Bore Optimisation and Impedance Modelling of Brass Musical Instruments

Alistair C.P. Braden

A thesis submitted in fulfilment of the requirements

for the degree of Doctor of Philosophy

to the

University of Edinburgh

2006

Abstract

The input impedance of a brass musical instrument is a good representation of its resonance characteristics. Methods of calculating input impedance for a known instrument shape, or bore-profile, are reviewed, and an extension to existing theory for bent waveguides is given.

These input impedance methods form the basis for consideration of the inverse problem; to find a bore-profile with given impedance characteristics. Such problems can be formulated as bore reconstruction – finding an unknown bore from its impedance, and performance optimisation – altering certain characteristics of a known bore.

The inverse problem is solved by means of optimisation, using either genetic algorithms or the Rosenbrock direct-search method. A number of new techniques are used to improve convergence speed by minimising both the size of the search space and the number of design variables. These techniques are incorporated into an elegant object-oriented instrument representation, allowing convenient and flexible problem definition and forming the basis of an integrated application in C++.

Experimental impedance measurements on a series of trombones are conducted, and typical results presented. These data are used to formulate optimisation targets. The optimiser is then used for bore reconstruction and performance optimisation to solve real-world trombone design problems.

Declaration

I declare that this thesis is composed by myself, and that the work therein is my own.

Alistair Braden

Acknowledgements

I offer sincere thanks to Murray Campbell, my supervisor, for his continual support and guidance throughout this research project, and to John Chick for providing invaluable discussion.

Gratitude is due to other researchers with whom I have had discussions. Firstly, to Wilfried Kausel, whose work this thesis continues; my discussions with him in the early stages of this work had a great influence on its eventual direction. Noam Amir, Daniel Noreland, Vincent Pagneux and Simon Félix offered useful comments and insight. Discussions with Stefan Bilbao were of great help with some of the theory. Arnold Myers provided some valuable bore measurements. My copy of Jonathan Kemp's thesis is well-thumbed and provided much assistance.

My thanks also to those in the trombone world. Most notably, Chris Beaumont and the team at Michael Rath Brass Instruments gave kind assistance; the opportunity to measure their instruments was greatly valued. I have had many enlightening discussions on the internet-based Online Trombone Journal; special thanks to Chris Stearn, Timothy Richardson and John Sandhagen for their useful comments. Further thanks to all the local trombonists who allowed me to measure their instruments.

The motley crew in 4201 contributed a lot to this thesis, as much in terms of distraction as in scientific input. Big thanks to Mike, Sam, Darren, Rob, Seona, Dave, Lin, Alan, Francois and Darren's guitar. Thanks to the physics cricket team, E=MCC, of which I have had the singular pleasure of being captain.

Finally, thanks to Karen for her unwavering support and encouragement; may her future be everything she wants it to be.

This work was supported by the Engineering and Physical Sciences Research Council (EPSRC).

List of Figures

List of Tables

Contents

List of Symbols

Below is a list of the symbols used in this thesis. The list is not entirely exhaustive – certain symbols are used only for a very short time, are clearly defined in the text and are therefore omitted here. Due to the large number of symbols needed, some overlap of notation is necessary. While effort has been taken to minimise this, there are situations in which only highly counter-intuitive notation could resolve the clash. In the interests of readability, the former option was usually preferred, and such overlaps as exist are marked clearly in the table, are defined clearly in the text and describe very different quantities which cannot be confused. Additionally, the conventions of the cited references are preserved where possible.

α	Modal index (Ch. 3)
$\boldsymbol{\alpha}$	Design vector
β	Modal index
γ	Flare coefficient
$\delta_{\alpha\beta}$	Kronecker delta function
ε	Boundary-specific admittance (Ch. 3)
ε	Design tolerance (Ch. 5)
ζ	Bore discontinuity radius ratio
η	Shear viscosity component
Θ	Solid angle
κ	Bend curvature
λ_i	Sum of successful steps in direction i (Ch. 4)
Λ	Modal eigenmatrix (diagonal) (Ch. 3)
$\boldsymbol{\Lambda}$	Direction matrix for orthoganilisation procedure (Ch. 4)
μ	Modal index (Ch. 3)

μ_ϕ Tuning parameter of peak frequency objective function

μ_z Tuning parameter of peak magnitude objective function

ν Modal index (Ch. 3)

ν Half-width of window in full-curve objective function (Ch. 5)

ν_ϕ Half-width of peak frequency objective function

ν_z Tuning parameter of peak magnitude objective function

ξ Displacement (Ch. 2)

ρ Equilibrium air density (see Ap. D)

σ Modal rotational symmetry index

ς Weight function for combining objective functions

τ Dummy variable of integration

ϕ_i Peak frequency (Ch. 5)

ϕ Polar coordinate

Φ Radiation source polar coordinate

$\boldsymbol{\psi}$ Modal transverse eigenfunctions

ω Angular frequency

A Amplitude constant (Ch. 2)

A Normalisation function for ψ (Ch. 3)

$\mathbb{A}$ Design space

B Amplitude constant (Ch. 2)

B Shorthand matrix for modal derivation (Ch. 3)

C Shorthand matrix for modal derivation

D Modal spatial dependence matrix (diagonal)

$\mathcal{D}$ $D(d)$

E Shorthand matrix for modal derivation

F Bore-discontinuity matrix (Ch. 3)

F EFP reference frequency (Ch. 7)

G Shorthand matrix for radiation impedance derivation

H Shorthand matrix for modal derivation

I Identity matrix

J_i	Bessel-function of the first kind, order i
K	Modal axial wavenumber matrix (diagonal)
$\mathcal{K}$	Bulk modulus
L	Shorthand matrix for modal derivation
M	Shorthand matrix for modal derivation
N_e	Number of elements discretising a horn
N_m	Number of modes included in calculation
N_v	Number of design variables
N_{pk}	Number of peaks included in objective function
N_{pt}	Number of points in discretised impedance curve
$\mathcal{O}$	Objective function
$\mathbf{P}$	Modal pressure amplitude vector
Q	Shorthand matrix for radiation impedance derivation
R	Radiation source radial coordinate
$\mathcal{R}(x)$	Bore profile function
$S(x)$	Cross-sectional area function
T	Transmission matrix
U	Acoustic volume velocity
$\mathbf{U}$	Modal volume velocity amplitude vector
V	Volume
Y_i	Bessel-function of the second kind, order i
Z	Impedance matrix
Z_0	Input impedance matrix
Z_d	Output impedance matrix
∇	Gradient operator
^{T}	Matrix transverse (post-fix) operator
$'$	Differential operator
a	Rosenbrock tuning parameter (Ch. 4)
b	Bessel-horn coefficient (Ch. 2)
b	Rosenbrock tuning parameter (Ch. 4)

c	Speed of sound (see Ap. D)
d	Length of waveguide
$\mathbf{d}$	Directions of traversal
h	Height of spherical cap (Ch. 2)
h	Radiation distance to point in free space (Ch. 3)
i	Index
j	$\sqrt{-1}$
k	Wavenumber
k^*	Lossy wavenumber
$k_{\parallel\alpha}$	Axial cut-off wavenumber of mode α
$k_{\perp\alpha}$	Transverse cut-off wavenumber of mode α
l_i	Step size in direction i
m	Nodal diameters index
n	Nodal circles index
$p(\mathbf{x})$	Acoustic pressure scalar field
q	Dummy index (radiation impedance)
r	Radial coordinate
r_v	Ratio of pipe radius to boundary layer thickness
s	Toroidal axial coordinate
t	Time coordinate
$\mathbf{v}(\mathbf{x})$	Acoustic particle velocity vector field
$v(x)$	Axial component (scalar) of acoustic particle velocity field of plane-wave
x	Axial coordinate
$\mathbf{x}$	Position vector
$z(x,\omega)$	Impedance
z_0	Input impedance
z_c	Characteristic impedance
z_c^*	Lossy characteristic impedance
z_r	Radiation impedance

Chapter 1

Introduction

1.1 Motivation: Musical Instruments

Music is notoriously difficult to define in rigorous scientific terms, but a reasonable, if incomplete, definition is 'a series of rhythmic and harmonically-related sounds produced to evoke an emotional response'. Musicians use their instruments to create pleasing sounds; the better the musician, the more pleasant the sound. However, the quality of music is also dependent on the instrument; one can play a Mozart Horn Concerto using a mouthpiece and a piece of hose, but it will not sound as pleasant, and will therefore be less musical, than the same piece played on a finely-crafted horn. Instruments are a necessary tool for producing good music, and, as is the case in general, good-quality tools facilitate better results.

The quality of instruments, like the quality of music, is a subjective judgement made by the performers and the audience, and musicians strive to produce the best possible performance. They will use the instrument that they perceive will best allow them to achieve their musical goals, which may vary depending on the music performed, and will certainly vary from player to player. There is therefore a clear motivation for manufacturers to produce the best instruments, because inferior models will be more difficult to sell. While forces such as brand recognition will have an effect, among professional players the instrument market is strongly driven by quality.

From an engineering standpoint, a cheap, mass-produced violin is almost indistinguishable from a Stradivarius, but to a violinist there is a world of difference. Musicians are highly skilled, work to extremely small tolerances, and possess very fine judgement; they are there-

fore sensitive to very small changes in equipment. Instrument craftsmen have been constantly improving their designs for centuries, and there are only small differences between excellent and merely very good instruments; a musician will favour (and buy) the excellent one. It is therefore difficult to improve on modern instruments, but manufacturers are constantly trying, and succeeding, to do so. Another issue is that general trends in taste change over time, a process which is complicit with the development of new models.

It is not possible to define, in any defensible and quantitive way, the properties of a perfect instrument, or indeed the meaning of 'perfect', even within the bounds of stringent assumptions about a specific player's tastes and requirements. It is reasonable for a given musician to state that under given criteria a given instrument is the best he has played, but the inherently qualitative nature of this judgement cannot rule out in principle the possibility of a better one. There is therefore always room for improvement, for a given meaning of improvement, regardless of the quality of the existing instruments.

When compared to almost any other manufacturing industry, the process of instrument design has not changed much for centuries. A prototype of a potential design will be built and tested in a costly trial-and-error process, requiring much time and materials to produce what may potentially be a very bad instrument. Manufacturers have accumulated a vast amount of knowledge and intuition to help them choose the next design to try, but this is a fundamentally unscientific approach – while it is proven beyond doubt as being effective, it may yet benefit from being augmented with scientific techniques. Given that the quality of the end result is judged subjectively by musicians, the design process cannot be entirely objective and is therefore dependent on the human expertise of the manufacturers and players. Computers will never be able to design instruments by themselves, but they are potentially an invaluable tool to assist human designers; this is also the case in countless other industries.

When considering brass instruments, such as trumpets, trombones, horns and the like, there are several ways in which acoustical science can assist instrument design; we focus here on two. First, a computational way of evaluating the playing properties of an instrument would allow theoretical designs to be tested, thus allowing unsatisfactory designs to be identified without the time and expense of construction. Second, the computer can try many different designs to find the best match for desired criteria, which can then be built and

tested in the traditional manner. Given that it can try many more designs much faster than can a human designer, it may be able to suggest solutions that a human may not otherwise have tried.

This work focusses on the development of computational tools useful for augmenting the design process of the manufacturers of brass instruments. We will give an overview of the relevant acoustics, which we will use to define an objective method for judging instruments. We will then specify how a computer might choose which instruments to try, implement the method computationally, and then explore the capabilities of the resulting software.

1.2 Acoustics of Brass Instruments

In the simple case of a cylindrical pipe open at both ends, the *acoustic pressure*, being the difference between the local air pressure inside the instrument and the ambient mean atmospheric pressure, is, to a first approximation, zero at both ends. The acoustic pressure within the pipe may vary with time, and will therefore form standing waves with nodes at the ends and integer numbers of half-wavelengths in the pipe. The set of possible standing waves forms a *harmonic series*; the first harmonic (that with one half-wavelength) is the *fundamental*, and the higher harmonics have frequencies of integer multiples of this fundamental. It is much easier to drive standing waves at any, or all, of these frequencies, or *resonances*, than at other frequencies.

Now consider a similar pipe closed at one end. The acoustic pressure is required to be zero only at the open end, so the standing wave solutions have an odd number of quarter-wavelengths in the pipe. The fundamental will be half the frequency of that for the open pipe, and the only the odd harmonics will be present.

The assumption of having pressure nodes at the open ends is somewhat crude, as is does not allow sound to be radiated from the instrument. More detailed consideration of open ends is given in Chapter 3, but the simplified model is sufficiently accurate for the purposes of this argument.

When brass instruments are played, the musician closes the mouthpiece end with his lips, and blows through them. This causes the lips to commence a self-sustained oscillation, with a small aperture between the lips rapidly opening and closing, thus inducing an oscillatory

volume velocity, which may drive the standing waves described. The oscillation of the air column couples non-linearly with that of the lips, forming a stable note and encouraging the lips to oscillate at a resonance of the instrument ('slotting'); skilled players can overcome this and allow the lips to oscillate at frequencies away from resonances ('lipping'). A small fraction of the energy of the standing waves is radiated at the open end of the instrument, and is audible as sound.

We now introduce the concept of *input impedance*, being the frequency-dependent quotient of pressure and volume-velocity at the plane of the mouthpiece; Figure 1.1 shows an input impedance plot typical of brass instruments. Each peak corresponds to a resonance, with taller peaks denoting a stronger response, and therefore greater ease of maintaining a standing wave at that frequency. The pitch of the radiated sound is that of the excitation, but the lips may additionally provide excitation at harmonics of this frequency. The magnitude of the impedance at the excitation frequency and at its harmonics will affect the amplitude of oscillation at these frequencies, and therefore the timbre of the sound. A cylindrical brass instrument, such as a trumpet or trombone, is a closed pipe, and we might therefore expect the resonances to be the odd harmonics only. However, the flaring bell section shifts the resonances into approximately a complete harmonic series. Consequently, when the instrument is excited at one of its resonances, it will give strong support to oscillations at this frequency and also its harmonics, thus producing a pleasant sound rich in harmonic content.

In practice, the resonances of brass instruments are close to being harmonically related. However, the first resonance is often very flat in comparison to the fundamental of the harmonic series which closely matches the other resonances. This resonance gives a weak response and is not played, but it is possible to excite the instrument at the 'fictitous' fundamental of the harmonic series and receive a strong response, producing a sound which has a very low fundamental component (a 'pedal note').

1.3 Objective Judgements

Musicians judge instruments in subjective ways, which may depend on the player's individual technique, personal preferences, and the type of music in which he specialises. Different musicians may give different judgements of the same instrument; while such methods are effective

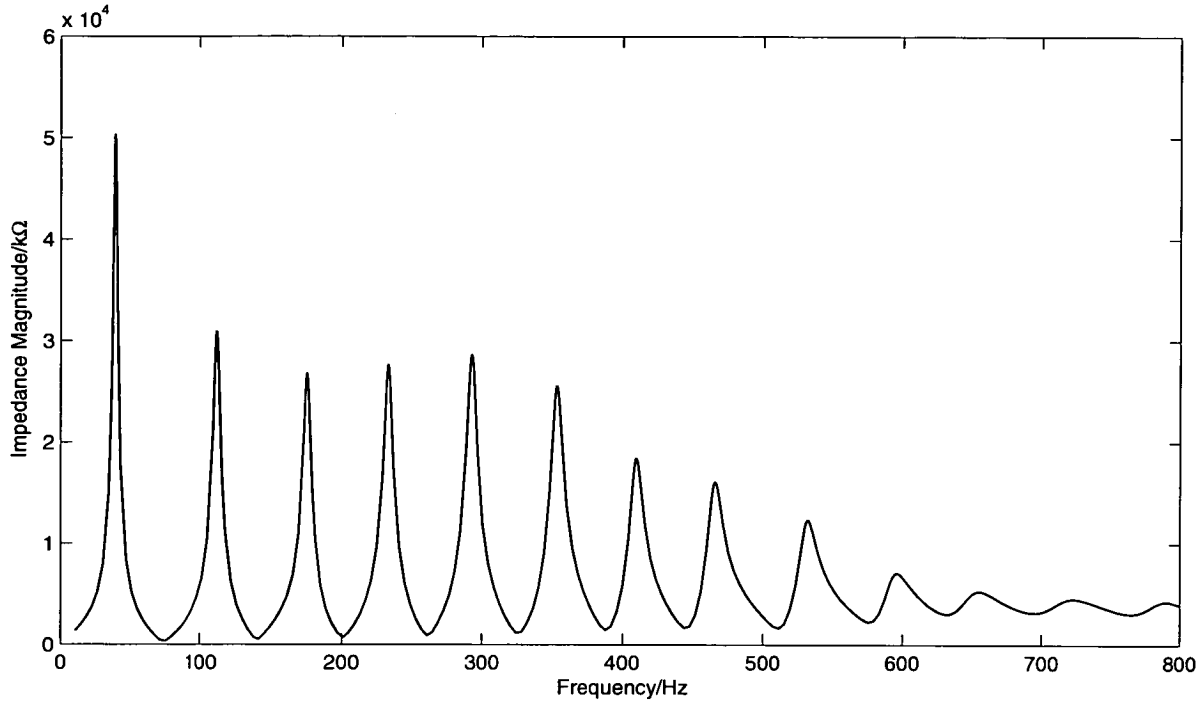


Figure 1.1: Input impedance of typical trombone.

from the musician's point of view, a scientific approach requires an objective method. We have seen that a good approximation of the behaviour of an instrument, at least in the linear regime, is given by its input impedance, which is therefore a good candidate for our objective measure. It has the additional benefits of allowing the use of established experimental techniques for measuring real instruments, and numerical techniques for theoretical instruments; the agreement between the two is very good, but not perfect as we shall see in Chapter 8.

The relationship between impedance and playing characteristics is understood only in broad terms. It is generally easy to identify bad instruments from impedance features like badly-aligned resonances and non-smooth peaks, but it is not yet possible to distinguish the good from the very good. Analysis of a wide range of instruments of varying quality, identifying common features, will give insight into what a quantitative description of a good instrument may be. It is possible, indeed likely, that there are features of an instrument not present in its impedance which affect its quality as judged by a musician. Until these features are identified and understood, we must neglect them for our purposes. Any process which produces objective data describing a fundamentally subjective system must, at some stage, involve a human judgement; clearly the more data available, the more informed this judge-

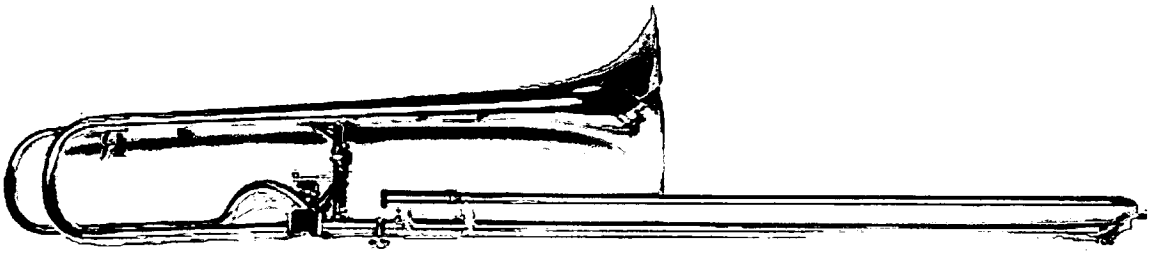


Figure 1.2: A modern trombone.

ment may be. In this case, we must acknowledge that the objective data are incomplete; this merely affects how well-informed is the judgement, not whether the judgement is necessary. A better understanding of the subjective properties of input impedance is useful, but is not essential to the process.

Input impedance models for straight general horn shapes were considered, with a plane-wave assumption, in [1, 2], and with the inclusion of higher modes in [3, 4, 5]. Bent tubes were investigated in [6, 7] and modelled in [8, 9, 10, 11, 12]. The effects of losses at the walls of the instrument were covered in [13, 14, 15], and the impedance at the radiating end in [16, 17, 18].

1.4 Brass Instrument Design

To a first approximation, a brass instrument is a long piece of narrow, straight, cylindrical or slowly-flaring metal tubing terminated by a rapidly-flaring bell. In practice, there will be certain discontinuities in the bore connecting sections of different radius, the tubing will have a number of bends, and the cross-section may not always be exactly circular. A cup-shaped mouthpiece is used for playing. The radius of the instrument can be expressed as a function of position down the axis of the instrument, termed the *bore-profile*, on which the performance of the instrument is strongly dependent. Although there are other design features, such as choice of valves, wall material and thickness, placement of braces etc., there is, to the author's knowledge, no scientific method for accounting for these choices, so we discount them and focus on input impedance, which is dependent only on geometry.

A designer can identify desired impedance features to give a quantitative target, and the

fitness of a potential design can be evaluated by comparing its impedance to this target. Calculating the impedance of a given bore profile and comparing to a target is sufficiently fast that a computer can try many different designs in a short time. This process can be combined with optimisation algorithms, which attempt to find an optimal solution with the smallest number of tries.

Techniques have already been developed to use optimisation algorithms to try many different bore profiles in search of certain impedance characteristics. The first attempt, by Kausel [19], is commercially available and has already been successfully used by manufacturers [20]. Noreland [21] uses a more advanced algorithm and develops ways to represent instruments numerically. Both these methods model only the plane-wave, and use relatively simple representations.

The primary objective of this work is to develop computer software that includes input impedance calculations, functions for evaluating the fitness of a design, and optimisation techniques in such a way as to provide a tool useful to assist instrument manufacturers with their design processes. This will comprise developing a sophisticated instrument representation, allowing flexibility of use and improving optimisation speed, and using more accurate physical models.

1.5 Thesis Aims

The aims of this thesis are:

1. To review existing methods for input impedance calculation, and to derive appropriate methods for bent tubing.

2. To develop the underlying techniques used in brass instrument optimisation, particularly the instrument representation.

3. To investigate the comparative efficacy of the Rosenbrock and genetic algorithms for brass instrument optimisation.

4. To develop integrated and easy-to-use software, written in C++, to perform optimisation and other related tasks.

5. To conduct experimental measurements of real trombones to gain insight into the impedance properties of good instruments.

6. To use the optimiser to reconstruct an instrument from its impedance.

7. To use the optimiser to modify existing instruments to alter certain impedance properties.

1.6 Summary of Thesis

Chapter 2 gives an overview of the physics of brass instruments, and considers in detail the linear propagation when only one (plane) acoustic mode is considered. The model is extended in Chapter 3 to include higher modes in both straight and bent tubes, giving more accurate results by considering motion in the transverse directions. Optimisation algorithms are laid out in general in Chapter 4, with particular focus on the Rosenbrock and genetic algorithms; their application to brass instrument design is considered in Chapter 5, along with a detailed examination of the techniques developed for this work. The methods by which these components are implemented and combined in the development of computer software are considered in Chapter 6. An experimental study of the input impedance of trombones is conducted in Chapter 7, providing target data for the optimiser. A numerical comparison of the various methods for calculating impedance is conducted in Chapter 8. The capabilities of the optimisation software are examined in Chapter 9.

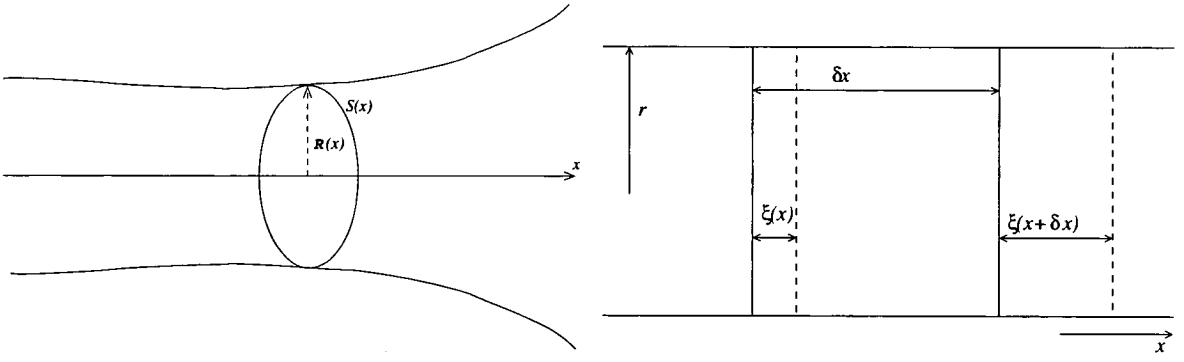
Chapter 2

Plane-wave Propagation in Straight Acoustic Horns

In this chapter we will derive results for evaluating the input impedance of a general horn under the plane-wave assumption.

2.1 Webster Horn Equation

Consider a horn of arbitrary radius profile $\mathcal{R}(x)$, where x is the axial coordinate [22], having cross-sectional area $S(x) = \pi\mathcal{R}(x)^2$ (figure 2.1a). A small volume $V = S\delta x$ is contained between two planes perpendicular to the axis (figure 2.1b), separated by length δx. These planes are subject to displacements $\xi(x)$ and $\xi(x + \delta x)$ respectively, giving change in volume



(a) Horn with arbitrary bore profile.

(b) Displacement in a volume element.

Figure 2.1: Diagrams of components of a general horn.

$$\begin{aligned} \delta V &= [S(x+\delta x)(\delta x + \xi(x+\delta x)) - \xi(x)S(x)] - S(x)\delta x \\ &\rightarrow \delta x \frac{\partial}{\partial x}(S(x)\xi) \end{aligned} \tag{2.1}$$

(requiring $S(x)$ differentiable), or fractional change

$$\frac{\delta V}{V} = \frac{1}{S}\frac{\partial}{\partial x}(S\xi). \tag{2.2}$$

We define the bulk modulus $\mathcal{K}$ quite generally by

$$p = -\mathcal{K}\frac{\delta V}{V}, \tag{2.3}$$

where p is the acoustic pressure. For air under the conditions that concern us [2], we may use $\mathcal{K} = \rho c^2$, where ρ is the equilibrium air density and c the speed of sound. Combining with (2.2) gives

$$p = -\rho c^2 \frac{1}{S}\frac{\partial}{\partial x}(S\xi). \tag{2.4}$$

The force imparted on the gas between the planes by the gas to the left is $Sp(x)$ and from the right is $S(p(x) + \frac{\partial p}{\partial x}\delta x)$, giving a net force of $-S\frac{\partial p}{\partial x}\delta x$. We equate this to the product of particle acceleration and the mass $\rho S\delta x$ by Newton's 2$^{\text{nd}}$ Law

$$-S\frac{\partial p}{\partial x}\delta x = \rho S \delta x \frac{\partial^2 \xi}{\partial t^2}, \tag{2.5}$$

giving

$$-\frac{1}{\rho}\frac{\partial p}{\partial x} = \frac{\partial^2 \xi}{\partial t^2}. \tag{2.6}$$

Differentiate twice with respect to t and substitute into equation (2.4) and to give

$$\frac{1}{S}\frac{\partial}{\partial x}\left(S\frac{\partial p}{\partial x}\right) = \frac{1}{c^2}\frac{\partial^2 p}{\partial t^2}, \tag{2.7}$$

which is the Webster horn equation; alternatively,

$$\begin{aligned}\frac{\partial^2 p}{\partial x^2} + \left(\frac{1}{S}\frac{\partial S}{\partial x}\right)\frac{\partial p}{\partial x} &= \frac{1}{c^2}\frac{\partial^2 p}{\partial t^2}, \\ \frac{\partial^2 p}{\partial x^2} + \frac{\partial p}{\partial x}\frac{\partial}{\partial x}(\ln(S)) &= \frac{1}{c^2}\frac{\partial^2 p}{\partial t^2}.\end{aligned} \tag{2.8}$$

Note that if $S(x)$ is constant with respect to x, then this reduces to the 1-dimensional wave-equation.

Assuming the motion is simple harmonic, we have $p(x,t) = p(x)e^{j\omega t}$, where ω is the angular frequency, giving $\partial^2 p/\partial t^2 = -\omega^2 p$. This yields

$$\frac{\partial^2 p}{\partial x^2} + \frac{\partial p}{\partial x}\frac{\partial}{\partial x}(\ln(S)) + k^2 p = 0, \tag{2.9}$$

where $k = \omega/c$ is the wavenumber. This is a (Helmholtz) horn equation for a horn of general shape, and is based on the following assumptions:

1. The horn is straight and has rigid, smooth walls
2. The fluid is inviscid and there is no friction with the walls of the horn
3. Pressure variations are infinitesimal
4. The pressure is uniform over the wavefront - i.e. is planar.

Given this general horn equation and a known bore profile $S(x)$, we differentiate the logarithm of this profile to produce a horn equation for that specific geometry. In general, this is not analytically soluble, but there are certain special cases with such solutions.

2.2 Cylinders

Consider a cylindrical waveguide of length d with bore radius $\mathcal{R}_0$ such that $S(x) = S_0$ for $x \in [0, d]$. The throat and mouth (by which we mean the input and output ends) are located at axial coordinates x_0 and x_1. From (2.9), the horn equation is then

$$\frac{\partial^2 p}{\partial x^2} + k^2 p = 0, \tag{2.10}$$

with solution

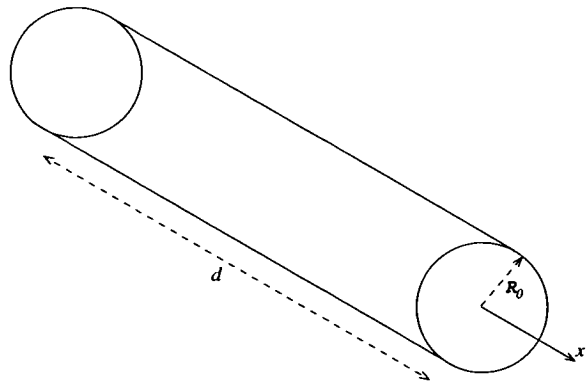


Figure 2.2: Schematic of cylindrical duct, length d and radius $\mathcal{R}_0$.

$$p(x) = Ae^{-jkx} + Be^{jkx}, \tag{2.11}$$

where A and B are the forward- and backward-moving amplitudes. This implies that the motion is sinusoidal in space as well as in time. Using momentum conservation

$$v(x) = \frac{j}{\rho\omega}\frac{\partial p}{\partial x} \tag{2.12}$$

gives acoustic particle velocity

$$v(x) = \frac{k}{\rho\omega}(Ae^{-jkx} - Be^{jkx}). \tag{2.13}$$

Although the particle velocity is strictly a vector field in three-dimensional space $\mathbf{v}(\mathbf{x})$, under the plane-wave assumption there is a component only in the axial direction and it is a function of axial coordinate only, so it therefore reduces to a scalar field $v(x)$. Define the impedance at axial coordinate x as being

$$z(x) = \frac{p(x)}{S(x)v(x)}. \tag{2.14}$$

Strictly speaking, this should be notated $z(x,\omega)$ as it is (importantly) a function of the angular excitation frequency (likewise $p(x,\omega), v(x,\omega)$), but this is omitted for clarity. Using $k = \omega/c$, the impedance at the mouth x_1 is

$$z_1 = z_c\frac{Ae^{-jkx_1} + Be^{jkx_1}}{Ae^{-jkx_1} - Be^{jkx_1}}. \tag{2.15}$$

where $z_c = \rho c/S_0$ is the characteristic impedance. Noting that $x_0 = x_1 - d$, we have the impedance at the throat x_0

$$\begin{aligned} z_0 &= z_c \frac{Ae^{-jk(x_1-d)} + Be^{jk(x_1-d)}}{Ae^{-jk(x_1-d)} - Be^{jk(x_1-d)}} \\ &= z_c \frac{\cos(kd)(Ae^{-jkx_1} + Be^{jkx_1}) + j\sin(kd)(Ae^{-jkx_1} - Be^{jkx_1})}{\cos(kd)(Ae^{-jkx_1} - Be^{jkx_1}) + j\sin(kd)(Ae^{-jkx_1} + Be^{jkx_1})} \\ &= \frac{\cos(kd)z_1 + jz_c\sin(kd)}{z_c^{-1}j\sin(kd)z_1 + \cos(kd)}, \end{aligned} \tag{2.16}$$

which is the lossless input impedance of a cylinder as shown in [1] table 2, equation 1, and also equation 8.23 on page 179 of [2]. We introduce a dissipative model from Keefe [13, 2], which models the effects at the wall by means of a boundary layer. This replaces lossless characteristic impedance z_c and wavenumber k with lossy z_c^* and k^* respectively:

$$\begin{aligned} z_c^* &= z_c[(1 + 0.369r_v^{-1}) - j0.369r_v^{-1}], \\ k^* &= k[1.045r_v^{-1} + j(1 + 1.045r_v^{-1})], \end{aligned} \tag{2.17}$$

where

$$r_v = \sqrt{\frac{\rho\omega S}{\eta\pi}} \tag{2.18}$$

is the ratio of the pipe radius to the boundary layer thickness, and η is the shear viscosity component – setting this to zero gives the lossless case of $z_c^* = z_c$ and $k^* = jk$. Noting that $\cos(i\theta) \equiv \cosh(\theta)$ and $\sin(j\theta) \equiv j\sinh(\theta)$, we have

$$z_0 = \frac{\cosh(k^*d)z_1 + z_c^*\sinh(k^*d)}{\frac{1}{z_c^*}\sinh(k^*d)z_1 + \cosh(k^*d)}, \tag{2.19}$$

which is the lossy input impedance of a cylindrical duct (cf. [1] Table 2, Equation 2).

