explained in Section 2.3.6.1, namely n=l+m-1, and thus avoids wraparound difficulties.

DATA SET (B) Noisy Model Data

This set was the same as Set (A) but with 1% Gaussian noise superimposed on g. The percentage of Gaussian noise is calculated using the following quotient

\[
100 \times \frac{\text{rms difference between noisy and noiseless g}}{\text{rms value of g}}
\]

In this way, the stability of each deconvolution method in the presence of noise could be tested. This is a particularly important test, as noise can adversely affect the numerical stability. Gaussian noise is a more realistic approximation to experimental noise than uniformly distributed noise.

DATA SET (C) Real Experimental Data

Measurements of a straight trumpet were used. See Figure 5.2. Here, n=2048, m=746 and l=1303.
Figure 5.1 Noiseless Model Data. (a) Input Pulse, \( f \), (b) Impulse Response, \( h \) (Synthesised From Model Bore Data of Figure 5.20), (c) Response, \( g \) Obtained by Convolution of \( f \) and \( h \). (NOTE: Above Plots all have Different Vertical Scales.)
Figure 5.2 Real Experimental Data. (a) Incident Pulse, $f$, (b) Response of Straight Trumpet.
DATA SET (D) Smoothed Real Experimental Data

Here, the f and g of Set (C) were smoothed using the method of 5-point smoothing (Su, 1971). This was to see whether smoothing would reduce the numerical instabilities caused by noise.

To assess the success of each method of deconvolution, the computed reflectance, h was compared with the known h. For data sets (A) and (B) the true h was known; for sets (C) and (D) it was not, so a "best approximation" to h had to be used instead. The error was quantified by calculating the rms error and the maximum error. A qualitative judgement was made by visual inspection of superimposed plots of the known and calculated h's. As a double check, the calculated h was convolved with f, and the result compared with the known g, again using rms error, maximum error and inspection of plots.

Results for each of the three deconvolution methods will be summarised and discussed in turn.

5.2.2. Time Domain Deconvolution

This method was explained in Section 2.3.1.

For all four data sets, this method produced results consisting of oscillations of ever-increasing amplitude. This was because the first element of each f, f(0) was equal or close to 0, corresponding to the foot of the input impulse. As f(0) is the denominator at each step of the deconvolution, this rendered the procedure unstable. When the first element had a sufficiently high value, the results were correct. Two attempts were made to overcome this problem of ill-conditioning.

5.2.2.1. Regularisation

First, a simple form of regularisation was tried (Louver et al, 1969).

Here, the normal convolution equation

\[ g(n) = \sum_{i=1}^{n} f(n-i+1) h(i) \]  

(2.8)
is replaced by

\[ (1 + e) g(n) = \sum_{i=1}^{n} (f(n-i+1) h(i) + e f^{-1}(n-i+1) h(i)) \quad (5.1) \]

where \( e \) is an arbitrary regularisation parameter. \( e \) is chosen to give the optimum compromise between instability and precision. As \( e \) tends to zero, the instability increases, but as it becomes larger, the solution moves further away from the true solution. A value of \( e \) which gave a reasonable solution could not be found. The introduction of \( e \) certainly delayed the onset of the oscillations; and the larger \( e \), the greater the delay, but the oscillations were always present no matter how large \( e \) was (\( e > 1000 \)).

5.2.2.2. Input Data Shift

The second attempt to overcome the instability was more successful. It consisted of shifting \( f \) so that its first element was no longer at the foot of the impulse, but a short way up the leading edge of the impulse. Thus, the first element takes on a higher value and the method is stabilised. In the process, the shape of \( f \) has been slightly modified, and consequently one would expect a certain amount of distortion in \( h \). Figure 5.3 illustrates this for data set (A). Here, \( f \) has been shifted by 25 samples to start at its peak - the resulting deconvolved \( h \) is distorted, with loss of the finer details, but nonetheless bears a fair resemblance to the true \( h \). It is certainly a better approximation to \( h \) than the former unstable oscillatory solution.

For the real data of data set (C), a stable solution was obtained when \( f \) was shifted by two or more samples. (A shift of seven samples corresponded to starting at the impulse peak.) Clearly, one would expect the deconvolved \( h \) to become increasingly distorted as the extent of the \( f \)-shift increases. The following trends were observed as the shift increased from two to six samples.

1. \( h \) became smoother and less noisy, but this was accompanied by a loss of fine detail. This is illustrated in Figure 5.4 where the calculated \( h \)'s for \( f \)-shifts of 2, 3 and 6 samples are compared. As the \( f \)-shift
Figure 5.3 Comparison of (a) True Model $h$ with (b) $h$ Calculated by Time Domain Deconvolution, $f$ having Been Shifted by 25 Samples to Start at its Peak. (The Resulting $h$ has Been Shifted back by 25 Samples for Ease of Comparison.)
Figure 5.4 Comparison of h's for Real Data Calculated By Time Domain Deconvolution, f having Been Shifted by (a) 2 Samples (b) 3 Samples (c) 6 Samples.
increases, the noise at the start is suppressed, but the salient features
lose their amplitude.

(2) \( g \) becomes progressively less accurate. Table 5.1 shows how the error in \( g \) increases.

**TABLE 5.1**

Dependence of Accuracy of \( g \) on extent of f-shift

<table>
<thead>
<tr>
<th>f-shift (samples)</th>
<th>RMS error in ( g ) (x10^{-2})</th>
<th>Maximum error in ( g ) (x10^{-2})</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.526</td>
<td>6.71</td>
</tr>
<tr>
<td>3</td>
<td>0.729</td>
<td>7.71</td>
</tr>
<tr>
<td>6</td>
<td>1.290</td>
<td>12.20</td>
</tr>
</tbody>
</table>

When this f-shifting technique was applied to data set (D), it transpired
that smoothing the real data caused the deconvolved \( h \)'s to be smoother
throughout, but necessarily entailed a further loss of detail. However, the
"best approximation" to the true \( h \) was chosen to be that obtained using data
set (D) with an f-shift of 2. For additional smoothness it was also subjected
to the 5-point smoothing procedure. This result was used as a reference \( h \) for
the remainder of the deconvolution tests. Although it was known to be dis-
torted, it acted as a useful guideline, particularly for the visual inspection
of a calculated \( h \), in that one could see immediately where any gross distor-
tions lay.

Clearly, the major disadvantage of time domain deconvolution is its ins-
tability when initial values of \( f \) are low. This renders it impractical in most
cases. However, with a shift of \( f \) by a few samples, a useful first approxima-
tion to the deconvolved result can be obtained.

5.2.3. Frequency Domain Deconvolution

5.2.3.1. Software

As explained in Section 2.3.2, frequency domain deconvolution involves
division of the Fourier transforms of \( g \) and \( f \). For data sets (A) and (B),
512-point FFT's were used. For data sets (C) and (D), 2048-point FFT's were used. As each data set already tapered to zero, windowing before the FFT's was unnecessary. Spectral division was accomplished by direct division of complex numbers.

5.2.3.2. Results

For data set (A), Fourier transform division reproduces the true h almost perfectly, with negligible numerical error. The addition of noise results in a noticeable ripple superimposed throughout the whole deconvolved h (see Figure 5.5). The errors in the corresponding spectrum occur mainly at higher frequencies. This is shown in Figure 5.6, where the difference between the spectra of the true and calculated h's is plotted. The effects of noise are greater at higher frequencies because that is where the spectral values of f and g are lower. (This point was mentioned in Section 2.3.2.3). Thus, just a small amount of noise in g — barely noticeable visually — has been enough to affect the quality of the deconvolved result severely.

The effects of noise in the real data of data set (C) are clearly seen in Figure 5.7, which compares the computed h with the reference h described in Section 5.2.2. The calculated h, although it retains the essential features of the true h, has a high level of superimposed noise, and is non-causal. This can be explained by the corresponding spectrum (Figure 5.8) which becomes unstable at high frequencies instead of tapering smoothly to zero. Figure 5.9 shows the reason — the spectra of f and g both have low values at high frequencies. When the above h is convolved with f, the resulting g is inaccurate, particularly at the start (see Table 5.2 and Figure 5.12).

Three attempts have been made to overcome the problem of high frequency noise.

5.2.3.3. Spectral Windowing

In this method, the spectrum H(ω) of h is windowed to remove the unstable
Figure 5.9 Comparison of (a) True Model $h$ with (b) $h$ Calculated by Frequency Domain Deconvolution of Data Set (B), i.e. Model Data with 1% Gaussian Noise Added to $g$. 
Figure 5.6 Difference Between the Spectra Corresponding to Figure 5.5.
Figure 5.7 Comparison of (a) Best Approximation of True $h$ of Straight Trumpet with (b) $h$ Computed by Frequency Domain Deconvolution. Note noisiness and non-causality of (b).
Figure 5.8 Spectrum Corresponding to Figure 5.7(b). Note Instability at High Frequencies.
Figure 5.9: Spectra corresponding to Figure 5.2(a) and (b): Dashed, g
Continuous. Note that both have low values at high frequencies.
high frequency region. The problem of finding the optimum cut-off frequency is not straightforward. Noise affects the whole spectrum; there is not a well-defined frequency where the true h stops and the noise begins. Therefore, windowing will remove high frequency information as well as noise, so a loss of detail in h will result. Several different cut-off frequencies were tried. Spectral windowing always led to an improved result. Generally, the lower the cut-off, the greater the improvement in h. A rectangular window gave better results than a tapering window. Figure 5.10 compares the results of rectangular spectral windows with cut-offs at samples 500 and 256. (The maximum frequency corresponds to sample 1024.) The latter is seen to be less noisy throughout and less non-causal; but at the same time, some high frequency information has been lost as evidenced by the loss in amplitude of the salient spikes. After convolution with f, the h's with a lower cut-off also produced more accurate g's, as shown in Table 5.2, and by visual inspection of plots.

<table>
<thead>
<tr>
<th>Spectral Window Cut-Off</th>
<th>RMS error in g (x10^-2)</th>
<th>Maximum error in g</th>
</tr>
</thead>
<tbody>
<tr>
<td>No spectral window</td>
<td>1.600</td>
<td>.210</td>
</tr>
<tr>
<td>525</td>
<td>1.203</td>
<td>.162</td>
</tr>
<tr>
<td>500</td>
<td>1.040</td>
<td>.140</td>
</tr>
<tr>
<td>500</td>
<td>1.012</td>
<td>.136</td>
</tr>
<tr>
<td>256</td>
<td>0.698</td>
<td>.088</td>
</tr>
</tbody>
</table>

These g's are still inaccurate, despite the improvement.

5.2.3.4. Input Data Shift

The second method of dealing with the problem of high frequency noise is to shift the input data, f. This is the same method as was suggested in Section 5.2.2.2 for time domain deconvolution. Here, when the impulse within f is started just after the foot, it becomes narrower and high frequencies are artificially introduced. Thus, values at high frequencies will increase, and
Figure 5.10 Results of Applying Rectangular Windows to Spectrum of Figure 5.6. (a) Cut-Off at Sample 500, (b) Cut-Off at Sample 256.
so noise will have less effect. This is illustrated in Figure 5.11, which shows in detail the high frequency region of the spectra corresponding to various shifts of \( f \). For no shift or a shift of only one sample, high frequency values are low. For a shift of two samples, high frequency values have increased. This results in a causal \( h \) with a much lower level of superimposed noise. The \( g \) calculated by convolution with \( f \) bears a closer resemblance to the true \( g \), as is shown when the following result is compared with those in Table 5.2.

\[
\text{RMS error in } g = 5.66 \times 10^{-3} \\
\text{Max error in } g = 6.70 \times 10^{-2}
\]

Figure 5.12 compares the above \( g \) with the true \( g \) and a \( g \) obtained through spectral windowing (Figure 5.10(b)).

5.2.3.5. Smoothing

Using the smoothed data of data set (D) resulted in a reduction in the non-causality and superimposed noise level in the calculated \( h \), but the resulting \( g \) was only improved marginally.

5.2.3.6. Discussion

Deconvolution by Fourier transform division has been seen to work perfectly with noiseless model data constituting a well-behaved analytic function. It is, however, susceptible to noise. Noise introduces a random fluctuation so that the function is no longer analytic (Cooper, 1977), and thus deconvolution produces a non-causal and highly noisy result. Sometimes, the resulting noise level was so high that the salient features of the deconvolved result were completely obscured. Smoothing of data has been found to reduce the non-causality and noise level. Windowing of the \( h \)-spectrum produces further improvement, but the greatest improvement is observed when the input data is shifted by two samples so that its high frequency components are increased. Here, causality is restored. The improvement is unquestionable,
Figure 5.11 Expanded Upper Frequency Region of f-Spectrum of Figure 5.9, Shifted by 0 (Continuous), 1 (Dashed) and 2 (Dotted) Samples.
Figure 5.12: Comparison of (a) True $g$ of Straight Trumpet with $g$'s Obtained by Convolution of $f$ with the Following Frequency Domain Deconvolved $h$'s: (b) $h$ Spectrally Windowed After Frequency Domain Deconvolution, and (c) $h$ Obtained When $f$ was Shifted by 2 Samples Beforehand.
although the result is known to be distorted. The result obtained using this latter method is very close numerically to that obtained when using time domain deconvolution on the same shifted data. This is because the two methods are carrying out the same function, that is inverting the impulse response, but different algebraic techniques are used.

In Section 5.3, a further attempt to overcome the problem of high frequency noise is made - using the Gerchberg algorithm.

5.2.4. Constrained Iterative Deconvolution

5.2.4.1. Method and Software

This method of deconvolution was introduced in Section 2.3.3. Further detail of this particular method will now be given.

Discrete convolution of \( f \) and \( h \) to give \( g \) may be expressed by the following matrix equation.

\[
\begin{pmatrix}
    g_0 \\
    . \\
    . \\
    . \\
    . \\
    g_{N-1}
\end{pmatrix}
= \begin{pmatrix}
    f_0 & 0 \\
    f_1 & f_0 \\
    . & . \\
    . & . \\
    . & . \\
    f_{M-1} & f_0
\end{pmatrix}
\begin{pmatrix}
    h_0 \\
    . \\
    . \\
    . \\
    . \\
    h_{N-M}
\end{pmatrix}
\]

As the first few terms of \( g \) and \( f \) may be dominated by noise, these can be disregarded, leading to the following restricted set of equations

\[
\begin{pmatrix}
    g_p \\
    . \\
    . \\
    . \\
    . \\
    g_{p+M-N}
\end{pmatrix}
= \begin{pmatrix}
    f_p & f_0 \\
    . & . \\
    . & . \\
    . & . \\
    . & . \\
    f_{M-1} & f_p
\end{pmatrix}
\begin{pmatrix}
    h_0 \\
    . \\
    . \\
    . \\
    . \\
    h_{N-M}
\end{pmatrix}
\]

or

\[ g = Fh \]
where $f_p$ is the peak value of the input pulse. Iterative deconvolution is facilitated in the following way. Let

$$g = (\mu I + F - \mu I)h$$

where $\mu$ is a constant and $I$ is the Identity Matrix. Then

$$h = [\mu(I + (1/\mu)F-I)]^{-1}g$$

(5.4)

$$= 1/\mu[I + (1/\mu)(F - \mu I)]^{-1}g$$

(5.5)

By applying the Taylor expansion formula, the $k$th and $(k+1)$th estimates of $h$ can be written as follows.

$$h^{(k)} = 1/\mu[1 - (1/\mu)(F-\mu I) + \ldots + (-1/\mu)^k(F-\mu I)^k]g$$

(5.6a)

$$h^{(k+1)} = 1/\mu[1 - (1/\mu)(F-\mu I) + \ldots + (-1/\mu)^{k+1}(F-\mu I)^{k+1}]g$$

(5.6b)

By simple rearrangement, $h^{(k+1)}$ can be expressed in terms of $h^{(k)}$ in the following way.

$$h^{(k+1)} = h^{(k)} + 1/\mu(g - Fh^{(k)})$$

(5.7)

For convergence, the modulus of all eigenvalues of the matrix $(F-\mu I)$ must be less than 1, i.e.

$$-1 < \left|\frac{\lambda - \mu}{\mu}\right| < 1$$

(5.8a)

where $\lambda$ is an eigenvalue of matrix $F$, i.e.

$$\mu > \lambda/2 > 0 \quad \text{or} \quad \mu < \lambda/2 < 0.$$  

(5.8b)

This means that for convergence, all eigenvalues of $F$ must have the same sign. Clearly then, there are certain input signals for which convergence will not be obtained. Values of $\mu$ must also be carefully chosen to satisfy the above criteria.

The software flow chart is given in Figure 5.13. It can be seen that $h$ is always constrained to have 1 elements. A printout of the main parts of the programme "DECONVIT2" appears in Appendix B.
Figure 5.13
Flow Chart of Constrained Iterative Deconvolution Program ("DECONVIT2")

- Read in f, g.
- Choose μ.

Form $h_{(0)}$ from g, i.e. shift g so that its origin is at its peak value, and its duration is 1 elements. Multiply all g-values by $(1/\mu)$.

Convolve the whole of f with $h_{(k)}$: $f \ast h_{(k)}$.
Calculate the difference $(g - f \ast h_{(k)})$.
Shift this difference so that its origin lies at the position of $f_p$ (the peak value of f), and its duration is 1 elements.

Calculate $h_{(k+1)} = h_{k} + (1/\mu)(g - f \ast h_{(k)})$.

Convergence?
Yes: Stop
No:

- Calculate $h_{(k+1)} = h_{k} + (1/\mu)(g - f \ast h_{(k)})$.
Various values of \( n \) were tried for each \( f \). For each \( f \) there is one value of \( n \) which gives the optimum compromise between rate of convergence and accuracy. Here, \( f \) was a positive-going pulse, so values of \( \lambda \) were positive. Therefore \( \mu > \lambda/2 > 0 \) applied. When \( 1/\mu \) was too small, the rate of convergence was extremely slow. When \( 1/\mu \) was too large, the \( \mu > \lambda/2 \) criterion was not satisfied, so the procedure did not converge – it diverged.

5.2.4.2. Results and Discussion

For data set (A), the procedure converged to the correct result. Here the optimum value of \( \mu \) was 58.82.

For the noisy data of data set (B), the result was still converging after 300 iterations. The calculated \( h \) had a noticeable ripple superimposed. This ripple was about the same size as the corresponding ripple in the frequency domain deconvolution result of Section 5.2.3.2, but stopped after \( l=440 \) elements because of the applied constraint. The size of the ripple did not change as the number of iterations increased.

For the real data of data set (C), \( f \) was such that perfect convergence could not be obtained. This was probably because the \( f \) had a less steep decay than the modelled \( f \) of data set (A).

The effect of this deconvolution method was tested nonetheless. Data set (D) was used. The optimum value of \( \mu \) was found to be 9.09. After the first few iterations, \( h \) appeared to have developed some extra ripples around points corresponding to the salient features of the curve. As the number of iterations increased, the amplitude of these ripples became larger, suggesting a diverging result – see Figure 5.14. However, the corresponding \( g \)'s, obtained by convolution with \( f \), became closer numerically and visually to the true \( g \), apart from a small ripple at a few positions which increased slightly in amplitude. The \( g \)'s obtained in this way were far closer to the true \( g \) than those obtained from the time or frequency domain deconvolutions (Sections 5.2.2 and 5.2.3.4). Here, after 50 iterations,
Figure 5.14 Comparison of (a) Best Approximation of True $h$ of Straight Trumpet with Results of Constrained Iterative Deconvolution (b) After 10 Iterations, (c) After 50 Iterations.
RMS error in g = $3.05 \times 10^{-3}$
Max error in g = $7.17 \times 10^{-2}$.

5.2.5. Final Assessment of Deconvolution Methods

The strengths and weaknesses of the deconvolution methods studied can now be assessed.

Time domain deconvolution in its 'most straightforward form is ill-conditioned when initial values of $f$ are low. The problem can be overcome partially by shifting $f$ slightly to increase its initial values, but distortion in $h$ automatically results.

Frequency domain deconvolution is fast, but is severely affected by noise when high frequency spectral values are low. By shifting $f$, high frequency values are artificially increased; the problems of instability are decreased, but distortion is introduced in the process.

The constrained iterative method suffers from two major disadvantages. First, it is extremely slow on large data sets. Secondly, there are some input data, $f$, for which convergence cannot be obtained. The measured data fitted into this category.

All deconvolution methods are susceptible to noise. Cepstral deconvolution, explained in Section 2.3.4.1, would have been affected by high frequency spectral noise in a similar way to frequency domain deconvolution. In addition, there would have been the problem of phase unwrapping which potentially could cause further errors (Bröel and Kjaer, 1981). For these reasons, a detailed investigation of cepstral deconvolution was not pursued.

In Section 2.3.1, another form of time domain deconvolution was mentioned which involved SVD and regularisation (Sondhi and Resnick, 1983). This method may produce better results for real data than those investigated in this work, but time has not permitted a proper study of it. Further investigation of it is recommended.

Considering the disruptive effects of noise on deconvolution, it could be argued that more effort should have been made to reduce the experimental noise
whilst measurements were being made. A greater number of averages could achieve this. Normally 50 averages were used. Better results may be achieved with 1000 averages, but one would have to ensure that ambient conditions remained the same throughout the extended measuring time. The minimum possible time required for 1000 averages is about twenty minutes; this would be far too long for an instrument being measured in the factory, but measurement time is not as important in a research context.

A spark source is a particularly noisy source — in this respect, an alternative source such as a loudspeaker may have been preferable. Experience has shown that lengthening the source tube reduces the amount of high frequency noise — possibly a longer source tube would have led to improved results, but inevitably more high frequency detail would be lost.

Experimental noise can never be removed completely. The results of frequency domain and constrained iterative deconvolution on data set (B) (Sections 5.2.3.2 and 5.2.4.2) show that the smallest amount of noise results in a much larger noise level in the deconvolved result. Therefore one expects that the reflectance calculated from non-ideal data by some method of deconvolution will never be completely correct.

Smoothing of real data serves to reduce the effects of noise, but not remove them completely. At the same time it results in loss of detail. In this respect it behaves similarly to regularisation in which the optimum compromise between stability and precision must be found.

Recent work at the University of Surrey by Watson (1986, unpublished) shows that frequency domain deconvolution can produce reasonable results on measured brass instrument data when the bandwidth is narrow. He used a loudspeaker source; the bandwidth was just 5 kHz and the signal was averaged three times. A possible explanation for his success is that most of the noise lay outside his 5 kHz bandwidth and was therefore not present. In the investigations in this work, the bandwidth was much larger at 23 kHz.
To date, the "best approximation" of the reflectance of the straight trumpet used for data set (C) has been obtained with either time or frequency domain deconvolution, \( f \) having been shifted (Section 5.2.2). This result is known to contain distortion; however, the frequency domain deconvolution result using the non-shifted \( f \) contains worse distortion because of the superimposed noise and non-causality.

An accurate estimate of reflectance was required as input to the bore reconstruction algorithm of Section 5.4. As this was not available, the bore reconstruction algorithm has not been used successfully with real data yet.

5.3. The Gerchberg Algorithm

The discrete Gerchberg algorithm, as described in Sections 2.3.5.1 and 2.3.5.3, has been written in double precision. In principle, it should be possible to use it to restore the distorted high frequency section of the deconvolved spectrum.

The algorithm was tested on the following data:

1. Noiseless model data: the reflectance of data set (A) (Section 5.2.1) synthesised by the method given in Section 5.4.1.2.

2. As (1) above, but with varying amounts of Gaussian noise added.

3. Deconvolved real data: the result of frequency domain deconvolution (no \( f \)-shift) given in Section 5.2.3.2, and illustrated in Figures 5.7 and 5.8.

For (1) and (2), \( N=512 \). For (3), \( N=2048 \). In each case, values of \( q \) and \( k \) were varied, but so that the constraint \( q + 2k + 1 \leq N \) was satisfied.

The accuracy of the results was judged using rms error, maximum error and visual inspection of plots, as for the earlier deconvolution investigations.

[*] \( q \) and \( (2k+1) \) are the sizes of the unknown portions of the time sequence \( (f_0^1) \) and the spectrum \( (F_0^1) \) respectively - see Section 2.3.5.3. \( f_0^1 \) and \( f_2^1 \) are forced to zero; \( F_0^1 \) and \( F_2^1 \) are forced to their known values.
5.3.1. Results for Noiseless Model Data

For the noiseless model data, the Gerchberg algorithm produced a highly smoothed version of the true upper frequency end of the spectrum. It followed the general periodic structure of the true spectrum, but the detailed structure was not reproduced. In the time domain, a ripple was superimposed throughout the region which was not constrained to be zero. This is because of the original truncation of the spectrum which induces convolution with a sinc function in the time domain. The magnitude of the ripple decreases with increasing number of iterations; but so far it has not been removed completely.

The acceleration procedure of Jones (1986) always improved the result for noiseless data. The numerical improvement can be seen in Table 5.3, which gives the frequency and time domain rms and maximum errors for four different combinations of q and k values. Here, the acceleration procedure was applied after five normal Gerchberg iterations.