2.3 Cones

Consider a conical waveguide, (fictitous) apex at the origin, solid angle Θ, with throat and mouth located at coordinates x_0 and x_1 respectively along the central axis x of the cone,

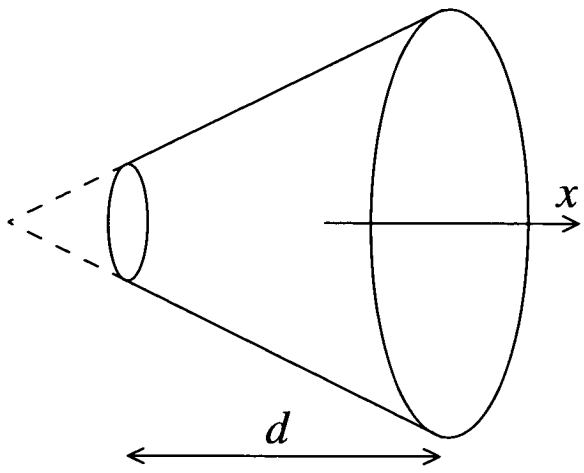


Figure 2.3: Schematic of conical horn section, axial length d.

separated by distance d. The cross-sectional area is $S(x) = \Theta x^2$, and the mouth and throat radii are $\mathcal{R}_0$ and $\mathcal{R}_1$. From (2.9), we have

$$\frac{\partial^2 p}{\partial x^2} + \frac{\partial p}{\partial x}\frac{\partial}{\partial x}(\ln(\Theta x^2)) + k^2 p = 0, \tag{2.20}$$

giving

$$\frac{\partial^2 p}{\partial x^2} + \frac{2}{x}\frac{\partial p}{\partial x} + k^2 p = 0, \tag{2.21}$$

which is the horn equation for a conical horn with plane waves. The complete solution is

$$p(x,t) = \left(A\frac{e^{-jkx}}{x} + B\frac{e^{jkx}}{x}\right)e^{i\omega t}. \tag{2.22}$$

Momentum conservation (2.12), recalling that $p(x,t) = p(x)e^{j\omega t}$, relates the acoustic particle velocity to the pressure by

$$\frac{\partial v(x,t)}{\partial t} = -\frac{1}{\rho}\frac{\partial p(x,t)}{\partial x}, \tag{2.23}$$

where ρ is the equilibrium density of the medium; separating the time-dependence $e^{j\omega t}$ we have

$$\frac{\partial v(x,t)}{\partial t} = -\frac{1}{\rho}\frac{\partial p(x)}{\partial x}e^{j\omega t},$$

$$v(x,t) = -\frac{1}{j\rho\omega}\frac{\partial p(x)}{\partial x}e^{j\omega t}, \tag{2.24}$$

from which we can now ignore the time-dependence in the usual manner. The pressure and particle velocity are then

$$p(x) = \frac{1}{x}[(A+B)\cos(kx) - j(A-B)\sin(kx)], \tag{2.25}$$

$$v(x) = -\frac{1}{j\rho\omega}\left(\frac{1}{x^2}[(A+B)\cos(kx) - j(A-B)\sin(kx)] + \frac{k}{x}[(A+B)\sin(kx) + j(A-B)\cos(kx)]\right). \tag{2.26}$$

The impedance at the mouth x_1, area $S_1 = S(x_1)$ is

$$z_1 = \frac{1}{S_1}\frac{\rho\omega[(A-B)\sin(kx_1)) + j(A+B)\cos(kx_1)]}{\frac{1}{x_1}[(A+B)\cos(kx_1) - j(A-B)\sin(kx_1)] + k[(A+B)\sin(kx_1) + j(A-B)\cos(kx_1)]}, \tag{2.27}$$

It is shown in appendix A that this gives the input impedance

$$z_0 = \frac{\left(\frac{x_1}{x_0}\cos(kd) - \frac{1}{kx_0}\sin(kd)\right)z_1 + jz_c\frac{x_0}{x_1}\sin(kd)}{jz_c^{-1}\left[\left(\frac{1}{k^2x_0^2} + \frac{x_1}{x_0}\right)\sin(kd) - \left(\frac{d}{kx_0^2}\right)\cos(kd)\right]z_1 + \frac{x_0}{x_1}\left(\frac{1}{kx_0}\sin(kd) + \cos(kd)\right)}. \tag{2.28}$$

This is the lossless input impedance of a conical horn in terms of its output impedance. This expression is identical with Table 2, Equation 3 in [1], and indeed with Equation 8.51 on page 191 of [2]. Applying the lossy model (2.17) under the approximation that $S = \pi(\mathcal{R}_0 + \mathcal{R}_1)^2/4$ gives

$$z_0 = \frac{\left(\frac{x_1}{x_0}\cosh(k^*d) - \frac{1}{k^*x_0}\sinh(k^*d)\right)z_1 + z_c^*\frac{x_0}{x_1}\sinh(k^*d)}{jz_c^{*-1}\left[\left(\frac{x_1}{x_0} - \frac{1}{k^{*2}x_0^2}\right)\sinh(k^*d) + \left(\frac{d}{k^*x_0^2}\right)\cosh(k^*d)\right]z_1 + \frac{x_0}{x_1}\left(\frac{1}{k^*x_0}\sinh(k^*d) + \cosh(k^*d)\right)}, \tag{2.29}$$

which is the lossy input impedance of a conical waveguide (cf. [1] Table 2, Equation 4).

2.4 Transmission Matrices

We have a formula for expressing the input impedance of cylindrical and conical waveguides in terms of their output impedance. Defining the acoustic volume velocity $U_0 = S_0 v_0$ gives input impedance $z_0 = p_0/U_0$, and we express (2.16), (2.19), (2.28) or (2.29) as

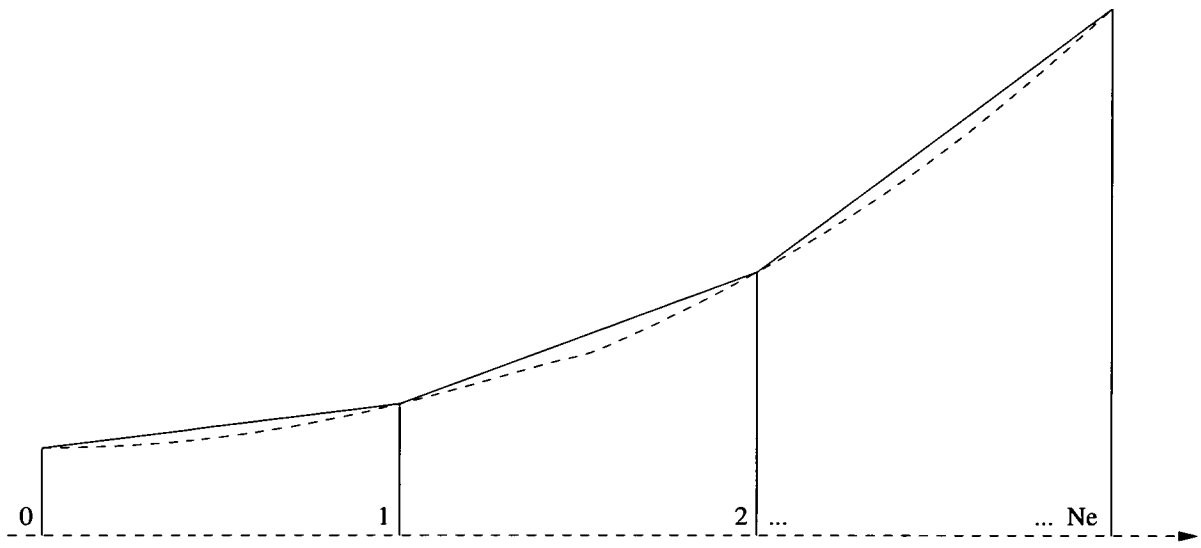


Figure 2.4: Discretisation of bore with conical sections.

$$z_0 = \frac{a_0 z_1 + b_0}{c_0 z_1 + d_0}, \tag{2.30}$$

or as

$$\frac{p_0}{v_0} = \frac{a_0 p_1 + b_0 v_1}{c_0 p_1 + d_0 v_1}. \tag{2.31}$$

Then we put the coefficients a, b, c, d into a matrix

$$T_0 = \begin{pmatrix} a_0 & b_0 \\ c_0 & d_0 \end{pmatrix}, \tag{2.32}$$

from which we deduce

$$\begin{pmatrix} p_0 \\ U_0 \end{pmatrix} = T_0 \begin{pmatrix} p_1 \\ U_1 \end{pmatrix}, \tag{2.33}$$

where all terms are frequency-dependent; we term T_0 a *transmission matrix.* Now consider the case where a second waveguide is attached to the first, such that the input of waveguide 2 is the output of waveguide 1. Given the output of section 2, we can solve for the input of waveguide 1:

$$\begin{pmatrix} p_0 \\ U_0 \end{pmatrix} = T_0 T_1 \begin{pmatrix} p_2 \\ U_2 \end{pmatrix}. \tag{2.34}$$

If N_e waveguides are placed end-on-end in a similar fashion, we have

$$\begin{pmatrix} p_0 \\ U_0 \end{pmatrix} = \prod_{i=0}^{N_e-1} T_i \cdot \begin{pmatrix} p_{N_e} \\ U_{N_e} \end{pmatrix}. \tag{2.35}$$

Then, if the radiation impedance z_r at the output end is known, we have

$$\begin{pmatrix} p_0/U_{N_e} \\ U_0/U_{N_e} \end{pmatrix} = \prod_{i=0}^{N_e-1} T_i \cdot \begin{pmatrix} z_r \\ 1 \end{pmatrix}. \tag{2.36}$$

Using this method, we can approximate a horn of general shape as a discretised series of truncated cylinders and cones placed end-on-end. Each section of the horn has an associated transmission matrix, which is dependent on its dimensions. The matrix product of these transmission matrices is then used, along with a known radiation impedance, to find the pressure and volume velocity at the input of the horn, from which the input impedance follows. This calculation must be performed once for each frequency, as the transmission matrices and radiation impedance are, in general, frequency-dependent. Noreland [21] gives a proof that, as $N_e \to \infty$, the approximate transmission matrix T of the piecewise-conical horn converges to the exact transmission matrix found from solving the Webster horn equation directly.

2.5 Bessel Horns

We are now able to derive transmission matrices for Bessel Horns (see Appendix B). Consider a waveguide

$$\mathcal{R}(x) = b(-x)^{-\gamma}, \tag{2.37}$$

$$S(x) = \pi b^2 (-x)^{-2\gamma}, \tag{2.38}$$

i.e. a Bessel-horn lying to the left of the origin with flare increasing from left-to-right. This definition is valid in the domain $(-\infty, 0)$, but in practice will only be needed in a subset of this domain. From (2.9),

$$x^2\frac{\partial^2 p}{\partial x^2} - 2\gamma x\frac{\partial p}{\partial x} + k^2x^2p = 0, \tag{2.39}$$

which, by (C.2) has solution

$$p(x) = x^{\gamma+1/2}\left[AJ_{\gamma+1/2}(kx) + BY_{\gamma+1/2}(kx)\right], \tag{2.40}$$

giving

$$v(x) = -\frac{k}{j\rho\omega}x^{\gamma+1/2}[AJ_{\gamma-1/2}(kx) + Bx^{\gamma+1/2}Y_{\gamma-1/2}(kx)]. \tag{2.41}$$

This solution is general for all values of the order $\gamma + 1/2$. Then, following the same recipe as the conical case above, we find:

$$z_0 = \frac{az_1 + b}{cz_1 + d}, \tag{2.42}$$

where

$$\begin{aligned}
a &= [J_{\gamma+1/2}(kx_0)Y_{\gamma-1/2}(kx_1) - Y_{\gamma+1/2}(kx_0)J_{\gamma-1/2}(kx_1)] \\
b &= jz_c\left(\frac{x_0}{x_1}\right)^{-2\gamma}[J_{\gamma+1/2}(kx_0)Y_{\gamma+1/2}(kx_1) - Y_{\gamma+1/2}(kx_0)J_{\gamma+1/2}(kx_1)] \\
c &= -\frac{j}{z_c}[Y_{\gamma-1/2}(kx_0)J_{\gamma-1/2}(kx_1) - J_{\gamma-1/2}(kx_0)Y_{\gamma-1/2}(kx_1)] \\
d &= \left(\frac{x_0}{x_1}\right)^{-2\gamma}[Y_{\gamma-1/2}(kx_0)J_{\gamma+1/2}(kx_1) - J_{\gamma-1/2}(kx_0)Y_{\gamma+1/2}(kx_1)],
\end{aligned} \tag{2.43}$$

which is an expression for the transmission matrix of a Bessel-horn. This line of reasoning is mentioned briefly in [2] but is not taken to this conclusion.

Validation: Conical Bessel-horn

To check the above transmission-matrix, we substitute $\gamma = -1$, to give a conical horn profile. Using the results (C.4) we simplify (2.43) term-by-term

$$a = \frac{2}{k\pi\sqrt{x_0x_1}}k\left[\cos(kd) - \frac{1}{kx_1}\sin(kd)\right]$$

$$
\begin{aligned}
b &= \frac{2}{k\pi\sqrt{x_0x_1}}\left(\frac{x_0}{x_1}\right)^2 jkz_c\sin(kd) \\
c &= \frac{2}{k\pi\sqrt{x_0x_1}}\frac{-jk}{z_c}\left[\left(\frac{1}{kx_0}-\frac{1}{kx_1}\right)\cos(kd)-\left(\frac{1}{k^2x_0x_1}+1\right)\sin(kd)\right] \\
d &= \frac{2}{k\pi\sqrt{x_0x_1}}\left(\frac{x_0}{x_1}\right)^2 k\left[\cos(kd)+\frac{1}{kx_0}\sin(kd)\right].
\end{aligned} \tag{2.44}
$$

If we substitute these terms back into (2.42), cancel over $2/\pi\sqrt{x_0x_1}$ and multiply by x_1/x_0 we get (2.28), thus showing that the conical transmission-matrix is indeed a special case of the Bessel-horn matrix.

2.6 Exponential Horns

As with Bessel horns, we can derive transmission matrices for exponential horns (see Appendix B). Consider a waveguide of the following shape:

$$
\begin{aligned}
S(x) &= S_0e^{\gamma x}, \\
\mathcal{R}(x) &= \sqrt{\frac{S_0}{\pi}}e^{\gamma x/2},
\end{aligned} \tag{2.45}
$$

i.e. an exponential horn with flare constant γ. From (2.9),

$$
\frac{\partial^2 p}{\partial x^2}+\gamma\frac{\partial p}{\partial x}+k^2p=0, \tag{2.46}
$$

which is a linear, second-order equation with constant coefficients and solution dependent on the relative values of γ and k^2. Note that this is equivalent to the governing equation of damped simple harmonic motion. We are considering only a propagating mode, so we may safely assume $k^2 > \frac{1}{4}\gamma^2$ (analagous to under-damping in SHM). Then

$$
p(x) = Ae^{\lambda_1 x}+Be^{\lambda_2 x}, \tag{2.47}
$$

where λ_1 and λ_2 are roots of the auxiliary equation

$$
\lambda^2+\gamma\lambda+k^2=0, \tag{2.48}
$$

giving

$$\lambda_1, \lambda_2 = -\frac{1}{2}\gamma \pm j\sqrt{k^2 - \frac{\gamma^2}{4}} \tag{2.49}$$

from which

$$\begin{aligned} p(x) &= Ae^{-\frac{1}{2}\gamma x}e^{-j\alpha x} + Be^{-\frac{1}{2}\gamma x}e^{j\alpha x}, \\ &= e^{-\frac{1}{2}\gamma x}[(A+B)\cos(\alpha x) - j(A-B)\sin\alpha x], \end{aligned} \tag{2.50}$$

$$\begin{aligned} v(x) &= \frac{k}{j\rho\omega}e^{-\frac{1}{2}\gamma x}\left(\frac{\gamma}{2}[(A+B)\cos(\alpha x) - j(A-B)\sin(\alpha x)]\right. \\ &\left. +\alpha[(A+B)\sin(\alpha x) + j(A-B)\cos(\alpha x)]\right), \end{aligned} \tag{2.51}$$

where $\alpha = \sqrt{k^2 - \frac{\gamma^2}{4}}$. Proceeding as above, we find the mouth impedance

$$z_1 = \frac{1}{S_1}\frac{\rho\omega[(A-B)\sin(\alpha x_1)) + j(A+B)\cos(\alpha x_1)]}{\frac{\gamma}{2}[(A+B)\cos(\alpha x_1) - j(A-B)\sin(\alpha x_1)] + \alpha[(A+B)\sin(\alpha x_1) + j(A-B)\cos(\alpha x_1)]}. \tag{2.52}$$

This is very similar to the equivalent expression (2.27) for a cone, saving that α replaces k, and $\gamma/2$ replaces $1/x_0$ and $1/x_1$. We can therefore follow the same derivation, noting that

$$\frac{1}{S_1} = \frac{1}{S_0}e^{-\gamma d} \tag{2.53}$$

in this case, to find the input impedance

$$z_0 = \frac{k\left(\alpha\cos(\alpha d) - \frac{\gamma}{2}\sin(\alpha d)\right)z_1 + jz_c k^2 e^{-\gamma d}\sin(\alpha d)}{jz_c^{-1}\alpha^2\left[\left(\frac{\gamma^2}{4\alpha^2}+1\right)\sin(\alpha d) - \left(\frac{d\gamma^2}{4\alpha^2}\right)\cos(\alpha d)\right]z_1 + ke^{-\gamma d}\left(\frac{\gamma}{2}\sin(\alpha d) + \alpha\cos(\alpha d)\right)}. \tag{2.54}$$

Setting $x_0 = 0$ and $x_1 = d$ as in Appendix B gives

$$z_0 = \frac{k\left(\alpha\cos(\alpha d) - \frac{\gamma}{2}\sin(\alpha d)\right)z_1 + jz_c k^2 e^{-\gamma d}\sin(\alpha d)}{jz_c^{-1}\alpha\left[\frac{\gamma}{2}\cos(\alpha d) + \alpha\sin(\alpha d)\right]z_1 + ke^{-\gamma d}\alpha\cos(\alpha d)}, \tag{2.55}$$

or, eliminating explicit use of γ,

$$z_0 = \frac{k\left(\alpha\cos(\alpha d) - \frac{1}{d}\ln\left(\frac{\mathcal{R}_1}{\mathcal{R}_0}\right)\sin(\alpha d)\right)z_1 + jz_c k^2\left(\frac{\mathcal{R}_0}{\mathcal{R}_1}\right)^2\sin(\alpha d)}{jz_c^{-1}\alpha\left[\frac{1}{d}\ln\left(\frac{\mathcal{R}_1}{\mathcal{R}_0}\right)\cos(\alpha d) + \alpha\sin(\alpha d)\right]z_1 + k\left(\frac{\mathcal{R}_0}{\mathcal{R}_1}\right)^2\alpha\cos(\alpha d)}. \tag{2.56}$$

This can be represented as a transmission matrix in the familiar manner.

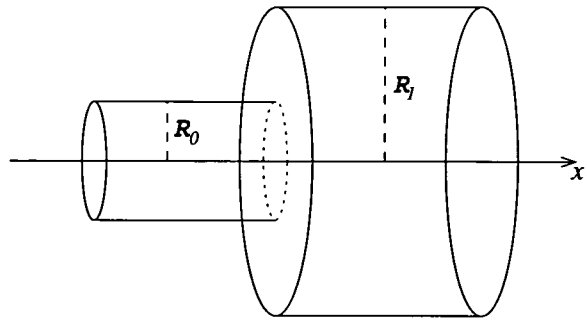


Figure 2.5: Schematic of bore discontinuity between two concentric cylinders of radius $\mathcal{R}_0$ and $\mathcal{R}_1$.

Validation: Cylindrical Exponential Horn

Setting $\gamma = 0$ gives $S(x) = S_0$, i.e. a cylinder, and $\alpha = k$. Substituting these values into (2.55), we find an expression identical to the transmission-line of a cylinder (2.16) of length d.

2.7 Bore Discontinuities

Consider junction of two ducts with radii $\mathcal{R}_0$ and $\mathcal{R}_1$. They meet such that they share a common central axis, with an abrupt jump in radius at the junction. Figure 2.5 shows this for two cylinders, but this argument holds for ducts of arbitrary geometries provided they are symmetric about the central axis. We seek how this jump will affect the propagation of a plane wave.

In order to preserve continuity, and recalling the the plane-wave assumption, the pressure to one side of the jump must equal that on the other $p^{(0)} = p^{(1)}$. Similarly, so must the volume velocity $U^{(0)} = U^{(1)}$, though it follows that the particle velocity will change such that

$$S^{(0)}v^{(0)} = S^{(1)}v^{(1)}. \tag{2.57}$$

It also follows that the input impedance is equal on either side

$$z^{(0)} = z^{(1)} \tag{2.58}$$

and that the bore discontinuity can be discounted from plane-wave impedance calculations. Note that the characteristic impedance $z_c = \rho c/S$ will differ on either side of the discontinuity,

and therefore z/z_c will likewise differ.

2.8 Spherical Waves in Cones

Until now, the wave propagation has been assumed to be planar. Flaring horns, in practice, exhibit waves with curved wavefronts [2]. We can approximate this behaviour by assuming instead that the wavefronts are spherical. Consider a point source in free space. The three-dimensional wave-equation is

$$\frac{\partial^2 p}{\partial t^2} = c^2 \nabla^2 p. \tag{2.59}$$

Again assuming sinusoidal time-dependence, we have

$$\nabla^2 p + k^2 p = 0. \tag{2.60}$$

Assuming rotational symmetry – i.e. spherical waves – gives the solution in terms of radial component r

$$\nabla^2 p = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 \frac{\partial p}{\partial r}\right), \tag{2.61}$$

from which

$$\frac{\partial^2 p}{\partial r^2} + \frac{2}{r}\frac{\partial p}{\partial r} + k^2 p = 0. \tag{2.62}$$

For a point-source radiating spherically, a rigid conical waveguide with walls perpendicular to the wave-fronts (i.e. lying on radial axes) will not affect the propagation. The above equation can therefore be considered as the Webster equation for a conical horn with spherical waves; although the equation is the same as that for the plane-wave case (2.21), the base assumptions, and therefore geometry, are different. Radial coordinate r is used in place of axial coordinate x, and the wavefront curves into and bulges out of a flaring cone at the throat and mouth respectively. For a spherical wavefront at given radial position r, the equivalent planar wavefront is located at axial coordinate x such that the two fronts would intersect at the walls of the tube (figure 2.6). Along the central axis of the cone, the distance between the spherical and planar wavefronts is

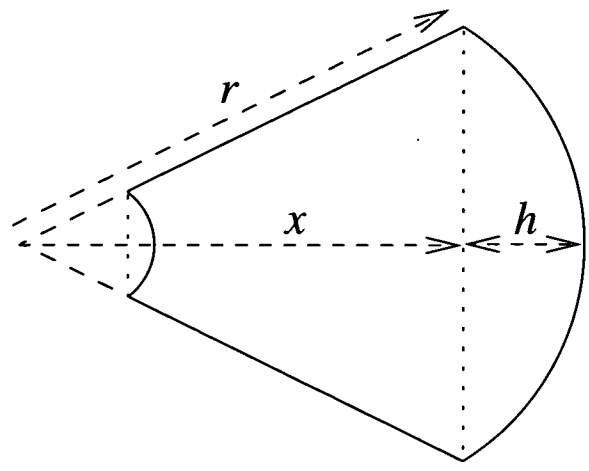


Figure 2.6: Spherical (solid) and plane (dotted) wavefronts in a conical duct.

$$h = r - \sqrt{r^2 - \mathcal{R}(x)^2}, \tag{2.63}$$

where $\mathcal{R}(x)$ is the bore radius at the mouth of the cone. Secondly, the surface area of the wavefront is that of a spherical cap of radius r, height h

$$S_i = 2\pi r h. \tag{2.64}$$

These two corrections are sufficient to modify the plane-wave derivation to describe the system with spherical waves. It should be noted that when two conical sections of different flare are joined together, the different locations of their respective fictitious apices results in a mismatch of the entry and exit wavefronts, with a corresponding 'missing volume' in between, thus introducing a systematic inaccuracy into the model (even within the bounds of its rather strong assumptions). This can be minimised by ensuring that the difference in flare between any two adjacent cones is small. A further problem occurs when a cone (with spherical waves) is modelled with a cylinder (plane); the bulging section of the wavefront protrudes into the cylinder, thus resulting in a 'doubly-defined' volume. This inaccuracy is avoidable only by the introduction of 'scattering regions' [23], which smoothly join the wavefronts; consideration of this is beyond the scope of this work, and the effect is small.

2.9 Radiation Impedance

The radiation at the open end of an instrument is here modelled [16] as a piston in an unflanged cylinder of radius $\mathcal{R}$

$$z_r = 0.5\frac{\omega^2\rho}{\pi c} + 0.6133j\frac{\omega\rho}{\pi\mathcal{R}}. \tag{2.65}$$

Chapter 3

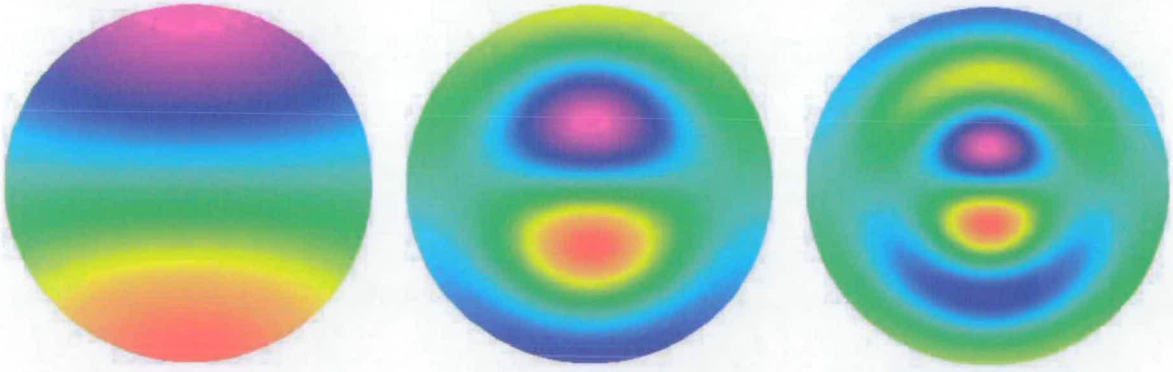
Multi-modal Propagation in Straight and Bent Acoustic Horns

3.1 Introduction

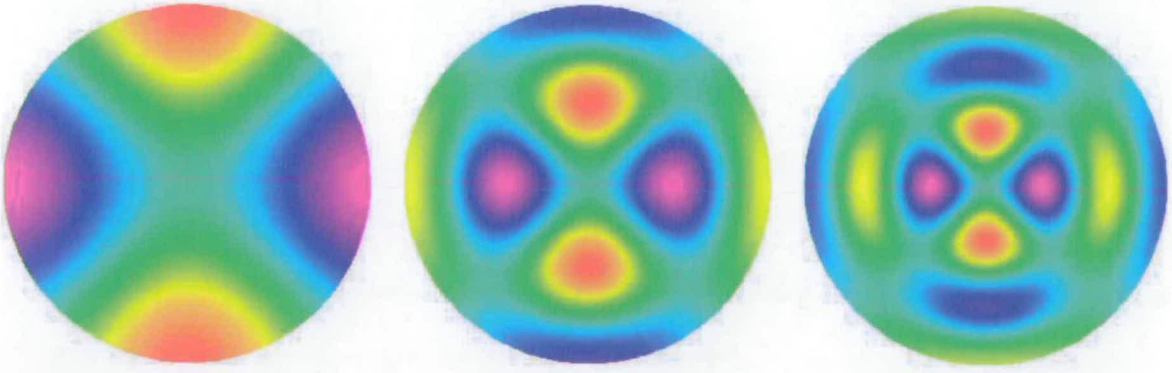
For a given excitation frequency, an acoustic duct has an infinite number of possible oscillatory pressure patterns, or *modes*. Assuming a circular cross-section, these modes can be described by a set of indices (m, n), representing a mode with m nodal diameters and n nodal circles, where the plane-wave is mode $(0, 0)$. Some of these modes are shown in a cross-section of the tube in figure 3.1. Each mode has an associated cut-off frequency below which it cannot propagate and is exponentially damped, or *evanescent*. In a general situation there will be a finite number of propagating modes and an infinite number of evanescent modes being excited; how many modes propagate depends on the excitation frequency in comparison to the cut-off frequencies of the various modes as determined by the geometry of the duct. These modes form an orthogonal basis, so any given pressure and velocity fields in the duct can therefore be decomposed into a sum of contributions from each of these modes.

Brass instruments generally feature tubing which is cylindrical in cross-section, is significantly longer that it is wide, and is excited generally at low frequencies (we will neglect non-linear effects in loud-playing). Consequently, only the plane-wave, which always has a cut-off frequency of zero, will propagate. However, when the radius of the instrument changes, energy is transferred between the modes through *mode conversion*. Energy is therefore lost

(a) (0,0) mode (b) (0,1) mode (c) (0,2) mode



(d) (1,0) mode (e) (1,1) mode (f) (1,2) mode



(g) (2,0) mode (h) (2,1) mode (i) (2,2) mode

Figure 3.1: Transverse pressure patterns ψ for (m, n) modes with m nodal diameters and n nodal circles.

from the propagating plane-wave and transferred into evanescent higher modes, thus affecting the overall acoustic properties of the instrument. Brass instruments feature rapidly-flaring bell sections in which a significant amount of mode conversion takes place; this process likewise occurs at the radiating end. Any accurate treatment of such instruments must take account of this effect.

We saw in Chapter 2 how a duct (or horn) of general shape can be discretised into a series of short sections for numerical calculation. With the multimodal model we use a piecewise-constant discretisation, approximating a general horn as a series of short cylindrical sections joined by small concentric bore discontinuities. Consideration is given to bent ducts, modelled as toroidal bends of constant curvature and tube radius. A series of short such bends joined with bore discontinuities can likewise be used to discretise more general bends. Derivation of a radiation impedance model is also necessary to provide a boundary condition at the open end.

Kemp [5] gives a thorough review of the modal decomposition technique set out by Pagneux, Amir & Kergomard [3, 4]. This technique is again reviewed here, presented in an alternative derivation, to give a starting point for a review of the corresponding theory for bent ducts, as presented by Félix and Pagneux [8, 9, 10, 11], which is then adapted and extended for our purposes.

3.2 Cylinders

3.2.1 Matricial Derivation

Consider a cylindrical tube of length d, radius $\mathcal{R}_0$ and cross-sectional area S_0. The walls are rigid, the tube contains air of density ρ and speed of sound c, and the angular excitation frequency is ω. We define $p(\mathbf{x})$ as the acoustic pressure (scalar field) and $\mathbf{v}(\mathbf{x})$ as the particle velocity (vector field). From the equations of mass and momentum conservation, we have

$$\nabla.\mathbf{v} = \frac{-j\omega}{\rho c^2} p, \tag{3.1}$$

$$j\omega\mathbf{v} = \frac{-1}{\rho}\nabla p. \tag{3.2}$$

Applying the gradient operator in cylindrical coordinates $r \in [0, \mathcal{R}_0]; \phi \in [0, 2\pi]$ and $x \in [0, d]$

$$\nabla = \left(\frac{\partial}{\partial r}, \frac{1}{r}\frac{\partial}{\partial \phi}, \frac{\partial}{\partial x} \right) \tag{3.3}$$

and eliminating radial and angular components v_r and v_ϕ gives

$$\frac{\partial v_x}{\partial x} = \frac{1}{\rho c}\frac{1}{jk}\left(k^2 p + \frac{1}{r}\frac{\partial p}{\partial r} + \frac{\partial^2 p}{\partial r^2} + \frac{1}{r^2}\frac{\partial^2 p}{\partial \phi^2} \right), \tag{3.4}$$

$$\frac{\partial p}{\partial x} = -jk\rho c v_x, \tag{3.5}$$

where $k = \omega/c$ is the wavenumber. We introduce the separation of variables

$$p(r, \phi, x, t) = \sum_{\alpha=0}^{\infty} P_\alpha(x)\psi_\alpha(r, \phi)e^{j\omega t}, \tag{3.6}$$

$$v_x(r, \phi, x, t) = \frac{1}{S_0}\sum_{\alpha=0}^{\infty} U_\alpha(x)\psi_\alpha(r, \phi)e^{j\omega t}, \tag{3.7}$$

where P_α and U_α are the pressure and volume velocity amplitudes in the axial direction, and α is an index we will define below. Unlike Kemp we do not assume an axisymmetric excitation and therefore must include nodal-diameter modes; for simplicity we do, however, restrict our solution to those cases with a single plane of symmetry along the central axis (i.e. a duct which is bent only in one axis). From the transverse parts of (3.5), ψ_α are therefore the eigenfunctions obeying the transverse eigenproblem

$$\frac{\partial^2 \psi_\alpha}{\partial r^2} + \frac{1}{r}\frac{\partial \psi_\alpha}{\partial r} + \frac{1}{r^2}\frac{\partial^2 \psi_\alpha}{\partial \phi^2} = -k_{\perp\alpha}^2 \psi_\alpha, \tag{3.8}$$

$\forall \alpha \in [0, \infty)$, where $k_{\perp\alpha}$ is the transverse wavenumber of mode α, with the homogeneous Neumann boundary condition on the (rigid) walls

$$\left.\frac{\partial \psi_\alpha}{\partial r}\right|_{r=\mathcal{R}_0} = 0. \tag{3.9}$$

The orthogonality relation is

$$\int_0^{\mathcal{R}_0}\int_0^{2\pi} \psi_\alpha \psi_\beta \, r d\phi dr = \delta_{\alpha\beta}, \tag{3.10}$$

or, in inner product notation,

$$\langle \psi_\alpha, \psi_\beta \rangle = \delta_{\alpha\beta}. \tag{3.11}$$

Solving for ψ gives

$$\psi_\alpha(r,\phi) = A_\alpha J_m\left(\frac{\gamma_{mn} r}{\mathcal{R}_0}\right)\sin\left(m\phi + \frac{\sigma\pi}{2}\right), \tag{3.12}$$

$$A_\alpha = \begin{cases} \dfrac{1}{\sqrt{S_0 J_0^2(\gamma_{0n})}} & \text{if } m = 0, \\ \dfrac{1}{\sqrt{(S_0/2)(1 - m^2/\gamma_{mn}^2) J_m^2(\gamma_{0n})}} & \text{if } m > 0, \end{cases} \tag{3.13}$$

where J_m is the Bessel-function of the first kind of order m, γ_{mn} is the $(n+1)^{\text{th}}$ zero of J'_m, and each value of α is defined to correspond to one set of values of (m, n, σ), ordered by their cut-off transverse wavenumbers $k_{\perp\alpha} = \gamma_{mn}/\mathcal{R}_0$, giving ψ_0 as the plane-wave (see table 3.1). m is the number of nodal diameters (Kemp sets this to zero), n the number of nodal circles. The index β corresponds to the modal indices (μ, ν, ς).

The third index σ, ς is a symmetry index with value 0 or 1. It has the effect of rotating the transverse pressure pattern through a right-angle about its central axis, thus describing two indepedent sets of modes (when $m > 0$). We have chosen to restrict consideration to those special cases with a single plane of symmetry along the central axis; in effect, we are choosing a value of σ and holding it constant throughout. This assumption is, in effect, roughly similar to Kemp's assumption of $m = 0$ throughout. We may therefore drop use of σ as an index for brevity, but will maintain treatment of its (constant) value.