TABLE 5.3
Effect of Jones Acceleration Procedure After 5 Normal Gerchberg Iterations

<table>
<thead>
<tr>
<th></th>
<th>Frequency domain rms error (x10^-2)</th>
<th>Frequency domain maximum error</th>
<th>Time domain rms error (x10^-2)</th>
<th>Time domain maximum error (x10^-2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. q=261, k=35</td>
<td>After 5 normal Gerchberg iterations 3.24</td>
<td>.193</td>
<td>.202</td>
<td>1.30</td>
</tr>
<tr>
<td></td>
<td>After subsequent acceleration        2.85</td>
<td>.201</td>
<td>.178</td>
<td>1.20</td>
</tr>
<tr>
<td>2. q=250, k=80</td>
<td>After 5 normal Gerchberg iterations 5.42</td>
<td>.250</td>
<td>.338</td>
<td>3.07</td>
</tr>
<tr>
<td></td>
<td>After subsequent acceleration        4.93</td>
<td>.224</td>
<td>.307</td>
<td>3.08</td>
</tr>
<tr>
<td>3. q=240, k=127</td>
<td>After 5 normal Gerchberg iterations 7.51</td>
<td>.290</td>
<td>.469</td>
<td>4.78</td>
</tr>
<tr>
<td></td>
<td>After subsequent acceleration        6.70</td>
<td>.236</td>
<td>.412</td>
<td>4.79</td>
</tr>
<tr>
<td>4. q=190, k=35</td>
<td>After 5 normal Gerchberg iterations 3.24</td>
<td>.192</td>
<td>.203</td>
<td>1.30</td>
</tr>
<tr>
<td></td>
<td>After subsequent acceleration        2.92</td>
<td>.204</td>
<td>.178</td>
<td>1.11</td>
</tr>
</tbody>
</table>

Graphical representations of results showed that after the acceleration
procedure, the upper region of spectrum became less smooth and more like the true jagged spectrum. This is illustrated in Figure 5.15 for the third case. The improvement introduced by the acceleration procedure is obvious, but the Gerchberg algorithm has not succeeded in restoring the upper frequency data perfectly. The corresponding time domain plots are shown in Figure 5.16.

5.3.2. Results for Model Data with Added Noise

The same smoothed upper end of spectrum was obtained as with the noiseless case. This smoothed spectrum was further from the true noiseless spectrum than the initial noisy spectrum. Therefore the Gerchberg algorithm, instead of helping to overcome the problem of high frequency noise, made the result worse.

The performance of Jones' acceleration technique in the presence of noise varied, depending on the amount of noise and the q and k values. For q=240 and k=127, the technique produced an improvement for 1%, 5% and 7% noise; but for 10% noise the technique produced a highly unstable result. (Noise percentages here are as defined in Section 5.2.1).

For q=190 and k=35, the acceleration technique produced a worse result with only 1% noise. Thus, it is clear that the performance of the acceleration procedure is susceptible to noise.

5.3.3. Results for Deconvolved Real Data

For these data the Gerchberg algorithm diverged and the acceleration technique produced highly unstable results. The divergence is demonstrated in Figures 5.17 and 5.18. These compare the results after one iteration and after fifty iterations in the time and frequency domains respectively, using q=1303 and k=372. Interestingly, the g obtained by convolution of f with the result became closer to the true g as the number of iterations increased!

Despite the overall divergence of the algorithm, the result after the first iteration was an improvement on the original deconvolved result for all
Figure 5.15 Results of Application of Gerchberg Algorithm to Noiseless Model Data with $q=240$, $k=127$. Comparison of the Original Data (Continuous), Result After 5 Gerchberg Iterations (Dashed) and Improved Result After Jones' Acceleration Procedure (Dotted).
Figure 5.16 Inverse Fourier Transforms Corresponding to Figure 5.15. (a) Original Data, (b) Result After 5 Gerchberg Iterations, (c) Result After Jones' Acceleration Procedure.
Figure 3.17 Demonstration of How Gerchberg Algorithm Diverges when Applied to Real Deconvolved Data with $q=1303$, $k=372$. (a) After 1 Iteration, (b) After 50 Iterations. (NOTE: Compare with Figure 5.7.)
Figure 5.18 Spectra Corresponding to Figure 5.17. After 1 Iteration (Continuous) and After 50 Iterations (Dotted). (NOTE: Compare with Figure 5.8, the Spectrum before Application of Gerchberg Algorithm.)
the combinations of q and k that were tried. The improvement was greater the higher the value of k.

Earlier, in Section 5.2.3, other methods of improving the frequency domain deconvolution result were tried, namely spectral windowing, input data shift and smoothing. One iteration of the Gerchberg algorithm is almost the same as spectral windowing, except that it includes the extra step of time-limiting which removes the non-causality, thus producing a better result. The single Gerchberg iteration result is superior to the result after smoothing, but not as good as the result after the input data shift of two samples (Section 5.2.3.4).

5.3.4. Discussion

Papoulis (1975) showed that, for continuous time signals, the Gerchberg algorithm converges to the true solution in the absence of noise. This has also been shown to be true for discrete time signals (Jain and Ranganath, 1981; Jones, 1986). The results for noiseless discrete time data (Section 5.3.1) do not at first sight appear to support these findings. A possible explanation may be deduced from Jones (1986), namely a slow rate of convergence. The rate of convergence is determined by the largest eigenvalues of the matrix \( \frac{1}{N} \sigma_{11} \sigma_{11}^{+} \) (defined in Section 2.3.5.3). A library programme has been used to calculate these eigenvalues, and reveals many eigenvalues close to 1; thus, the number of iterations required to obtain convergence may be computationally unfeasible. Although the result of Section 5.3.1 appeared to be changing little after about fifty iterations, it may well be that after several hundred more iterations, its character would have significantly changed. The results of Jones' acceleration technique would add support to this suggestion. Jones (1986) states that his technique will give a good approximation to the correct result if there are only a few eigenvalues close to 1. Indications are that there were more than a few, and this could explain why the acceleration technique, although it improved the result considerably,
did not produce a perfect result.

The performance of the Gerchberg algorithm in the presence of noise will now be considered. Gerchberg (1974) recognised from the start that his algorithm became unstable in the presence of noise and could diverge. No divergence was observed in the results of Section 5.3.2. This could have been because the levels of added noise were relatively low, so that divergence did not begin until many iterations later. For higher noise levels, divergence was observed to occur during the first thirty iterations. For the real data of Section 5.3.3, divergence occurred after just one iteration.

Abbiss et al (1983a, 1983b) have developed a regularised form of the Gerchberg algorithm in an attempt to stabilise it in the presence of noise. The POCs algorithm, described in Section 2.3.5.2, would be another alternative to the unregularised Gerchberg algorithm.

Although the Gerchberg procedure has not succeeded in restoring the high frequency information of the deconvolved reflectance correctly, it has improved the accuracy of the result slightly. It may well be that a different restoration algorithm would produce an even better result. Further investigation of this would appear to be worthwhile.

5.4. Bore Reconstruction

5.4.1. Theory

Various techniques for reconstruction of cross-sectional area from transient response were discussed in Section 2.4. A new method proposed by Jones (Jones, 1983; Duffield and Jones, 1985), whereby bore shape may be recovered from the reflectance, will now be described. The bore is modelled as n cylindrical elements, each of length Δ, propagating plane waves. Damping is assumed to be independent of frequency.

The theory is based on z-transform relationships. Similar techniques have already been used in other fields, for instance, lattice filters (Rao and
Kailath, 1984), and geophysics (Robinson and Durran, 1986, Section B.3). The method will now be explained in some detail, and it will be seen that it can also be used to predict the reflectance of a given bore shape.

5.4.1.1. Essence of the Method

The steps involved in obtaining a rational function representation of the transient response will now be explained.

(1) Obtain expressions for the forward and backward waves in the kth element along the tube

At each element k, there are two incident waves—one forward $f_k^B(t)$ and one backward $f_k^B(t)$—and two outgoing waves—one forward $g_k^B(t)$ and one backward $g_k^B(t)$. This is illustrated in Figure 5.19.

![Figure 5.19](image)

**Figure 5.19**

Schematic Representation of the Reflection and Damping of Forward and Backward Travelling Waves in the Region $k \Delta < x < (k+1) \Delta$.

For a reflection coefficient of $r_k$, the reflection relations between the four
waves are summed up in the following matrix equation.

\[
\begin{pmatrix}
    g_k^B(t) \\
    f_k^F(t) \\
    g_k^F(t)
\end{pmatrix} =
\begin{pmatrix}
    r_k & (1 - r_k) & f_k^F(t) \\
    (1 + r_k) & -r_k & f_k^B(t) \\
    1 & 0 & f_k^B(t)
\end{pmatrix}
\begin{pmatrix}
    f_k^F(t) \\
    f_k^B(t) \\
    f_{k+1}^B(t)
\end{pmatrix}
\]

(5.9)

A damping factor of \( d_{k+1} \) gives rise to further relations.

\[
f_{k+1}^F(t) = d_{k+1}^F g_k^F(t - \Delta/v)
\]

(5.10a)

\[
f_k^B(t) = d_k^B g_{k+1}^B(t - \Delta/v)
\]

(5.10b)

where \( v \) is velocity, and \( \Delta/v \) is sampling interval.

This model is the same as that of Descout et al (1976) (see Section 2.4.3).

(2) Use the z-transform to obtain a recurrence relation for \( H_k(z) \), the z-transform of the reflectance

If \( f_k^B(t) = 0 \) and \( f_k^F(t) = 0 \) for \( t < 0 \), then

\[
f_k^F, B[m] = g_k^F, B[m] = 0 \quad \text{for} \ m < k,
\]

(5.11)

where \( f(m \Delta/v) \) is denoted by \( f[m] \). Therefore, the z-transforms are given by

\[
F_k^F, B(z) = \sum_{m=k}^{\infty} z^{-m} f_k^F, B[m]
\]

(5.12a)

\[
G_k^F, B(z) = \sum_{m=k}^{\infty} z^{-m} g_k^F, B[m]
\]

(5.12b)

The z-transform of the reflectance of the part of the system in the region \( x > k \Delta \) is given by

\[
H_k(z) = G_k^B(z) / F_k^F(z)
\]

(5.13)

\( H_0(z) \) corresponds to the z-transform of the reflectance of the whole system measured at the input.

Equations (5.9) and (5.10) when combined and rewritten in terms of z-transforms give [m]-

\[
\begin{pmatrix}
    G_k^B(z) \\
    z^{-1} F_k^F(z)
\end{pmatrix} =
\begin{pmatrix}
    r_k & (1 - r_k) & F_k^F(z) \\
    (1 + r_k) & -r_k & d_k z^{-1} G_{k+1}^B(z)
\end{pmatrix}
\]

(5.14)
Combination of equations (5.13) and (5.14) gives the recurrence formula

\[ H_k(z) = \frac{r_k + z^{-2} d_{k+1}^2 H_{k+1}(z)}{1 + r_k z^{-2} d_{k+1}^2 H_{k+1}(z)} \]  \hspace{1cm} (5.15)

(3) Represent \( H_k(z) \) as a rational function \( P_k(w) / Q_k(w) \)

Since \( H_n(z) = r_n \), it may be seen that \( H_k(z) \) is a rational function \( P_k(w) / Q_k(w) \) with \( w = z^{-2} \). Substitution into equation (5.15) gives

\[ P_k(w) = \frac{r_k + w d_{k+1}^2 P_{k+1}(w)}{1 + r_k w d_{k+1}^2 P_{k+1}(w)} \] \hspace{1cm} (5.16)

\( P_k(w) \) and \( Q_k(w) \) are polynomials of degree \( (n-k) \). They can be written

\[
P_k(w) = p_0^{(k)} + p_1^{(k)} w + \ldots + p_{n-k}^{(k)} w^{n-k}
\] \hspace{1cm} (5.17a)

\[
Q_k(w) = q_0^{(k)} + q_1^{(k)} w + \ldots + q_{n-k}^{(k)} w^{n-k}
\] \hspace{1cm} (5.17b)

\( q_0^{(k)} \) is always taken to be 1.

Equations (5.16) and (5.17) form the means whereby the following two operations can be performed.

(1) Reflectance synthesis - for a given bore shape and its damping properties, the reflectance is calculated.

(2) Bore reconstruction - the bore shape and the damping profile are calculated from the reflectance.

\( H_0(z) \) is defined by the following equations,

\[ H_0(z) = P_0(w) / Q_0(w) \] \hspace{1cm} (5.18)

\[ \begin{align*}
&= \sum_{j=0}^{\infty} h(j) w^j \\
&= h(j) * q_j^{(0)}
\end{align*} \hspace{1cm} (5.19)\]

where \( (h(j)) \) is the discrete reflectance. Thus, \( P_0(w) = H_0(z) Q_0(w) \), which corresponds to the following convolutional sum in the time domain,
\[ h(n + 1) + h(n) q_1 + \ldots + h(0) q_n = p_0 \]
\[ h(n+1) + h(n) q_1 + \ldots + h(0) q_n = 0 \]
\[ h(2n) + h(2n-1) q_1 + \ldots + h(n) q_n = 0 \]

5.4.1.2. Reflectance Synthesis

For the known bore shape, \( r_j \) is calculated using equation (2.42) (Section 2.4.3). If the damping \( d_{k+1} \) is also known, the reflectance \( h(j) \) may be synthesised as follows.