Physically, we are assuming that the bends all take place in the same plane; at any given cross-section of the duct there will be a line of symmetry normal to that plane, and we can choose, without loss of generality, the coordinate system such that this plane corresponds to either $\phi = 0$ or $\phi = \pi/2$. By so doing, any given pressure pattern can be represented using linear combinations of the subset of ψ having a constant value σ. Were the coordinates not so aligned, or the system had bends other than in that plane, then linear combinations of ψ with both values of σ would be required to describe a general pressure profile. This physical assumption does not significantly restrict the applicability of the analysis for the purposes of this thesis, and reduces the number of indices required.

Having given due consideration to σ, we return to the other indices. By lumping m, n

α	m	n	γ_{mn}
0	0	0	0.00
1	1	0	1.84
2	2	0	3.05
3	0	1	3.83
4	3	0	4.20
5	4	0	5.32
6	1	1	5.33
7	5	0	6.42
8	2	1	6.71

Table 3.1: Correlation between values of α and m, n, ordered by γ_{mn}.

into α, and μ, ν into β, and by holding σ constant, we are able to produce the following derivation using matrices; without such lumping, tensor analysis would be required (this is left for future work).

Having established the transverse behaviour, we can define infinite column vectors $\mathbf{P} = (P_\alpha)_{\alpha \geq 0}$, $\mathbf{U} = (U_\alpha)_{\alpha \geq 0}$ and $\boldsymbol{\psi} = (\psi_\alpha)_{\alpha \geq 0}$, omitting terms of $e^{j\omega t}$ for clarity, such that

$$p = \boldsymbol{\psi}^{\boldsymbol{T}} \boldsymbol{P}, \tag{3.14}$$

$$v_x = \frac{1}{S_0} \boldsymbol{\psi}^{\boldsymbol{T}} \boldsymbol{U}. \tag{3.15}$$

We can now project equations (3.4) & (3.5) over the orthogonal function basis $\boldsymbol{\psi}$ and solve for $\mathbf{P}$ and $\mathbf{U}$. Starting with the projection of (3.5),

$$\frac{1}{\rho c} \left\langle \frac{\partial p}{\partial x}, \psi_\alpha \right\rangle = -jk \left\langle v_x, \psi_\alpha \right\rangle . \tag{3.16}$$

Substituting the definitions (3.6) & (3.7) and rearranging, noting that $P_\alpha(x)$ is a function of x only, gives

$$\frac{1}{\rho c} \sum_\beta \frac{\partial P_\beta}{\partial x} \left\langle \psi_\beta, \psi_\alpha \right\rangle = -\frac{jk}{S_0} \sum_\beta U_\beta \left\langle \psi_\beta, \psi_\alpha \right\rangle . \tag{3.17}$$

Using the orthogonality relation (3.10) gives

$$\frac{1}{\rho c} P'_\alpha = -\frac{jk}{S_0} U_\alpha, \tag{3.18}$$

from which

$$\mathbf{P}' = -jkz_c\mathbf{U}, \tag{3.19}$$

where $z_c = \rho c/S_0$ is the characteristic impedance. Proceeding in a similar fashion, we project (3.4) term-by-term. The left-hand-side gives

$$\left\langle \frac{\partial v_x}{\partial x}, \psi_\alpha \right\rangle = \frac{1}{S_0} U'_\alpha, \tag{3.20}$$

while the first term of the right-hand-side gives

$$\frac{1}{\rho c}\frac{1}{jk} \left\langle k^2 p, \psi_\alpha \right\rangle = \frac{1}{\rho c}\frac{1}{jk} k^2 P_\alpha, \tag{3.21}$$

and the remaining terms give, recalling equation (3.8),

$$\frac{1}{\rho c}\frac{1}{jk} \left\langle \left(\frac{1}{r}\frac{\partial p}{\partial r} + \frac{\partial^2 p}{\partial r^2} + \frac{1}{r^2}\frac{\partial^2}{\partial \phi^2} \right), \psi_\alpha \right\rangle = \frac{1}{\rho c}\frac{1}{jk} - k_{\perp\alpha}^2 P_\alpha. \tag{3.22}$$

Recombining gives

$$\mathbf{U}' = \frac{1}{jkz_c} K\mathbf{P}, \tag{3.23}$$

where K is a square matrix defined in terms of axial wavenumber $K_{\alpha\beta} = k_{\parallel\alpha}^2 \delta_{\alpha\beta}$, where

$$k^2 = k_{\parallel\alpha}^2 + k_{\perp\alpha}^2 \tag{3.24}$$

We define the *impedance matrix* Z such that $\mathbf{P} = Z\mathbf{U}$, and use $\mathbf{P}' = Z'\mathbf{U} + Z\mathbf{U}'$ along with (3.19) and (3.23) to find:

$$Z' = -jkz_c I - \frac{1}{jkz_c} ZKZ, \tag{3.25}$$

which is the matrix Ricatti equation given in [3], but is, in this case, only of passing interest. We use (3.19) and (3.23) to give

$$\mathbf{P}'' + K\mathbf{P} = \mathbf{0}, \tag{3.26}$$

which has solution

$$\mathbf{P} = D(x)\mathbf{\Xi_1} + D^{-1}(x)\mathbf{\Xi_2}, \tag{3.27}$$

where $D_{\alpha\beta}(x) = e^{jk_{\|\alpha}x}\delta_{\alpha\beta}$ is a spatially-dependent diagonal matrix and $\mathbf{\Xi_1}, \mathbf{\Xi_2}$ are constant column-vectors, determined from the initial conditions, which will subsequently be eliminated. Defining diagonal matrix Λ with elements $1/jk_{\|\alpha}$, differentiating (3.27) and substituting (3.19) gives:

$$Z^{-1}\mathbf{P} = -\frac{1}{jkz_c}\Lambda^{-1}(D(x)\mathbf{\Xi_1} - D^{-1}(x)\mathbf{\Xi_2}). \tag{3.28}$$

Define

$$H = \frac{1}{jkz_c}K\Lambda, \tag{3.29}$$

which, noting that $K = -\Lambda^{-2}$, gives

$$H = -\frac{1}{jkz_c}\Lambda^{-1}, \tag{3.30}$$

from which

$$Z^{-1}\mathbf{P} = H(D(x)\mathbf{\Xi_1} - D^{-1}(x)\mathbf{\Xi_2}). \tag{3.31}$$

We equate (3.27) and (3.31) over $\mathbf{P}$ and evaluate at the input ($x = 0$)

$$(\mathbf{\Xi_1} + \mathbf{\Xi_2}) = Z_0H(\mathbf{\Xi_1} - \mathbf{\Xi_2}) \tag{3.32}$$

and output ($x = d$) ends

$$(\mathcal{D}\mathbf{\Xi_1} + \mathcal{D}^{-1}\mathbf{\Xi_2}) = Z_dH(\mathcal{D}\mathbf{\Xi_1} - \mathcal{D}^{-1}\mathbf{\Xi_2}), \tag{3.33}$$

where $\mathcal{D} = D(d)$ and Z_0, Z_d the impedances at the input (unknown) and output (known) ends. We can eliminate over $\mathbf{\Xi_2}$ and $\mathbf{\Xi_1}$ to get

$$(I - Z_0H) = (I + Z_0H)\mathcal{D}(I + Z_dH)^{-1}(I - Z_dH)\mathcal{D}, \tag{3.34}$$

where I is the identity matrix. Define

$$E = \mathcal{D}(I + Z_d H)^{-1}(I - Z_d H)\mathcal{D} \tag{3.35}$$

to give solution

$$Z_0 = (I - E)(I + E)^{-1}H^{-1}, \tag{3.36}$$

which is an expression for the input impedance in terms of the output impedance and geometry of the duct.

3.2.2 Validation and Discussion

The matrices derived are infinite to account for the infinite basis of orthogonal modes. In practice, we must truncate these matrices to finite size N_m; the more modes included, the more accurate the result, and the higher computation cost (Chapter 8). By considering the scalar case where only one mode is included – i.e. the plane wave – we can derive a familiar result. In this case

$$E = e^{-2jkd}\frac{1 - Z_d H}{1 + Z_d H}, \tag{3.37}$$

and $H = 1/z_c$; noting that

$$\begin{aligned} \mathcal{D}^{-1} + \mathcal{D} &= e^{-jkd} + e^{jkd} \\ &= 2\cos(kd), \end{aligned} \tag{3.38}$$

$$\begin{aligned} \mathcal{D}^{-1} - \mathcal{D} &= e^{-jkd} - e^{jkd} \\ &= -2j\sin(kd), \end{aligned} \tag{3.39}$$

we evaluate the input impedance

$$Z_0 = \frac{Z_d\cos(kd) + jz_c\sin(kd)}{jz_c^{-1}Z_d\sin(kd) + \cos(kd)}, \tag{3.40}$$

which we recognise as the plane-wave result (2.16).

The derivation given is that of [3], extended to give a solution of the Ricatti equation – the cited paper gives the Ricatti equation for a varying-radius duct and uses numerical methods to solve for Z_0; here we take the cylindrical case and solve analytically. Kemp [5] adopts a different approach, and while it produces a solution which is intuitively more satisfying, his recipe is more difficult to apply to more complicated cases.

Although the derivation given is for a cylinder, which is axisymmetric by definition, it does not assume that the oscillation is axisymmetric. Physically, this implies that either the overall duct system or the excitation is non-axisymmetric, or both. The simplest example of a non-axisymmetric duct is that of the toroidal bend, which is common in brass instruments; the model of any system which includes at least one such bend must account for the non-axisymmetric oscillation throughout the length of the horn, much of which may consist of axisymmetric ducts.

We now move to consider such a toroidal duct. Consideration of the above matricial derivation allows us to tackle the bent case having already dealt with many of the difficulties faced.

3.3 Toroidal Bends of Constant Radius

Consider a tube of constant circular cross-section S_0 with radius $\mathcal{R}_0$. The tube is bent toroidally, such that the central axis is a constant radius κ^{-1} from an imaginary origin, where κ is termed the curvature. The length along the axis is d. We proceed along similar lines to the derivation above. In the toroidal coordinate system $r \in [0, \mathcal{R}_0]$, $\phi \in [0, 2\pi]$ and $s \in [0, d]$ shown in figure 3.2, we have (using unit vectors $\mathbf{u}$)

$$d\mathbf{r} = dr\hat{\mathbf{u}}_\mathbf{r} + rd\phi\hat{\mathbf{u}}_\phi + (1 - \kappa r\cos\phi)ds\hat{\mathbf{u}}_\mathbf{s}, \tag{3.41}$$

from which we have the gradient and divergence operators, using dummy scalar and vector functions f and $\mathbf{F}$,

$$\begin{aligned} \nabla f &= \frac{\partial f}{\partial r}\hat{\mathbf{u}}_\mathbf{r} + \frac{1}{r}\frac{\partial f}{\partial \phi}\hat{\mathbf{u}}_\phi + \frac{1}{1-\kappa r\cos\phi}\frac{\partial f}{\partial s}\hat{\mathbf{u}}_\mathbf{s} \\ \nabla\cdot\mathbf{F} &= \frac{\partial F_r}{\partial r} + \left(\frac{1}{r} - \frac{\kappa\cos\phi}{1-\kappa r\cos\phi}\right)F_r + \frac{1}{r}\frac{\partial F_\phi}{\partial \phi} \end{aligned} \tag{3.42}$$

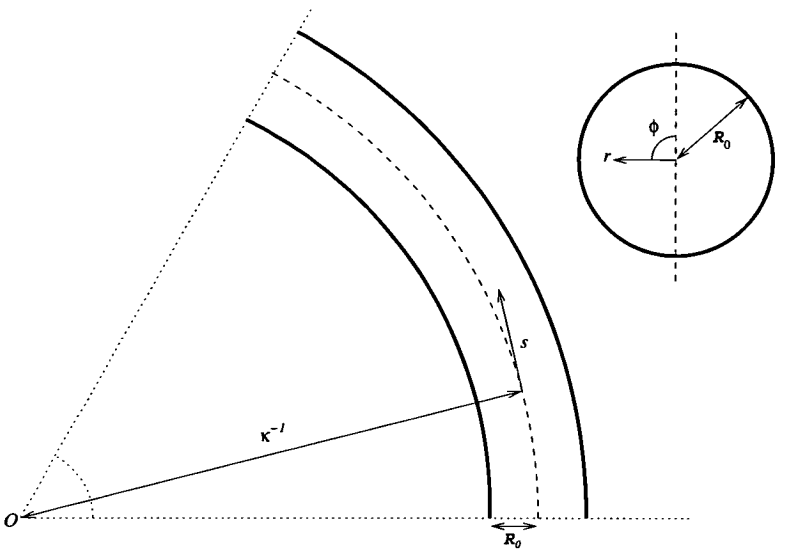


Figure 3.2: Schematic of toroidal bend, curvature κ, tube radius $\mathcal{R}_0$, length d along toroidal axis s.

$$+\frac{\kappa \sin\phi}{1-\kappa r\cos\phi}F_\phi + \frac{1}{1-\kappa r\cos\phi}\frac{\partial F_s}{\partial s}. \tag{3.43}$$

We apply this to (3.1) & (3.2), and eliminate components v_r and v_ϕ to give

$$\begin{aligned}\frac{\partial v_s}{\partial s} &= \frac{1}{\rho c}\frac{1}{jk}\left(k^2(1-\kappa r\cos\phi)p + (1-2\kappa r\cos\phi)\frac{1}{r}\frac{\partial p}{\partial r} + \right.\\ &\left. (1-\kappa r\cos\phi)\frac{\partial^2 p}{\partial r^2} + (1-\kappa r\cos\phi)\frac{1}{r^2}\frac{\partial^2 p}{\partial \phi^2} + \kappa\sin\phi\frac{1}{r}\frac{\partial p}{\partial \phi}\right), \qquad (3.44)\\ \frac{\partial p}{\partial s} &= -(1-\kappa r\cos\phi)jk\rho c v_s. \qquad (3.45)\end{aligned}$$

Projecting this over the same basis ψ (3.12) gives

$$\begin{aligned}\mathbf{U}' &= \frac{1}{jkz_c}(C+KB)\mathbf{P}, \qquad (3.46)\\ \mathbf{P}' &= -jkz_cB\mathbf{U}, \qquad (3.47)\\ Z' &= -jkz_cB - \frac{1}{jkz_c}Z(C+KB)Z, \qquad (3.48)\end{aligned}$$

where

$$K_{\alpha\beta} = \left(k^2 - \frac{\gamma_{mn}^2}{\mathcal{R}_0^2}\right)\delta_{\alpha\beta}, \tag{3.49}$$

$$B_{\alpha\beta} = \int_0^{\mathcal{R}_0}\int_0^{2\pi}(1-\kappa r\cos\phi)\psi_\alpha\psi_\beta\, r d\phi dr, \tag{3.50}$$

$$C_{\alpha\beta} = A_\alpha A_\beta \frac{\gamma_{mn}}{\mathcal{R}_0}\kappa M_{m\mu}\int_0^{\mathcal{R}_0} r J_{m+1}\left(\frac{\gamma_{mn}r}{\mathcal{R}_0}\right)J_\mu\left(\frac{\gamma_{\mu\nu}r}{\mathcal{R}_0}\right)dr$$
$$-A_\alpha A_\beta m\kappa(M_{m\mu}-L_{m\mu})\int_0^{\mathcal{R}_0} J_m\left(\frac{\gamma_{mn}r}{\mathcal{R}_0}\right)J_\mu\left(\frac{\gamma_{\mu\nu}r}{\mathcal{R}_0}\right)dr, \tag{3.51}$$

$$M_{m\mu} = \int_0^{2\pi}\cos\phi\sin\left(m\phi+\frac{\sigma\pi}{2}\right)\sin\left(\mu\phi+\frac{\sigma\pi}{2}\right), \tag{3.52}$$

$$L_{m\mu} = \int_0^{2\pi}\sin\phi\sin\left(m\phi+\frac{\sigma\pi}{2}\right)\sin\left(\mu\phi+\frac{\sigma\pi}{2}\right), \tag{3.53}$$

which can be simplified to

$$K_{\alpha\beta} = k_\parallel\delta_{\alpha\beta} \tag{3.54}$$

$$B_{\alpha\beta} = \delta_{\alpha\beta} - A_\alpha A_\beta\kappa M_{m\mu}\int_0^{\mathcal{R}_0} r^2 J_m\left(\frac{\gamma_{mn}r}{\mathcal{R}_0}\right)J_\mu\left(\frac{\gamma_{\mu\nu}r}{\mathcal{R}_0}\right)dr, \tag{3.55}$$

$$M_{m\mu} = \frac{\pi}{2}\delta_{|m-\mu|,1}(1+(\delta_{\sigma1}-\delta_{\sigma0})(\delta_{m1}-\delta_{m0})), \tag{3.56}$$

$$L_{m\mu} = \frac{\pi}{2}(\delta_{\mu-m,1}(1-(\delta_{\sigma1}-\delta_{\sigma0})\delta_{m0}) - \delta_{m-\mu,1}(1+(\delta_{\sigma1}-\delta_{\sigma0})\delta_{\mu0})). \tag{3.57}$$

Taking the Ricatti equation (3.48), we apply a matrix equivalent of the change of variables commonly used for simplifying Riccati equations [24, 25]

$$Z = -\mathbf{P}'\mathbf{P}^{-1}\left(-\frac{1}{jkz_c}(C+KB)\right)^{-1}, \tag{3.58}$$

giving the transformed Jacobi differential equation

$$\mathbf{P}'' + B(C+KB)\mathbf{P} = \mathbf{0}, \tag{3.59}$$

which is a matrix equivalent of the simple-harmonic motion equation. This can be found more directly from (3.46) and (3.47), thus giving a physical interpretation to what would otherwise be an arbitrary change of variables.

Now assume that $B(C+KB)$ is diagonalisable, with eigenvalues λ_α^2 and the corresponding eigenvectors being the columns of the square matrix X. Then

$$B(C+KB) = -X\Lambda^{-2}X^{-1}, \tag{3.60}$$

where Λ is a diagonal matrix with elements $1/j\lambda_\alpha$. (3.59) then becomes

$$\mathbf{P}'' - X\Lambda^{-2}X^{-1}\mathbf{P} = \mathbf{0}, \tag{3.61}$$

which, exploiting the spatial independence of B, C, K, and therefore X, gives

$$(X^{-1}\mathbf{P})'' + (-\Lambda^{-2})(X^{-1}\mathbf{P}) = \mathbf{0}. \tag{3.62}$$

This equation is now of the same form as (3.26), with solution

$$\mathbf{P} = X(D(s)\mathbf{\Xi_1} + D^{-1}(s)\mathbf{\Xi_2}), \tag{3.63}$$

where $D(s)$ is a spatially-dependent diagonal matrix with elements $e^{j\lambda_\alpha s}$. Differentiating (3.63) with respect to s and substituting (3.47) gives

$$Z^{-1}\mathbf{P} = \frac{-1}{jkz_c}B^{-1}X\Lambda^{-1}(D\mathbf{\Xi_1} - D^{-1}\mathbf{\Xi_2}). \tag{3.64}$$

Define

$$H = \frac{-1}{jkz_c}B^{-1}X\Lambda^{-1} \tag{3.65}$$

to give

$$Z^{-1}\mathbf{P} = H(D(s)\mathbf{\Xi_1} - D(s)^{-1}\mathbf{\Xi_2}), \tag{3.66}$$

which is in the same form as (3.31). We proceed as before (3.31)-(3.36); defining

$$E = \mathcal{D}(X + Z_dH)^{-1}(X - Z_dH)\mathcal{D}, \tag{3.67}$$

we have

$$Z_0 = (X - XE)(H + HE)^{-1}. \tag{3.68}$$

This gives the input impedance of a toroidal bend in terms of its known output impedance. Such a result, when combined with an appropriate bore discontinuity model, can be used to discretise a bend of varying radius and curvature.

We note the solution found previously for straight ducts (3.27) is a special case with $\kappa = 0$, $X = I$ and $\lambda_\alpha = k_{\|\alpha}$. Once again, it can also be shown that evaluating the above result in a 1-mode (scalar) case results in the classical plane-wave result, demonstrating that the propagation of the plane-wave is independent of bends in the duct.

3.4 Bore Discontinuities

Any general system must include nodal-diameter modes throughout. Kemp [5] assumes axi-symmetry and excludes these modes; here we are considering instruments which may have bends, and therefore a more general formulation including these modes is necessary. We follow Kemp's derivation and generalise it to include these modes. Readers are referred to that work for a derivation from first principles; here we give a brief description of the early part, and continue in detail from where this derivation diverges.

3.4.1 Derivation

Consider a join between two tubes with circular cross-section of radii $\mathcal{R}_0$ and $\mathcal{R}_1$. They meet concentrically – i.e. they share the same central axis (figure 2.5). They need not have the same curvature properties. The impedance matrices on either side of the jump are $Z^{(0)}$ and $Z^{(1)}$. The pressure field on the plane of the dicontinuity is similarly described by $\mathbf{P}^{(0)}$ and $\mathbf{P}^{(1)}$, and these must be equal on the shared area. We assume a rigid (Neumann) boundary condition on the annular perpendicular section surrounding the discontinuity. We start from equation 2.97 of [5]

$$Z^{(0)} = FZ^{(1)}F^T, \tag{3.69}$$

where, from B.15 of [5], adapting to our notation,

$$F_{\alpha\beta} = \int_0^{\mathcal{R}_0} \int_0^{2\pi} \psi_\alpha^{(0)} \psi_\beta^{(1)} r d\phi dr, \tag{3.70}$$

with $\psi^{(0)}$ and $\psi^{(1)}$ the transverse pressure profiles either side of the jump. The pressure profile of the circle radius $\mathcal{R}_0$ is the same on both sides. Given (3.12) we have

$$F_{\alpha\beta} = A_\alpha A_\beta \int_0^{2\pi} \sin\left(m\phi + \frac{\sigma\pi}{2}\right) \sin\left(\mu\phi + \frac{\sigma\pi}{2}\right) d\phi \int_0^{\mathcal{R}_0} r J_m\left(\frac{\gamma_{mn} r}{\mathcal{R}_0}\right) J_\mu\left(\frac{\gamma_{\mu\nu} r}{\mathcal{R}_1}\right) dr. \tag{3.71}$$

Consider first the ϕ integral. This vanishes if $m \neq \mu$ – i.e. modes couple only with modes with the same number of nodal diameters. Then,

$$\int_0^{2\pi} \cos(m\phi)^2 d\phi = \begin{cases} 2\pi & \text{if } m = 0 \\ \pi & \text{if } m > 0, \end{cases} \tag{3.72}$$

giving

$$\int_0^{2\pi} \cos(m\phi)\cos(\mu\phi) d\phi = \pi\delta_{m\mu}(\delta_{m0} + 1). \tag{3.73}$$

likewise

$$\int_0^{2\pi} \sin(m\phi)\sin(\mu\phi) d\phi = \pi\delta_{m\mu}\delta_{m0}, \tag{3.74}$$

therefore, assuming that $\sigma = 1$ when $m = 0$ without loss of generality,

$$\int_0^{2\pi} \sin\left(m\phi + \frac{\sigma\pi}{2}\right) \sin\left(\mu\phi + \frac{\sigma\pi}{2}\right) d\phi = \pi\delta_{m\mu}(\delta_{m0} + 1). \tag{3.75}$$

In considering the r integral in (3.71), we can now assume $m = \mu$ without loss of generality

$$F_{\alpha\beta} = A_\alpha A_\beta \pi\delta_{m\mu}(\delta_{m0} + 1) \int_0^{\mathcal{R}_0} r J_m\left(\frac{\gamma_{mn} r}{\mathcal{R}_0}\right) J_m\left(\frac{\gamma_{m\nu} r}{\mathcal{R}_1}\right) dr. \tag{3.76}$$

Using a standard integral (C.6), we have

$$F_{\alpha\beta} = A_\alpha A_\beta \pi\delta_{m\mu}(\delta_{m0} + 1)\mathcal{R}_0^2 \left(\frac{\gamma_{m\nu}\zeta J_m(\gamma_{mn}) J_{m-1}(\gamma_{m\nu}\zeta) - \gamma_{mn} J_{m-1}(\gamma_{mn}) J_m(\gamma_{m\nu}\zeta)}{\gamma_{mn}^2 - \zeta^2\gamma_{m\nu}^2}\right), \tag{3.77}$$

where $\zeta = \mathcal{R}_0/\mathcal{R}_1$. This is the tranformation matrix for a bore-discontinuity, including the nodal-diameter modes.

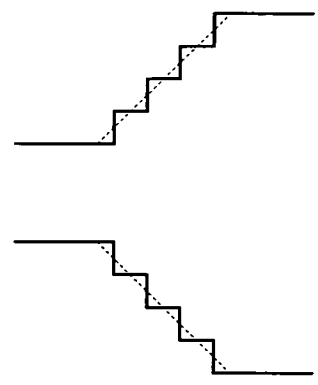


Figure 3.3: Diagram demonstrating small jumps required for discretisation of a general horn using short cylinders and bore jumps.

3.4.2 Validation

Consider the special case where $m = \mu = 0$ (i.e. omitting the nodal-diameter modes). We have

$$F_{\alpha\beta} = 2A_\alpha A_\beta \pi \mathcal{R}_0^2 \left(\frac{\gamma_{0\nu}\zeta J_0(\gamma_{0n}) J_{-1}(\gamma_{0\nu}\zeta) - \gamma_{0n} J_{-1}(\gamma_{0n}) J_0(\gamma_{0\nu}\zeta)}{\gamma_{0n}^2 - \zeta^2 \gamma_{0\nu}^2} \right). \tag{3.78}$$

We note that $J_{-1}(\gamma_{0n}) = 0$ and use equation (C.7) to find

$$F_{\alpha\beta} = 2A_\alpha A_\beta \pi \mathcal{R}_0^2 \left(\frac{\gamma_{0\nu}\zeta J_0(\gamma_{0n}) J_1(\gamma_{0\nu}\zeta)}{\zeta^2 \gamma_{0\nu}^2 - \gamma_{0n}^2} \right), \tag{3.79}$$

which, using the results (3.13) for A, becomes

$$F_{\alpha\beta} = \frac{2J_1(\gamma_{0\nu}\zeta)\gamma_{0\nu}\zeta}{J_0(\gamma_{0\nu})(\zeta^2\gamma_{0\nu}^2 - \gamma_{0n})}. \tag{3.80}$$

This is the result given by Kemp, of which the result (3.77) is a generalisation.

3.4.3 Small Discontinuities for Discretisation

It is convenient to approximate a general bore profile as a series of short cylinders, connected with bore discontinuities. These discontinuities may be arbitrarily small. We here briefly examine the behaviour of the bore discontinuites in cases of ζ close to 1.

Consider, for general $m > 0$, the case with no discontinuity – i.e. $\zeta = 1$. We expect this to give $F = I$, the identity matrix. First, we look at the off-diagonal terms ($\alpha \neq \beta$, i.e. $n \neq \nu$). From equation (C.8) follows

$$\gamma_{mn}\frac{J_{m-1}(\gamma_{mn})}{J_m(\gamma_{mn})} = m, \tag{3.81}$$

which, when substituted into (3.77), gives

$$\begin{aligned} F_{\alpha\beta} &= \frac{1}{\sqrt{1-\frac{m^2}{\gamma_{mn}^2}}}\frac{1}{\sqrt{1-\frac{m^2}{\gamma_{m\nu}^2}}}\left(\frac{\gamma_{m\nu}\frac{J_{m-1}(\gamma_{m\nu})}{J_m(\gamma_{m\nu})} - \gamma_{mn}\frac{J_{m-1}(\gamma_{mn})}{J_m(\gamma_{mn})}}{\gamma_{mn}^2-\gamma_{m\nu}^2}\right) \\ &= 0. \end{aligned} \tag{3.82}$$

Evaluating the diagonal terms will require a limiting case; analytically this is very involved, but numerically the limit tends to 1, excepting a region very close to $\zeta = 1$ where numerical errors dominate. For the purposes of discretising a general horn this need not concern us as the stiff region is sufficiently small.

3.5 Radiation Impedance

We use here the model by Zorumski [17]. Kemp reviews this model and presents a more convenient derivation, but omits the nodal-diameter modes; as before we follow his derivation and generalise it. Consider a circular duct, radius $\mathcal{R}_0$ and cross-sectional area S_0, terminated by a rigid infinite baffle and radiating into free space. Given the pressure and velocity profiles at the radiating plane, we can calculate the influence that the velocity at each source point (R, Φ) on the plane has on the pressure field a distance h into the free space. From this, we calculate an expression for the coupling between the pressure and velocity modes at the radiating plane, and, from there, derive an expression for the radiation impedance.

We take as our starting point equation 3.14 from [5], which is still general enough for our purposes. Here we are integrating across the opening area twice – once to get the pressure field due to the sum of all source elements, and again to isolate a single modal pressure amplitude component. Adapting to our index notation, we have

$$Z_{\alpha\beta} = \frac{j\omega\rho}{2\pi S_0^2}\int_0^{2\pi}\int_0^{\mathcal{R}_0}\int_0^{2\pi}\int_0^{\mathcal{R}_0}\psi_\alpha(R,\Phi)\psi_\beta(r,\phi)\frac{e^{-jkh}}{h}rRdRd\Phi drd\phi, \tag{3.83}$$

where R and Φ are dummy variables representing integration over the source elements, and

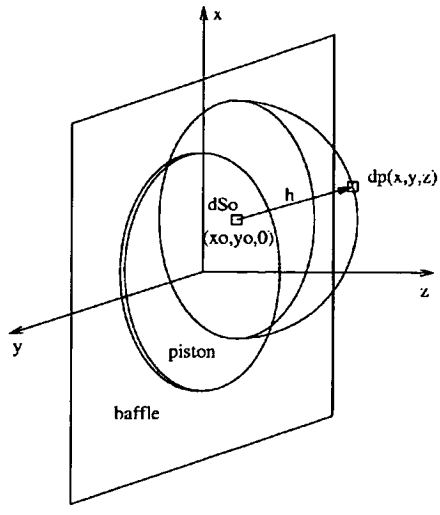


Figure 3.4: Radiation of a plane piston in an infinite baffle (Kemp, figure 3.1).

$$h = \sqrt{r^2 + R^2 - 2rR\cos(\phi - \Phi)}. \tag{3.84}$$

We have the transverse pressure modes

$$\psi_\alpha(r,\phi) = A_\alpha J_m\left(\frac{\gamma_{mn} r}{\mathcal{R}_0}\right)\sin\left(m\phi + \frac{\sigma\pi}{2}\right), \tag{3.85}$$

from which

$$\begin{aligned} Z_{\alpha\beta} &= \frac{j\omega\rho}{2\pi S_0^2} A_\alpha A_\beta \int_0^{2\pi}\int_0^{\mathcal{R}_0}\int_0^{2\pi}\int_0^{\mathcal{R}_0} \sin\left(m\Phi + \frac{\sigma\pi}{2}\right)\sin\left(\mu\phi + \frac{\sigma\pi}{2}\right) \\ &\quad \times J_m\left(\frac{\gamma_{mn}R}{\mathcal{R}_0}\right) J_\mu\left(\frac{\gamma_{\mu\nu}r}{\mathcal{R}_0}\right)\frac{e^{-jkh}}{h} r R dR d\Phi dr d\phi. \end{aligned} \tag{3.86}$$

We can eliminate h using Sonine's integral ([26] p.416, eq. 4),

$$\frac{e^{-jkh}}{h} = k\int_0^\infty \tau(\tau^2 - 1)^{-\frac{1}{2}} J_0(\tau kh) d\tau, \tag{3.87}$$

where τ is a dummy variable of integration, and Neumann's addition formula ([26] p.358, eq. 1)

$$J_0(\tau kh) = \sum_{q=-\infty}^{\infty} J_q(\tau kr) J_q(\tau kR) e^{jq(\phi-\Phi)}, \tag{3.88}$$

with dummy index q. Substituting into (3.86) gives

$$\begin{aligned} Z_{\alpha\beta} &= \frac{j\omega\rho}{2\pi S_0^2} A_\alpha A_\beta \\ &\times \int_0^{2\pi}\int_0^{\mathcal{R}_0}\int_0^{2\pi}\int_0^{\mathcal{R}_0} \sin\left(m\Phi + \frac{\sigma\pi}{2}\right)\sin\left(\mu\phi + \frac{\sigma\pi}{2}\right) J_m\left(\frac{\gamma_{mn}R}{\mathcal{R}_0}\right) J_\mu\left(\frac{\gamma_{\mu\nu}r}{\mathcal{R}_0}\right) \\ &\times k\int_0^\infty \tau(\tau^2-1)^{-\frac{1}{2}} \sum_{q=-\infty}^{\infty} J_q(\tau kr) J_q(\tau kR) e^{jq(\phi-\Phi)} d\tau\, r R dR d\Phi dr d\phi. \end{aligned} \tag{3.89}$$

Separating and rearranging the integrals gives

$$\begin{aligned} Z_{\alpha\beta} &= \frac{j\rho c}{2\pi S_0^2} A_\alpha A_\beta \\ &\times \sum_{q=-\infty}^{\infty} \int_0^{2\pi} \sin\left(m\Phi + \frac{\sigma\pi}{2}\right) e^{-jq\Phi} d\Phi \int_0^{2\pi} \sin\left(\mu\phi + \frac{\sigma\pi}{2}\right) e^{jq\phi} d\phi \\ &\times k^2 \int_0^\infty \tau(\tau^2-1)^{-\frac{1}{2}} \int_0^{\mathcal{R}_0} R J_m\left(\frac{\gamma_{mn}R}{\mathcal{R}_0}\right) J_q(\tau kR) dR \int_0^{\mathcal{R}_0} r J_\mu\left(\frac{\gamma_{\mu\nu}r}{\mathcal{R}_0}\right) J_q(\tau kr) dr d\tau. \end{aligned} \tag{3.90}$$