From equation (5.16), it can be seen that \( P_k(w) \) and \( Q_k(w) \) satisfy

\[
P_k(w) = r_k Q_{k+1}(w) + w d_{k+1}^2 P_{k+1}(w) \quad (5.20a)
\]
\[
Q_k(w) = Q_{k+1}(w) + r_k w d_{k+1}^2 P_{k+1}(w) \quad (5.20b)
\]
with \( P_n(w) = r_n \) and \( Q_n(w) = 1 \).

Equations (5.17) and (5.20) lead to the following iterative equations for coefficients \( p_j^{(k)} \) and \( q_j^{(k)} \) of polynomials \( P_k(w) \) and \( Q_k(w) \) respectively.

\[
p_0^{(k)} = r_k \quad (5.21a)
\]
\[
q_0^{(k)} = 1 \quad (5.21b)
\]
For \( j=1,\ldots,(n-k-1) \)

\[
p_j^{(k)} = r_k q_j^{(k+1)} + d_{k+1}^2 p_{j-1}^{(k+1)} \quad (5.21c)
\]
\[
q_j^{(k)} = q_j^{(k+1)} + r_k d_{k+1}^2 p_{j-1}^{(k+1)} \quad (5.21d)
\]

With \( p_0^{(n)} = r_n \) and \( q_0^{(n)} = 1 \), \( k \) is initially set to \( (n-1) \). The iterative
157

equations are applied progressively, decreasing \( k \) by 1 each time, until \( k \) reaches zero. This gives \( \{p_0\} \) and \( \{q_0\} \). \( \{h(j)\} \) can then be found by solving equations (5.19). Note that this method can be used to predict \( \{h(j)\} \) for \( j > 2n \) [*]. The programme "SYNTH", which performs reflectance synthesis, is reproduced in Appendix C.

5.4.1.3. Bore Reconstruction

The inverse problem of calculating \( \{r_k\} \) and \( \{d_{k+1}\} \) from \( \{h(j)\} \) is more complicated.

First, \( \{p_{j=0}^{n}\} \) and \( \{q_{j=0}^{n}\} \) are found from \( \{h(j)\} \) for \( j = 0 \). The duration of \( h \) corresponds to the time taken for the incident wave to travel from the system input to the far end and back twice. Thus, when \( h \) is found by deconvolution of real measurements \( \{g(j)\} \) and \( \{f(j)\} \), the response \( g \) will have to be measured for \( j > 2n \) (see Section 2.3.6.1).

\( \{q_j\} \) is found by solving the latter \( n \) equations of equation (5.19) using a fast bordering method (Duffield and Jones, 1985). \( \{p_{j=0}^{n}\} \) is then computed using the first \( (n+1) \) equations of equation (5.19). From equation (5.16), it is seen that polynomials \( P_k(w) \) and \( Q_k(w) \) will also satisfy

\[
P_{k+1}(w) = \frac{P_k(w) - r_k Q_k(w)}{w d_{k+1}^2 (1 - r_k^2)} \quad \text{(5.22a)}
\]

\[
Q_{k+1}(w) = \frac{Q_k(w) - r_k P_k(w)}{(1 - r_k^2)} \quad \text{(5.22b)}
\]

A new polynomial, \( S \) is defined as

\[
S_k(w) = d_k^2 P_k(w) \quad \text{(5.23)}
\]

\[
= s_0^{(k)} w + s_1^{(k)} w + \ldots + s_{n-k}^{(k)} w^{n-k}
\]

so that equations (5.22) become

[*] In Section 2.4.3, the simpler recursive method of Descout et al (1976) for computing the reflectance of a given bore shape was described.
By setting $w=0$ in equation (5.24a) and equating the coefficients of $w^{n-k}$ in equation (5.24b), the following expressions are obtained.

For $k=0$ to $n$:  
$$s_0^{(k)} - r_k d_k^2 q_0^{(k)} = 0$$  
(5.25a)

For $k=0$ to $(n-1)$:  
$$q_{n-k}^{(k)} - \left( \frac{r_k}{d_k^2} \right) s_{n-k}^{(k)} = 0$$  
(5.25b)

Thus, if polynomial coefficients $\{s_j^{(k)}\}$ and $\{q_j^{(k)}\}$ are known, $r_k$ and $d_k^2$ can be calculated. This in turn enables calculation of $\{s_j^{(k+1)}\}$ and $\{q_j^{(k+1)}\}$ using the following iterative equations derived from equations (5.24).

For $j=0, \ldots, (n-k-1)$:  
$$s_j^{(k+1)} = \frac{s_j^{(k)} / d_k^2 - r_k q_j^{(k)}}{1 - r_k^2}$$  
(5.26a)

$$q_j^{(k+1)} = \frac{q_j^{(k)} - \left( \frac{r_k}{d_k^2} \right) s_j^{(k)}}{1 - r_k^2}$$  
(5.26b)

NOTE: $r_k \neq 1$

$d_0$ is set to 1. Thus, $r_k$ and $d_k^2$ can be calculated progressively until $k = (n-1)$, and the product $\left( r_n d_n^2 \right)$ is calculated using equation (5.25a). The bore cross-sectional area is then computed from $r_k$ using the transmission line equation given in Section 2.4.3. It should be noted that the above algorithm can be simplified so that it no longer computes damping, by forcing $d_k$ to 1 throughout. Then it would become essentially the same as Goodwin's multiple reflection algorithm (see Section 2.4.5).

By examining the above equations, several potential sources of instability in the bore reconstruction algorithm come to light. Combining equations (5.25) gives
\[ d_k^4 = \frac{s_j^{(k)} s_{n-k}^{(k)}}{q_{n-k}} \quad (5.27) \]

\[ q_j^{(k)} = 1 \]

When \( p_j^{(0)} \) and \( q_j^{(0)} \) take on low values (see, for example, Figures 5.21 and 5.22), values of \( s_j^{(k)} \) and \( q_j^{(k)} \) may become very small, and when numerical errors are present, the above quotient becomes unstable; indeed, computations have sometimes given \( d_k^4 \) a negative value, although by definition it must be positive. In this instance, \( d_k^4 \) would have to be reset to a more plausible value — either 1 or \( d_{k-1}^4 \). This, however, seems rather arbitrary and would affect the accuracy of the corresponding \( r_k \). From equations (5.25)

\[ r_k = s_0 / d_k^2 \quad (5.28) \]

When the value of \( q_{n-k}^{(k)} \) is below some limit \([*]\), normally \( 1 \times 10^{-12} \) (the limit of numerical precision), it is treated as zero. Consequently, division by zero occurs in equation (5.19), and so \( d_k^4 \) cannot be defined. Here again, \( d_k^4 \) would have to be assigned some value, as above, in order for the iterative process to continue.

Because of numerical inaccuracies, \( d_k^4 \) has sometimes been computed to be greater than one. As this is physically impossible, it must be reset to one. Errors in \( d_k^2 \) will lead to errors in \( r_k \), such that \( |r_k| \) may assume values greater than one, again physically impossible. In these cases, \( r_k \) is reset to \( p_k^{(0)} \), which may be considered as a first approximation to \( r_k \). Such constraints reduce numerical inaccuracies but cannot remove them completely.

When, in the simplified algorithm, \( d_k^4 \) is forced to one, many of the above errors are avoided and a more stable algorithm may be anticipated.

The results of tests on the bore reconstruction algorithm will now be given.

\([*] \) When dealing with noisy data, the best choice of this limit is not easy.
5.4.2. Performance of Bore Reconstruction Algorithm

5.4.2.1. Test Data

For a controlled test of the bore reconstruction algorithm, the reflectance data is generated using the reflectance synthesis algorithm (Section 5.4.1.2). The model bore profile used, having \( n = 128 \), is shown in Figure 5.20. \( \{ p_j \}, \{ q_j \} \) and \( \{ h(j) \} \) were generated for the two cases of

1. no damping \( (d_k = 1) \),
2. a simple damping profile.

\( h \), for case (1) corresponds to the \( h \) used in the deconvolution tests (Data Set A), and can be seen in Figure 5.1, Section 5.2.1. The bore reconstruction algorithm generated \( \{ p_j(0) \}, \{ q_j(0) \}, \{ r_k \}, \{ d_k \} \) and \( \{ a_k \} \), where \( a_k \) denotes radius. These calculated values were compared with the true values by comparing rms and maximum errors in each case; thus, the accuracy of each stage of the algorithm was checked.

After this, the effects of distorting \( h \) in various ways were tested.

5.4.2.2. Results

In the absence of damping the non-simplified algorithm reproduced the original bore profile with high precision. When the damping coefficients were forced to one, so simplifying the calculations, the final result was even more accurate.

When damping was present, the algorithm did not work so well. With all \( d_k \) set to a value close to one, eg. .99, the original radius was reproduced correctly, but with less numerical accuracy than with no damping. The accuracy decreased as the values of \( d_k \) decreased, ie. as the extent of the damping increased. A possible reason is that the greater the damping, the more quickly \( h \) is reduced to small values. Very low values in the latter part of \( h \) could promote greater numerical inaccuracies in the ensuing calculations. For this reason, in the remainder of the tests no damping was incorporated.
Figure 5.20 Model Bore Profile (Radius in cm Versus Sample Number) Used to Test Bore Reconstruction Algorithm.
In the previous sections, it has been seen that deconvolution of measured data produces inaccurate h's. Therefore, it was necessary to test the stability of the bore reconstruction algorithm in the presence of noise. Addition of the smallest amounts of Gaussian noise to h resulted in significant numerical inaccuracies in p and q, which led to progressively larger errors in \( r_k \) and \( d_k^2 \), and the reconstructed radius. Figures 5.21 and 5.22 show how the calculated p and q were distorted when 1% Gaussian noise was added to h. Despite the error, most of the significant features of the originals were retained. The resulting \( r_k \) and radius were grossly distorted unless the damping coefficients were forced to one. Even \( 1 \times 10^{-6} \) Gaussian noise produced a badly distorted bore profile.

To try to pinpoint the source of instability, just one element of h was modified slightly. This was enough to produce noticeable errors throughout p and q. The algorithm was also rerun starting with \( p_j^{(0)} \) and \( q_j^{(0)} \) and proceeding straightaway to the calculation of \( r_k \) and \( d_k^2 \). Here the input data was the true p and q with just one element of one or the other slightly modified. This also produced a distorted bore profile.

5.4.2.3. Discussion

The bore reconstruction algorithm works extremely well for idealised data, but its performance is severely degraded by the presence of noise, however small its extent. Therefore, in its present form, it cannot be used effectively on deconvolved measured brass instrument data.

Possible reasons for the instability of the algorithm were suggested in Section 5.4.1.3. It may well be that some form of regularisation would improve the results, although this would necessarily entail a loss in precision.

The problem of noise has not been emphasised by investigators of vocal tract area-function recovery methods (Section 2.4). Their deconvolved data may have had negligible noise levels, possibly because of the narrower bandwidth, or their versions of the inversion algorithm may have been highly stable in
Figure 5.21 Results of Analysis on Synthesised \( \{ h(k) \} \) with 1% Gaussian Noise Added. (a) Original Synthesised \( \{ p_k \} \), (b) \( \{ p_k \} \) After Analysis.
Figure 5.22 Results of Analysis on Synthesised \( h(k) \) with 1% Gaussian Noise Added. (a) Original Synthesised \( \{ q_k \} \), (b) \( \{ q_k \} \) After Analysis.
5.4.3. Reflectance Synthesis Algorithm

5.4.3.1. Performance

Ideally the performance of this algorithm would be tested by comparing the synthesised $h$ with the true $h$. However, the true $h$ is not known, only a distorted version of it obtained by deconvolution of measured data. The generated $h$ is therefore convolved with the measured $f$ to produce a $g$ which can be compared directly with the measured $g$.

The response of a straight trumpet was generated as follows. Using calipers, its internal diameter was measured at various points along the bore as necessary, to obtain the variation of internal diameter with distance along the axis. For reflectance synthesis, these data had to be converted to equally-spaced samples of radius, and so the axial distance corresponding to one sample had to be determined. To do this, an approximation of the velocity of sound in the trumpet had to be made. The best approximation was the velocity of sound in the narrow 0.95 centimetre internal diameter source tube. This was found using the results of the source tube attenuation measurements of Section 3.1.3.5 by comparing pulse timings with distance travelled at each of the three microphone positions. The resulting velocity was 33.4 centimetres/millisecond. The diameter in centimetres was plotted by hand against axial distance in centimetres. Diameter data was read off from the graph every 0.364 centimetres of axial distance, which corresponded to one sample, and the resulting data (363 samples) was typed into the computer. Division by two gave radius variation, and five-point smoothing was applied to the parts of the data which might have become slightly uneven because of human error in the aforementioned process.

The precise position at which the bell reflection occurs depends on frequency (Benade, 1976, Section 20.5), and is therefore a variable distance.
inside the bell. It was arbitrarily assumed that it occurred two centimetres inside the bell, so that only 357 of the 363 radius "samples" were used for the synthesis. The assumption of no damping was also made. The resulting synthesised reflectance is shown in Figure 5.23. Also shown for comparison is the "best approximation" to the true h obtained by time-domain deconvolution (Section 5.2.2), and the measured radius file.

When the synthesised reflectance was convolved with the known input impulse, the resulting response bore considerable similarity to the experimentally measured response; Figure 5.24 shows this. This demonstrates the suitability of the model used (Section 5.4.1). The processes of reflectance synthesis and convolution are both well-conditioned and stable in the presence of noise, unlike the corresponding inverse problems of bore reconstruction and deconvolution. The success of the synthesis shows that the bore reconstruction algorithm would also give good results if it could be made less sensitive to noise.

Close inspection of Figures 5.23 and 5.24 reveals some differences. The sizes of some of the salient features vary. This could easily be because of inaccuracies in the diameter measurements. Alternatively, as the features are larger in the synthesised results, it could be caused by the lack of incorporation of damping.

A variation in the position of the features has also been noted. As time increases, features in the synthesised data occur progressively later than corresponding features in the measured data. This is because velocity variation in the real trumpet is not taken into account in the model which assumes the velocity to be uniform throughout. This point will be examined more closely in the next section.

5.4.3.2. Velocity Variation

A first approximation to velocity variation can be made using the following formula for the velocity of sound in a cylindrical duct of infinite length
Figure 5.23 Results of Reflectance Synthesis. (a) Measured Radius Data of Straight Trumpet, (b) Reflectance Synthesised From (a), (c) Best Approximation to True Reflectance Obtained Using Time Domain Deconvolution.
Figure 5.24 Comparison of Synthesised and Measured Responses. Response Obtained by Convolution of Measured Input Pulse with Synthesised Reflectance (Figure 5.23(b)) (Dashed), and Experimentally Measured Response of the Same Trumpet (Continuous).
v = c \left[ 1 - \frac{.162}{(a \sqrt{f})} \right] \quad (5.29)

where v is the velocity in the duct, a is the duct radius, f is the frequency in kHz and c is the velocity of sound in air. Here,

\[ c = 331.6 \sqrt{(1 + T/273)} \quad (5.30) \]

where T is temperature in °Celsius. In this way velocity is calculated as a function of radius. Velocity also depends on temperature, which is assumed uniform throughout. A frequency of 590 Hz was arbitrarily chosen, although the pulse contains all frequencies up to about 20 kHz, so propagation will be dispersive. Figure 5.25 shows computed velocity variation along the bore. It is not an accurate picture because the conditions attached to equation (5.29) are not satisfied. However, it is clear that velocity will tend to increase as radius increases, and will tend to remain constant in regions where the radius does not change. Consequently one would expect the features of the synthesised response to become progressively later in time than those of the measured response, which is what actually occurred (see Figures 5.23 and 5.24).

The above suggests a use to which the synthesis algorithm may be put. It could be used to investigate the velocity variation along bores of varying taper by comparing positions of features in the synthesised and measured responses. It should be noted that the value of velocity used in the synthesis (see previous section) must be correctly chosen in order for this comparison to be valid.

5.4.3.3. Discussion

The reflectance synthesis algorithm is stable and produces fairly realistic results. However, the model suffers from some limitations. It takes no account of variation of velocity along the instrument bore. It also suffers from all the limitations of a transmission line model - principally, that wave motion is not planar and that losses occur. The need for the inclusion of
Figure 5.25 Computed Approximation of Velocity Variation Along the Bore of the Straight Trumpet (Shown in Figure 5.23(a)).
frequency-dependent losses is demonstrated in Figure 5.23. The bell reflection in the synthesised response is thinner and more pointed than that of the measured response, indicating that the latter has experienced more high frequency attenuation, dispersion, and loss by radiation from the bell.

The synthesised reflectance could be used in various ways. It could be input into the time domain model of McIntyre et al (1983), which could then simulate mouthpiece pressure and volume flow rate (see Section 2.2.2). It can also be used to compute input impedance, as will be explained in Section 6.5.3.

5.5. Overall Summary of Chapter 5

In this chapter, careful investigations of three different processes — deconvolution, signal restoration and bore reconstruction — have been reported. None of these techniques have proved completely successful when applied to real data because all three are unstable in the presence of noise.

Three methods of deconvolution were investigated systematically. Time domain deconvolution was unstable for low initial values of input signal. Constrained iterative deconvolution would not converge for real data, because the measured input impulse did not satisfy certain eigenvalue criteria. The noise present in real data rendered the frequency domain deconvolution result noisy and non-causal. Various techniques to reduce the effects of noise were tried. The most effective was to shift the input data by two samples, but this introduced another type of distortion.

The technique of restoration was originally used in optics for image enhancement. Here its intended use was to correct the high frequency information which had become distorted by noise during deconvolution. However, the Gerchberg algorithm is unstable in the presence of noise. It diverged after a number of iterations dependent on the noise level. For the deconvolved real data, the noise level was so high that divergence occurred after just one iteration; however, one iteration did improve the result slightly.
The bore reconstruction procedure was more sensitive to noise than either deconvolution or restoration. It produced a perfect result for idealised data, but the smallest amount of noise led to grossly distorted results.

These investigations have served to illustrate the types of problems that may arise when attempting to solve inverse problems. Sometimes, the problems are so great that the final solution is nothing like what it should be. Solving inverse problems is not as straightforward as it might appear at first.

At the outset of the present work, such severe problems had not been anticipated, since the relevant literature did not warn of them. It may well be that the inversion algorithms (area function recovery) used by other investigators (Sondhi and Resnick, 1983; Lefèvre et al, 1983) were particularly stable versions, since no problems of instability in the presence of noise were reported; if so, no information about how this stability was achieved was given.

Sarkar et al (1981) explained that inverse problems are far more difficult than the direct problem. They are normally ill-posed, such that small errors in the original data lead to much larger errors in the final solution. In addition, the problem may not have a unique solution (Helmholtz, 1853). Regularisation can be used to transform the ill-posed problem into a well-posed one, but necessarily involves a loss of precision. Using the direct solution rather than the inverse solution where possible would be far more satisfactory. The direct process of reflectance synthesis has proved more successful than its inverse, bore reconstruction. If the limitations of the simple model employed can be overcome, the synthesis process will have many potential uses, one of which is outlined in the next chapter.
6.1. Introduction

In this chapter a general assessment will be made of the links which exist between the bore and the acoustics of brass instruments and how these may aid instrument design procedures. The assessment will take into account work carried out by others, and particularly the work presented in this thesis.

Section 6.2 contains a résumé of the achievements of the research. Subsequent sections discuss the extent to which these achievements may enhance progress towards the goal of designing a "better" instrument, as defined in Chapter 1. In Section 6.3, current instrument development procedures are reviewed, and their limitations exposed. Ideas on how transient response measurements may help in instrument design are presented in Section 6.4. Section 6.5 contains an assessment of links which already exist between bore geometry and acoustics, and ideas about how these may be strengthened using the transient response. Finally, in Section 6.6, ideas for future work are summarised.

6.2. Summary of Achievements of Research

The aims of the research were outlined at the end of Chapter 1. The extent to which the first four of these goals have been attained will now be discussed. The fifth goal will be dealt with in the remainder of the chapter.

(1) The experimental arrangement devised by Goodwin (1981), for measuring transient responses, has undergone considerable refinement in both apparatus and software. It is now capable of making accurate measurements of transient response, specifically reflectance (with non-ideal impulse) as defined in Section 2.2.3. Major advances include the introduction of a source tube, a method of effectively improving impulse consistency, and a technique for com-
pensating for changes in ambient temperature. By-products of this work were a method of investigating pulse attenuation in tubes and information about the inconsistency of spark sources.

(2) A technique for detecting physical differences between instruments by comparing their measured transient responses has been developed; the combination of visual inspection of superimposed response plots and analysis of arithmetical difference data has proved most effective. Small faults, which were missed by normal factory inspection methods, have been detected, for instance, a missing sleeve in a cornet. The potential benefits of this technique to the brass instrument industry are vast.

(3) A survey of different methods of deconvolution was made, and a systematic study of three of these - time domain, frequency domain and constrained iterative deconvolution - was carried out. The performance of all three was poor with non-ideal data. For future work, an investigation of Sondhi's method of deconvolution (SVD with regularisation) is recommended.

A by-product of these tests was an investigation of the performance of the Gerchberg restoration algorithm. Results confirmed that this algorithm tended to diverge in the presence of noise. Testing this data with an alternative restoration algorithm, eg. POCs, is a possible course for future work.

(4) A novel method of relating bore shape to reflectance has been developed by Jones (1983). The bore is modelled as successive cylindrical elements. Plane wave propagation and frequency-independent losses are assumed. Using z-transform relationships, the reflectance of a given bore profile may be synthesised. The reflectance synthesis algorithm was found to be stable and gave quite good results, subject to the limitations of the model used. The inverse algorithm for predicting bore shape from reflectance worked perfectly for idealised noiseless data, but was highly unstable in the presence of noise. Thus, it is of no practical use in its present form. Further work to make the
algorithm more stable in the presence of noise is recommended.

6.3. Existing Development Methods

Many factors, besides musical quality, govern the development of a new model. Fashion trends play a part, for instance, over the last fifty years, there has been a general change from small- to large-bore trombones. Instrument designs are also affected by players' tastes and prejudices. Competitors' models exert an influence. Mechanical working of slides and/or valves, position of stays, dimensions, balance, comfort, appearance, weight, finish, and even colour are all of importance. Other more down-to-earth factors include time constraints, cost constraints and technological limitations which govern the ease with which a particular model can be manufactured.

With a few exceptions [*], when a new model is being developed, one or more prototypes are produced. These may be based on the current model, or may be copies of a competitor's design. The suitability of a prototype is tested by inviting a professional musician to give his or her opinion. (The approval of a well-known player is a good selling point!) If the player is dissatisfied with some aspect of the quality, the prototype is modified and re-assessed, and so the process goes on (often for a long time!). This traditional design process is illustrated in Figure 6.1. Unfortunately, this process may sometimes have to be hurried when the manufacturer is under pressure to produce a new model quickly.

Two major problems exist with the above scheme. First, the manufacturer does not know what physical changes will bring about the desired improvement in musical quality. His decision about the alteration is largely empirical, made on the basis of his own experience and craftsman's intuition. This is clearly unsatisfactory. A relationship between the geometry of the instrument

[*] It should be noted that for several years, Dr. Richard Smith at Boosey and Hawkes has been using the more refined instrument assessment techniques to be discussed in Section 6.5.1.
bore and the subjective quality would be of great use here. It would mean that the musical quality of a given bore shape could be predicted. In Section 6.5, we discuss attempts that have been made to develop this relationship, and possible paths for future development.
The second problem is that it is an unfortunate possibility that even after all the fine adjustments and assessments of the prototype, the final production model may be an incorrect copy of the prototype. This is because at present there is no way of checking the exactness of the copy. Also, in the factory, the original prototype has been known to get lost, damaged, or even sold so that no permanent record of the original design remained. The next section contains a discussion of how impulse measurements overcome these problems, and how they can be of further use in the design and development of new instruments.

6.4. How Impulse Measurements Enhance Design and Development Procedures

6.4.1. Precise Copies Ensured

Using the impulse apparatus, the transient response of the new prototype is measured and remains permanently stored on the computer. The responses of production models are measured and compared with the prototype response; the two responses should be the same. The equipment does not ensure that the production model will be an exact copy; but it can show whether an attempt at copying is exact, and if it is not, where it is not, and whether it is too big or too small. Such a facility has not been available to brass instrument manufacturers before.

The loss of the original prototype would no longer be a problem. Its response will always be available on the computer (with back-up copies), so production models can be compared with it even years after its original design.

6.4.2. Comparison of Old and New Models

A prototype may be compared with the old model to see how and where it differs; for instance, a new Boosey and Hawkes prototype B-flat euphonium was compared with the old "Sovereign" model. Each instrument was measured twice - once with all valves up and once with all valves down. The measurements showed
that the Sovereign mouthpipe was shorter than the prototype's, so its record had to be shifted by two samples to make comparison easier. Comparison of the prototype with a second Sovereign euphonium showed that these both had the same mouthpipe. For some reason the mouthpipe of the first Sovereign was different. Later it was discovered that the first Sovereign had been fitted with an "Imperial" mouthpipe instead of a Sovereign mouthpipe. This incident served to confirm that the impulse apparatus is able to detect whether the wrong mouthpipe has been fitted unintentionally to an instrument. In Section 4.5.5.2, the equipment also detected that the wrong mouthpipe had been fitted to a trumpet.

Figure 6.2 shows the "valves up" comparison of the two instruments. Apart from mouthpipe differences and the anticipated differences in the valve group region, it appears that in a few positions between the valve group and the bell the prototype record is smoother. The instrument designer had actually sought to make the instrument smoother at these positions, so the measurements confirm that he was successful. These differences do not show up clearly in the corresponding "valves down" comparison.