We can now simplify the integral terms in ϕ and Φ, which differ from those in [5] due to the additional modes. Consider the integrals

$$\begin{aligned} \int_0^{2\pi} \cos(\mu\phi) e^{jq\phi} d\phi &= \frac{1}{2}\int_0^{2\pi} (e^{j\mu\phi} + e^{-j\mu\phi}) e^{jq\phi} d\phi \\ &= \begin{cases} 2\pi & \text{if } q = \mu = 0, \\ \pi & \text{if } q = \pm\mu \neq 0, \\ 0 & \text{if } q \neq \pm\mu, \end{cases} \\ &= \pi(\delta_{q\mu} + \delta_{q(-\mu)}) \end{aligned} \tag{3.91}$$

and

$$\int_0^{2\pi} \sin(\mu\phi) e^{jq\phi} d\phi = \pi j(\delta_{q\mu} + \delta_{q(-\mu)})(1 - \delta_{q0}), \tag{3.92}$$

which give

$$\int_0^{2\pi} \sin\left(\mu\phi + \frac{\sigma\pi}{2}\right) e^{jq\phi} d\phi = \pi(\delta_{q\mu} + \delta_{q(-\mu)})(\delta_{\sigma 1} + \delta_{\sigma 0} j(1 - \delta_{q0})). \tag{3.93}$$

Define

$$Q_{m\mu}(q) = \int_0^{2\pi} \sin\left(m\Phi + \frac{\sigma\pi}{2}\right) e^{-jq\Phi} d\Phi \int_0^{2\pi} \sin\left(\mu\phi + \frac{\sigma\pi}{2}\right) e^{jq\phi} d\phi, \tag{3.94}$$

which, with (3.93), becomes

$$\begin{aligned} Q_{m\mu}(q) &= \pi^2(\delta_{q\mu} + \delta_{q(-\mu)})(\delta_{qm} + \delta_{q(-m)})(\delta_{\sigma 1} + \delta_{\sigma 0} j(1 - \delta_{q0})) \\ &= \pi^2(\delta_{qm} + \delta_{q(-m)})^2 \delta_{m\mu}(\delta_{\sigma 1} + \delta_{\sigma 0} j(1 - \delta_{q0})). \end{aligned} \tag{3.95}$$

This vanishes if $m \neq m$, and also if $q \neq \pm m$; we can therefore substitute into (3.90) and discard terms in the sum to leave only $q = m$ and $q = -m$, requiring $m = \mu$.

$$\begin{aligned} Z_{\alpha\beta} = {} & \frac{j\rho c}{4\pi^2 S_0} A_\alpha A_\beta \sum_{q=-m,m} Q_{m\mu}(q) \int_0^\infty \tau(\tau^2 - 1)^{-\frac{1}{2}} \frac{k\sqrt{2}}{\mathcal{R}_0} \times \\ & \int_0^{\mathcal{R}_0} r J_m\left(\frac{\gamma_{mn} r}{\mathcal{R}_0}\right) J_q(\tau k r) dr \frac{k\sqrt{2}}{\mathcal{R}_0} \int_0^{\mathcal{R}_0} r J_m\left(\frac{\gamma_{mnu} r}{\mathcal{R}_0}\right) J_q(\tau k r) dr d\tau. \end{aligned} \tag{3.96}$$

Define

$$G_{mnq}(\tau) = \frac{k\sqrt{2}}{\mathcal{R}_0} \int_0^{\mathcal{R}_0} r J_m\left(\frac{\gamma_{mn} r}{\mathcal{R}_0}\right) J_q(\tau k r) dr. \tag{3.97}$$

to give

$$Z_{\alpha\beta} = \frac{j\rho c}{4\pi^2 S_0} A_\alpha A_\beta \sum_{q=-m,m} Q_{m\mu}(q) \int_0^\infty \tau(\tau^2 - 1)^{-\frac{1}{2}} G_{mnq}(\tau) G_{m\nu q}(\tau) d\tau. \tag{3.98}$$

Given that $q \in \mathbb{Z}$, and that $J_{-q}(x) = (-1)^q J_q(x)$ (equation C.7), we have

$$G_{mn(-q)}(\tau) = (-1)^q G_{mnq}(\tau), \tag{3.99}$$

giving

$$Z_{\alpha\beta} = \frac{j\rho c}{4\pi^2 S_0} A_\alpha A_\beta \left[Q_{m\mu}(m) + Q_{m\mu}(-m)\right] \int_0^\infty \tau(\tau^2 - 1)^{-\frac{1}{2}} G_{mnm}(\tau) G_{m\nu}(\tau) d\tau. \tag{3.100}$$

Then we assume, without loss of generality, that $\sigma = 1$ when $m = 0$, and find

$$Q_{m\mu}(m) + Q_{m\mu}(-m) = \begin{cases} (2\pi)^2\delta_{m\mu} & \text{if } m = 0, \\ 2\pi^2\delta_{m\mu}(\delta_{\sigma 1} + j\delta_{\sigma 0}) & \text{if } m > 0, \end{cases}$$
$$= (1 + \delta_{m0})2\pi^2\delta_{m\mu}(\delta_{\sigma 1} + j\delta_{\sigma 0}). \quad (3.101)$$

Substitute into (3.100) to find

$$Z_{\alpha\beta} = \frac{j\rho c}{2S_0} A_\alpha A_\beta (1 + \delta_{m0})\delta_{m\mu}(\delta_{\sigma 1} + j\delta_{\sigma 0}) \int_0^\infty \tau(\tau^2 - 1)^{-\frac{1}{2}} G_{mn}(\tau) G_{m\nu}(\tau) d\tau \quad (3.102)$$

where $G_{m\mu} \equiv G_{m\mu m}$. Note that putting $m = \mu = 0$ gives eq. 3.21 from [5]. The integral (3.97) for G can be solved analytically using a standard integral ([28] p.146)

$$G_{mn}(\tau) = \sqrt{2} \left(\frac{\tau J_m(\gamma_{mn}) J_{m-1}(\tau k\mathcal{R}_0) - \frac{1}{k\mathcal{R}_0}\gamma_{mn} J_{m-1}(\gamma_{mn}) J_m(\tau k\mathcal{R}_0)}{\left(\frac{\gamma_{mn}}{k\mathcal{R}_0}\right)^2 - \tau^2} \right). \quad (3.103)$$

We now simplify the integral in (3.102), which is real for $0 < \tau < 1$ and imaginary for $1 < \tau < \infty$. We split the integral into real and imaginary parts and substitute $\tau = \sin\chi$ and $\tau = \cosh\xi$ respectively.

$$Z_{\alpha\beta} = \frac{j\rho c}{2S_0} A_\alpha A_\beta (1 + \delta_{m0})\delta_{m\mu}(\delta_{\sigma 1} + j\delta_{\sigma 0})$$
$$\times \left[\int_0^{\frac{\pi}{2}} \sin\chi (\sin^2\chi - 1)^{-\frac{1}{2}} G_{mn}(\sin\chi) G_{m\nu}(\sin\chi) \cos\chi d\chi \right.$$
$$\left. + \int_0^\infty \cosh\xi (\cosh^2\xi - 1)^{-\frac{1}{2}} G_{mn}(\cosh\xi) G_{m\nu}(\cosh\xi) \sinh\xi d\xi \right]. \quad (3.104)$$

Now $\sin^2\chi - 1 = \cos^2\chi$ and $\cosh^2\xi - 1 = \sinh^2\xi$, giving

$$Z_{\alpha\beta} = \frac{\rho c}{2S_0} A_\alpha A_\beta (1 + \delta_{m0})\delta_{m\mu}(\delta_{\sigma 1} + j\delta_{\sigma 0}) \quad (3.105)$$
$$\times \left(\int_0^{\frac{\pi}{2}} \sin\chi G_{mn}(\sin\chi) G_{m\nu}(\sin\chi) d\chi + j \int_0^\infty \cosh\xi G_{mn}(\cosh\xi) G_{m\nu}(\cosh\xi) d\xi \right).$$

The first integral can be evaluated using Simpson's rule, and the second (truncated to be finite) by Gaussian quadrature.

3.6 Wall Losses

3.6.1 Straight Ducts

In real instruments there is a loss of acoustic energy to heat through viscous and thermal losses in a thin boundary layer beside the walls of the tube. In the straight case, this is modelled as a modification of the boundary condition (3.9) found by imagining that there is a flow into the (rigid) walls [14]:

$$\left.\frac{\partial \psi_\alpha}{\partial r}\right|_{r=\mathcal{R}_0} = -\frac{jk}{k^*_{\perp\alpha}}\varepsilon_\alpha \psi_\alpha, \tag{3.106}$$

where ε_α is the boundary specific admittance at the wall for mode α. From this we find the lossy transverse cutoff wavenumber

$$k^*_{\perp\alpha} \simeq \frac{\gamma_{mn}}{\mathcal{R}_0} + j\frac{k\gamma_{mn}}{\gamma^2_{mn} - m^2}\varepsilon_\alpha, \tag{3.107}$$

which gives the lossy axial wavenumber

$$k^{*2}_{\parallel\alpha} \simeq k^2 - \frac{\gamma_{mn}}{\mathcal{R}_0} - \frac{2jk}{\mathcal{R}_0}\varepsilon_\alpha. \tag{3.108}$$

We use the following simplified expression for the boundary admittance under standard conditions at frequency f

$$\varepsilon_\alpha = \left[\left(1 - \sin^2\theta_\alpha\right) 2.03 \times 10^{-5} + 0.95 \times 10^{-5}\right](1+j)f^{1/2}. \tag{3.109}$$

When the mode α is propagating, $\sin^2\theta_\alpha > 0$ and θ_α can be thought of as the angle of incidence with the side wall. When the mode is evanescent, $\sin^2\theta_\alpha < 0$ and θ_α has no direct physical interpretation [14, 15]. Calcuting the values of θ_α using the Neumann conditions gives

$$\varepsilon_\alpha = \left[\left(\frac{\gamma^2_{mn} - m^2}{k^2\mathcal{R}_0^2}\right) 2.03 \times 10^{-5} + 2.98 \times 10^{-5}\right](1+j)f^{1/2}. \tag{3.110}$$

Since the tubes we are modelling are significantly longer than they are wide, we neglect the effect on the transverse wavenumbers $k^*_{\perp\alpha}$ and consider only the effect on the axial wavenumbers $k^*_{\parallel\alpha}$. Essentially we are using the lossless (Neumann) boundary condition to

model the transverse motion, and the lossy condition to model the axial motion. Any truly rigorous treatment would maintain the same boundary conditions throughout.

3.6.2 Bent Ducts

Treatment of wall losses in bent ducts is altogether more involved. The derivations given in [10, 12] work under the boundary condition

$$\left.\frac{\partial \psi_\alpha}{\partial r}\right|_{r=\mathcal{R}_0} = -jk\varepsilon\psi_\alpha. \tag{3.111}$$

Comparing this with equation (3.106), we note the key difference of the lack of index on ε – i.e. the admittance here is not dependent on the mode index, and is therefore the same for all modes. The boundary admittance model used for straight ducts defines this (simulated) admittance in terms of the incidence angle θ_α, which is dependent on index α. It is therefore not possible to use the boundary condition (3.111) in conjuction with the above admittance model. Even were this not the case, further complications are caused by the bend. The incidence angle will vary through the length of the bend, and will vary around the circumference of the bend (i.e. depend on angle ϕ). Such complications will make such a derivation, which has not yet been successfully attempted, rather difficult.

Chapter 4

Optimisation Algorithms

4.1 Design Optimisation Problems

Optimisation is the process of finding the minimum of an *objective function* $\mathcal{O}(\boldsymbol{\alpha})$ and the values of the *design variables* $\boldsymbol{\alpha}$ which give this minimum. In general, $\boldsymbol{\alpha}$ may represent a set of any number of variables of any kind, and $\mathcal{O}$ may be any scalar function and need not be analytical. Often many evaluations of the objective function are required, each time with different arguments, before an optimum is found. A problem may also specify *constraints*, restricting the possible values that $\boldsymbol{\alpha}$ can take. In addition, algorithms generally include a set of parameters to be chosen by the user; a degree of 'tuning' is necessary to find the values best suited to the problem in hand.

More formally, define a space $\mathbb{A}^{N_v}$ with N_v dimensions, which we call the *design space*. This represents all possible solutions, good or bad. Define $\boldsymbol{\alpha} \in \mathbb{A}^{N_v}$ as a set of N_v design variables $\boldsymbol{\alpha} = (\alpha_1, \alpha_2 \ldots \alpha_{N_v})$ representing a single possible solution. The objective function $\mathcal{O} : \mathbb{A}^{N_v} \mapsto \mathbb{R}^1$ maps from its domain, being the design space, to a single real number of arbitrary range. Thus $\mathcal{O}(\boldsymbol{\alpha})$ is used to evaluate the merit, or *fitness*, of each solution $\boldsymbol{\alpha}$. That set of values $\boldsymbol{\alpha}^* = \arg\min_{\boldsymbol{\alpha} \in \mathbb{A}^{N_v}} \mathcal{O}(\boldsymbol{\alpha})$ which gives the minimum value $\mathcal{O}(\boldsymbol{\alpha}^*)$ is the *optimum solution*. Note that minimisation and maximisation problems are equivalent and are interchangeable by substitution of $-\mathcal{O}$ for $\mathcal{O}$.

An *optimisation algorithm* is a process which systematically searches the multi-dimensional space of possible designs to find a minimum value of the objective function. Objective function minima may be either *global*, being the lowest value possible in the design space, or *local*,

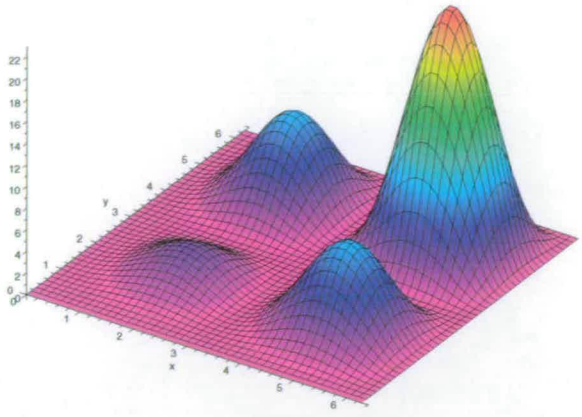


Figure 4.1: A function of two variables, with a global maximum and several local maxima.

being the minimum of a contiguous sub-space but not the entire design space. In essence, the algorithm is a method to choose which point in the design space to 'try' next given knowledge of the previous tries. On the most simplistic level, it may simply try every possible point on a given grid in the space and select the best; in general, algorithms are more sophisticated than this and can find optima with far fewer tries. Various different optimisation algorithms exist, including *direct search*, *gradient-based* and *stochastic* methods. Direct search methods require only the objective function, and navigate the search space based solely on known values; they tend to be robust, but converge slowly when close to an optimum. Gradient-based methods require the objective function to have a known derivative, and attempt to speed the convergence by using knowledge of this gradient; they converge more quickly than direct search methods, but again tend to get 'stuck' in local optima. Stochastic processes are based on random numbers and use a non-determininistic approach; they can be good at finding global optima, but, due to their inherently random nature, repeated attempts at the same problem may yield different solutions.

Design optimisation is a process by which the design of a real object is described as a set of numerical parameters $\boldsymbol{\alpha}$ and a function $\mathcal{O}(\boldsymbol{\alpha})$ is chosen to judge the merit of a possible design given specified goals; by minimising $\mathcal{O}$ we find the design of the object which best matches these goals.

4.2 Gradient-Based Algorithms

4.2.1 Simple Gradient Methods

Steepest Descent

The most simple gradient-based optimisation algorithm is *steepest descent.* Put simply, if we take a small step of length l away from a point $\boldsymbol{\alpha}^*$ in a space, the greatest possible change in the objective function $\mathcal{O}(\boldsymbol{\alpha})$ will be in the direction of the gradient vector. We therefore choose a step such that

$$\boldsymbol{\alpha} = \boldsymbol{\alpha}^* - l\nabla\mathcal{O}(\boldsymbol{\alpha}), \tag{4.1}$$

where

$$\nabla\mathcal{O}(\boldsymbol{\alpha}) = \begin{bmatrix} \frac{\partial\mathcal{O}}{\partial\alpha_1} \\ \frac{\partial\mathcal{O}}{\partial\alpha_2} \\ \vdots \\ \frac{\partial\mathcal{O}}{\partial\alpha_n} \end{bmatrix}. \tag{4.2}$$

By repeating this process iteratively, we expect to reach a local minimum. Note that the step-length l is not fixed between steps, though choice is not trivial: a small value will result in slow convergence when starting far from an optimum, even with a favourable objective function $\mathcal{O}$; too large a step size can result in the optimum being 'jumped over', and even in other local optima being explored.

In practice, this algorithm can converge very slowly to local optima, particularly if the curvature in different directions varies greatly. It is also very unsuited to finding global optima in regions with many local optima, as it will readily become 'stuck' in local optima.

Newton-Raphson

The Newton-Raphson method is a more sophisticated algorithm, using the second derivative in addition to the gradient vector.

$$\boldsymbol{\alpha} = \boldsymbol{\alpha}^* - l\nabla\mathcal{O}(\boldsymbol{\alpha}^*)H^{-1}\left(\mathcal{O}(\boldsymbol{\alpha}^*)\right), \tag{4.3}$$

where H is the Hessian matrix

$$H(\mathcal{O}(\boldsymbol{\alpha})) = \begin{bmatrix} \frac{\partial^2 \mathcal{O}}{\partial \alpha_1^2} & \frac{\partial^2 \mathcal{O}}{\partial \alpha_1 \partial \alpha_2} & \cdots & \frac{\partial^2 \mathcal{O}}{\partial \alpha_1 \partial \alpha_n} \\ \frac{\partial^2 \mathcal{O}}{\partial \alpha_2 \partial \alpha_1} & \frac{\partial^2 \mathcal{O}}{\partial \alpha_2^2} & \cdots & \frac{\partial^2 \mathcal{O}}{\partial \alpha_2 \partial \alpha_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 \mathcal{O}}{\partial \alpha_n \partial \alpha_1} & \frac{\partial^2 \mathcal{O}}{\partial \alpha_n \partial \alpha_2} & \cdots & \frac{\partial^2 \mathcal{O}}{\partial \alpha_n^2} \end{bmatrix}. \tag{4.4}$$

Geometrically, the Newton-Raphson method approximates the function with a quadratic, and then steps towards the minimum of that quadratic.

Problems with the method include the Hessian being expensive to compute, and being required to be invertible; numerical instabilities can also arise if the matrix is nearly invertible. As a local method, it still suffers from the same local optima problems as the steepest descent method.

Quasi Newton-Raphson

These methods are in principle similar to the Newton-Raphson method, but use approximations to the Hessian matrix instead of the exact value. A popular example is that of Broyden, Fletcher, Goldfarb & Shanno, the *BFGS method.* It approximates the Hessian with matrix B^{-1}, refining it such that

$$\lim_{i \to \infty} B_i = H^{-1}. \tag{4.5}$$

The method is composed of four steps:

1. A direction of step $\mathbf{d}^{(k)}$ is determined such that

$$\mathbf{d}^{(k)} = -B(\mathcal{O}(\boldsymbol{\alpha}))\nabla(\mathcal{O}(\boldsymbol{\alpha})). \tag{4.6}$$

2. An optimal step size $l^{(k)}$ is obtained by minimising $\mathcal{O}(\boldsymbol{\alpha}^{(k)} + l\mathbf{d}^{(k)})$ using a line or grid search.

3. The next step is taken such that

$$\boldsymbol{\alpha}^{(k+1)} = \boldsymbol{\alpha}^{(k)} + l^{(k)}\mathbf{d}^{(k)}. \tag{4.7}$$

4. Define $\mathbf{F} = \nabla\mathcal{O}^{(k+1)} - \nabla\mathcal{O}^{(k)}$ and $\mathbf{G} = l^{(k)}\mathbf{d}^{(k)}$. Then we refine the matrix B according to

$$B^{(k+1)} = B^{(k)} + \frac{GG^T}{F^TG} - \frac{B^{(k)}FF^TB^{(k)}}{F^TB^{(k)}F} + F^TB^{(k)}FSS^T, \tag{4.8}$$

where

$$S = \frac{G}{G^TF} - \frac{B^{(k)}F}{F^TB^{(k)}F}. \tag{4.9}$$

4.2.2 Trust Region Methods

Newton-Raphson-based methods rely on the objective function being globally approximated by a quadratic function; this is often not a valid assumption. Trust region methods restrict the approximation to a within a local hyper-sphere of radius $\Delta^{(k)}$. The radius is increased or decreased according to the closeness of the quadratic approximation as measured by

$$\rho^{(k)} = \frac{\mathcal{O}(\boldsymbol{\alpha}^{(k)}) - \mathcal{O}(\boldsymbol{\alpha}^{(k)} + l^{(k)}\mathbf{d}^{(k)})}{Q^{(k)}(0) - Q^{(k)}(l^{(k)}\mathbf{d}^{(k)})}, \tag{4.10}$$

where Q is the quadratic approximation. If $\rho^{(k)}$ is close to 1, then the approximation is good and we increase $\Delta^{(k)}$; if it is close to 0 then we decrease $\Delta^{(k)}$.

Levenberg-Marquardt

This popular algorithm is best applied to minimising a function which is the sum of squares of nonlinear functions

$$\mathcal{O}(\boldsymbol{\alpha}) = \sum_{i=1}^{n} f_i(\boldsymbol{\alpha})^2. \tag{4.11}$$

It searches in the direction $\mathbf{d}$ such that

$$(J_k^TJ_k + \lambda_k I)d_k = -J_k^Tf_k, \tag{4.12}$$

where λ_k are non-negative 'damping parameters', and J the Jacobian matrix.

4.2.3 Applicability to Problem

Gradient-based methods, by definition, require explicit computation of derivatives of the objective function with respect to the design variables. While the potential performance benefits offered as a result are appealing in principle, in practice work must be done to formulate the problem appropriately. Noreland [21] has done precisely that; by careful choice of formulation, the problem is expressed in terms of the gradient of an objective function based on the plane-wave impedance calculations like those in Chapter 2.

The work in this thesis may in principle follow a similar path to Noreland's. However, the primary aims of this thesis are to explore the possibilities of higher-order instrument representations, and to develop and implement more accurate multi-modal models. In order to attempt this with a gradient-based method, derivatives of the models in Chapter 3 with respect to the design variables would have to be derived; given the complexity of these models, such derivations are outside the scope of this thesis. We will hence consider only algorithms which do not require explicit derivatives. It is plausible that future work will address this particular problem, and allow the gradient methods to be used with the multi-modal models.

4.3 Rosenbrock Algorithm

4.3.1 Review of the Procedure

The Rosenbrock algorithm [29, 30, 31] is a direct search method, meaning that it searches the target space without explicit calculation of the gradient of $\mathcal{O}$ with respect to $\boldsymbol{\alpha}$. Broadly, the algorithm takes a series of steps in orthogonal directions, proceeding from the last successful step. At intervals, it changes the directions in which it moves to point in those directions most likely to yield successes – this allows it to navigate, for example, ridges diagonal to its coordinate system much more effectively than a simple direct search.

We seek a row-vector $\boldsymbol{\alpha} = (\alpha_1, \alpha_2 \ldots \alpha_{N_v})$ of N_v design variables (each lying on a co-ordinate axis of an N_v-dimensional space) to give the optimum $\min \mathcal{O}(\boldsymbol{\alpha})$ of the objective function. Throughout the procedure we retain $\boldsymbol{\alpha}^*$ and $\mathcal{O}^* = \mathcal{O}(\boldsymbol{\alpha}^*)$ to denote the best values found so far; an initial 'guess' at $\boldsymbol{\alpha}$ must be made to give a starting value of $\mathcal{O}^*$.

Define an $N_v \times N_v$ matrix D, whose columns are N_v mutually-orthogonal directions $\mathbf{d}_i$

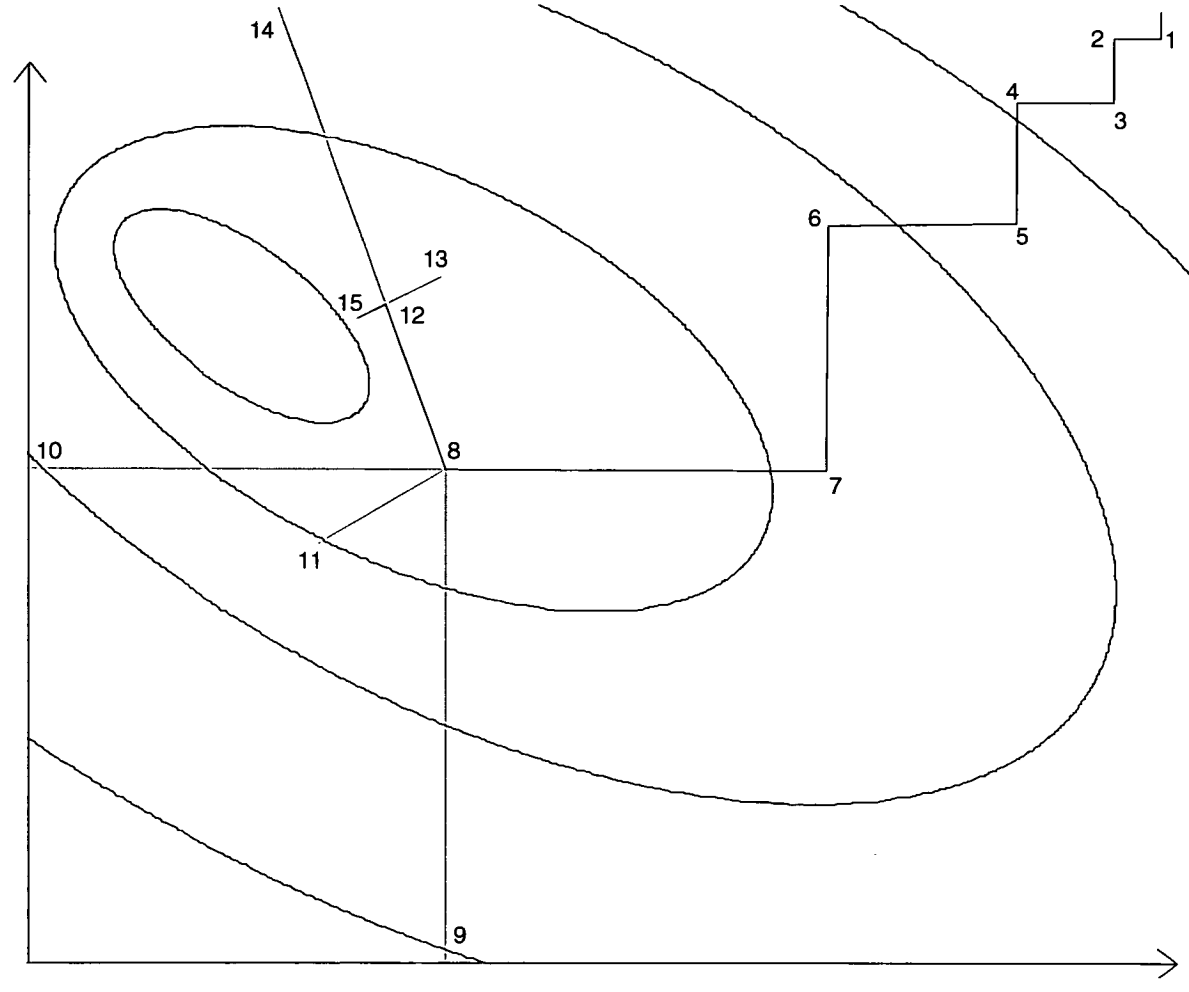


Figure 4.2: Example of Rosenbrock navigating a contour plot of a function of two variables, completing two stages. Black and grey lines denote successful and failed steps respectively.

such that $|\mathbf{d}_i| = 1$ (usually these are set as the coordinate directions for the inital stage). Define $\mathbf{l} = (l_1, l_2 \ldots l_{N_v})$ as a row-vector of step-sizes to be taken in the directions $\widehat{\mathbf{d}}_i$ (using $\widehat{\ }$ to denote a unit-vector), and two tuning parameters $a > 1$ and $0 < b < 1$. Then

- A step of length l_i is taken in direction $\widehat{\mathbf{d}}_i$ from point $\boldsymbol{\alpha}^*$, giving

$$\boldsymbol{\alpha} = \boldsymbol{\alpha}^* + l_i \widehat{\mathbf{d}}_i^T \tag{4.13}$$

- $\mathcal{O}(\boldsymbol{\alpha})$ is evaluated. If a success is found, i.e. $\mathcal{O}(\boldsymbol{\alpha}) \leq \mathcal{O}(\boldsymbol{\alpha}^*)$, then l_i is increased by a factor of a and $\boldsymbol{\alpha}^*$ set to equal $\boldsymbol{\alpha}$. If a failure, the step is discarded and l_i multiplied by $-b$.

- This procedure is repeated for each dimension $i = 1 \ldots N_v$.

This process constitutes a *phase*, and is repeated until both a success and a failure are found in each direction; at this point the *exploratory stage* is complete. The direction vectors are then modified as follows, using the notation $\widehat{\mathbf{d}}^{(0)}$ to denote directions used in the 0^{th} stage.

If λ_i is a sum of all successful steps in direction $\widehat{\mathbf{d}}_i^{(0)}$, we define a matrix with columns

$$\boldsymbol{\Lambda}_k = \sum_{i=k}^{N_v} \lambda_i \widehat{\mathbf{d}}_i^{(0)}, \tag{4.14}$$

in the special case that the directions $\widehat{\mathbf{d}}_i^{(0)}$ are the coordinate vectors, we have

$$\Lambda = \begin{bmatrix} \lambda_1 & & & \\ \lambda_2 & \lambda_2 & & \\ \vdots & \vdots & \ddots & \\ \lambda_{N_v} & \lambda_{N_v} & \ldots & \lambda_{N_v} \end{bmatrix}. \tag{4.15}$$

The column $\boldsymbol{\Lambda}_1$ is a sum of all the successful steps taken in a stage, and is therefore a vector joining the start and end points of the exploratory stage. $\boldsymbol{\Lambda}_2$ is a sum of the successful steps in all directions except the first, and so on. We can use the Gram-Schmidt orthogonalisation procedure to gain a new set of directions

$$\widehat{\mathbf{d}}_1^{(1)} = \frac{\boldsymbol{\Lambda}_1}{|\boldsymbol{\Lambda}_1|},$$

$$\begin{aligned}
\mathbf{d}_i^{(1)} &= \mathbf{\Lambda}_i - \sum_{k=1}^{i-1} (\mathbf{\Lambda}_i \cdot \widehat{\mathbf{d}}_k^{(1)}) \widehat{\mathbf{d}}_k^{(1)}, \\
\widehat{\mathbf{d}}_i^{(1)} &= \frac{\mathbf{d}_i^{(1)}}{|\mathbf{d}_i^{(1)}|} \qquad \text{for} \quad 2 \leq i \leq N_v., \qquad (4.16)
\end{aligned}$$

The vector $\widehat{\mathbf{d}}_1^{(1)}$ points in the direction of progress in the previous exploratory phase; we expect this to be a good approximation to the direction of best advance. $\widehat{\mathbf{d}}_2^{(1)}$ is the best direction which is normal to $\widehat{\mathbf{d}}_1^{(1)}$ and so on.

We seek an equivalent orthogonalisation procedure which avoids summation, and is therefore less computationally expensive. Gram-Schmidt is applicable for general $\mathbf{\Lambda}$; here we are interested only in those cases which satisfy (4.14). Combining this with (4.16), we find [31] the orthogonalisation procedure

$$\begin{aligned}
\widehat{\mathbf{d}}_1^{(1)} &= \frac{\mathbf{\Lambda}_1}{|\mathbf{\Lambda}_1|}, \\
\widehat{\mathbf{d}}_i^{(1)} &= \frac{\mathbf{\Lambda}_i |\mathbf{\Lambda}_{i-1}|^2 - \mathbf{\Lambda}_{i-1} |\mathbf{\Lambda}_i|^2}{|\mathbf{\Lambda}_{i-1}||\mathbf{\Lambda}_i| \sqrt{|\mathbf{\Lambda}_{i-1}|^2 - |\mathbf{\Lambda}_i|^2}} \qquad \text{for} \quad 2 \leq i \leq N_v. \qquad (4.17)
\end{aligned}$$

As we shall see, this is only of passing interest. Consider the case in which one or more λ is zero; for sake of argument, choose $\lambda_{p-1} = 0$ where $1 \leq p-1 < N_v$. Then

$$\begin{aligned}
\mathbf{\Lambda}_{p-1} &= \sum_{1=p-1}^{N_v} \lambda_i \widehat{\mathbf{d}}_i^{(0)}, \\
&= \sum_{1=p}^{N_v} \lambda_i \widehat{\mathbf{d}}_i^{(0)}, \\
&= \mathbf{\Lambda}_p, \qquad (4.18)
\end{aligned}$$

giving us two identical adjacent columns in the matrix Λ. It is clear that the denominator of $\widehat{\mathbf{d}}_p^{(1)}$ in (4.17) will be zero in this case, as will the term $\sqrt{|\mathbf{\Lambda}_{p-1}|^2 - |\mathbf{\Lambda}_p|^2}$; hence $\widehat{\mathbf{d}}_p^{(1)}$ is undefined. This is clearly unacceptable. (Note that, from (4.14), the denominator can only be zero when some $\lambda = 0$).

Palmer resolves this difficulty by rearranging $\widehat{\mathbf{d}}_i$ such that terms of λ_i cancel, leaving $\widehat{\mathbf{d}}_i^{(1)}$ determinate (unless $\lambda_i = 0 \; \forall i$, which is not possible with the Rosenbrock algorithm). In this vein, it can be shown that the following expression is equivalent to (4.17)

$$\widehat{\mathbf{d}}_k^{(1)} = \frac{\lambda_{k-1}\mathbf{\Lambda}_k - \widehat{\mathbf{d}}_{k-1}^{(0)}|\mathbf{\Lambda}_k|^2}{|\mathbf{\Lambda}_{k-1}||\mathbf{\Lambda}_k|}. \tag{4.19}$$

with the exception that evaluating this equation when $\lambda_{p-1} = 0$ yields $\widehat{\mathbf{d}}_p^{(1)} = -\widehat{\mathbf{d}}_{p-1}^{(0)}$ which is determinate in all non-trivial cases. Readers are referred to Palmer's paper for a complete derivation.

With our new orthogonal set of direction vectors, we commence another exploratory phase, and $\mathbf{s}$ is reset to default values (see section 4.3.2). The entire process, alternating exploratory phases with orthogonalisation, repeats until certain termination criteria have been met (see section 4.3.3).