6.4.3. Comparison with Competitors' Models

Instruments may be compared with the corresponding models made by competitors. This has been done for euphoniums, trumpets and cornets.

Euphoniums

A Boosey and Hawkes "Sovereign" B-flat euphonium was compared with the equivalent Yamaha model. It was suspected that Yamaha had attempted to copy the Besson/Boosey and Hawkes valve compensating system. The measurements confirmed that this was so. Figure 6.3 shows the "valves up" comparison; Figure 6.4 shows the "valves down" comparison. Apart from the mouthpipe differences, these plots reveal some other differences. In Figure 6.3, the Yamaha record has an extra spike (positive then negative) in the part corresponding to the
Figure 6.2 Valves Up Comparison of Two B-Flat Euphoniums. (a) Response of "Sovereign", (b) Response of New Prototype Euphonium. Positions at which Prototype is Smoother are indicated.
Figure 6.3 Complete Valves Up Comparison of Two Makes of B-Flat Euphonium. (a) Boosey and Hawkes "Sovereign", (b) Yamaha.
Figure 6.4 Valves Down Comparison of the Same Euphoniums as in Figure 6.3. (a) Sovereign, (b) Yamaha.
tubing between valves 3 and 4. Other than that, in both figures the Yamaha record appears to be generally smoother and flatter with lower amplitude spikes than the Sovereign record, suggesting smoother construction. One semi-professional player has said that the Yamaha instrument is easier to play. It may well be that smoothness of transient response is related to ease-of-blowing. (The relationship between transient response and subjective quality is discussed further in Section 6.5.4.) The Yamaha seems to be slightly shorter overall than the Boosey and Hawkes instrument, judging from the position of the bell reflection.

These sorts of differences would be difficult to detect without transient measurements.

Trumpets

A Boosey and Hawkes "Symphony" trumpet was compared with the corresponding Vincent Bach model (American). See Figure 6.5. The major arithmetical difference was at the mouthpipe ($3.4 \times 10^{-2}$), with less significant differences at the valve group ($1.9 \times 10^{-2}$). In fact, the two instruments were known to have different lengths of mouthpipe - 9 and 8.65 inches respectively - but this could not have been deduced from the responses. Possibly with further experience, such information could be gleaned.

Cornets

Comparison of two completely different makes of cornet resulted in a high rms difference ($2.65 \times 10^{-2}$) and a maximum arithmetical difference of $8.4 \times 10^{-2}$. See Figure 6.6. Here the differences were far more pronounced than in the above examples of euphoniums and trumpets.

6.4.4. Further Potential of the Apparatus as a Research Tool

In this section, further ways in which transient measurements could enhance research and development within the brass instrument industry are suggested.
Figure 6.5 Comparison of Two Makes of Trumpet. (a) Boosey and Hawkes "Symphony", (b) Vincent Bach, (c) Arithmetical Difference (Expanded Vertical Scale).
Fig. 6.6: Comparison of Two Makes of Cornet. (a) Boosey and Hawkes, (b) Unknown Make, (c) Arithmetical Difference.
When several different prototypes of a new model have been prepared, their transient responses can be measured and compared. If certain aspects of musical behaviour could be deduced from the transient response (see Sections 6.5.3 and 6.5.4 for possibilities), this would be a useful way of predicting the musical qualities of the prototypes. The transient responses would also show whether a serious fault, which the manufacturer was not aware of, was present in any of the prototypes — an unintended leak or constriction could be detected easily.

When the same model is manufactured over several years, subtle changes in manufacture may occur. After a period of time, the manufactured model may differ from the original prototype. Comparison of the transient responses of early and later versions of the same model will reveal any such differences. Of course, when the new instrument comparison technique is implemented in the factory, such variations will be avoided.

If each prototype of a new model is measured, a permanent record of how instruments have evolved over the years may be stored on computer. Such records may provide a useful reference for future workers.

Transient measurements may be used to investigate how or whether an instrument changes as it ages. If certain instruments are measured regularly throughout their "lifetimes", this may show up particular changes or faults which could be expected to arise through fair wear and tear, eg. a layer of deposit develops on the interior surface of the instrument. It could reveal some hitherto unknown weakness in the instrument which could then be remedied; for instance, the measurements could show that a particular model tends to develop a small leak in a particular region of the instrument after some years, suggesting a weakness in that region. The results could be used to determine what the lifetime of an instrument may be expected to be. They would also prove useful if an owner believes that the musical quality of his instrument has changed.
The measurements could also be used to keep in touch with the trends being followed by competing manufacturers. Measurements could help to explain why one brand of instrument sells better than another. However, one should bear in mind that there are many factors, eg. advertising, reputation, price, which influence sales more than the musical instrument itself.

6.5. Relationship between Bore Shape and Musical Quality

The problem of whether it is possible for the musical quality of an instrument to be predicted from its bore shape will now be considered.

(a) Bore Shape ---(Predict)--- > Musical Quality

If such a relationship could be achieved, it would be an extremely valuable tool for instrument design.

As yet no direct relationship between bore geometry and subjective quality has been formulated. The closest approach to a direct relationship has been clarified by R.A. Smith and Mercer (1979). They explain how small changes in the bore cross-section near the node or antinode of a standing wave can change the resonance frequencies. Resonance frequencies affect the intonation and the timbre (tone quality). Such bore perturbations can be caused by the essential mechanical parts of the instrument, eg. bends, valves, slide junctions and waterkeys. R.A. Smith and Daniell (1976) used the principles of bore perturbation theory to change selected resonance frequencies by desired small amounts. The above could be considered to be an indirect rather than direct link between bore shape and musical quality, via resonance frequency. Other investigators have found that the resonance characteristics are a measure of the quality of an instrument (Igarashi and Koyasu, 1953; Porvenkov and Sazarovskij, 1984).

Workers at the National Physics Laboratory have also been attempting to quantify relationships between the instrument's geometrical parameters and its subjective and acoustic properties (Anthony and Cox, 1984), but it is believed that this work has now stopped.
Some attempts have been made to forge an indirect link via input impedance. These will now be reviewed, and then suggestions for future lines of research will be made.

6.5.1. Indirect Relationship via Input Impedance

The proposed link is as follows.

(b) Bore Shape--(Predict)--Input Impedance--(Predict)--Musical Quality
(Intonation only)

The input impedance of a given bore can be calculated theoretically (Section 2.1.2.4). Alternatively, the input impedance could be measured using one of the methods given in Sections 2.1.2.2 and 2.1.2.3.

(c) Bore Shape--(Measure)--Input Impedance--(Predict)--Musical Quality
(Intonation only)

If subjective quality could be predicted from input impedance, the link would be complete. In Section 2.1.2.5, the prediction of intonation from input impedance using the "Sum Function" of Wogram (1972) was explained. However, a link between input impedance and other subjective qualities has not yet been found (Section 2.1.3). Therefore, the proposed schemes of (b) and (c) above break down.

Pratt (1978) and later Goodwin (1981) were the main researchers of the input impedance link, and both tried to use it to improve the traditional instrument design process shown in Figure 6.1 (Section 6.3). Pratt proposed a revised flow-chart shown in Figure 6.7 which replaced judgement by a professional player with two scientific evaluation procedures - measurement of input impedance and assessment of subjective quality. An example of the successful application of Pratt's subjective assessment techniques in the brass instrument industry was given in Section 2.1.3. The input impedance measurements have not, however, been adopted, as their usefulness was not apparent.

Goodwin (1981) modified this flow-chart again. See Figure 6.8. In Pratt's flow-chart, the first prototype model is not based on any scientific
Figure 6.7
Flow Chart Showing Modified Design Process Suggested by Pratt

Start

Construct prototypes

Measure prototypes' complex acoustic impedance

Submit prototypes to a panel of players for subjective assessment using SDS and MDS

Players satisfied with quality?

Yes → Enter next stage

No → Modify instrument using Perturbation Theory
Figure 6.8
Flow Chart Showing Modified Design Process Suggested by Goodwin

1. Design new bore shape
2. Calculate input impedance, etc.
3. Good intonation?
   - Yes: Submit to objective test methods based on existing SDS and MDS scales.
   - No: Modify bore shape
4. Satisfactory?
   - Yes: Construct prototype to spec.
   - No: Modify rating scales for objective tests.
5. Satisfactory?
   - Yes: Start production
   - No: Modify rating scales for objective tests.
recommendation. In Figure 6.8, Goodwin aims to predict the musical properties of the proposed bore shape, and thus check its suitability before it is made. He uses his "Look Back" method (see Section 2.1.2.4) to calculate the input impedance (Goodwin, 1981, Section 2.3), from which the quality of intonation can be deduced. The weak part of Goodwin's proposed scheme is the box labelled "Submit to objective test methods based on existing SDS and MDS scales". Clearly, this cannot be achieved because at present the subjective response cannot be determined from input impedance.

6.5.2. Indirect Relationship via Transient Response

The weakness of the link between input impedance and subjective quality means that an alternative route for the relationship between bore shape and musical quality must be sought. From the context of this thesis, the obvious choice is to use transient response instead of input impedance. This would result in the following indirect relationship.

(d) Bore Shape--(Predict)--->Transient Response--(Predict)--->Musical Quality

For the link to be realised, the two constituent links must be forged, namely the link between bore shape and transient response, and the link between transient response and musical quality. Each of these will be discussed in turn.

6.5.3. Link between Bore Shape and Transient Response

One advantage of replacing input impedance with transient response is immediately obvious. The transient response as a function of time at the input of the instrument is closely related to the physical structure of the instrument, whereas the input impedance as a function of frequency is not. Simple inspection of a plot of the transient response immediately shows the precise locations of bore discontinuities; there is an increase in bore cross-sectional area where the response decreases, and an area decrease where it increases. This simple direct relationship has been utilised as the basis of the instrument bore comparison technique of Chapter 4. One cannot look at a
plot of input impedance and immediately see where bore discontinuities occur. Therefore, any link between transient response and subjective quality will be far more like a direct link between bore structure and subjective quality than will a link between input impedance and subjective quality.

As the impulse apparatus can be used to measure the transient response of an instrument, the following link can also be considered.

(e) Bore Shape→(Measure)→Transient Response→(Predict)→Musical Quality

Alternatively, the transient response of a given bore shape may be predicted theoretically (see (d) above). Chapter 5 dealt with the theoretical link between transient response and bore shape. In Section 5.4.3, a way of synthesising the reflectance of a brass instrument was derived, and the results were shown to be fairly accurate when compared with experimental measurements (see Figure 5.24). An additional advantage to this approach is that input impedance can be calculated from reflectance using the following convolutional relationship derived by Sondhi and Resnick (1983).

$$2 \frac{r(t)}{r(t)} = z(t) - \int_0^t z(\tau) r(t - \tau) d\tau$$  \hspace{1cm} (6.1)

where \( r(t) \) is the reflectance and \( z(t) \) is the input impedance in the time domain. This can be forward Fourier transformed to give

$$Z(\omega) = \frac{2R(\omega)}{1 - R(\omega)}$$  \hspace{1cm} (6.2)

where \( Z(\omega) \) and \( R(\omega) \) are the Fourier transforms of input impedance and reflectance respectively. One can compensate for the absence of the mouthpiece simply by modelling it as a mass-spring-damper lumped element, and using the method of Elliott (1979, Section 2.4.7) to adjust \( Z(\omega) \) [*].

Using equation (6.2), the input impedance of a straight trumpet has been calculated from its synthesised reflectance (see Figure 5.23, Section

[*] It would be interesting to compare the results of the two methods of computing input impedance, i.e. the \( Z(\omega) \) calculated using the above method, and the \( Z(f) \) computed using Goodwin's "Look Back" method (Section 2.1.2.4).
5.4.3.1). (The mouthpiece correction was not included.) The first 4 kHz are shown in Figure 6.9. The characteristic resonance frequencies are clearly visible. Normally, all trumpet resonances have disappeared by 1 kHz. Their presence here can be accounted for by the absence, in the model of Section 5.4, of frequency-dependent damping and radiation loss.

Figure 6.10 compares reflectance and input impedance in the time domain. The input impedance is of much greater duration, which agrees with the prediction of Schumacher (1981) (Section 2.2.3).

When input impedance is calculated from reflectance, any relationship between input impedance and musical quality (Sections 2.1.3 and 6.5.1) can also be utilised. The main use of input impedance in this respect is that the intonation of a proposed bore shape could be predicted.

6.5.4. Links between Transient Response and Subjective Quality

The idea of a link between transient response and subjective quality is not new, but has received little attention as yet. Benade (1969) states that "...short constrictions in a bore give early partial reflections which also damage tone onset...". In other words, he is proposing a link between transient response and the subjective quality of responsiveness (or "ease-of-blowing"). Edwards (1978) has explained that players use the word "responsiveness" to describe the transient behaviour of the instrument; a responsive instrument allows the player to start a note easily, or change from one note or sound to another easily.

Bowsher (1976) reported investigations into the possible relationship between the transient behaviour of trombones and their quality. On this occasion Benade (1976) proposed that a "stuffy" instrument is one that suffers from premature reflections. Kent (1956) has also said that "unwanted reflections within the instrument cause it to be 'stuffy', or hard to blow. They reduce the efficiency of the instrument and cause the tone quality to suffer." The term "stuffy" is a subjective description of the response (or "feel") of
Figure 4.9 Input Impedance of Straight Trumpet Computed from its Synthesised Reflectance.
Figure 6.10 Comparison in the Time Domain of (a) Synthesised Reflectance and (b) Input Impedance Computed From It.
the instrument, the opposite of stuffy being "freeblowing". (The correlation between "stuffiness" and "poor responsiveness" varies from player to player.)

The results of Section 4.5.5.2 would confirm the existence of a direct relationship between transient response and stuffiness; the trumpet with a constriction in the bell bend was reported as stuffy, the trumpets without this constriction were not reported as stuffy. The fact that the cigarette lighter in the euphonium bell caused stuffiness (Section 4.5.5.1) would also add support. It was suggested, in Section 6.4.3, that the smoothness of the transient response may be related to ease-of-blowing.

So far, experiments to investigate the effects of bore changes, and therefore transient response changes, on subjective acoustics have been inconclusive. The work of R.A. Smith and Daniell (1976), mentioned at the start of Section 6.5, could be extended further. They established links between bore perturbations and changes in resonance frequencies, which can be related to changes in pitch and tone quality. The bore perturbations could be represented by the corresponding change in transient response which could be determined either by measurement or theoretical calculation. Well-defined changes in transient response could lead to a strengthening of the links between transient response and musical qualities.

Krüger (1979) has investigated the relationship between the playing quality of brass instruments and aspects of their measured transient response, principally the bell reflection (Section 2.2.2).

Pratt (1978, unpublished work) found that a trombone with a severely dented tuning slide could not be distinguished musically from an undamaged version of the instrument. A possible reason here was that the person who undertook the subjective testing was not used to the particular instrument. A more experienced player would be more sensitive to differences. The difference in the measured input impedances of the two instruments was negligible (Pratt, 1978, Section 4.9.4). Therefore, input impedance was not a useful physical
measurement in this instance. A careful measurement of transient response using the new equipment would detect and locate the dent easily.

Goodwin (1981, Section 5.4) postulated "that premature reflections affected the transient properties of the instrument in a musically significant way"; to investigate this he used a French Horn with two sets of slides - one with standard ends, and one with ends chamfered to a knife-edge. Subjective tests showed that differences were not significant. Goodwin's simple measurements of the transient response of the horn in the two different states did not detect the change in bore geometry. With the refined impulse apparatus, these differences could almost certainly have been detected.

### 6.5.5. Discussion of Future Subjective Experiments

Further research in this area is strongly recommended. It seems probable that Goodwin's experiment would have been more conclusive had the bore modification been greater. Future subjective experiments to investigate the effects of bore constrictions could be designed along the following lines.

1. **Start with a set of identical instruments.** (Their similarity could be verified using the impulse apparatus.) Add constrictions of varying size at one particular point in each instrument. (Again, the correctness of the position of the constriction could be verified using the impulse equipment.) Carry out subjective tests to see whether experienced players [*] can distinguish between the musical quality of the instruments. This will show what size of constriction causes a significant difference in subjective quality.

2. **Repeat (1) at different positions** to see whether at certain positions, constrictions have a greater effect on quality, and whether the size of constriction required to produce a perceptible change in subjective quality varies depending on the position of the constriction.

[*] It is important that the musician doing the tests is able to detect small differences in quality. Experience has shown that it is not necessarily the top players who are best able to do this.
(3) Starting with a set of identical instruments, place a single constriction at a different position in each instrument except one. Determine whether each instrument can be distinguished from the unmodified instrument. If each constriction is the same size, this will show at which positions constrictions are most likely to affect quality. This could be used as a means of consolidating the results of experiment (2).

(4) Investigate the effect of having more than one constriction. (Interaction of constriction effects could be a complex problem.)

One could investigate the subjective effects of leaks in a similar way.

An alternative approach would be to investigate the effects of different bore sizes, and profiles of bells and mouthpipes. R.A. Smith (1986) has already begun investigations of this type on trumpets, carrying out blindfold tests to determine each individual player's preferences.

It seems appropriate at this stage to point out that musical quality will not be solely dependent on bore discontinuities. Many and various factors influence the musical performance, as illustrated by the complicated model of the control structure governing the player/instrument interaction given in Bowsher (1983). Therefore, the problem must not be oversimplified. However, further research into the relationship between transient response and subjective quality would appear to be worthwhile.

6.5.6. Potential Use of Indirect Transient Response Relationship in Instrument Design

If the transient response proves to be a more useful way of predicting instrument quality than input impedance, the instrument design procedure could be improved greatly. The flow-chart given in Figure 6.11 could be utilised.

The first stage is an attempt to eliminate those factors which may prejudice a player's judgement of the musical quality of the instrument, eg. action of slides/valves, appearance, finish, weight, balance, etc. Market research should be carried out beforehand to discover exactly what physical
Figure 6.11
Flow Chart Showing Proposed New Instrument Design Procedure

Run visual and touch tests (not musical or aural) to determine which non-musical features suit players best.

Design Prototype bore shape

Synthesise reflectance of proposed or modified bore shape. Calculate input impedance - either via reflectance or directly using Goodwin's "Look Back" Method. From input impedance calculate "Sum Function".

Predict subjective quality using relationships between reflectance and subjective quality (not yet fully developed). Predict intonation from Sum Function.

PREDICTED quality and intonation good enough? No Evaluate quality and/or intonation faults and modify bore shape accordingly.

YES Make prototype to exact specification of proposed bore shape and measure its transient response.

Submit to a panel of players for subjective assessment of intonation and quality using SDS and MDS.

ACTUAL quality and intonation good enough? No Keep permanent record of relationship between measured response and subjective judgements for future reference.

YES Start production, measuring transient response of each manufactured instrument and testing it against measured transient response of prototype to ensure consistency of final product.
characteristics the players would like their instruments to have. In this way a picture of the most popular or "optimum" instrument may be built up. (Of course, this instrument would change with time, in accordance with predominant fashion trends.) All prototypes should be made to correspond with this optimum instrument, but their bore dimensions may be varied experimentally in order to find the bore shape which will give optimum musical quality.

The manufactured prototypes will be furnished with bore profiles whose theoretically predicted intonation and subjective quality are good. One problem is that it is difficult to ensure that the manufactured bore is an exact copy of the one specified theoretically, because of limitations in the technique by which bore diameters are measured. An exact copy would be a particular problem with the tuba, for the reasons explained earlier (Section 4.5.3).

Intonation and quality of the manufactured prototypes are then assessed subjectively. Ideally the results should agree with those predicted theoretically. The prototype which performs best in these tests will then be the new instrument model.

This rigorous instrument design procedure has a more rational basis than the procedures currently followed in the brass instrument industry.

When proper production of the new model begins, the impulse equipment ensures that the production model is an exact copy of the prototype and is consistent. A musician could also play each instrument as a double check. This two-part quality assurance test would be a good marketing point.

It should be noted that in some senses there is no such thing as the "ideal" or "optimum" instrument. Players' tastes vary. Each player will have his own preference. An instrument which one musician regards as perfect in all respects may be considered by another to have poor tone quality and to be unattractive in appearance. However, for mass produced instruments, this problem will always be present.
6.6. Summary of Suggestions for Future Work


(2) Look for ways of making the bore reconstruction algorithm (Section 5.4) more stable in the presence of noise.

(3) Investigate the variation of velocity along the bore of the instrument; transient measurements could help. Then velocity variation could be incorporated in the reflectance synthesis algorithm (Section 5.4.3.2).

(4) Investigate the attenuation in the flaring parts of brass instruments using the method of Section 3.1.3.5.

(5) Compare the input impedances computed theoretically using a time domain method (Section 6.5.3) and a frequency domain method (Section 2.1.2.4).

(6) Try an alternative to the Gerchberg restoration algorithm, eg. POCs (Section 2.3.5.2).

(7) Extend the bore perturbation work of R.A.Smith and Daniell (1976) (Section 6.5.4), so that changes in resonance frequencies of brass instruments can be related to changes in transient response.

(8) Carry out the subjective experiments outlined in Section 6.5.5, to investigate links between bore geometry and subjective quality.

6.7. Overall Summary of Chapter 6

A relationship between musical quality and bore shape is necessary in order that instrument designers can be sure that their modifications to instruments will bring about a musical improvement. There are indications that the transient response may be more helpful in this respect than input impedance. Further investigation of links between transient response and subjective quality is required.

It is recommended that current instrument design procedures be modified to include impulse measurements as well as subjective assessments. The measurements provide a useful way of comparing a prototype with other versions of
the same instrument, and ensure that a production model is an exact copy of the prototype.
APPENDIX A

SUMMARY OF COMPUTER PROGRAMMES AND THEIR FUNCTIONS

(1) Data Acquisition

Main Routine: IMPULSE5 (in NOVA FORTRAN 5, loadline: IMP5LD)

Measures transient response (see Figure 3.18 for flowchart and Section 3.2.1).

IMPULSE5 calls the following subroutines:

- MIKECAL5 - microphone calibration
- FUND5 - calibration aid
- IDEFIPKAR - aid for inspection of incoming impulses
- INTERPEAK - interpolation to determine the precise timing of impulse peak
  (see Section 3.2.2)
- IPEAKMODE - summarises information regarding the timing of the previous
  10 impulses
- IAUTORATE - changes the sampling rate to compensate for ambient tempera­
  ture variation (see Section 3.2.4)
- ITEMPCHECK - subroutine relating ambient temperature, source tube velo­
  city and sampling rate.

(2) Response Comparison Software

INORMSHIFT (loadline: INOSHLD) - standardises a measured response, ie. shifts,
subtracts background level, normalises and creates ROI (see Sections 3.3.3 and
4.2.2).

ISUBTRACT (loadline: ISUBLD) - finds arithmetical difference between two
responses

ISUBTRACT2 (loadline: ISUB2LD) - finds arithmetical difference between two
responses, lists on screen the highest differences in order of decreasing size
and in order in which they occur along the instrument; calculates root mean square difference; finds position of first significant difference and the number of signi-
significant differences.

(3) Plotting (HP Plotter)

IMRECPL (loadline: IMRPLD) - plots one or more transient responses (whole or part), using any horizontal or vertical scale; one, two or three plots per A4 page.

YANDPSPEC (loadline: YANDPLD) - plots one or more spectra; can cope with different sampling rates; can plot either real/imaginary or magnitude/phase (subroutine FREQCONV converts data from real/imag to mag/phase).

(4) Deconvolution

(a) Time Domain Deconvolution

Main Routine: DECONTIME (loadline: DETIMELD)

Carries out either normal or regularised time domain deconvolution. DECONTIME calls the following subroutines:

FOPEN/BINREAD/BINWRITE - read/write single or double precision binary data

ASCREAD/ASCREWRITE - read/write ASCII data

DCONVDP - carries out normal time domain deconvolution

DCONVDPREG - carries out regularised time domain deconvolution

CONVDP - carries out convolution (double precision)

DFUNCTION - finds rms difference between true g and computed g

SHIFT - shifts result if necessary

(b) Frequency Domain Deconvolution

Main Routine: FREDECONV (loadline: FREDLD)

FREDECONV calls the following subroutines:

SHIFT - explained above

FOPEN/BINREAD2/BINWRITE2 - read/write single or double precision binary spectrum
BINREAD/BINWRITE - explained above

PYFFT32 - double precision FFT, size varies up to 2048 complex points.

(c) Constrained Iterative Deconvolution

Main Routine: DECONVIT2 (loadline: DEC2LD)

See Figure 5.13 for an explanatory flow chart and Section 5.2.4. DECONVIT2 calls the following subroutines:

ASCREAD/ASCWRITE, FOPEN/BINREAD/BINWRITE, DFUNCTION, CONVDP - all explained above.

Most of the main routine DECONVIT2 appears in Appendix B.