Observation

Palmer's modification to the orthogonalisation procedure was motivated by cases in which some $\lambda_i = 0$; Rosenbrock was aware of this difficulty ([29], p. 177) and states that it is avoided in his method. This is, however, not strictly true. The Rosenbrock method ensures that it has moved in each direction over the course of a stage, but the total movement in a given direction may yet be zero if the optimiser has returned to a previous position in one direction after movement in another – in two dimensions this might reasonably be described as a 'U-turn'. The optimiser may take a successful step l_i in direction $\mathbf{d}_i$, followed by some failed steps in this direction (interleaved with steps in other dimensions, some of which will be successful). The step size may subsequently become $-l_i$ (using, say, $a = 2$ and $b = 0.5$ this is not uncommon) and another successful step is taken in direction $\mathbf{d}_i$. Now $\lambda_i = 0$, because two steps of equal magnitude have been taken in this direction, one positive and one negative, and ≥ 1 successes and failures have been found. If no further successes are found in this direction before the end of stage, λ_i retains its value of 0, and can therefore cause problems in the orthogonalisation procedure unless dealt with using Palmer's method.

4.3.2 Constrained Optimisation & Step Sizes

In some optimisation problems, the design variables must be constrained in order to ensure feasible results. For example, lengths must be held to be strictly positive to ensure the design is physically realistic; likewise upper and lower bounds may be placed on variables to exclude

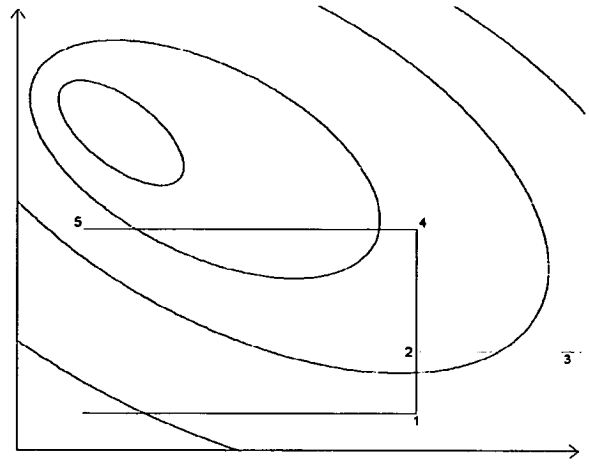


Figure 4.3: Contour plot of a function of two variables with demonstration of a U-turn taken by Rosenbrock.

realistic but impracticable designs.

Rosenbrock [29] gives a boundary-layer method for considering constraints, but this was stated with the expectation that optima are likely to lie on, or beyond, the constraints of one or more variables, as is often the case in linear programming problems. In problems where the constraints are arranged such that the optima are expected to lie away from the boundary regions, there is no need for numerical consideration of a boundary layer; in such cases we can set the objective function to a maximum value for all $\boldsymbol{\alpha}$ outside the feasible region without further computation. We define the constraints

$$\begin{aligned} \alpha_i^{\min} < \quad \alpha_i \quad < \alpha_i^{\max}, \\ \boldsymbol{\alpha}^{\min} < \quad \boldsymbol{\alpha} \quad < \boldsymbol{\alpha}^{\max}, \end{aligned} \tag{4.20}$$

and define the N_v-dimensional feasible region enclosed as $\mathbb{A}$. This region is convex.

The feasible region may have some dimensions with boundaries which are several orders of magnitude larger than others. Choosing step-sizes l is therefore not trivial, as a given step size may be very small in one dimension and very large in another – while the optimisation algorithm is capable of dealing with this, there are more efficient solutions. We can normalise the finite-sized space $\mathbb{A}$ into another space $\mathbb{A}'$, such that the constrained region is of unit size in each direction, by the simple normalisation

$$\alpha_i' = \frac{\alpha_i - \alpha_i^{\min}}{\alpha_i^{\max} - \alpha_i^{\min}}, \tag{4.21}$$

where $\boldsymbol{\alpha} \in \mathbb{A}$ and $\boldsymbol{\alpha}' \in \mathbb{A}'$; the optimiser then moves in space $\mathbb{A}'$, but de-normalises into $\mathbb{A}$ to compute the objective function (and termination criteria). It is then reasonable to set the step sizes to a common value at the beginning of each stage; a good choice for the initial stage has been found to be $l_i^{(0)} = 0.1\ \forall i$, which corresponds to 10% of the total size of the constrained space.

Various values were tried for the step size at the beginning of subsequent stages; it was found that a peculiarly effective method was to leave the step sizes unchanged from one stage to the next, i.e. $l_i^{(1)} = l_i^{(0)}$. Conceptually, this is difficult to justify because there is no direct correlation between a step of size $l_i^{(0)}$ in direction $d_i^{(0)}$ and in direction $d_i^{(1)}$. We retain the effectiveness while resolving this difficulty by setting every new step size to be the mean absolute value of the previous step sizes $\forall k$.

$$l_k^{(1)} = \frac{1}{N_v} \sum_{i=1}^{N_v} \left| l_i^{(0)} \right|. \tag{4.22}$$

We expect the vectors to point towards the direction of best progress, so the step sizes should all be positive at the beginning of a stage.

4.3.3 Termination Criteria

There are several candidate criteria for termination of the procedure; we consider them each in turn. First, we must consider *when* these criteria might be applied. There are three sensible possibilities: after each single evaluation, after each phase (i.e. one step in each direction) and after each stage (i.e. immediately before the orthogonalisation procedure). We note the possibility that, even in an otherwise successful optimisation, a single phase may yield no change to the design (e.g. if the initial direction vectors all pointed away from a single optimum) and care must be taken when applying convergence criteria here.

The most simple criterion is to count the total evaluations of $\mathcal{O}$, and terminate the program once a specified maximum has been reached, allowing an upper bound to be placed on the total computation time. This may be applied after each evaluation, but doing so after each single phase allows each design variable to receive equal consideration.

Placing some criterion on the step size, for example $\max(l_i) < \varepsilon$, may seem intuitively appealing, but this is of little use as $\mathbf{l}$ dictates the size of the *next* step to be taken, and is not therefore a direct representation of the progress of the optimisation during the preceding phase or stage. When no further improvements are found by movement in a given direction, the step sizes will become progressively larger, while at the same time progress is being made in other dimesions, with the corresponding step sizes becoming smaller. Hence the optimiser may be very close to an optimum and yet have some very large step sizes.

Comparing $\mathcal{O}^*$ against some earlier, retained value, terminating if difference $\delta < \varepsilon$ is a reasonable method when applied at the end of a stage. The Rosenbrock procedure tends to converge slowly and somewhat sporadically when it is close to an optimum; this criterion may be useful for coarse approximations but will not be sufficiently reliable for fine work.

We consider now terminating if the total steps taken in a each direction $\mathbf{d}_i$ are all small, i.e. $\max(\lambda_i) < \varepsilon$, after a completed stage. This is a good measure of how far the optimiser is 'moving' around the target space, but suffers from the vectors $\mathbf{d}_i$ being linear combinations of the design variables and therefore having no direct physical interpretation. We instead simply subtract $\boldsymbol{\alpha}^*$ from some previously retained value (say $\boldsymbol{\alpha}^*$ after the previous stage) to give μ_i, being the distances travelled in each *coordinate* direction during that stage (note that $\boldsymbol{\mu}$ and $\boldsymbol{\lambda}$ are equivalent if $\mathbf{d}$ are the coordinate directions). We now have a measure of how much each design variable has been changed – if the design variables all represent lengths then we can set a length ε as a tolerance, so when the largest changes are below this tolerance we deem the optimisation to have converged.

4.4 Genetic Algorithms

4.4.1 Overview

Genetic algorithms [32] are examples of stochastic optimisation; they converge to optima by simulating processes found in evolutionary biology. There are several distinct types of genetic algorithm, differing in details of implementation; we focus here on a *simple genetic algorithm* (in practice this is no less complicated than other types, but is thus named as it was the first type to be developed). An initial population of individuals (i.e. possible solutions $\boldsymbol{\alpha} \in \mathbb{A}$) is randomly generated, or otherwise determined, with the values of each

design variable α_i called chromosomes. The fitness of each individual is evaluated. A new generation of individuals is introduced into the population to replace the existing individuals (The fittest, elite, members of the population are not replaced). The chromosomes of the new individuals are found by mating two 'parent' individuals from the previous generation; thus the 'child' will share certain characteristics with each of its parents. The process by which parents are paired together is called selection, and is commonly based on the fitness data. Additionally, random perturbations, or mutations, may be applied to introduce new 'genetic material' into the population. The process is then repeated, with one generation's children becoming the parents of the next, until the termination criteria have been met.

4.4.2 Selection, Crossover & Mutation

In order to produce a new generation, members of the current generation must mate; selection is the process by which these mates are paired up. For each child to be generated, the parents are randomly selected. Under roulette-wheel selection, the probability of a parent being chosen is directly proportional to its fitness relative to the population - thus the fittest parents will have the most children.

Crossover is the process by which two parents' chromosomes are combined to produce a child. The method described here is array uniform crossover. If one parent is the 'father' and the other the 'mother', for each of the N_v chromosomes (design variables), there is a prescribed probability that the child will inherit its father's version of this chromosome; it otherwise inherits that of its mother. The child will therefore have a mixture of its mother's and father's chromosomes.

Each of the chromosomes of each new child is subjected, with prescribed probability, to a perturbation, or 'mutation', to the value of the chromosome, the size of which is randomly-determined within a prescribed range. While crossover jumbles existing genetic material into different combinations, mutation introduces new material into the population. Either of these processes may result in a child which is more or less fit than its parents. Two parents which are not the fittest may produce the fittest offspring due to a fortuitous combination of chromosomes.

Chapter 5

Optimisation of Brass Instruments

5.1 Introduction

The calculation of the input impedance of a brass instrument from its bore-profile has been covered in Chapters 2 & 3. The inverse problem of finding the bore of an instrument from its impedance can be solved using optimisation techniques. This requires an optimisation algorithm such as those in Chapter 4, a suitable geometric representation of the instrument as a set of design variables to be optimised, and an objective function with which to judge the merit of each potential match.

To the author's knowledge, two previous attempts have been made at this inverse problem. The first, by Kausel [19] represents the instrument using a piece-wise linear interpolation of points along the bore ('point-wise'), and optimises using either the 0^{th}-order Rosenbrock search algorithm, or (unsuccessfully) an incremental genetic algorithm. This method has been incorporated into a commercially-available product, and has been successfully used by a Swiss manufacturer to design trumpet leadpipes [20]. While practical use of this kind is most encouraging, the method certainly has room for improvement. The second attempt, by Noreland [21, 33], uses a point-wise representation of the second derivative of the bore radii, and the gradient-based Levenberg-Marquardt algorithm to give smoother bore-profiles. Both of these techniques use only the plane-wave model. Similar techniques have been applied to the optimisation of xylophones [34, 35], and to leak detection [36].

Optimisation techniques of this kind can be applied either to *bore reconstruction* problems, which find an unknown bore from a known impedance curve, or *performance optimisation*

problems, which improve the design of a given bore so that its impedance matches desired characteristics (such as intonation or impedance envelope). While these problems are conceptually quite different, when considered from an optimisation standpoint they can be solved in very similar ways.

Instrument manufacturers have an enormous amount of accumulated expertise regarding the design and construction of high-quality instruments, and newly-designed models are produced frequently. The design process is costly and time-consuming, as it generally involves building prototypes which may or may not be an improvement over existing models. An easy-to-use computational design tool would be an invaluable asset as it would allow a great many potential designs to be tried very cheaply and quickly. Given a successful and fast optimisation technique, an accurate physical model, accurate measurements of the bore and impedance of real instruments, and an understanding of the correlation between impedance properties and playing characteristics, a tool can be created which is of genuine use for brass instrument design; while human expertise in design and testing will always be paramount, such a computational tool may give valuable assistance.

In this chapter we consider the concepts behind the optimisation technique presented in this thesis. In Chapter 6 we consider the practical implementation of the technique, and we explore its capabilities in Chapter 9.

5.2 Problem Definition

5.2.1 Objective Functions

In order to set up an optimisation problem, we must have a quantitative method of evaluating possible designs. Input impedance magnitude is an objective acoustic property, contains much useful information about the performance of the instrument, is experimentally measurable, and can be calculated numerically; this combination of features allows us to centre our optimisation around it. The peak locations (frequency, magnitude) have the greatest influence on the performance, though other properties, such as the width and shape of peaks, also have an effect. We wish to find the instrument shape whose impedance curve best matches certain targets; this target will be defined in terms of another impedance curve and/or certain specific characteristics regarding the impedance peaks.

An objective function must be chosen to compare the impedance properties of each 'tested' instrument to target impedance data. In the manner of Chapter 4, we denote the instrument geometry symbolically by $\boldsymbol{\alpha}$ – we will return to the details of this later. Three objective functions are used here in different combinations depending on the task, each giving scores between 0 (a perfect match) and 1 (no match). Typical tuning-parameter values, found from experience, are given. The objective functions are formulated as for a minimisation problem, though in practice a percentage score (maximisation) may prove more intuitive, and conversion to such an output is trivial. The first function is a windowed linear least-squares comparison of the entire (discrete) impedance magnitude curve with N_{pt} points

$$\mathcal{O}_1(\alpha) = \frac{1}{N_{pt}} \sum_{i=1}^{N_{pt}} \begin{cases} \dfrac{(\delta z_i)^2}{\nu^2} & \text{if } \delta z_i < \nu, \\ 1 & \text{if } \delta z_i \geq \nu, \end{cases} \tag{5.1}$$

where $\delta z_i = |z(f_i) - \bar{z}(f_i)|$, i.e. the difference between tested impedance magnitude z and target $\bar{z}$ measured at a series of frequencies f_i, and ν is the half-width of the window (usually $100 - 300k\Omega$). The windowing eliminates the case where the score of a certain very bad feature can overwhelm the more subtle difference between another feature being good and very good, thus improving overall convergence. Both Kausel and Noreland use this least-squares approach, but without the windowing. The second function is a windowed-Gaussian comparison, scoring only the peak-frequencies

$$\mathcal{O}_2(\alpha) = \frac{1}{N_{pk}} \sum_{i=1}^{N_{pk}} \begin{cases} \dfrac{1 - \exp\left(\dfrac{-\mu_\phi(\delta\phi_i)^2}{\nu_\phi^2}\right)}{1 - \exp(-\mu_\phi)} & \text{if } \delta\phi_i < \nu_\phi, \\ 1 & \text{if } \delta\phi_i \geq \nu_\phi, \end{cases} \tag{5.2}$$

where $\delta\phi_i = |\phi_i - \bar{\phi}_i|$, i.e. the difference between tested peak-frequency ϕ_i and target peak-frequency $\bar{\phi}_i$ of peak i, the window half-width ν_ϕ (typically 10Hz), μ_ϕ is a 'strictness' parameter (typically 20 – see figure 5.1) and N_{pk} the number of impedance peaks being tested. The function joins smoothly at the window bounds $\phi_i \pm \nu_\phi$ to take a value of 1. The gradient of the Gaussian function is much steeper at a moderate distance from the target than the equivalent least-squares function, and consequently offers improved convergence speed in this region.

A very similar function $\mathcal{O}_3$ is also used for the peak-heights, using difference $|z(\phi_i) - \bar{z}(\bar{\phi}_i)|$

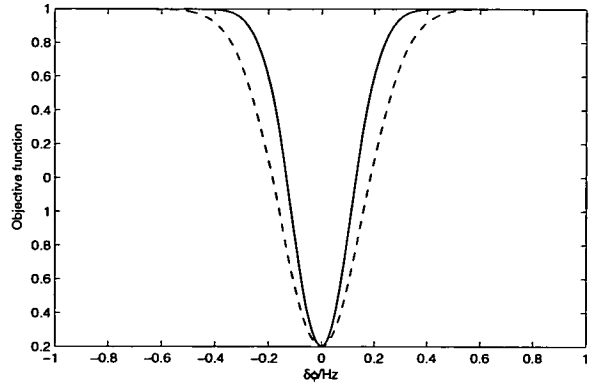


Figure 5.1: Gaussian objective function, half-window size 1, with strictness parameter $\mu = 20$ (solid) and $\mu = 40$ (dashed).

and magnitude-dependent window half-width $\nu_z z(\phi_i)$, with parameters ν_z and μ_z (typically 0.05 and 10 respectively). We then define an overall objective function

$$\mathcal{O}(\alpha) = \frac{\varsigma_1 \mathcal{O}_1(\alpha) + \varsigma_2 \mathcal{O}_2(\alpha) + \varsigma_3 \mathcal{O}_3(\alpha)}{\varsigma_1 + \varsigma_2 + \varsigma_3}, \tag{5.3}$$

where $\varsigma_{1...3}$ are weights which can take any real value, but are generally 0 or 1. Later we refer to the choice of objective function as e.g. $\{0, 1, 0\}$, denoting values of ς. We then seek $\boldsymbol{\alpha}$ to maximise $\mathcal{O}(\boldsymbol{\alpha})$.

5.2.2 Instrument Representation

Analogy

In the most general terms, there are two basic approaches to design problems: *top-down* and *bottom-up*. Top-down design considers the problem as a whole and divides it into a series of tasks which are then, in turn, divided into sub-tasks, and so on until the specification is detailed enough to give a complete design. Bottom-up design starts with the specific placement of the most basic components (often within an outline of the overall project) and builds up, section by section, until the design is complete. The most efficient choice of approach depends on the problem.

Consider two example projects: an architect designing a house, and a writer authoring a book. The architect may begin by specifying the locations of the walls and roof, followed by the windows and doors, followed by plumbing and other fixings, and so forth – only

at the end of the process are the finest details specified. At the most fundamental level, consideration of the size and location of the individual bricks is the last part of the design process, and even this may be delegated to a builder. The architect would not begin by choosing a total number of bricks and considering how best to place them to suit his overall goals. This is an example of a problem that is suited to top-down design – while it is possible to use a bottom-up approach, it is clearly not the most effective method. Now consider the writer. Having decided on a rough overall structure, she writes each chapter in detail before proceeding to the next. Clearly there will be some interconnection between the chapters, but her primary concern is ensuring that words and sentences flow continuously. She would not begin by laying out the meaning of every chapter, paragraph and sentence in the book before returning to fill in the words. This is an example of a problem that is suited to bottom-up design; again, it is possible to use the opposite approach, but it is clearly less efficient.

In both examples, the end result is simply a complicated combination of unit elements (bricks or words), the correct placement of which is essential. The problems are not different in principle, but in practice they are solved in quite different ways. Consider all the possible ways of arranging of a number of bricks. The vast majority of these would be described as 'a pile' and not as 'a wall' or 'a house', and are therefore useless to the architect. However, the same is true in the case of words – the vast majority of possible permutations of words are nonsensical, which is clearly undesirable for writers. The manner in which these useless cases are best avoided may dictate the approach used – the architect specifies the wall first and arranges the bricks accordingly, whereas the writer arranges the words into sentences and builds them into chapters.

Brass Instruments

Returning to the problem in hand, we observe the previous attempts at solving it – we recall the distinction between bore reconstruction and performance optimisation. In both Kausel's and Noreland's methods, the instrument is represented point-wise as a large number of points along the horn (assuming axi-symmetry); the coordinates of these points are the design variables. In Kausel's method, no attempt is made to avoid inappropriate designs, so many outlandish shapes are possible; for example, [19] fig. 6 cannot reasonably be described

as 'a brass instrument'. Noreland's method uses the second derivative of the geometry to ensure smooth transitions from point to point. Nevertheless, 'wiggly' shapes are common ([21] fig. 14); while this is clearly an improvement, it can still give unsatisfactory solutions.

Using the above analogies, it may be argued that Kausel's method is akin to searching for a wall within all the possible permutations of bricks, whereas Noreland's method ensures that each brick is placed next to the previous one in a (locally) orderly fashion (as the writer might with her words), but again does not guarantee a (global) wall. These are bottom-up approaches; while they are both successful, they may not be the most efficient solution. Here we present a top-down approach to the same problem, which has proven to be most effective at providing feasible instrument shapes and avoiding 'jagged' or 'wiggly' designs.

Implementation

In general, the convergence rate of optimisation algorithms is strongly dependent on the number of design variables N_v and the size of the design space; by reducing each of these properties without compromising effectiveness, we can improve the performance of the optimisation. Here we lay out a technique that both reduces the number of design variables and constrains the design space as much as possible by trading off generality that is not relevant to the problem in hand. Care must be taken to avoid pre-determining the eventual solution by constraining the space too restrictively; the optimiser must be given enough freedom to explore many reasonable solutions.

Let us consider the design of a trombone from a top-down approach. Neglecting the mouthpiece, it can be divided into two basic sections: the slide and the flaring bell section. To a first approximation, the slide is a long cylindrical tube, the geometry of which can be described completely by its length and radius. Likewise, a simple, higher-order way of approximating a trombone bell is with a Bessel-horn (see appendix B), defined as

$$r(x) = b(-x)^{-\gamma}, \tag{5.4}$$

where r is the bore radius a distance x along the axis of instrument, γ is a flare constant and b a constant usually defined in terms of specified length and input/output radii. This simple analytical equation gives a good first approximation of real instrument bells; given

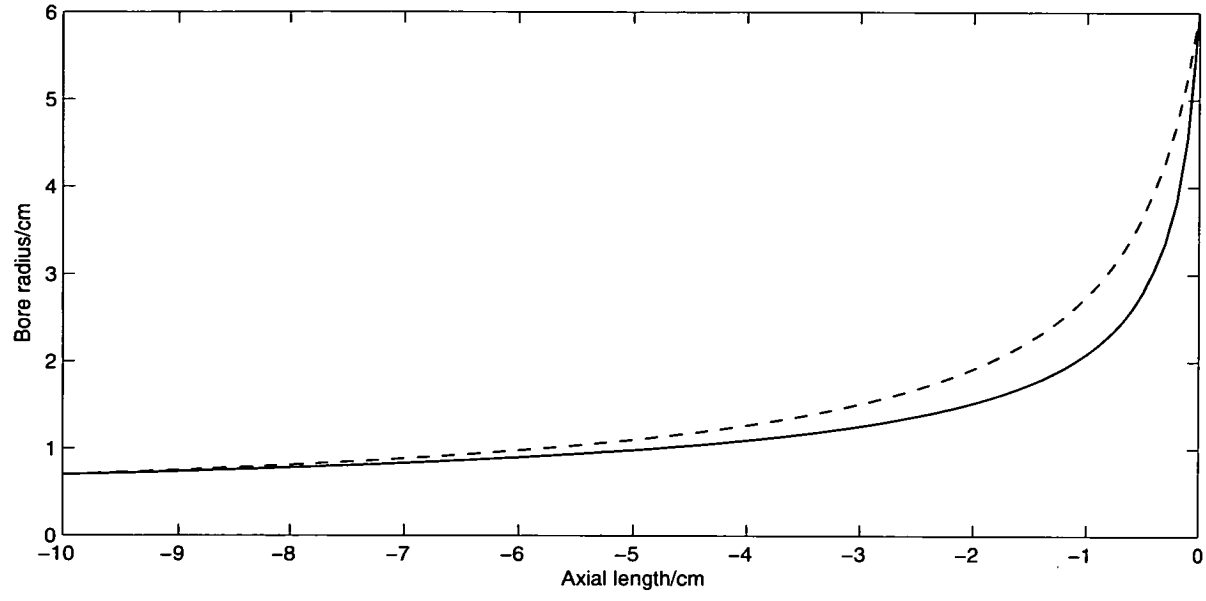


Figure 5.2: Two Bessel-horns with different flare parameters.

appropriate constraints on length, radius and flare, we can define a space of possible shapes, all of which resemble trombone bells. This, combined with the cylindrical slide section, uses 5 design variables to give a design space which consists entirely of reasonable (albeit simple) trombone shapes. By confining the design space in this manner, we have eliminated a large number of unacceptable possible solutions which human judgement would reject out of hand. Bessel-horns, however, are not exact matches for real trombone bells, so the design space here is too limited to solve the problems for which it was intended.

Continuing in this manner, we can expand on the description of the bell to allow a greater freedom of design. By describing the bell as a number of shorter Bessel-horns placed end-on-end [37], we can achieve much closer matches to real instruments. We term this *piece-wise Bessel-horn interpolation.* As with other piecewise interpolation techniques, we must allow a discontinuity of gradient at the joins between pieces, without which the two pieces would be indistinguishable from a single piece and would therefore confer no additional freedom. Cases with small gradient discontinuities are realistic, but some unrealistic designs with large gradient discontinuities do exist into the space (e.g. figure 9.2); while this is not ideal, they are proportionately far less frequent than for other representations. In practice, the shape of a real trombone bell can be very closely described by 5 Bessel-horns of different flares, giving $N_v = 12$, where one variable is the length of the sections, 6 describe the input and

output radii of the sections and 5 give their flare coefficients (point-wise approaches typically require $N_v \sim 100$). Use of this higher-order description has therefore conferred significant reduction in the number of design variables, and therefore an increase in optimisation speed when compared to pointwise representations. Examples are given in Chapter 9. Note that, under the plane-wave assumption, the Bessel-horn transmission matrix (2.43) may be used to reduce computation time.

In this way, we can form descriptions of the complete general shape of the instrument to be designed [38]; we term this a *design template*, and it is a representation of the detailed geometry of the instrument without the exact dimensions (roughly speaking, the designer knows the topology of the design without knowing the topography). A template can be constructed that can describe all trombones fairly closely; however, in practice it is more effective to have more detailed templates for more specific problems, such as tenor and bass trombones.

Approaching optimisation problems in this manner trades off generality to improve convergence. One of the advantages of the pointwise approach was the suitability to any given problem with little or no prior knowledge of the instrument shape – so-called 'black-box' problems. The approach given here assumes that the rough shape of the instrument is already known; while this is less attractive in principle, the disadvantage is negligible in practice. Given that the primary objectives are the reconstruction and optimisation of brass instruments, it is not unreasonable to assume that the class of the instrument (e.g. trumpet, trombone) is already known, and therefore that the rough shape is also known; we can thus choose a template from prior knowledge of these instruments. We set the constraints to be large enough that the design space includes all known instruments of that class, so that it is still general enough to include all solutions to the specific problem in hand, but removes many of the unreasonable solutions that another, more general, approach would include. This approach favours pragmatism over rigorous generality, but it should be noted that we can construct templates to consider general problems in the same (point-wise) way as [19], so, rather than being lost, this generality is merely put aside unless needed. Results are given in Chapter 9.

Given the reliance on templates, a user must have the ability to define new templates

easily. The task of implementing this computationally is far from trivial; however, an elegant solution is given in Chapter 6, along with other practical details of the optimisation technique.

5.2.3 Templates

In the previous section, we defined a template as being an abstract description of the geometry of an instrument without the exact dimensions. We will expand on this definition here, and give some example templates.

We define, as our base unit, an *element*. Each element represents part of the instrument. There are two types of element: single and list. Single elements describe a section of the instrument, often with simple geometry; examples include cylinders, cones, bore discontinuities and Bessel-horns. Single elements are defined in terms of variables dictating their individual properties (e.g. length, radius, flare constant) – an instrument may then be constructed from these elements. The properties of one element may depend on those of another – for example, two adjacent cones may be defined such that the input radius of one is the output radius of the other to ensure a continuous join. List elements describe more complex parts of the instrument as a list of single elements. They have no additional properties, but are merely a convenient way of treating a distinct section of the instrument as a single element; for example, a mouthpiece is well-suited to being represented as a list element consisting of cone elements, so it may then be moved easily between instruments. An entire instrument is itself a list element, and may recursively contain other list elements.

Symbolically, once a template has been defined, we can specify values of the design variables in order to construct an instrument $\boldsymbol{\alpha}$. We return to our simple example and model a trombone as a cylindrical slide $\alpha^{(1)}$ and a Bessel-horn $\alpha^{(2)}$; the trombone $\boldsymbol{\alpha}$ is then a column-vector of elements such that $\boldsymbol{\alpha}^T = [\boldsymbol{\alpha}^{(1)}, \boldsymbol{\alpha}^{(2)}]$. Continuing in more detail, the slide has a length $d^{(1)}$ and a radius $r^{(1)}$, giving $\boldsymbol{\alpha}^{(1)} = [d^{(1)}, r^{(1)}]$, and the Bessel-horn has length $d^{(2)}$, input radius $r^{(1)}$ (i.e. a smooth connection), output radius $r^{(2)}$, and flare $\gamma^{(2)}$, giving $\boldsymbol{\alpha}^{(2)} = [d^{(2)}, r^{(1)}, r^{(2)}, \gamma^{(2)}]$. Combining this gives $\boldsymbol{\alpha}^T = [d^{(1)}, r^{(1)}, d^{(2)}, r^{(2)}, \gamma^{(2)}]$ – so we have constructed a set of $N_v = 5$ numerical design variables from our template (note that $r^{(1)}$ appears only once).

The template approach has several advantages, which we examine more closely in Chapter

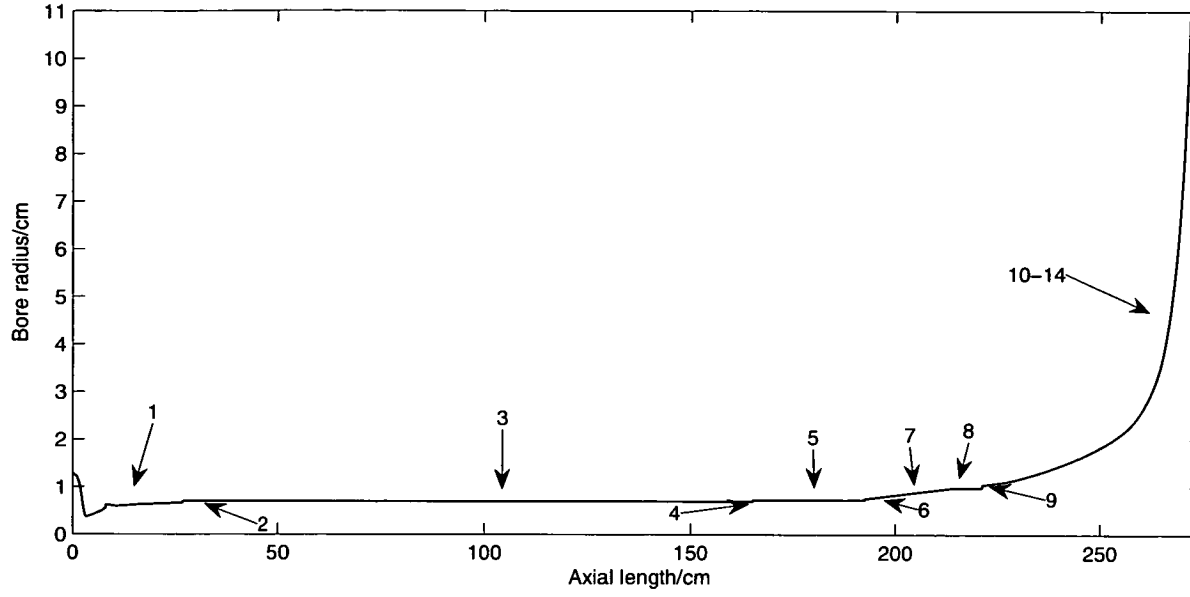


Figure 5.3: Bore-profile generated by template from table 5.1 with each element numbered.

6. Point-wise approaches describe the instrument as a series of points relative to a coordinate origin; if, for example, one section were to be lengthened, all other points further down the instrument would have to be modified accordingly. With the elemental method, elements can be modified, removed or added without affecting the rest of the instrument. This self-contained approach has been proven to be convenient for defining and modifying templates.

Example

Detailed measurements were taken of the geometry of a real trombone. A template was constructed, and specific values were chosen to approximate this bore profile (table 5.1, figure 5.2.3). This template contains 14 elements and 22 design variables; the mouthpiece and lead-pipe section is specified as preset, as is the input radius; all other parameters are variables. For optimisation, wide constraints are placed on each the design variables, such that all reasonable designs are included. Note that the flares of the shorter Bessel-horn sections differ somewhat from the flare of the closest equivalent single Bessel horn with $\gamma = 0.7$.

No.	Element	length	radius	flare	Description
1	LIST	26.145	1.286..0.664		Mouthpiece & lead-pipe
2	JUMP		0.664..0.695		
3	CYLR	139.439	0.695		Main slide
4	JUMP		0.695..0.725		
5	CYLR	26.910	0.725		Neck-pipe
6	JUMP		0.725..0.755		
7	CONE	21.990	0.755..0.990		Tuning-slide
8	CYLR	6.732	0.990		Sleeve
9	JUMP		0.990..1.053		
10	CONE	5.336	1.053..1.105		Bell
11	BESS	15.042	1.105..1.507	1.260	
12	BESS	15.042	1.507..2.228	0.894	
13	BESS	9.685	2.228..4.118	0.494	
14	BESS	5.593	4.118..10.860	1.110	

Table 5.1: Templated description of a trombone (units/cm).

5.2.4 Design Tolerances

When applying the Rosenbrock algorithm to trombone design, we must ensure that the value of tolerance ε is appropriate for all the design variables – the largest of the distances travelled in each coordinate direction must be less than this tolerance value for the optimisation to terminate. Michael Rath Trombones [39] specify that a tolerance of 0.01mm is ample for their design purposes, so no further precision on any direct physical parameter is necessary. However, we use here the flare coefficients for Bessel horns as design variables; these generally take values between 0 and 1.5. In order to meet the specified tolerance on the dimensions of instruments, we must examine how the flare parameter affects these dimensions. The software stores length values in centimetres, so the tolerance in these units is therefore 1×10^{-3}cm.

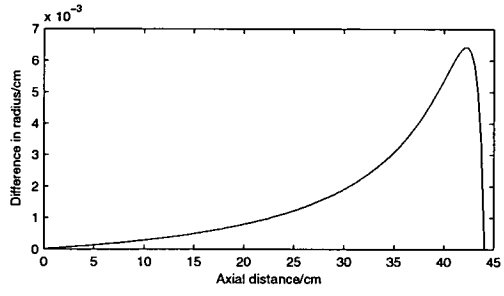


Figure 5.4: Difference in bore between two Bessel horns, demonstrating effect of changing flare constant from 0.7 to 0.701.