(d) Other

CONVOLVDP (loadline: CONVLD) - discrete convolution (uses subroutine CONVDP)

PYTIMFRE32 - reads data, FFT's it (using subroutine PYFFT32) and writes away resulting spectrum

SMOOTHSTA (loadline: SMOULD) - 5- or 7-point smoothing of data

WINSPEC (loadline: WSPECLD) - spectral windowing (see Section 5.2.3.3)

IRANADD (loadline: IRANLD) - add random noise with either uniform or Gaussian distribution to data file

(5) Restoration: Gerchberg Algorithm

Main Routine: PYGERCHBC (loadline: PYGBCLD)

PYGERCHBC calls the following subroutines:

PYFFT32 - explained above

PGSORTHOG - Gram-Schmidt Orthogonalisation (part of Jones' Acceleration Procedure)

PCOEFF - forms polynomials from Gram-Schmidt results

PNEWRAPH - solves above polynomials using Newton-Raphson method

PFINDW - forms vectors through which the normal Gerchberg result is improved
PCORRECT - accelerate normal Gerchberg result by applying the above vectors

FOPEN/BINREAD/BINWRITE - explained above

(6) Bore Reconstruction

Main Routine: BOREC (formerly DECONVIT4) - loadline: BORECLD (formerly DEC4LD)

See Section 5.4.1.3. BOREC calls the following subroutines:

TOEPLITZ - finds q from h by solving linear equations of the Toeplitz type

CONVDP - convolves h with q to find p

FOPEN/BINREAD/BINWRITE - explained above

ASCREAD/ASCWRITE - explained above

ANALYSIS - finds reflection and damping coefficients from p and q

AREACALC - computes area function from reflection coefficients

(7) Reflectance Synthesis

(a) Main Routine: SYNTH (loadline: SYNLD)

See Section 5.4.1.2. SYNTH calls the following subroutines:

REFCALC - computes reflection coefficients from area function

DAMPCALC - forms damping function - a constant, or proportional to either length of element or its wall area

DCONVDP - computes h from p and q by time domain deconvolution.

The first part of the routine SYNTH appears in Appendix C.

(b) Other

REFTOIMP (loadline: RTOILD) - converts reflectance to input impedance using the equation in Section 6.5.3

VELTUBE (loadline: VELD) - calculates first approximation to velocity variation along tube of varying cross section (Section 5.4.3.2).
(0) Miscellaneous

FILEEDIT - creates or modifies a data file (ASCII or binary)

ISCAN - allows user to survey contents of single precision data file

ISCANDP - allows user to survey contents of double precision data file

RRCOMPARE - compares contents of two single precision data files

RRCOMPDP - compares contents of two double precision data files

IBGMODEL2 - generates n sets of m random numbers in specified range, sums them, and calculates rms value at each stage (see Section 3.2.5).
APPENDIX B
PRINTOUT OF THE MAIN PARTS OF THE PROGRAMME "DECONVIT2" (CONSTRAINED ITERATIVE DECONVOLUTION)

double precision h(0:2048), f(0:2048), g(0:2048), fch(0:2048)
double precision hactual(0:2048)
double precision amax,gmax,amu,fmax,hdiff,hdiff2,hmax,hmax1
double precision sd,rms

cccc ccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccc

c Routine to iteratively deconvolve, ie. find h (1 elements) from g (n elements) and f (m elements).
cccc ccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccc

c (1) Identify how many elements in each array, read in arrays c from relevant files.

isd=0
print #,'is g noisy (l=yes) ?'
read *,in
if(in.ne.l)go to 64
print *,'what is standard deviation (sd) of added noise ?'
read *,sd
print *,'do you want the opportunity to stop the program when
print *,'rms error between f*h & g reaches sd (l=yes) ?'
read *,isd
64 continue

(2) locate positive maximum of g (& f), & form h0.

nitwrite=0
nit=0

------------------------------------ a i n  l o o p ------------------------------------------

90 if(nit.eq.nitwrite)print *, 'after which iteration do you next want
1 to write h-data away ?'
   if(nit.eq.nitwrite)read *,nitwrite
   if(nitwrite.le.nit)go to 90

(3) convolve f & h by calling subroutine convdp. [fch=f*h]
call convdp(n,f,h,fch)
nit=nit+1
if(nit.ne.nitwrite)go to 91
print *, 'write away final f*h for comparison with g (l=yes) ?'
   read *,iw
   if(iw.ne.1)go to 91
92 print *, 'write away in binary (2) or ascii (1) ?'
   read *,ibin
   if(ibin.ne.1.AND.ibin.ne.2)go to 92
   if(ibin.eq.1)call ascwrite(fch,(n-1))
   if(ibin.eq.2)call binwrite(fch,n)
(4) Form difference array between fch and g and store in fch.
   do 60 i=0,(n-1)
   fch(i)=g(i)-fch(i)
60 continue

(5) Apply 'dfunction' to this difference array
   call dfunction(1,n,fch,ifmax,rms)
   if(isd.ne.1) go to 65

   Opportunity to stop if rms between g & f*h = sd.
   if(rms.gt.sd) go to 65
   print *, 'opportunity to stop program because rms is approx equal to sd . . . '
   nitwrite=nit
   isd=0
   pause

65 continue

(6) Form next approx'n to h by adding fch to previous approx'n
   do 70 i=0,(n-1)
70 h(i)=h(i)+fch(i)*mu

Extra bit to calc amax within the initial 1 elements only.
   amax=0.d0
   iamax=0
   do 130 i=0,(n-1)
   if(abs(fch(i)).gt.amax) amax=abs(fch(i))
   if(abs(fch(i)).eq.amax) iamax=i
130 continue

Print *, 'iteration number ',nit,' complete.'
   amax=mu*amax
   print *, 'maximum difference between old and new h = ', amax
   print *, 'at iamax = ', iamax,'(this takes mu into account)'

Find rms error between true h (hactual) and calculated h (h)
   if(icomp.ne.1) go to 54
   hdiff=0.d0
   hmax=0.d0
   hmaxl=0.d0
   do 34 i=0,(n-1)
   hmax1=damax1(hmax1,hactual(i))
   hdiff2=dabs(hactual(i)-h(i))
   hmax=damax1(hmax,hdiff2)
   if(hmax.eq.hdiff2) ihmax=i
   hdiff=hdiff+hdiff2*hdiff2
34 continue

   hdiff=hdiff/dfloat(1)
   hdiff=dsqrt(hdiff)
   print *, 'rms error between hactual & h = ', hdiff
   print *, 'max " " " = ', hmax1,' at ', ihmax
   print *, 'where hactual(i) = ', hactual(ihmax)
   hmax=max(hmax,100.d0/hmax1)
   print *, 'max value of hactual(i)=', hmax1, 'so % error=', hmax1,' %'
54 continue

Print *, '*******************************************************'
   if(nit.ne.nitwrite) go to 90

(7) Option of writing away new h
   print *, 'write away latest h (1=yes) ?'
   read *,iw
if(iw.ne.1) go to 56
56   print *, 'write away in ascii (1) or dp binary (2) ?'
    read *, ibin
    if(ibin.ne.1.AND.ibin.ne.2) go to 56
    if(ibin.eq.1) call ascwrite(h,(n-1))
    if(ibin.eq.2) call binwrite(h,n)
55   continue
   c (8) option of finishing program
   print *, 'maximum difference between HNEW & HOLD = ', amax
   print *, 'stop program (1=yes) ?'
   read *, istop
   if(istop.ne.1) go to 90
   c-------------------end of main loop -------------------------
   print *, 'end of program'
   stop
   end
APPENDIX C

FIRST PART OF THE MAIN ROUTINE "SYNTH" (REFLECTANCE SYNTHESIS)

double precision ax(0:1025),ra(0:1025)
double precision re(0:1024),d(0:1024)
double precision h(0:1024)
double precision pc(0:1024),qc(0:1024)
double precision pcdash(0:1024),qcdash(0:1024)
double precision t,temp,pdiff,qdiff,pi

pi=4.0*datan(1)

14 print *, 'duct elements range from i=0 to i=n ..... n ='
read *,n
if(n.gt.1024)go to 14

(1) option of reading area data (ax) or radius data (ra)
11 print *, 'read in x-sec area (1) or radius data (2) ?'
print *, '(if neither is required type 0)'
read *,iar
if(iar.eq.0)go to 13
if(iar.eq.1)go to 12
if(iar.ne.2)go to 11

(2) option of calculating reflection coefficient data or reading it
17 print *, 'read in reflection coefficient data (1) or calculate it using
1refcalc (2) ?'
read *,iar
if(iar.eq.1)go to 18
if(iar.ne.2)go to 17
c calculate reflection coefficients from x-sec areas by calling subroutine
 c refcalc
    call refcalc(ax,re,n)
g o to 19
1 8 continue
5 2 print *, 'is reflection coefficient data in ascii (1) or binary (2) ?'
   read *, ibin
   if (ibin.eq.1) call ascread(re,n)
   if (ibin.eq.1) go to 19
   if (ibin.ne.2) go to 5 2
   call binread(re,(n+1))
1 9 continue

(3) option of reading damping coefficient data or calculating it
2 1 print*, 'read in damping coefficient data (1) or calculate it using
   ldampcalc (2) ?'
   read *, iar
   if (iar.eq.2) go to 2 2
   if (iar.ne.1) go to 2 1
6 8 print *, 'is damping coefficient data in ascii (1) or binary (2) ?'
   read *, ibin
   if (ibin.eq.1) call ascread(d,n)
   if (ibin.eq.1) go to 6 1
   if (ibin.ne.2) go to 6 8
   call binread(d,(n+1))
6 1 continue
   go to 2 3
2 2 continue

(4) calculate damping coefficients from radius data by calling subroutine
 c dampcalc.
    call dampcalc(ra,d,n,pi)
2 3 continue

(5) step through all k-values from (n-2) to 0, calculating pc's & qc's.
do 10 k=(n-2),0,-1
   pc(0)=re(n-1)
   pc(1)=d(n)*d(n)*re(n)
   qc(0)=1.0
   qc(1)=re(n-1)*re(n)*d(n)*d(n)

1 0 continue

---------------------------end of main loop----------------------------------------

(6) form pcdash & qcdash (from 0 to (n-k) ) from previous pc & qc
   pcdash(0)=re(k)
   qcdash(0)=1.0
   do 20 j=1,(n-k-1)
      pcdash(j)=re(k)*qc(j)+d(k+1)*d(k+1)*pc(j-1)
      qcdash(j)=qc(j)+re(k)*d(k+1)*d(k+1)*pc(j-1)
2 0 continue
   pcdash(n-k)=d(k+1)*d(k+1)*pc(n-k-1)
   qcdash(n-k)=re(k)*d(k+1)*d(k+1)*pc(n-k-1)

(7) replace old pc's & qc's with newly-formed pcdash's & qcdash's.
do 30 j=0,(n-k)
   pc(j)=pcdash(j)
   qc(j)=qcdash(j)
3 0 continue
1 0 continue

---------------------------end of main loop----------------------------------------

(8) form first n values of h by deconvolution using subroutine...
    call dconvd(n,qc,pc,h)

further deconvolution necessary to form values of h from h(n+1) to h(2n)
do 90 i=(n+1),(4*n)
t=0d0
do 91 j=1,n
temp=qc(j)*h(i-j)
t=t+temp
91 continue
h(i)=-t
90 continue

C (9) option of writing away h, p & q in ascii or binary.

C (10) option of comparing calculated p,q with known p,q

stop
end
REFERENCES


Agulló, J. and Badrinas, J. (1985) "Improving the Accuracy of the Capillary-Based Technique for Measuring the Acoustic Impedance of Wind Instruments," Acustica 59, 76-83


Benade, A.H. (1976b) Private Discussion with Dr. J.M. Bowsher, 91st Meeting of the A.S.A., Washington, DC


Cooper, M.J. (1977) "Deconvolution: If in Doubt, Don't Do It," Physics Bulletin 28(10), 463-466


Fletcher, N.H. (1979) "Excitation Mechanisms in Woodwind and Brass Instruments," Acustica 43, 63-72


Helmholtz, H.L.F. (1853) Ann. Phys. Chem. (Ser. 3) 89, pp. 211-233 and 353-377 (In German)


Speech, Signal Processing ASSP-34(3), 624-626


Ducts," J. Acoust. Soc. Am. 74, 320-332


of Brass Instruments," Proc. 18th Acoust. Conf., Cesky Krumlov, Czechoslo-
vakia, pp. 91-94 (In German)

Instruments as a Tool for Assessment of Quality," Personal Communication

XII Journées d'Etude sur la Parole, Montreal: GALF, pp. 26-41 (In French)

ic Configurations from Acoustical Measurements," Acustica 52(4), 227-231 (In French)


Peckham, H.D. (1971) "Computers, BASIC, and Physics," Reading, Massachusetts: Addison-Wesley, Ch.6


Pratt, R.L. and Bowsher, J.M. (1979) "The Objective Assessment of Trombone Quality," J. Sound Vib. 65(4), 521-547


Rubchinskii, A.V. (1964) "Recovery of Breakdown Strength after a Spark Discharge." In Klyarfel'd, B.N. "Investigations into Electrical Discharges in Gases," Pergamon Press, pp. 56-91


Skudrzyk, E. (1971) "The Foundations of Acoustics," N.Y.: Springer-Verlag (Section 17.3.4) (First Published in 1954 in German)


Wogram, K. (1972) "A Contribution to the Measurement of the Intonation of Brass Instruments," Ph.D. Thesis, Carolo-Wilhelmina University, Braunschweig. (Originally in German, English Translation Available at University of Surrey)


Young, F.J. (1960) "The Natural Frequencies of Musical Horns," Acustica 10, 91-97