We choose two Bessel horns of dimensions typical of a complete trombone bell (length 44cm, radius 1.13 - 10.6cm) with flare constants of 0.7 and 0.701 respectively. Fig. 5.4 shows a plot of the difference in radius of these two bells throughout their length; we see that the maximum difference is 6×10^{-3}cm; this is the same order of magnitude as the tolerance, but a little too large. We repeat the test with a shorter section from the extreme bell (length 14.8cm, radius 4.4 - 10.6cm), such as that might be found in a 4-Bessel bell, with flares 0.7 and 0.701 - we find the maximum difference is 1.5×10^{-3}cm, which is only slightly too large. For most work a tolerance of $\varepsilon = 1 \times 10^{-3}$ for all variables is acceptable, but for fine work $\varepsilon = 5 \times 10^{-4}$ is needed to enforce the design tolerance rigidly with multiple Bessel-sections and $\varepsilon = 1 \times 10^{-4}$ with one.

Chapter 6

Design of Optimisation Software

6.1 Introduction

In previous chapters, details of optimisation algorithms, objective functions, instrument representations and impedance calculations have been laid out. In this chapter we consider the practical details of their implementation, with an emphasis on the programming problems faced.

A primary objective of this work is to develop software for optimising brass instruments; this software must be easy to use, fast, and flexible enough to be applied to a variety of problems. It is common practice for computational work in this field to produce a series of short ad-hoc programs in an interpreted language (e.g. MATLAB). While this has obvious advantages, it is not sufficient for our purposes; we instead attempt the more difficult task of a self-contained application in a compiled language. The software produced is a single executable application which requires no programming knowledge to use. Care has been taken to ensure that the code is properly laid out, commented, easy to read, and easy to modify and expand with minimal effort. The code is highly modular, and sections may be reused for future applications (e.g. [40]). Approximately 6000 lines of source code constitute the program, and represent a very substantial time investment.

ANSI C++, a modern object-oriented compiled language, was chosen for the source code. There exist C++ compilers for virtually any hardware platform, so the software is easily portable, and has already been used on both Linux and Windows-based PC's. The GCC [41] compiler (which is very widely available under a GNU license) was used during development.

At present, the software operates from the command-line, though a graphical user-interface (GUI) could be added easily; development of this is beyond the scope of this work. Instead, graphical output, such as impedance plots and bore profiles, is provided in a format readable by either the MATLAB or Maple mathematical software environments.

With interpreted languages it is common practice to 'hard-code' input values (e.g. optimisation parameters) into the program itself, so the program can evaluate only one special case and is modified for each desired case; with compiled languages this is more inconvenient and is considered inelegant. A considerable amount of programming effort has ensured that all input data can be provided by the user in plain-text files, so the program requires no modification to attempt a given problem. Given the complex nature of the data structures used, this is not a trivial task, but it allows software to be applicable to the general case, and makes it considerably easier for a new user to learn how to use the software.

The description of the software is far from exhaustive, and omissions are intentionally made for clarity. This thesis is written from the point of view of a physicist rather than a computer programmer; while a large proportion of the work pertaining to this thesis has been invested in the programming of this software, a detailed analysis is not relevant within these bounds, and would require a lot of background to be given for the average physicist to comprehend it in full. The main underlying concepts are laid out, along with the fundamental aspects of the program design. Any reader interested in a greater amount of detail should consult the source code on the CD-ROM which accompanies this thesis.

6.2 Instrument Representation

In order to best implement the template-based instrument representation described in chapter 5, the program requires an object-oriented design. We give an overview of the key concepts of object-orientation sufficient to give details of the template implementation.

6.2.1 Object Orientation

This is a modern programming paradigm common in the development of large-scale applications. Rather than the source code being a list of functions or instructions, it is instead organised into *objects*, being self-contained data structures with all functions necessary for

their own maintenance and operation. Objects can be created, manipulated and destroyed by other objects.

Consider first the classic analogy of a dog, which we'll name 'Fido'. This dog is of a certain size, has four legs, can propel itself (walk), and can fetch sticks. Fido is like other dogs, but is not identical – other dogs may be larger for example, but will share the same inherent 'dogginess'. Dogs are mammals, and operate biologically in a similar way to other mammals. All mammals propel themselves in some way, but only dogs fetch sticks. All mammals have the property of size, but not all have four legs. A dog is an example of a mammal; Fido is an example of a dog, but also an example of a mammal. Fido has characteristics unique to dogs, and 'doggy' versions of characteristics common to all mammals.

In more general programming language, objects are described by (among others) the following terms:

Class: An overall description of a set of possible objects which share certain basic characteristics and behaviour, but differ in the details.

Base class: A high-level description of possible objects which share common features but may have significant differences, e.g. mammal.

Derived class: A more specific description of one type of object (e.g. dog) which also belongs to one or more base classes (dogs are mammals). The range of certain values may be more closely specified (e.g. size), and other values may be known exactly (e.g. number of legs).

Object: A single instance of an entity belonging to a class, e.g. Fido. All values describing the object are exactly specified. There may be many different objects of the same class (e.g. another dog called 'Rover') which may differ in certain ways (e.g. Rover may be a larger dog than Fido).

Member Function: An operation which an object of that class is capable of performing, either on itself or on other objects. For example, dogs have the 'fetch stick' member function.

Member Variable: A value describing a feature of the class, e.g. size is a member variable of mammal (and therefore dog), whereas 'preferred stick length' is a member only of dog.

Encapsulation: The inclusion of relevant functions within an object, so that the object may act by itself, e.g. Fido does not need to be told how to fetch a stick, he knows how to do that

by himself – the 'fetch stick' function is encapsulated within the dog.

Inheritance: The process by which objects of a derived class obtain the members defined in a base class. For example, all mammals, including dogs, can propel themselves. While dogs may move differently from, say, whales, the ability of propulsion is inherent to mammals, not just dogs. Therefore, we say that a dog's ability to move is inherited from its definition as a mammal.

Override: A derived class may replace functions defined in its base class with its own versions. For example, a dog class will override the 'propel self' function with one that specifies the details of its own style of movement; a whale may do likewise, and, while the details will be very different to the dog's function, the effect will be equivalent.

Polymorphism: The ability of a function to act on objects of various derived classes that share a common base class, which may or may not act differently depending on the identity of the derived class. For example, one might stroke a cat and a dog in a similar way, but a horse in a different way; the 'stroke' function is polymorphically applicable to all mammals (by contrast, the 'throw stick' function is applicable only to a dog; both of these would be member functions of class 'human').

6.2.2 Element Implementation

As we saw in chapter 5, the template representation is based on elements, which are sections of tubing which can be combined to build an instrument. Single elements describe a distinct section with an analytical description of its geometry, such as cylinders, cones and Bessel-horns. We now axpand this description to two categories: primitive and composite elements. A primitive element has simple geometry and an analytical solution for its input impedance, such as cylinders and (in the plane-wave approximation) cones. A composite element has no such analytical solution, but its impedance can be calculated numerically by approximating it as a series of primitive elements – e.g. a Bessel-horn is a single element, but its geometry must be approximated by cylinders or cones to obtain a (lossy) impedance result. In addition, a list element consists of a series of single elements – e.g. a mouthpiece may be a list element consisting of many cone elements.

In keeping with the principles of object-oriented design, we define a base class called

`HornElement`, which we will use to describe the general properties of elements. We then define a series of derived classes inheriting from this base class, each representing one of various single elements. Consider first `CylinderSection`. This class contains members describing the geometry of the cylinder and encapsulated member functions to calculate its own input impedance from a given output impedance. Other primitive elements, such as `ConeSection` and `BoreJumpSection` are defined in a similar way. Then we define a composite element `BesselHornSection`; in order to calculate its input impedance, it approximates itself as a series of cones and calls the appropriate ConeSection impedance methods, then returns the overall result.

We now have a set of classes describing the single elements; we now consider the list elements, for which we will use a linked-list data structure. In general terms, this is an efficient way of representing an ordered list of objects of a given class. Given that a list element (and, therefore, an instrument) is likely to contain several different types of element, we must define this linked-list polymorphically. Hence, we define `list<HornElement*>` – we set up a list of HornElements (actually pointers to HornElements, but we will not consider the details of pointers here), so that any object which inherits from HornElement may be included. However, we wish for list elements to include other list elements, so they must also inherit from HornElement. We have thus set up a recursive, polymorphic linked-list, which is the central part of the structure of our instrument representation. It is therefore very easy to add and remove elements from this list as required.

Given the linked-list structure we have set up, we now define the class `WindInstrument`, which we use to represent entire instruments. The class inherits from HornElement (and can therefore be included in list elements) and has, as a member, a list as described above. Its impedance member functions call, in turn, the impedance functions of each element in the list to obtain the impedance of the whole instrument. WindInstrument also contains member functions that can load and save instruments from stored text files, the implementation of which was not trivial. Due to the elegance of the template approach, complex instruments can be described (and modified) very easily in these text files, and it is significantly easier for a user to 'tinker' with instrument shapes than it is with a point-wise approach.

New element types can be defined and included very easily; this has obvious advantages

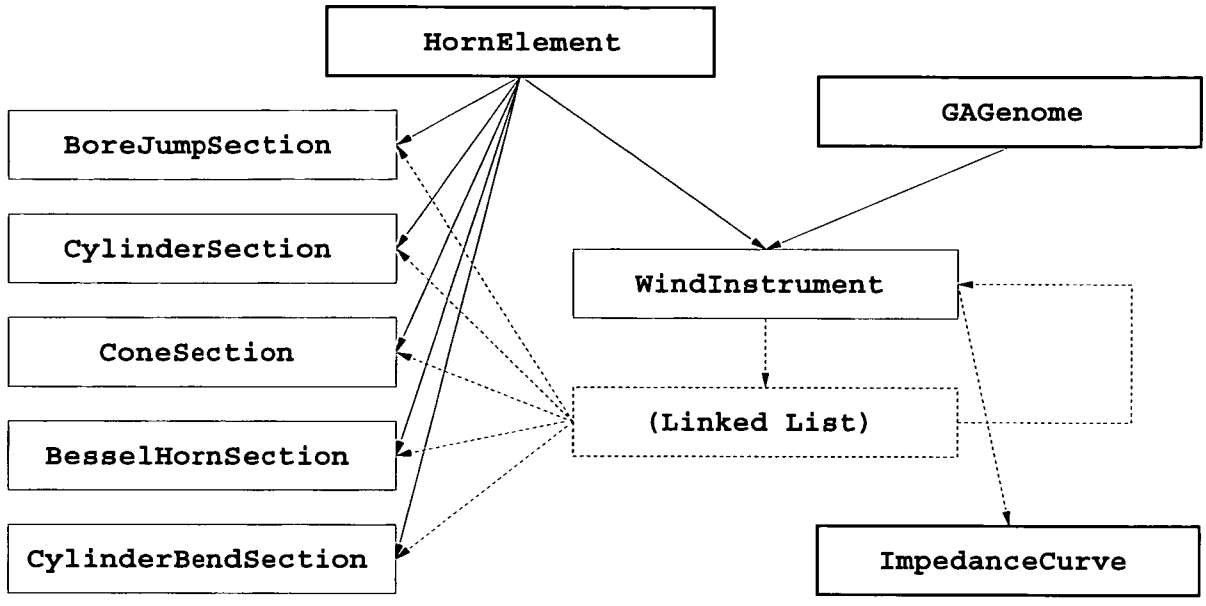


Figure 6.1: Class heirarchy, showing base classes (bold), inheritance (solid lines), derived classes (solid) and member data (dashed).

for future development. Indeed, at a later stage the author added the new element type `CylinderBendSection` with minimal difficulty – adding the capability for such objects to a point-wise representation would be troublesome, and inelegant at best (though implementing the input impedance calculations is equally difficult in either case).

As an example of the easy expansion of this software, we take the leak-detection optimisation problem attempted in [36]. In order to attempt similar problems, a `ToneHoleElement` class would need to be defined, and the relevant impedance model defined as a member function. Then, once an appropriate design template was constructed, the software would be able to attempt leak-detection problems with no further difficulties. Each of these tasks is rendered straightforward by the flexible program design; without it they would likely cause some considerable difficulty.

6.2.3 Template Implementation

It has been shown how the elemental description of an instrument provides a greater deal of flexibility than does a point-wise representation. However, the principal advantage becomes apparent when optimisation problems are considered. The concept of top-down design and its application to brass instruments has already been discussed, and the concept of templates

introduced. The object-orientation and elemental implementation described above are essential to the template representation; while they were difficult to implement, the flexibility gained allows other complex tasks, such as templates, to be achieved very easily. Without the sophisticated elemental representation, the templates would be very difficult to implement in a general way.

A user may easily specify, in a text file, a template for an instrument to be optimised. All that is required is to specify the element types required for each section, along with constraints or constant values for each geometric parameter. Each design variable is therefore specified ready for optimisation. Other saved list elements, such as a mouthpiece, may be included and held constant. Each design variable may specify dimensions for several elements, such as a bell with multiple Bessel-sections of equal length, or two sections always meeting smoothly. At present it is not possible to co-optimise over several different valve or slide positions, though the necessary expansion should be relatively straightforward.

The net result of this approach is to allow the user a fast and flexible method of defining optimisation problems, with the additional benefit of minimising the size of the design space. It is easily expandable: it works for any general template, is independent of the choice of objective function, and it can be used with any appropriate optimisation algorithm, including both the Rosenbrock and genetic algorithms used here, and also the Levenberg-Marquadt algorithm used in [18] (Noreland sets out how this algorithm may be used with paramaterised representations). The template approach can therefore be declared not only a success, but a good starting point for any future attempts at instrument optimisation.

6.3 Impedance Curves

Although the objective functions have been stated mathematically (5.1, 5.2), the practical implications of their implementation must be considered. Most notably, in discrete impedance curves the magnitude will be known only at certain values. Calculated impedance curves are commonly approximated by a series of points at regular frequency intervals. Finer grids give more accurate peak locations but are computationally more expensive; an alternative adaptive grid is offered here. A coarse grid, calculating impedance every 8.0Hz, is used to give a first approximation. The peak frequencies are then refined to a given tolerance by a

combination of the bisection method and inverse parabolic interpolation.

To seek numerically the extremum of a smooth function $f(x)$ known at three points $f(a) < f(b) > f(c)$ with the extremum bracketed by a and c, and $a < b < c$, the bisection method evaluates the function at two new points $f(d) = f(\frac{a+b}{2})$ and $f(e) = f(\frac{b+c}{2})$ equally spaced between the centre point and the right and left points. Then, $\max\{f(d), f(b), f(e)\}$ is chosen as the new centre point, and new points are inserted equidistant between it and the nearest known points to either side. This process continues recursively until the distances between known points, and therefore the peak location, is within a given tolerance ε.

A more effective but less robust method for solving the same problem is inverse parabolic interpolation [42]. Given $f(a) < f(b) > f(c)$ as above, a parabola is drawn through these three points. It is assumed that this parabola is a good approximation of the function; for any smooth function this is valid when close to a peak. The maximum of this parabola

$$x = b - \frac{1}{2}\frac{(b-a)^2[f(b)-f(c)] - (b-c)^2[(f(b)-f(a)]}{(b-a)[f(b)-f(c)] - (b-c)[f(b)-f(a)]} \tag{6.1}$$

will therefore be a good guess at the location of the maximum of f. This point, along with its two neighbours, can recursively be used to give another parabola, and so on until the tolerance criterion is met.

By using either of these peak-finding algorithms, an impedance curve can be represented in the region $(0, 2000)$Hz with peaks known to a tolerance of 0.00001Hz by $N_{pt} \sim 850$ points, compared with 2×10^8 points needed with a regular grid; these peak data are then passed to (5.2).

The HornElements have member functions for calculating impedance using either the plane-wave or the multi-modal method. A user may specify which model to use, along with how many modes to include. These functions return the values which together form an impedance curve; any subsequent treatment of the impedance curve is independent of the choice of model.

A class `ImpedanceCurve` is defined, which holds linked-lists of points on the curve. The objective functions are encapsulated; given a target curve and relevant tuning parameters, the impedance curve compares itself and returns the fitness value. Other objective functions can be defined and added easily. Saved impedance curves can be specified as targets, including

the output from BIAS [43].

6.4 Optimisation Algorithms

For the experiments using genetic algorithms, the well-established free program library GALib [32] was used. An implementation of the Rosenbrock algorithm was written from scratch.

GALib is an extensive programming library, allowing genetic algorithms to be set up and run easily. It contains classes to describe the genome (structure) of individuals and the whole population, and functions to handle the evolution of the population from one generation to the next. It is highly customisable for individual problems.

The capabilities of the template representation have already been discussed, but fundamentally the instrument $\boldsymbol{\alpha}$ is represented by a vector of real numbers. When setting up the genetic algorithm, rather than using the built-in `GA1DArrayGenome`, we gain greater freedom by having WindInstrument inherit from `GAGenome`, and having each HornElement encapsulate its own crossover and mutation operators. We therefore use the template representation throughout the genetic algorithm, without at any point needing to reduce the problem to an array of numbers. The same technique is not possible with Rosenbrock, because the movement of the algorithm around the design space depends on a vector of all the design variables. The template must be reduced to an array of design variables during a Rosenbrock optimisation.

Chapter 7

Experimental Input Impedance Measurements

Before we attempt instrument optimisation, it is useful to gain an understanding of the impedance curves of real instruments so that the features of high-quality instruments may be identified and desirable characteristics specified. The results given here are good examples of quality instruments, and may reasonably be set as design targets. Possible modifications are also discussed.

Additionally, these results may be used to attempt to identify correlations between impedance and playing characteristics; although similar results were also gained for a variety of other instruments (not shown), there are insufficient data for drawing firm conclusions, and any analysis of this kind should be treated as somewhat speculative. Nevertheless, when considered in conjunction with known performance characteristics of these instruments, certain patterns emerge which suggest causal links; these may form a basis for a more rigorous future study. This analysis neither affects nor compromises the usefulness of the impedance results for optimisation purposes (see also discussion in section 1.3), but forms a useful basis for interpretation.

Bertsch has made extensive studies into the correlation between input impedance and playing characteristics as determined through blind testing [44, 45, 46]. Plitnik & Lawson study the effect of varying the mouthpieces of French Horns [47]. Poirson [48] conducted research into the effect of mouthpiece depth on trumpet timbre, conducting a principal com-

ponent analysis to find that the impedance peak corresponding to the second harmonic of a played tone is highly correlated with its brightness.

7.1 Measurement Techniques

7.1.1 BIAS Equipment

For the impedance measurements, the Brass Instrument Analysis System, or BIAS, was chosen; the system offers a fast and convenient method of measuring a wide variety of instruments with their mouthpieces. It is based on the capillary method [49], where microphones are placed either side of a narrow, high-impedance capillary, which is then connected to the instrument.

Figure 7.1: The BIAS measuring head.

The version of the BIAS equipment available for these experiments was, however, not in prime factory condition and had some limitations for fine work. The impedance plot below 50Hz is unusable due to the very large amount of noise (omitted from the plots); likewise, higher frequencies ($>$ 800Hz) are also somewhat noisy. Certain impedance peaks, particularly those of lower frequency, can have some noise near the peak, adding a significant uncertainty into the location of the peak in both frequency and magnitude of $\pm$1Hz and $\pm$1MΩ respectively. There are additional problems with repeatability. Despite the calibration procedure,

results taken in different sessions on different days have large disagreements, particularly when variations of the temperature of the air and the instrument exist. This problem causes discrepancies of ±5c and ±1MΩ. Measurements of different instruments taken in the same session under the same conditions do not suffer from these problems, and can reasonably be compared to each other (i.e. plotted on the same axes), which is not strictly the case with those from different sessions.

After the initial submission of this work, an improved BIAS head in factory condition was made available. Some repeat measurements were made (see Appendix E). The results suffered from greatly reduced noise and measurement artifacts, were measured over a greater frequency range; repeatability problems similar to those found with the damaged old head were not observed. Nevertheless, for the purposes of the following analysis, the results were qualitatively similar, and so the remainder of this chapter presents the more extensive measurements taken with the old head.

7.1.2 Limitations of Impedance

While the input impedance of an instrument contains much useful information about its musical performance characteristics, it is not a complete description. It assumes linear behaviour, which in practice is attainable only at very low dynamics [50]. It is a steady-state (frequency-domain) measure, and therefore contains no information about the attack and decay transients of a played note; for both player and listener the articulation at the onset of a note has an important effect on its perception. The player's lips are very complex, and are in no way accounted for; given the strong coupling that exists between the lips and the air column, these effects cannot be neglected in any truly complete model of an instrument – indeed, players with different anatomies and techniques may interact differently with the same instrument. Impedance measurements are here taken only with the slide closed in first position; in practice the player will spend much of his time in other positions, which will have somewhat different acoustic properties.

In addition to the effects of the air column are those of the instrument itself. A long-established belief of manufacturers and players alike [39, 51] is the musically-significant effect of the material from which the instrument is constructed. There is no incontrovertible scien-

tific proof of this either way, despite a number of studies [52, 53, 54]; it is therefore reasonable to cite this as a possible limitation in the absence of evidence to the contrary. At any rate, structural resonances are unlikely to have any effect at the low amplitudes at which the instrument is excited during impedance measurement.

7.1.3 Harmonicity of Resonances

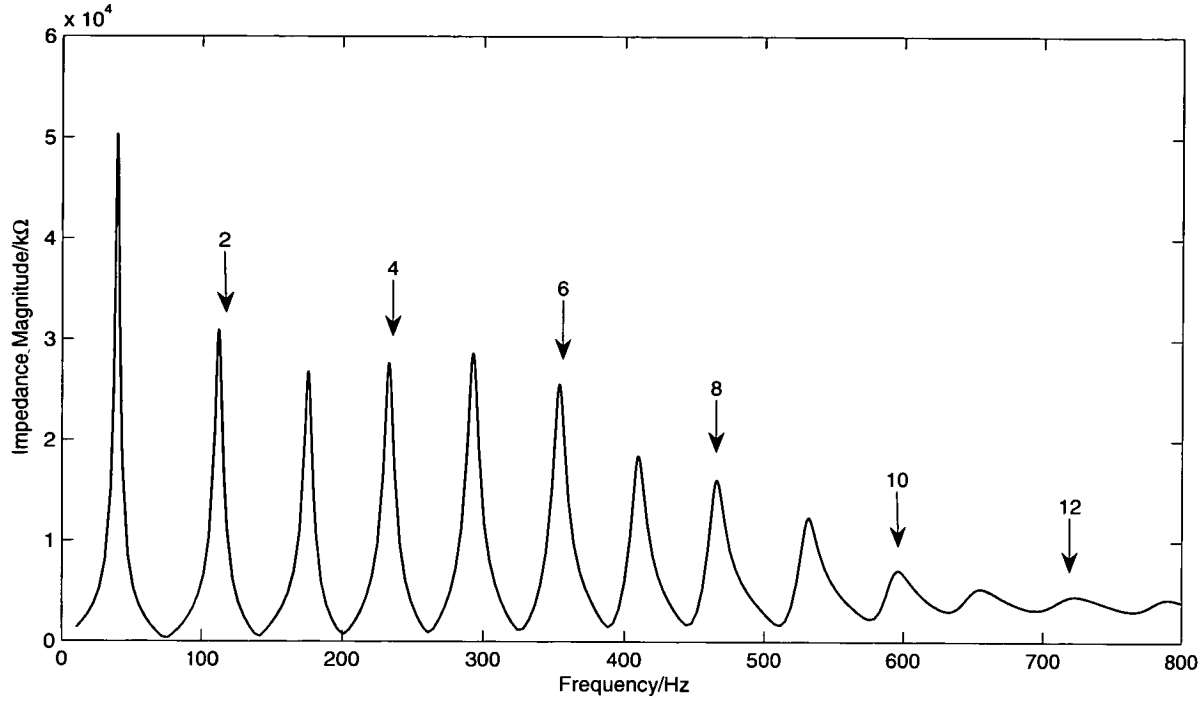


Figure 7.2: Theoretical impedance curve of a tenor trombone with slide closed. Peaks comprising the regime of $B\flat_2$ highlighted.

Benade [55] gives a thorough examination of why the characteristic shape of a brass instrument has arisen. Briefly, when the lips are playing a given note, they will excite the instrument at the frequency of that note and at integer multiples of that frequency. It is therefore advantageous to arrange some of the higher resonances of the instrument to coincide with these harmonics, thus allowing a cooperative multi-frequency 'regime' to be set up for that note, providing greater stability to the player and a more pleasant timbre. By repeating this process iteratively, we arrive at a series of resonances which coincide with a complete harmonic series (the very definition of Western European tonal music is based on the same integer relationships). Modern brass instruments have sets of resonances which

approximate this harmonic series, allowing regimes of several harmonically-related peaks to be employed over a wide and musically-useful tonal range. For example, $B\flat_2$ played on a trombone excites the resonances numbered 2, 4, 6 etc., represented by the corresponding impedance peaks (figure 7.2). The relative influence of each of the higher resonances on the regime is dependent on the dynamic level played; qualitatively, at pianissimo the excitation is nearly sinusoidal and only the sounded resonance has a significant influence, whereas at fortissimo (excluding 'brassy' cases [50]) the higher resonances have a much larger effect. It should be noted that some peaks (e.g. 4, 6 & 8) are part of the regimes of several sounded notes, whereas others (those of prime index) contribute only to one regime. Given the above reasoning, the following naturally follows:

Hypothesis: An 'ideal' instrument will have all its peaks lying exactly on a harmonic series in order to maximise the stability of the note to the player, and the harmonic content in the radiated sound.

We may reasonably suspect that this is somewhat oversimplified. From a purely musical standpoint, the objective is not to maximise harmonic content (or 'brightness'), but to maximise the pleasantness of the sound. While it is true that brass instruments with strong harmonic content are considered to be more pleasant than the 'dull' sounds yielded by similar tubes with non-harmonically-aligned resonances, it is not clear that 'pleasantness' (in the absence of a quantitative term) increases monotonically with brightness. Put simply, 'can the sound be too bright?'. The answer to this question may vary between different instruments, players, and musical styles.

Equivalent Fundamental Pitch

We now defined a quantitative measure of the harmonicity of a series of resonances. The peak-frequencies of an impedance curve can be plotted in terms of *equivalent fundamental pitch* (EFP) [56], by calculating, for each peak frequency f_i, the fundamental frequency of which it is an exact i^{th} harmonic, and the intonation of this equivalent fundamental pitch relative to an arbitrary reference frequency F. This gives a measure, in cents, of the individual harmonic alignment of the peak frequencies, and therefore how closely the peaks collectively match a harmonic series. An example EFP plot is given in figure 7.3; our hypothetical ideal

instrument would show a vertical line on this plot.

$$EFP(f_i) = \frac{1200}{\log(2)} \log\left(\frac{f_i}{iF}\right). \tag{7.1}$$

When plotting EFP for several different instruments on the same axes, it is often convenient to normalise by choosing $F = f_4/4$ for each instrument, thus shifting all plots such that each peak 4 lies on the vertical axis, effectively retuning the instruments so that each of the 4^{th} modes coincide (It is this fourth resonance which is generally used by musicians to tune their instruments). The shape of the plots is independent of F. The lowest peak of many brass instruments is generally extremely flat (some 7 semi-tones in the case of a trombone) and is not plotted here.

Strictly speaking, the points in an EFP plot should not be joined as the data are discrete and not continuous. Likewise, EFP is the dependent variable and should strictly be on the vertical axis. Both of these points are neglected for the sake of greater clarity.

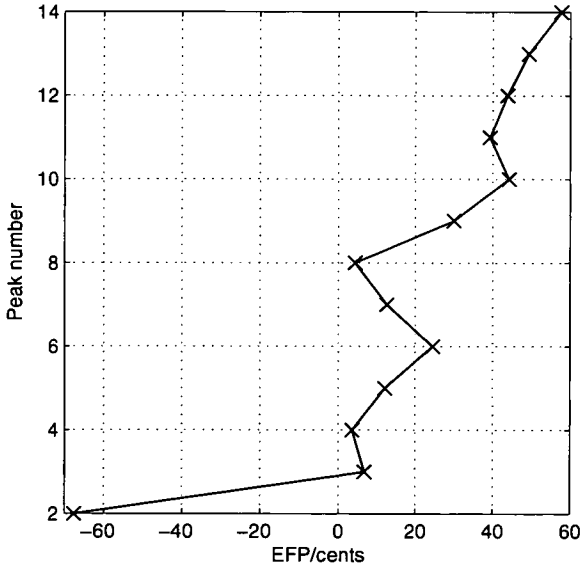


Figure 7.3: EFP plot of a typical large-bore tenor trombone, showing resonances 2-14 (F=58.3Hz).

7.2 Trombone Measurements

7.2.1 Medium-bore, Large-bore and Bass Trombones

Three different-sized trombones, all made by the same manufacturer, were measured for comparison, namely a Rath R10, a medium-bore jazz tenor trombone, an R4, a large-bore orchestral tenor trombone, and an R9, a bass trombone. Each was measured with an appropriate mouthpiece (namely a Rath S11, L5 and B1.5 respectively); due to different shank sizes it was not possible to measure each with the same mouthpiece – in any case these results would have little musical relevance as mouthpieces of such widely different sizes are seldom mixed between instrument types. The trombones were all measured with the slide closed in first position.

We start with a brief description of the general playing characteristics of these instruments, gained from detailed discussion with the manufacturers [39] and a number of professional players [51]; it is impossible to get a 'perfect' description of the instruments for any definition of the word as musicians will have differing preferences and abilities, but the terms stated here are those for which a clear consensus exists. There is no doubt, for example, that a bass trombone is more suited to playing low than a tenor, as this is the purpose for which it was designed.

In general terms, it is progressively more difficult to obtain a clean, accurate and pleasing sound on larger equipment (i.e. instruments and mouthpieces). Likewise, larger equipment gives greater assistance to low playing, and less to high playing; larger equipment provides a darker (less harmonic content), less 'penetrative' timbre.

Consider first the input impedance magnitude plot in figure 7.4, which overlays a measured impedance plot for each of the three trombones. The most striking difference is in the peak heights for the R10, which are consistently rather taller than for the larger instruments. The peak heights of the bass trombone (R9) are generally, but not always, somewhat shorter than those for the R4, particularly in the upper range of the instrument. The peaks of the R10 in the upper range are particularly tall, which may be linked with the comparative ease of playing in this register with this instrument.

From these magnitude plots are produced EFP plots (figure 7.6a) based on the peak frequencies in each plot. It is clear that the peak frequencies do not represent an exact

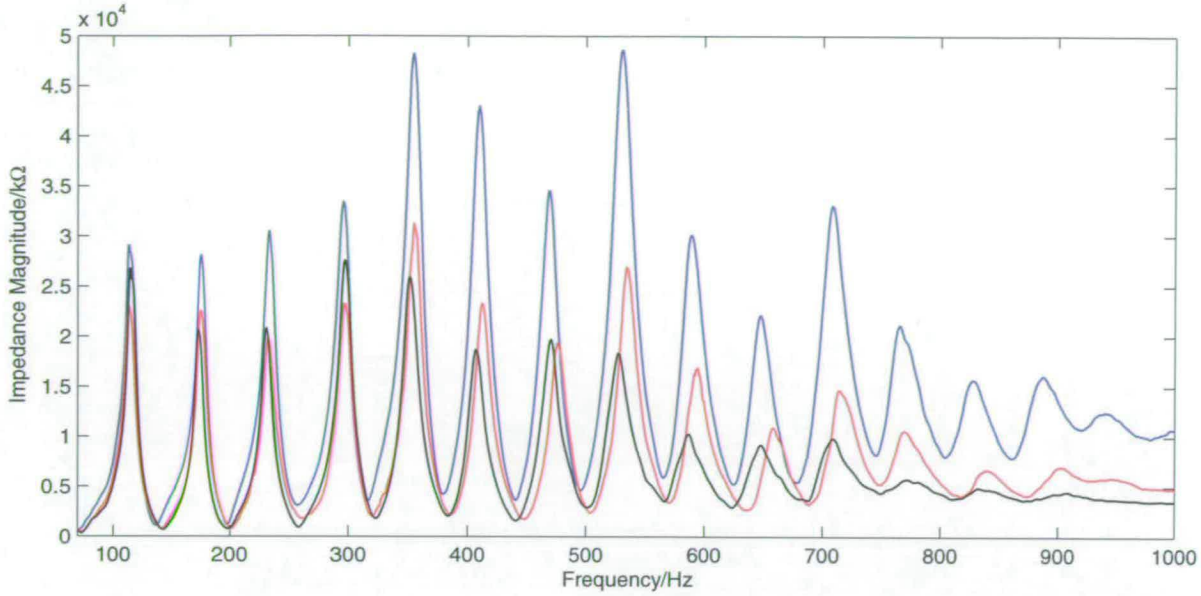


Figure 7.4: Measured impedance plots for Rath R10 (blue), R4 (red) & R9 (black).

harmonic series. For each trombone the second peak is rather flat; markedly less so for the bass trombone. Given the comparatively similar peak heights in the lower range between the R4 and R9, and the knowledge that the R9 (bass) would be generally preferred to the the R4 (tenor) for music written predominately in this register (which might reasonably be described as 'outperforming'), we can tentatively link this preference with the large difference in the alignment of peak 2, which is very important in lower-range playing, and suggest that closely-harmonic resonances make notes easier to play and produce a more pleasant sound. This would seem to support our earlier hypothesis; however, there is evidence to the contrary.

Each instrument shows a general right-hand diagonal trend on the EFP plot, denoting that higher peaks are tuned to a progressively sharper equivalent fundamental; the effect is lesser in the smaller trombone. Similar results were taken with the same BIAS equipment on French Horns [56], and demonstrate that B♭ horns have peaks which are considerably closer to a harmonic series. From these two studies, given that these instruments are considered of first-rate quality, we must conclude that inharmonicity of this kind is a desirable feature contributing to the characteristic sound of the trombone, and particularly the large-bore trombone. We can reject our hypothesis that brass instruments should have exactly harmonically-aligned peaks as being somewhat naïve, and state that the 'ideal' harmonicity may vary between instruments, and indeed sub-types of instruments.

It should be stated in passing that the inharmonicity mentioned in brass instruments is acoustically distinct from other inharmonic effects, such as 'octave-stretching' in pianos. Pianos have an inharmonicity of the overtones in their *radiated* sound, caused by stiffness effects in the strings; to minimise 'beating', piano tuners align the fundamental of one piano key with the slightly-sharp first overtone of the key one octave below. Thus two separate nearly-harmonic series are aligned to maximise the pleasantness of the sound. In a brass instrument, the overtones of the radiated sound are exact integer multiples (at least in the linear regime) due to mode-locking, and the inharmonicity of the *resonances* therefore is not related to octave-stretching.

7.2.2 Varying the Mouthpiece

The experiment in the previous section varies not just the instrument but the mouthpiece as well; here we isolate the effect of the mouthpiece. We choose one instrument, the R4, and use three different mouthpieces: the L6.5, which is considered too small for this trombone, the L5, a popular size, and the B1.5, which is a bass trombone mouthpiece and is considered far too large. From the impedance plots in figure 7.5, we see that this has a significant effect on the heights of the peaks. The small mouthpiece makes the upper peaks much taller, and the larger makes the lower peaks taller, and the upper peaks shorter to a lesser degree. While it is clear that the mouthpiece has an effect on the peak magnitudes, this effect is not sufficient to explain certain features of the results from the previous section (e.g. the peaks around 350Hz), allowing us to deduce that the instrument bore, as one might expect, also has an independent effect on the peak magnitudes.

Moving on to the EFP plots in figure 7.6b, we see that the different mouthpieces have caused some changes in alignment, but not to the extent of the differences seen between the different instruments; most notably the alignment of peak 2 is largely unaffected.

On the whole it is reasonable to suggest that taller and more closely-aligned impedance peaks contribute to the ease of playing and the brightness of the tone produced; determining exactly what the ideal is for a given instrument is not simple. Given that large-bore trombones were created specifically for their darker sound, it would appear that a degree of playing ease has been traded off in the process. Beginner trombonists are frequently told to avoid large-

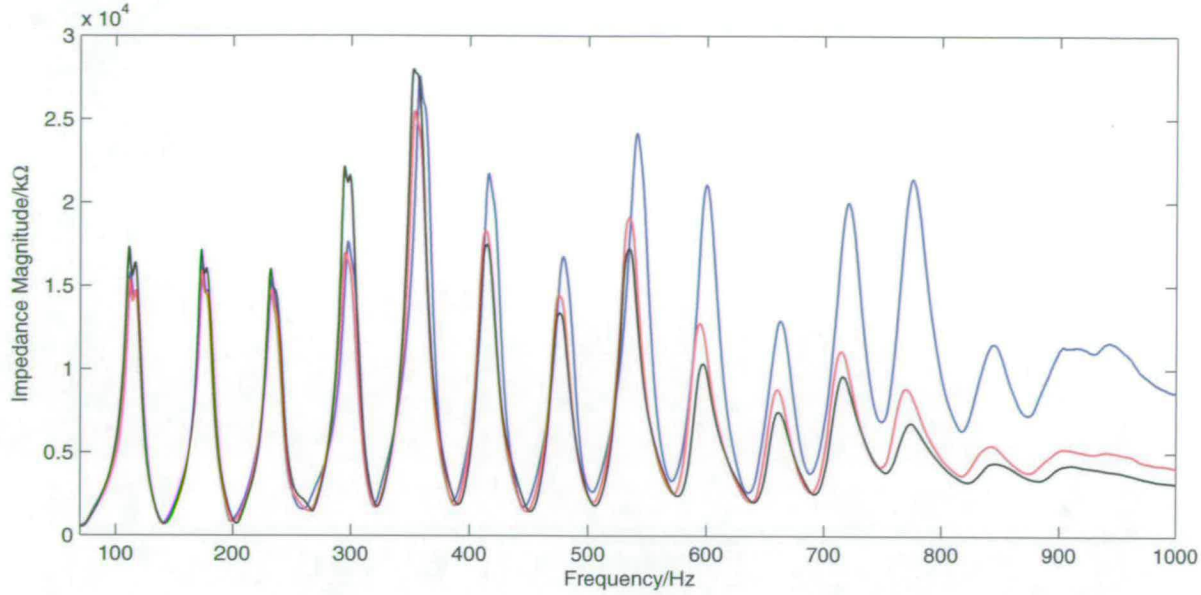


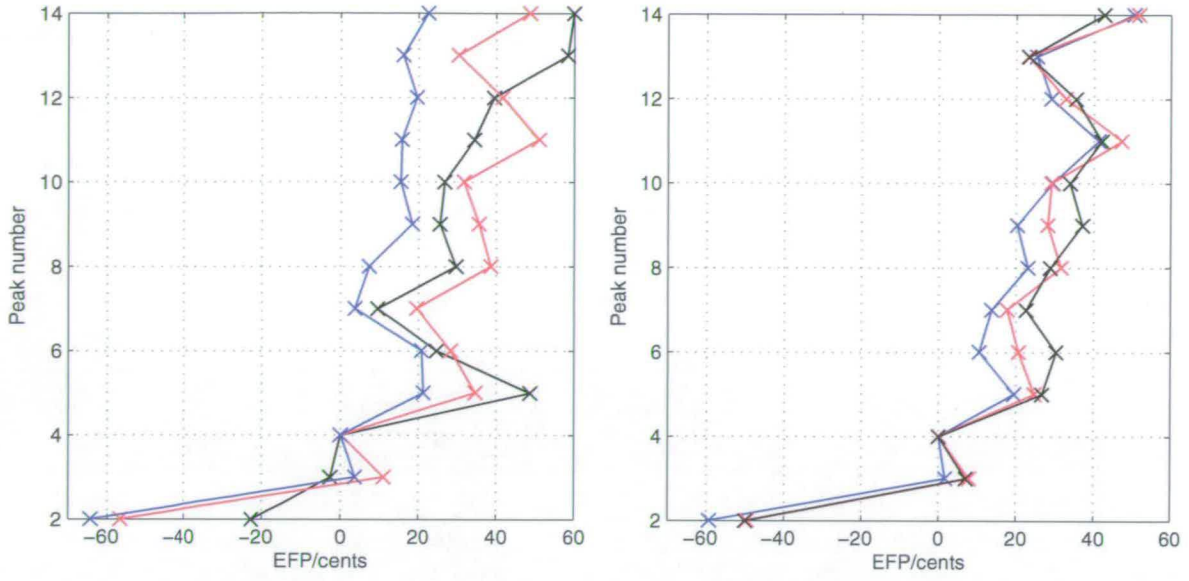
Figure 7.5: Measured impedance plots for R4 with mouthpiece L6.5 (blue), L5 (red) & B1.5 (black). Note measurement artefacts on the lower peaks.

bore instruments in the early stages of their development as they require a greater level of proficiency to play to a given standard.

7.2.3 Commentary – Room for Improvement?

Consider the regimes of B♭$_3$ (i.e. peaks 4, 8, 12 etc.) & D$_4$ (peaks 5, 10, 15 etc.) of the R9 – peak 8 has an equivalent fundamental that is +30c sharper than that of peak 4, and peak 12 is likewise +40c sharper; similarly peak 10 is -22c and peak 15 -26c compared to peak 5. It is reasonable to suggest that a player moving from the regime of peak 4 (+30, +40) to that of peak 5 (-22, -26) with the R9 will therefore have to make a larger adjustment to his embouchure than he would with the R10 (+7, +20 and -6, +1) as the two regimes differ much more in relative peak intonation. Players may be sensitive to this difference, and further investigation of this matter may prove valuable. In the case of the R9, it is reasonable to suggest that the sharp fifth peak was a necessary compromise in order to obtain the aligned second peak; we will explore this matter with the optimiser.

It is common for trombones to be fitted with valves to provide the option for extra tubing, most commonly 1m to lower the instrument from B♭ to F (bass trombones commonly have another such valve, often in G♭). A feature of these F-attachments found on all instruments



(a) R10 (blue), R4 (red) & R9 (black)

(b) R4 with mouthpiece L6.5 (blue), L5 (red) & B1.5 (black)

Figure 7.6: EFP plots for Rath trombones, normalised to 4^{th} resonance.

measured was a very flat (∼150c) second peak (figure 7.7); given that the F-attachment is most frequently used to access the low notes available on this resonance, there is clear room for improvement – trombones are well known to be 'stuffy' when using the valve. A similar problem exists in F/B♭ Horns [56]; this appears to be symptomatic of adding a significant length of cylindrical tubing to an instrument with tuned resonances. If it were possible to alter the bore profile of the additional tubing such that this resonance were more closely aligned without compromising the other features of its impedance, it is likely that an improvement in performance would result.

We now consider the question of why the inharmonicity present in a large-bore trombone is desirable. Consider a hypothetical trombone with many tall harmonic peaks (similar to the R10 in many respects). This instrument would have a large tonal range and a bright tone. If this brightness were undesirable, as it generally is in orchestral settings, then a designer might combat this by reducing the heights of the middle and upper peaks to reduce the harmonic content in the sound. However, this would have the knock-on effect of reducing the stability of the higher tones. In order, therefore, to reduce the brightness without affecting the tonal range, some inharmonicity is necessary to reduce the influence of the higher peaks on the lower tones. In reality the orchestral tenor trombone (R4) may be thought of as such a

compromise, featuring shorter, less harmonic peaks than the jazz trombone (R10) but having a similar tonal range.

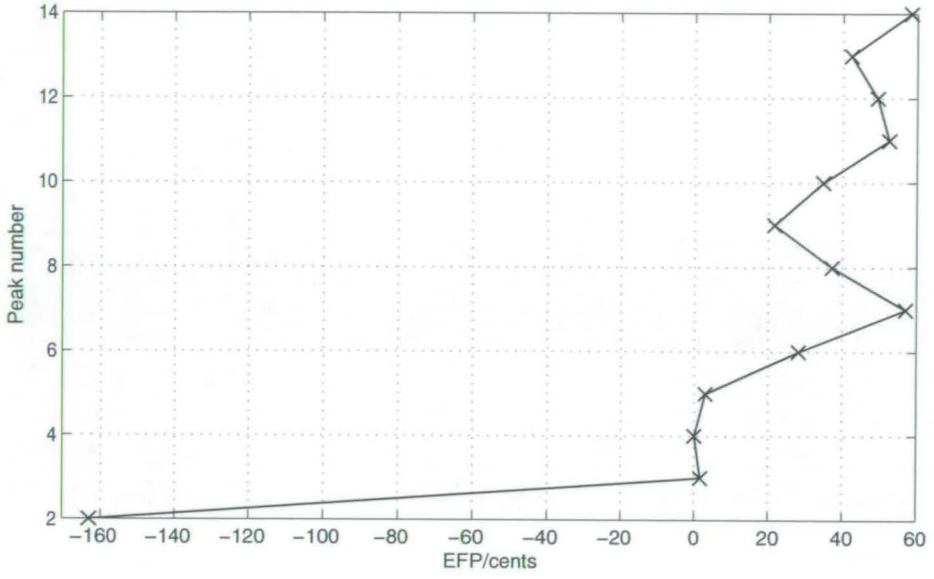


Figure 7.7: Measured EFP plot for R9 with F-valve engaged (F=43.0Hz).

Chapter 8

Numerical Results

In this chapter we apply the numerical models described in chapters 2 & 3 to 'virtual' instruments based on geometrical measurements of existing instruments. The models are compared, and convergence analyses conducted. Theoretical results are compared with the experimental results from chapter 7. Most of the results are based on a Conn 88H large-bore tenor trombone owned by the author.

8.1 Plane-wave and Spherical-wave Assumptions

To compare the effects of the plane-wave assumption (section 2.3) with those of the spherical-wave assumption (section 2.8) we compute the impedance of the 88H using both models, and compare with that from the multi-modal method. It should be noted that certain assumptions are not entirely consistent between calculations. The planar and spherical models share the same loss and radiation models, but these differ for the multi-modal method. In any case, the differing effects of the loss and radiation models can be shown to be much less significant than the geometrical effects under consideration; readers are referred to [57] for further detail.

The plane-wave model is a crude first approximation to the wave behaviour inside a flaring instrument. The spherical-wave model attempts to improve upon the plane assumption by modifying the shape of the wavefront to model more closely the real behaviour. The multi-modal method, by including more complicated oscillation patterns (figure 3.1), gives a more accurate picture than either of the single-mode models, and can be used to generate pressure plots showing the wavefronts, which are nearly spherical ([5] figure 4.6).

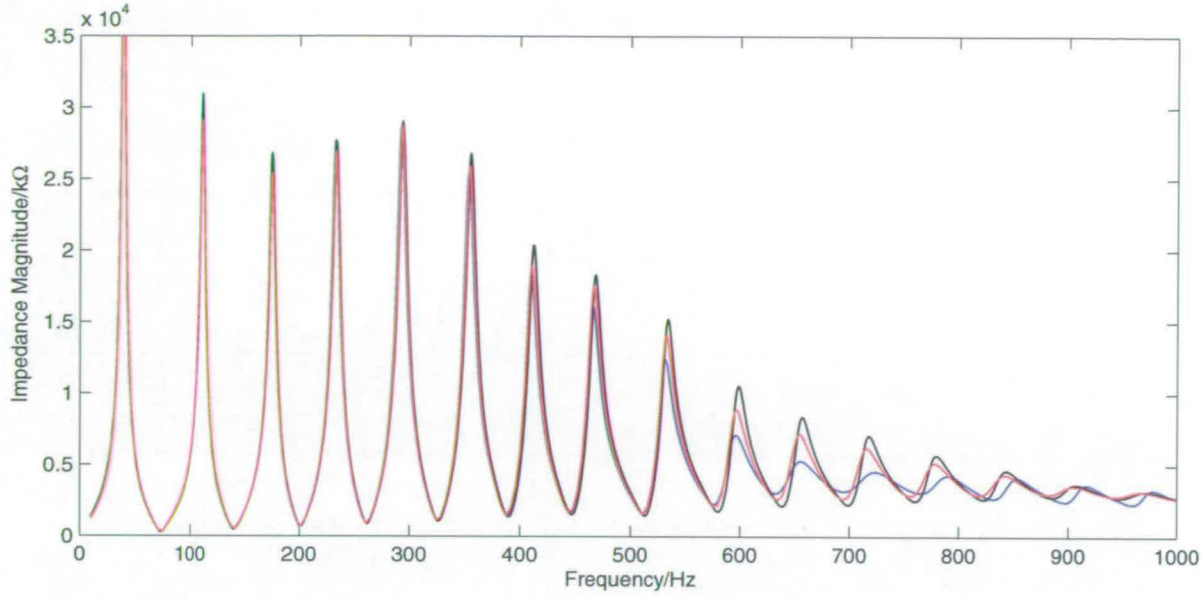


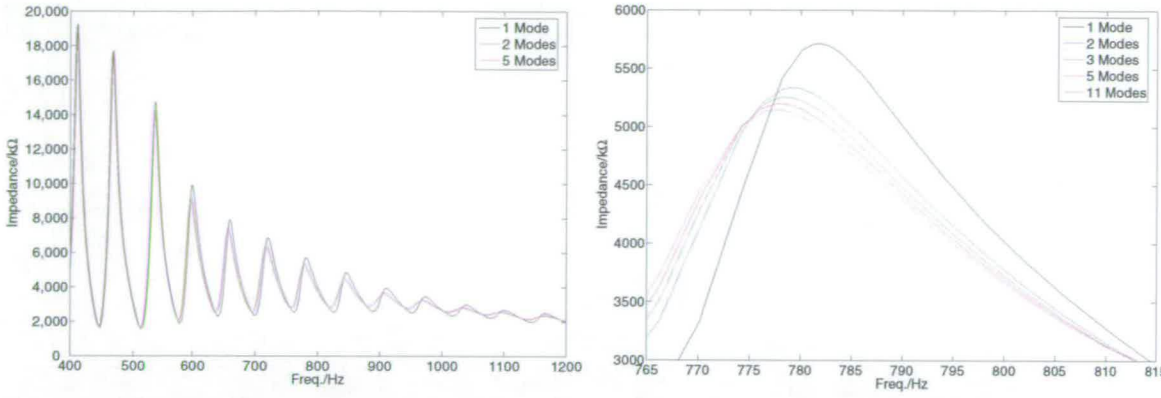
Figure 8.1: Impedance of 88H, with plane-waves (black), spherical waves (blue) and 11 modes (red).

We see from figure 8.1 that the multimodal (11 modes) plot lies between the two single-mode plots in magnitude; given the above commentary, we therefore conclude that the spherical assumption over-corrects the error in the plane-wave assumption, and over-estimates the effect of the higher modes.

Consider also figure 8.4, which compares the EFP plots for the 11-mode result and the spherical-wave result. The two theoretical results are very similar for peaks 1-10, but diverge sharply for higher peaks. Clearly the effect of the higher modes becomes at this point large enough to cause a significant change to the frequency of the peaks (the magnitudes being significantly affected from peak 8 upwards). We conclude that the spherical wave model is suitable for those cases where the frequency region of interest lies in the first 8-10 peaks of a trombone in B♭.

8.2 Effect of Higher Modes

The impedance of the 88H was calculated using the multi-modal method (see chapter 3), including several different numbers of nodal-circle modes. The flaring parts of the instrument were divided into sections of length 1mm (see [5] section 4.3 for a detailed convergence



(a) Wide view

(b) Peak 13

Figure 8.2: Plot of input impedance of 88H, with varying numbers of modes included.

analysis). We see in figure 8.2 that the inclusion of the second mode (the plane-mode being the first) has a significant effect, shifting the peaks to lower frequency and magnitude. The addition of further modes has a progressively smaller effect, rapidly becoming insignificant. This result concurs with those in [4] and [5].

8.3 Comparison of Theory and Experiment

Figure 8.3 shows a comparison of a plot measured experimentally with BIAS (including error bars representing the ±5c uncertainty found (section 7.1.1) from repeat measurements), and a plot calculated theoretically using 11 nodal-circle modes, respectively for the 88H and a virtual model of same based on bore measurements. As we can see, there is good agreement between the two; the peaks are of similar magnitude, envelope and frequency. However, none of these properties shows an exact match, and clear discrepancies exist. Figure 8.4 compares the EFP for the theory and experiment; we see a broad agreement (note the same reference frequency for all plots), but discrepancies of up to 40c, most notably for peak 5.

These discrepancies may be caused by any of the following:

- Inaccuracies in the results from the BIAS equipment, the shortcomings of which are discussed in section 7.1.1.
- Lack of a rigid infinite baffle. The theoretical model assumes such a baffle; no approximation was used in the measurement. Amir et al. [4] conducted experiments on a

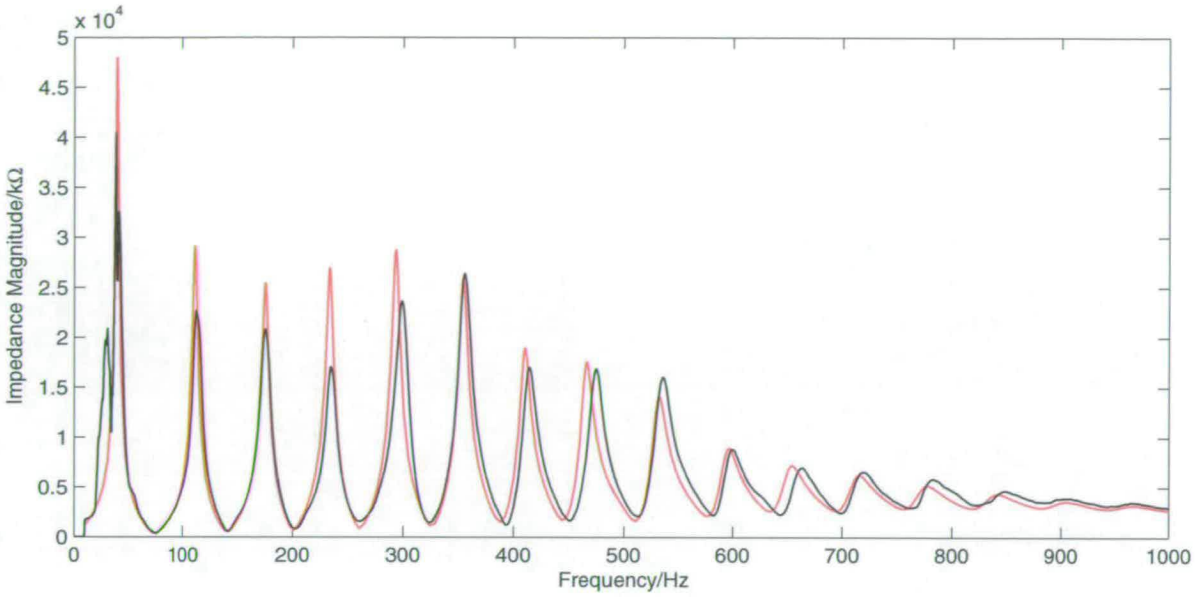


Figure 8.3: Impedance of 88H, experimental measurement (black), and calculated with 11 modes (red).

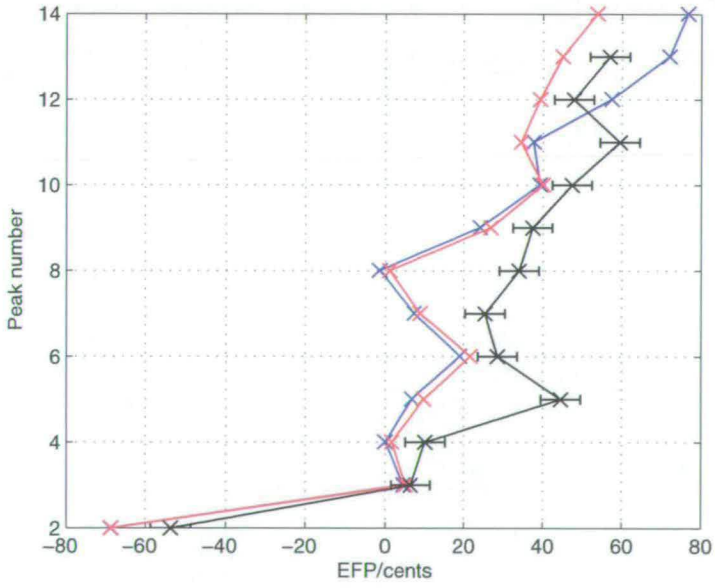


Figure 8.4: EFP of 88H (F=58.3Hz), experimental measurement (black), spherical wave (blue) and with 11 modes (red).

trombone bell with and without baffle, finding a small effect; once the rest of the trombone is included this effect is likely to be too small to explain fully the discrepancies found.

- Inaccuracies in the virtual bore-profile. The profile was carefully measured as closely as possible, but certain regions, such as the bends and the leadpipe, are difficult to measure accurately. Given that the impedance can be sensitive to small changes in bore in certain regions of the instrument (section 9.1 gives more detail), it is possible that this may have a significant effect.

- Shortcomings in the modelling. The majority of the theoretical model is discussed in chapter 3 and appears to be well-grounded. However, we cannot discount the possibility of an invalid assumption or approximation sufficient to render the model inaccurate. [4] shows a similar discrepancy but gives no explanation. Of the constituent parts of the model, that of the wall-losses is undoubtedly the least rigorous, and seems the most likely cause of any numerical discrepancies.

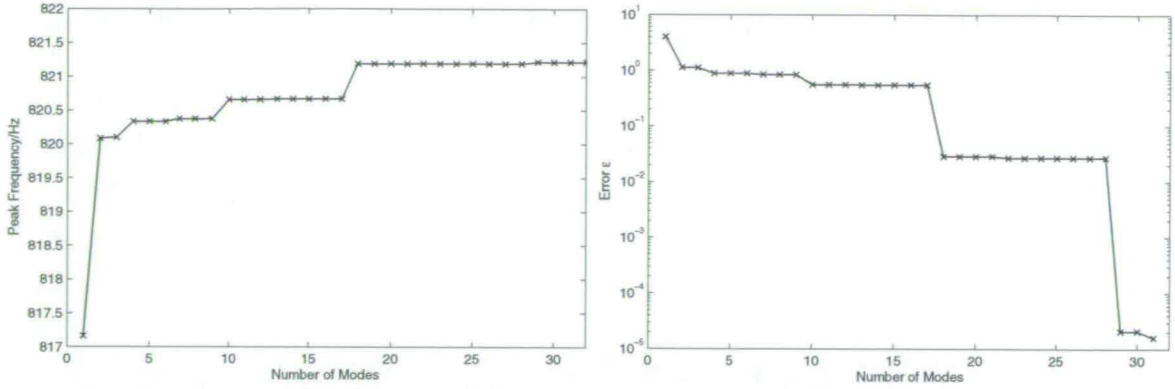
- Lack of bends in the model. This matter is considered in section 8.4; several bends are present in the experimental results but not the theory, and may potentially explain the discrepancies in part. However, the discrepancies in [4] were found in an entirely straight trombone bell section, indicating that bends cannot be the sole cause.

Further work is required to establish the exact causes of such discrepancies and to rectify any problems found.

8.4 Effect of Bends

Previously, numerical models have treated brass instruments as straight and neglected the bends which are present in the tubing of any real instrument. With the formulation derived in chapter 3, we are now in the position to perform a numerical investigation into the effects that these bends have, first in isolation, and then as part of a complete instrument.

Firstly, we conduct a series of experiments on various bend geometries; the dimensions are similar to those found in musical instruments, with obvious variations made for the purposes of experiment.



(a) First resonance frequency calculated with different numbers of modes

(b) Error value ϵ against number of modes included

Figure 8.5: Convergence of resonance of bent tube with increasing number of modes included.

8.4.1 Convergence Analysis

In order to establish the comparative effects of each of the possible modes included, a convergence analysis was performed. A bend with curvature and radius $\kappa^{-1} = \mathcal{R}_0 = 5.842\text{mm}$ and length $d = 100\text{mm}$, which resembles the tuning-slide of a trumpet, was defined. The input impedance of this bend was repeatedly calculated to include N_m from 1 to 32 modes, and the frequency of the first resonance peak was plotted. The bend was left with an open end using the radiation condition described in chapter 3. In all cases only the plane-mode was propagating in the frequency range of interest. The results converge to a limit as more modes are added; we see behaviour comparable with the convergence results in figure 3 of [9] (which are for the pressure field). After the first modes, certain later modes have a greater influence than others – we note that the $(0, n)$ modes, being those with only nodal circles (the 4^{th}, 10^{th}, 18^{th} and 29^{th} modes), have the greatest influence. This may, at first glance, suggest that the other modes are insignificant, but they too are dependent on the bend geometry and can have a very large effect, as we see below. We also plot error $\epsilon_{N_m} = |f_{N_m} - f_{32}|$ where f_{N_m} is the resonance frequency when N_m modes are accounted for. Note that the index α (see e.g. table 3.1) starts at zero, whereas the number of modes N_m starts at 1.

8.4.2 Effects of Bend Geometry

Here we will will use bend A_1 to denote a bend with $\kappa^{-1} = \mathcal{R}_0 = 5.842\text{mm}$ and axial length $d = 100\text{mm}$. This bend is an approximation of a tightly-wrapped helical tube and was cho-

Name	κ^{-1}/mm	Freq.(Hz)	Mag. (kΩ)	Description
$A1$	5.842	1011.539	401891.403	Helical bend
$A2$	31.83	821.248	995764.524	U-bend
$A3$	1.000×10^4	818.240	1011532.581	Near-straight bend
$A4$		818.240	1011532.596	Straight (all modes)
$A5$		818.240	1011532.596	Straight (nodal-circle)

Table 8.1: Geometry of bent tubes A having $d = 100.0$mm, $\mathcal{R}_0 = 5.842$mm, with first resonance as calculated with the model stated.

sen to exaggerate the effects of the curvature. Bend A_2 is a similar bent tube with bend radius 31.83mm (as used above, i.e. a U-bend), bend A_3 has very low curvature (i.e. almost straight), A_4 is the equivalent result for a straight cylinder using the straight-tube theory from chapter 3 including nodal-diameter modes, and A_5 is a straight cylinder without these modes. This allows us to vary the curvature while keeping the other geometrical parameters constant. All results are calculated with 32 modes, except A_5 which has only 5 (i.e. the same set of nodal-circle modes but omitting the nodal-diameter modes).

There is a ± 0.0005Hz error in peak location arising from the very fine, but finite, adaptive grid size used near the peaks. The straight and near-straight cases agree to 6 significant figures, which sufficiently close to be considered perfect agreement for our purposes. The straight case was measured both with 32 modes, including nodal-diameter modes, and, equivalently, with 4 nodal-circle modes and the plane-wave – there is again perfect agreement in both cases. As expected, taking account of the nodal-diameter modes does not affect the impedance of an axi-symmetric waveguide.

When comparing the bent tubes, we see that increasing the curvature of the bend causes the resonance to shift upwards in frequency and downwards in impedance magnitude.

Next another set of bends were investigated, having length $d = 100$mm and $\mathcal{R}_0 = 15$mm. B_1 has $\kappa^{-1} = 15$mm and B_2, B_3, B_4 have values of κ^{-1} equal to those of the previous set – these are relatively wide tubes with the same length and curvature properties as tube A.

We see that the larger radius tubes show a greater correction due to the bend, but exhibit

Name	κ^{-1}/mm	Freq.(Hz)	Mag. (kΩ)	Description
$B1$	15.00	877.032	14090.763	Helical bend
$B2$	31.83	779.815	25010.123	U-bend
$B3$	1.000×10^4	764.124	27516.289	Near-straight bend
$B4$		764.124	27516.291	Straight (all modes)
$B5$		764.124	27516.291	Straight (nodal-circle)

Table 8.2: Geometry of bent tubes B having $d = 100.0$mm, $\mathcal{R}_0 = 15.00$mm, with first resonance as calculated with the model stated.

Name	κ^{-1}/mm	Freq.(Hz)	Mag. (kΩ)	Description
$C1$	5.842	1041.273	379987.923	Helical bend
$C2$	31.83	820.952	996488.978	U-bend
$C3$	1.000×10^4	817.643	1013006.585	Near-straight bend
$C4$		817.643	1013006.584	Straight (all modes)
$C5$		817.643	1013006.584	Straight (nodal-circle)

Table 8.3: Geometry of bent tubes C having $d = 310.0$mm, $\mathcal{R}_0 = 5.842$mm, with first resonance as calculated with the model stated.

the same qualitative properties. A third set of tubes with length $d = 310$mm and both bend radius and radius ($\mathcal{R}_0 = 5.842$mm) as A was chosen, allowing us to observe the effect of the length of bend when compared with shorter but otherwise similar tubes; the length was chosen such that the second resonance of tube C approximately corresponds to the first resonance of tube A. We see that lengthening the bend has had little influence on the effect of the curvature.

Previous theoretical and experimental studies in the literature [6, 7] report an increase in the frequency of the resonance; this effect is seen here, with a similar order of magnitude of ~ 10cent.

8.4.3 Concluding Remarks

A numerical model of lossless toroidal bends has been formulated for input impedance calculations and the effect of these bends has been investigated. A small increase in resonance

frequency and decrease in impedance magnitude has been observed in agreement with previous studies, though this model is lossless and therefore not entirely realistic.

The inclusion of losses is discussed briefly in section 3.6.2, and is geometrically involved. Once losses are included in the model (lossy effects are stated to be greater in bent tubes; the results here make no account of any losses), we will have a better idea of the true effect of curvature. With the above results we may speculate that, in the case of brass instruments, the effect will be small enough to neglect even if losses are included.

A model of the 88H was modified to include the bends in the main and tuning slides and the impedance calculated. To ensure a controlled test, this result (with realistic bend curvature) was compared to the same instrument with very small (i.e. near-straight) curvature. The effect of the bends is then isolated, and proves to be very small, with peaks differing by $\sim$ 0.01Hz. A result including losses in the bend will be of greater interest – if it also demonstrates an effect of this magnitude, we will have evidence suggesting that the bends of the instrument are not directly significant. In practice some indirect effects, such as an alteration to the geometry of the tube as it is bent, may be musically significant.

8.5 Computation Times

A simple instrument was used as a benchmark, comprising a cylinder, a discontinuity and a cone. Repeated impedance calculations were made, and the computation time recorded. Varying numbers of modes were included in both the straight and bent models. Calculations were performed on a Pentium 4 PC running Linux.

As we see in figure 8.6, the inclusion of higher modes causes the computation time to increase roughly as the square of the number of modes, owing to the (generally dense) matrix multiplication taking place. The bent ducts take roughly six times as long to compute than do straight ducts, due to the more complex calculations (including finding the eigenvectors of an $N_m \times N_m$ matrix) required. For comparison, plane-wave transmission matrix calculations took 0.007s, whereas the corresponding single-mode calculations in the straight and bent multimodal methods took 0.29s and 8.04s respectively. This is partly due to the conical section being treated analytically in the transmission-line, but as a series of short cylinders in the multimodal methods.

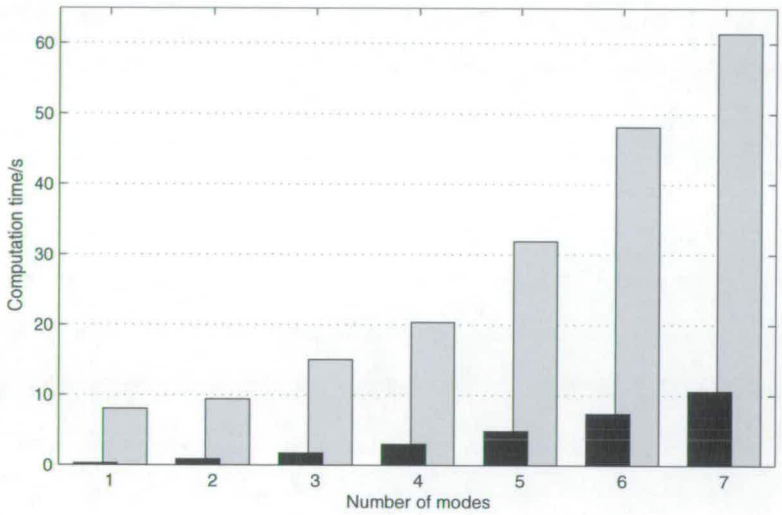


Figure 8.6: Computation time for simple benchmark instrument with varying number of modes included, using straight multimodal method (black) and bent method (grey).

Chapter 9

Optimisation Results

This chapter gives results found through the use of the optimiser described in Chapters 4, 5 & 6. The performance of two different optimisation algorithms is compared. The superior algorithm is then used to reconstruct instrument geometries from both theoretical and experimental results, and to modify the bore of a real trombone for the purposes of design optimisation. Certain problems found through use of the optimiser are discussed.

9.1 Visualising the Design Space

Given a method for impedance calculation and a suitable objective function (section 5.2.1), we can investigate the behaviour of the objective $\mathcal{O}$ as a function of the design variables $\boldsymbol{\alpha}$. It is desirable for this function to be as smooth and well-behaved as possible (e.g. figure 4.2), as this will allow the optimiser to operate effectively; a jagged, discontinuous, or 'bumpy' objective function provides a more difficult optimisation task.

A simplified trombone, consisting of a cylinder and a single Bessel-horn connected by a preset tuning-slide section, was created as a test subject. A target impedance curve of a more detailed virtual trombone was set (i.e. a perfect match was outside the design space). The design space consisted of 6 variables: Cylinder length and radius, bell mouth and throat radii, bell length and ball flare coefficient. The design space was then systematically traversed point-by-point along a regular grid, varying two variables at a time and holding the others constant. This provided plots of the objective as a series of functions of two variables. The objective function was restated as a maximisation problem with a percentage score for

convenient plotting. Each of the plots is a two-dimensional slice through a 6-dimensional space; it therefore represents a small part of the space, but allows us to get an overall picture of the behaviour of the objective function $\mathcal{O}$. To the author's knowledge, no previous analysis of this kind has been conducted on musical instruments.

Consider first figure 9.1a, which plots the objective as a function of the lengths of the two sections. This takes the form of a ridge, with a line of high values (red – a good match) dividing a region of low values (blue – a poor match). The same plot is also shown in 3-D in figure 9.1b. The clear shape of the ridge is to be expected; the impedance peak frequencies are strongly dependent on the total length of the instrument, and the ridge follows a line of constant total length. Some regions of the ridge have a higher fitness value than others; the relative lengths of the two sections affects the relative locations of the peaks, and therefore the fitness (compare with figure 8.19 in [2], which shows how the peak locations are affected by the relative lengths of a cylinder and a cone).

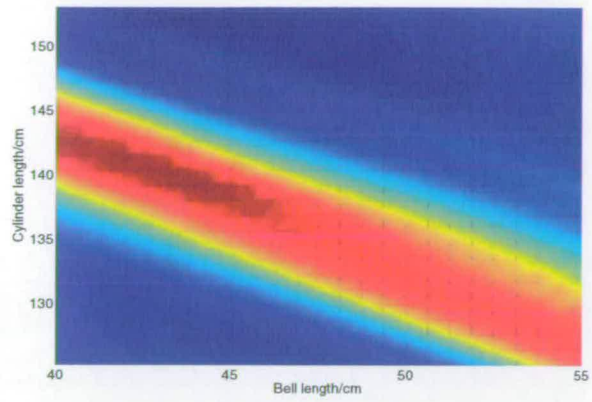
Figure 9.1c shows a similar basic ridge structure, but with two clear maxima at the top and bottom of the plot, and a saddle-like structure in between; likewise figure 9.1d has three peaks. Figure 9.1e shows a single, elongated peak, whereas figure 9.1f shows a somewhat more complicated structure; several peaks lie on a saddled ridge running right-left, intersected by a second diagonal ridge in the lower-left corner.

On the whole, we can conclude that the design space is relatively smooth, but possesses many local maxima.

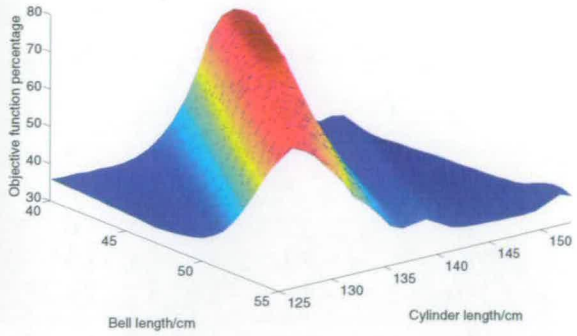
9.2 Bore Reconstruction Results

9.2.1 Problem Definition

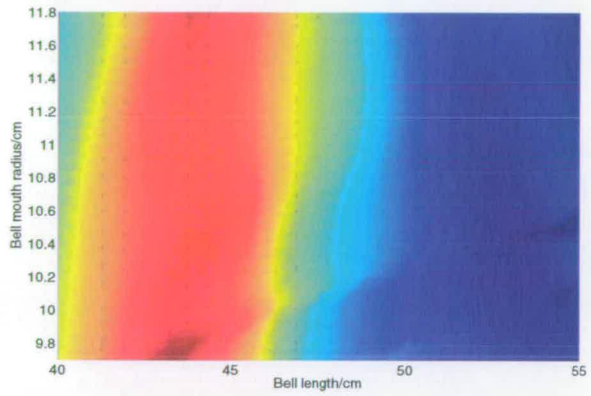
In order to set up and test the optimiser, some problems must be defined. The most straightforward problems are those we term *perfect-feasible*. A perfect result is known to be within the design space (and therefore feasible), so a good optimiser will be able to reproduce it to a close tolerance. Such problems can be defined by setting up an instrument template, specifying values for each design variable (often to give an approximation to a known real instrument bore) and calculating its impedance curve; this curve is then set as the target for the optimiser, which should then be able to reproduce the specified design values. In general



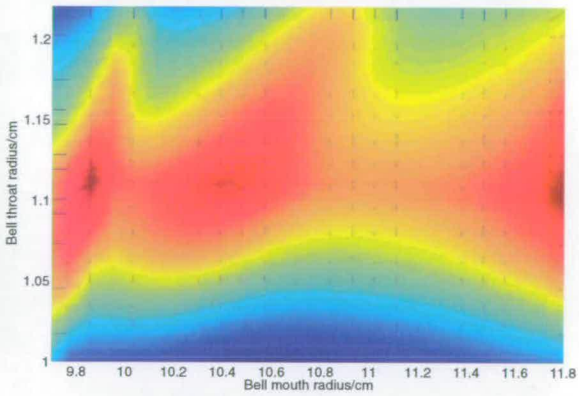
(a) Bell & Cylinder length (colour-plot)



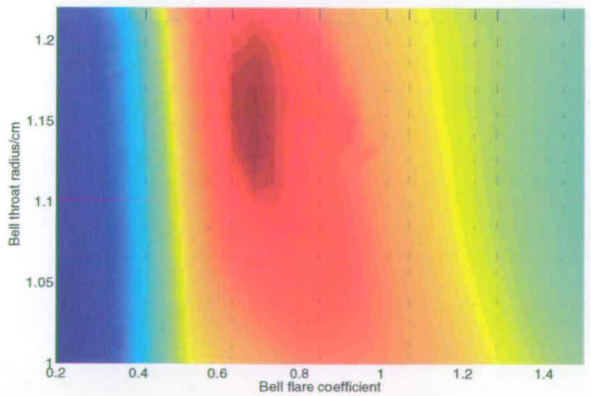
(b) Bell & Cylinder length (3D-plot)



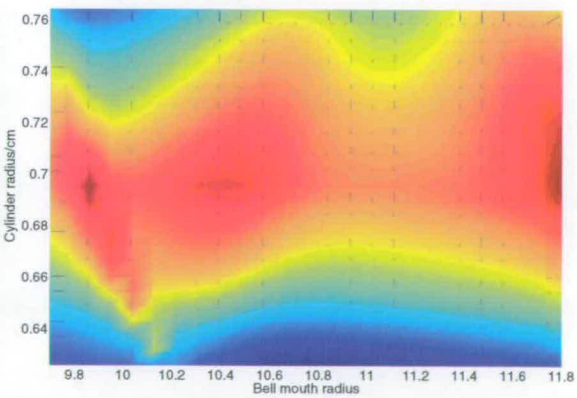
(c) Bell length & Mouth radius



(d) Throat & Mouth radius



(e) Throat radius & Bell flare



(f) Cylinder & Mouth radii

Figure 9.1: Plots of objective function values given by varying pairs of design variables.

usage, problems will not *a priori* be perfect-feasible (but may, in principle, be found *a posteriori* to be so), but such problems are a useful step for testing and tuning the optimiser; any results can be compared against the known 'correct' result, and discrepancies used to improve the procedure.

No constraints are placed on the smoothness of the bore. This is particularly relevant in a multi-Bessel bell, where a large discontinuity in gradient may arise as a result of the interpolation. In practice, we find that the optimiser performs better without such a constraint, and only returns unsatisfactory solutions with such gradient discontinuities in problems which it cannot otherwise solve (see section 9.4.1).

It would be desirable to have a proof of the uniqueness and one-to-one relationship between a bore and its impedance curve. In the absence of such a proof, we nevertheless neglect the possibility of two significantly different bores giving, to a close tolerance, the same impedance curve. In practice there is very little chance of such duplicate solutions occurring, so even without a uniqueness theorem we proceed under the assumption that solutions are unique.

Perfect-infeasible problems arise because of discrepancies between the design and the target; these discrepancies may arise in the geometry of the design, or in the objective function (i.e. impedance calculation). Theoretical problems may be based on calculated impedance curves of target instruments with a greater level of complexity than the template permits – the measured bore of a real instrument will be complex, and a template may approximate certain areas with more simple shapes. Perfect solutions are therefore not feasible (under the above uniqueness assumption) as a result of this approximation. Experimental problems set the experimentally-measured impedance of a real instrument as the target; a discrepancy arises if the theoretical and experimental results do not concur (as is presently the case), and likewise the simple-template discrepancy may also arise. Performance optimisation problems attempt to modify an existing instrument by specifying modifications to its impedance properties; the solution is unknown, and a satisfactory match may not exist.

Table 9.1 is an index of the bore reconstruction experiments performed.

Number	Description	Model	Target Data
I	Reconstruct simplified virtual trumpet	Spherical-wave	Theoretical
II	Reconstruct virtual trumpet	Spherical-wave	Theoretical
III	Reconstruct simplified virtual trombone	Spherical-wave	Theoretical
IV	Reconstruct virtual trombone	Spherical-wave	Theoretical
V	Reconstruct virtual trombone	2 modes	Theoretical
VI	Reconstruct virtual trombone	2 modes	Experimental

Table 9.1: Table of bore reconstruction experiments.

9.2.2 Genetic Algorithm

A virtual trumpet had its cylindrical section smoothed to remove 'kinks' due to the valves, tuning-slide etc. The optimisers were asked to reconstruct the bore of this instrument from its calculated impedance (experiment I – perfect-feasible). The kinks in the cylindrical section were then replaced, altering the target impedance (expt II – perfect-infeasible) accordingly. The instrument had a reconstruction prototype of a cylinder attached to a multi-Bessel bell; the bell length was a variable but each Bessel-section was an equal fraction of this length. Wide constraints were set on all variables and objective function $\{1, 1, 0\}$ was used (recalling the notation defined in section 5.2.1).

After considerable experimentation, the best results were found by optimising a 4-Bessel bell ($N_v = 11$) for two 50-individual generations, then, for each member of the population, splitting each Bessel-section into two equal parts, giving an 8-Bessel bell ($N_v = 19$), and optimising for a further 18 generations. Results were inconsistent, not always producing good matches when experiments were repeated with the same parameters. Experiment I (at best) gave a good match to both impedance and bore in ~5 minutes; experiment II gave very poor results (figure 9.2 gives an example of a failed optimisation). Although the GA performance is a significant improvement on that in [19], it is still neither stable nor capable enough to be practicable. As we shall see, the GA is again outperformed by Rosenbrock; given that GA's generally perform best in low-N_v problems such as that formulated here, we can conclude that further use of genetic algorithms for for brass instrument optimisation is not worthwhile.

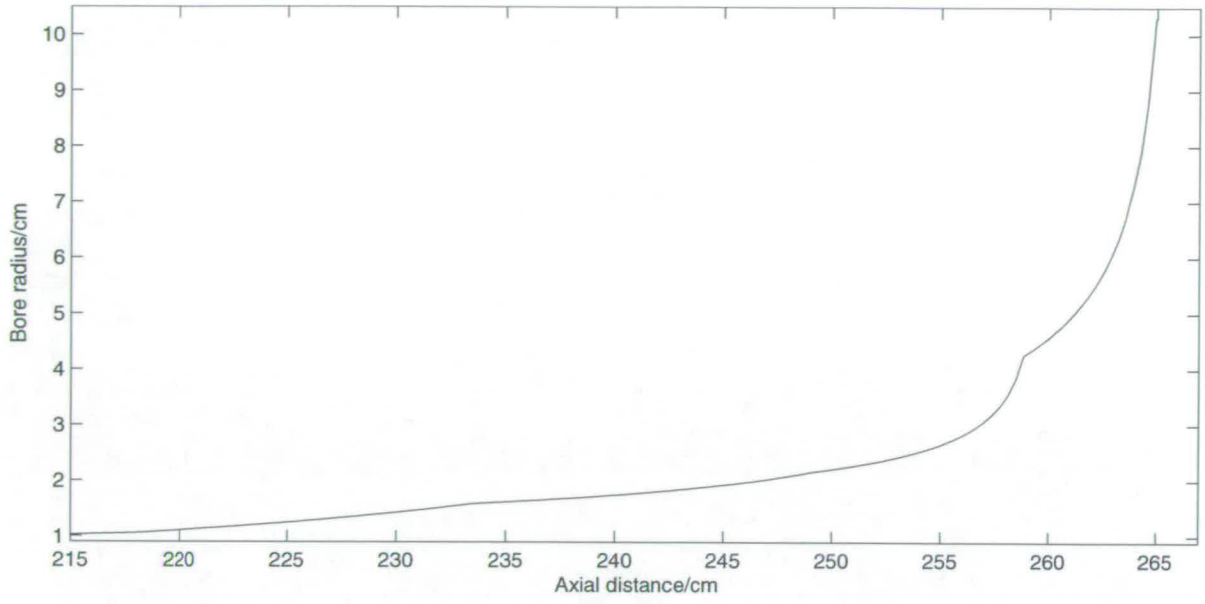


Figure 9.2: Example of an unsatisfactory bore-profile from a poorly-converging optimisation.

9.2.3 Rosenbrock

Experiment I was repeated using the Rosenbrock algorithm, with a 4-Bessel bell ($N_v = 11$). Results were in every way superior to the GA: a near-perfect match of both impedance and bore was found with total repeatability. Experiment II was then attempted, and is perfect-infeasible due to the kinks in the target's cylindrical section, which were intentionally not accounted for in the reconstruction template. In spite of this, an equally-good match of the bore (excepting the kinks in the cylinder) was found – the closest impedance match was still very close to the correct bore-profile, suggesting that this optimiser is capable of reconstructing an instrument's bell without any knowledge of its valve sections.

Having established the clear superiority of Rosenbrock over the genetic algorithm, we move on to more challenging problems. The template in table 5.1 was used to reconstruct the 88H. First, the perfect-feasible experiment III was conducted, with target impedance generated from a virtual instrument (including mouthpiece) fitting the template. A near-perfect match was found; the most significant discrepancy lay in the location of the discontinuities – this is examined in more detail in section 9.4.3, and was of order 3mm (figure 9.3a). The error in the bell section was generally less than 0.01mm, and was at worst 0.25mm at the extreme bell end – the model is very insensitive in this region. The optimiser is therefore declared a

success, having been proven capable of finding a very fine reconstruction of a perfect-feasible target, with the given problem of the discontinuities. Objective function $\{1, 1, 1\}$ was used, with very strict parameters $\nu = 100\text{k}\Omega$, $\nu_\phi = 10\text{Hz}$, $\mu_\phi = 20$, $\nu_Z = 0.05$, $\mu_Z = 10, N_{pk} = 25$, taking 1500 evaluations and approximately 15 minutes.

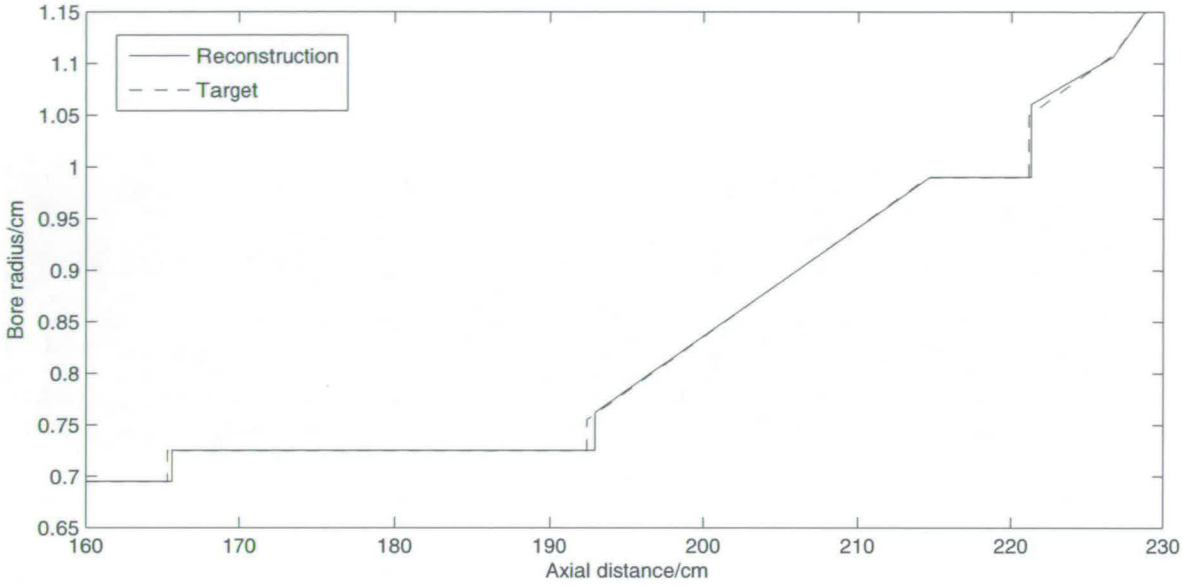
Experiment IV set as the target the impedance of a virtual 88H – i.e. a more detailed bore-profile than the template could model exactly, giving a perfect-infeasible problem. The reconstruction was only slightly less close than than for III, but still within acceptable tolerances. The discontinuous section, in which a perfect match was impossible due to the coarseness of the template in this area, had similar discrepancies (figure 9.4a), and the bell section had an error of generally less than 0.01mm, with a maximum error of 0.5mm. The reconstruction required 2037 evaluations and 20 minutes, using parameter $\nu = 300k\Omega$, with the others unchanged from III. This result is proof that the optimiser can reconstruct the geometry of a real instrument from calculated impedance data.

9.2.4 Reconstruction with Higher Modes

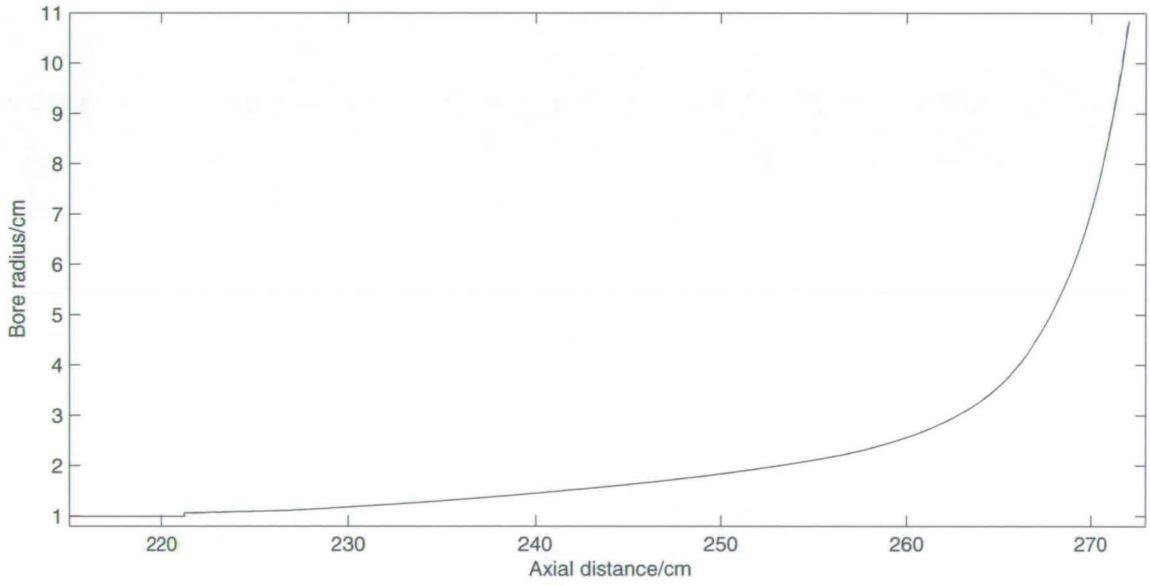
Experiment IV was repeated, this time with both target and candidate impedance curves calculated using the multi-modal method, with 2 modes included (experiment V). Results of comparable quality and geometry to the plane-wave case were found. This result is proof that the effectiveness of the optimiser is not significantly affected by changes to the model used for impedance calculation. The choice of 2 modes is arbitrary, but offers the best compromise between computation time and improved accuracy (figure 8.2). Due to the high number of cylindrical sections needed to model each conical section to sufficient accuracy, the computation time is very much larger than for the plane-wave; 2477 evaluations took a total of $\sim$ 20 hours of computation – i.e. approximately 50 times the corresponding reconstruction with the spherical waves. For the multimodal method to be sufficiently convenient for design work, the computation speed must be improved.

9.2.5 Reconstruction from Experimental Data

This is the ultimate challenge of any optimiser; to reconstruct the geometry of a complete instrument from experimentally-measured impedance data. In order for any designs output

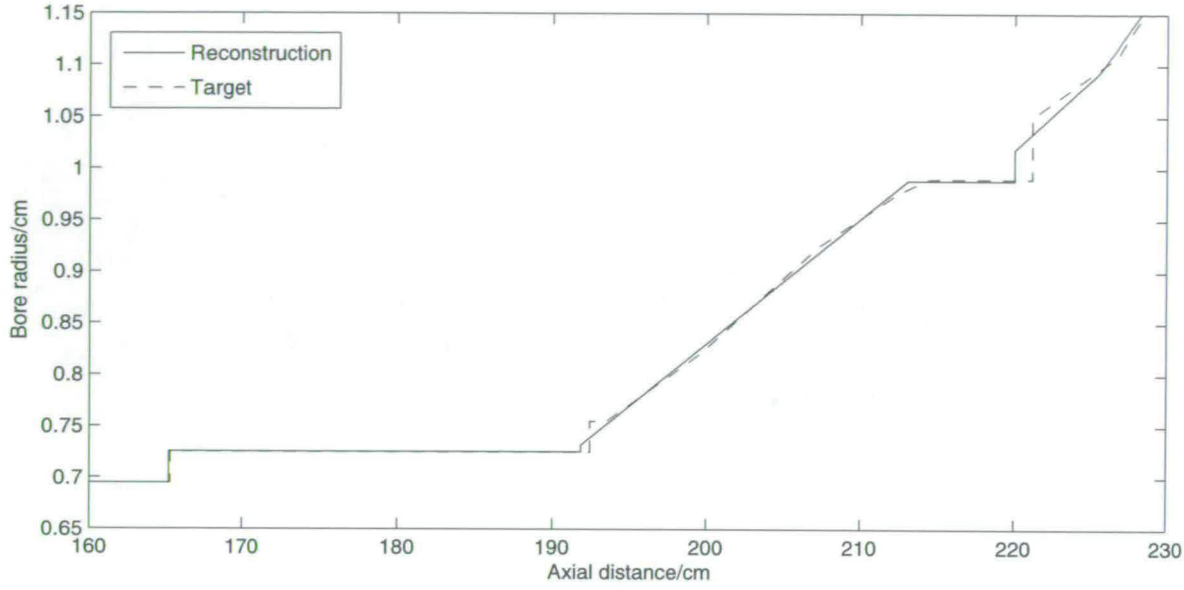


(a) Middle section

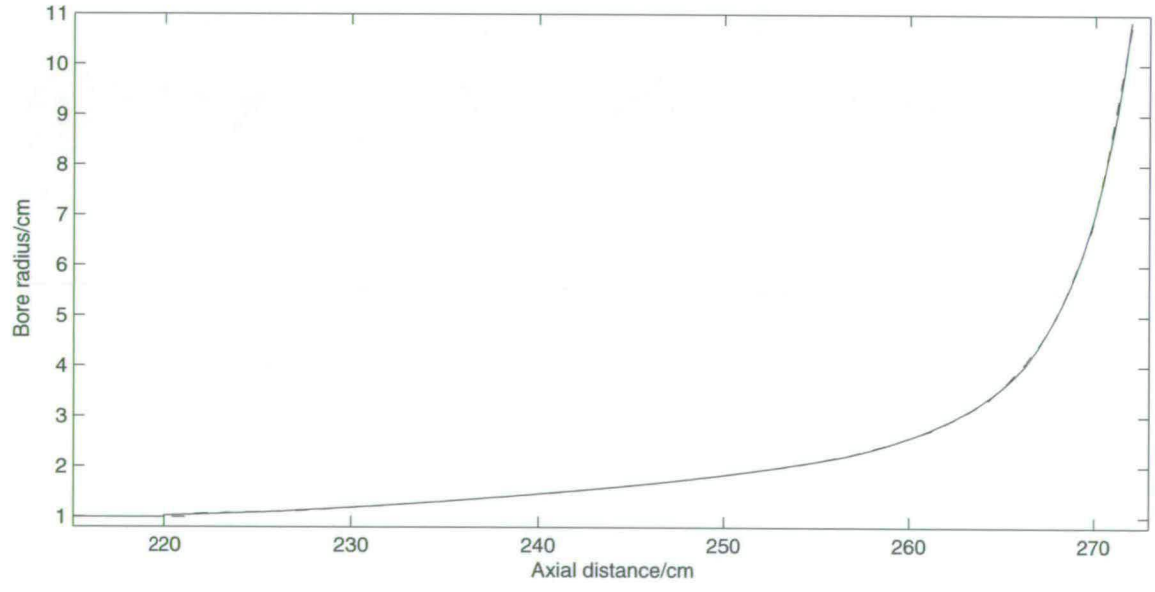


(b) Bell section

Figure 9.3: Parts of reconstructed 88H trombone approximation (expt III).



(a) Middle section



(b) Bell section

Figure 9.4: Parts of reconstructed virtual 88H (expt IV).

from the optimiser to have any relevance to real-world construction, the optimiser must move from the experimental to the theoretical domain without a significant loss of accuracy in the instrument geometry.

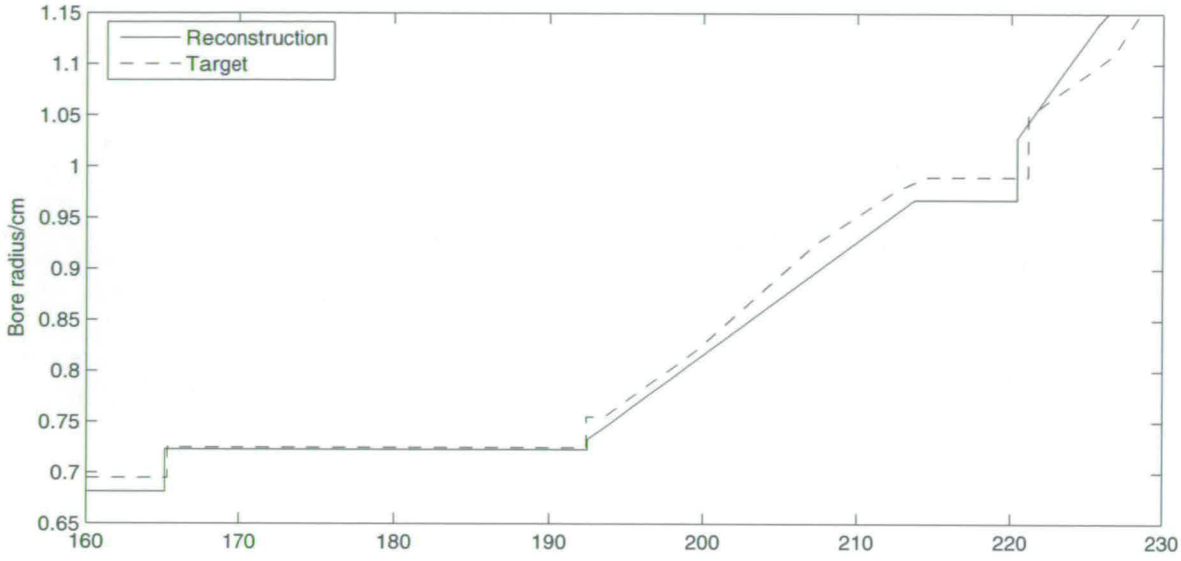
To test the capability of the optimiser to solve problems of this kind, a BIAS measurement of the 88H was set as a target, and the above trombone template used above was once again employed (experiment VI). The most successful result took 1463 evaluations and ran overnight, using objective function $\{1, 1, 0\}$ with parameters $\nu = 100\text{k}\Omega$, $\nu_\phi = 10\text{Hz}$, $\mu_\phi = 20$, $\nu_Z = 0.05$, $\mu_Z = 10$. 2 modes were used in the calculation. The error between the reconstructed and 'correct' geometry was, aside from the extreme bell region, less than 0.5mm, and no greater than 0.1mm for much of the length of the instrument; however, in the final 8cm of the bell, the reconstruction differed by ~1cm. The bell geometry is a smooth and plausible curve.

On the whole this is an encouraging result. The optimiser has successfully converged to a good, if not perfect, reconstruction of the entire instrument in a strongly perfect-infeasible problem. Given the discrepancy between theory and experiment (section 8.3), the inaccuracies found in the extreme bell region are to be expected. The remainder of the instrument was reconstructed to within a satisfactory tolerance. However, this tolerance is not fine enough to guarantee sufficient accuracy for detailed design work. Although the optimiser has proven to be sufficiently powerful for such problems, the theory-experiment discrepancy is a significant handicap, and until this discrepancy is resolved, or at least reduced, the required design tolerance is unattainable in problems of this kind.

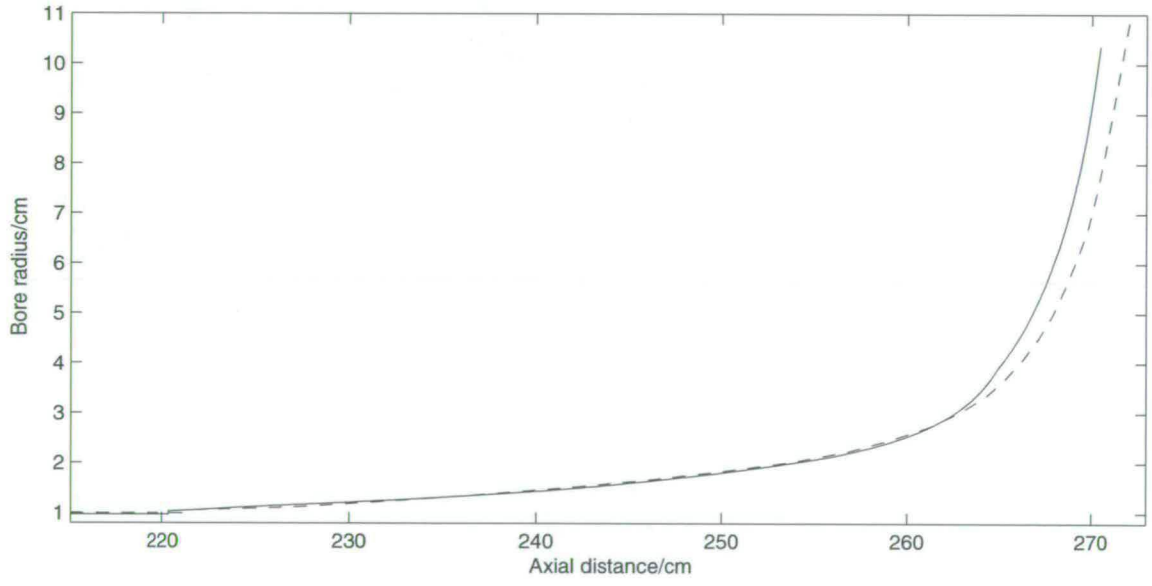
9.3 Performance Optimisation Results

Number	Description	Model	Target Data
VII	Tuning peak 2	Spherical-wave	Modified theoretical
VIII	Increasing height of peaks 8-10	Spherical-wave	Modified theoretical
IX	Tuning peak 2 of F-attachment	Spherical-wave	Modified theoretical

Table 9.2: Table of performance optimisation experiments.



(a) Middle section



(b) Bell section

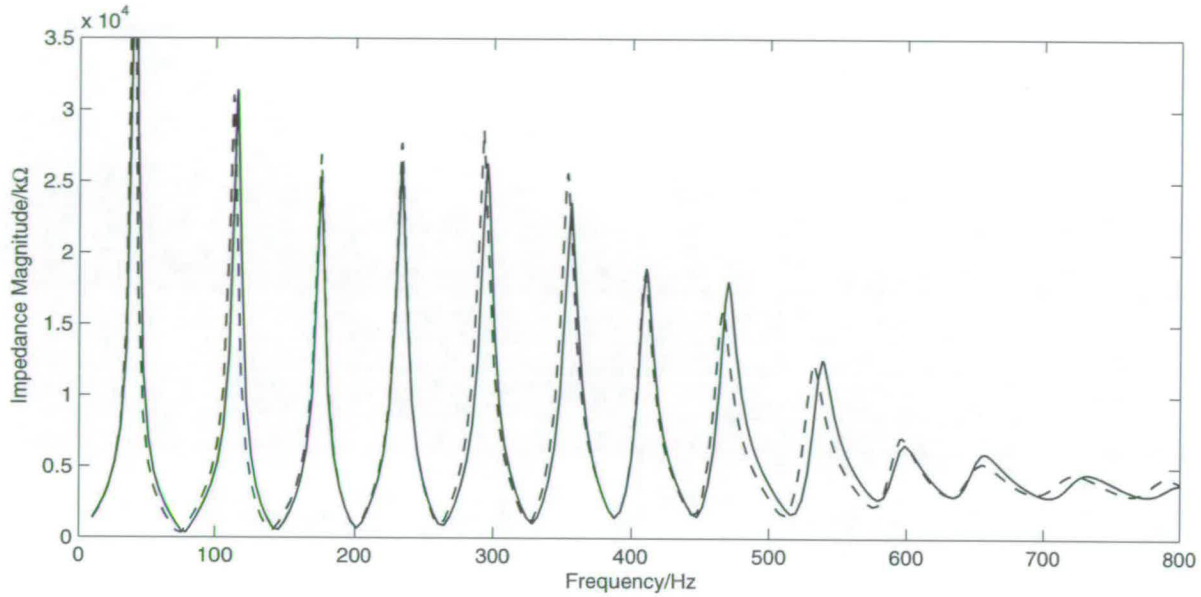
Figure 9.5: Parts of 88H trombone reconstructed from BIAS data (expt VI).

9.3.1 Tuning Peak 2

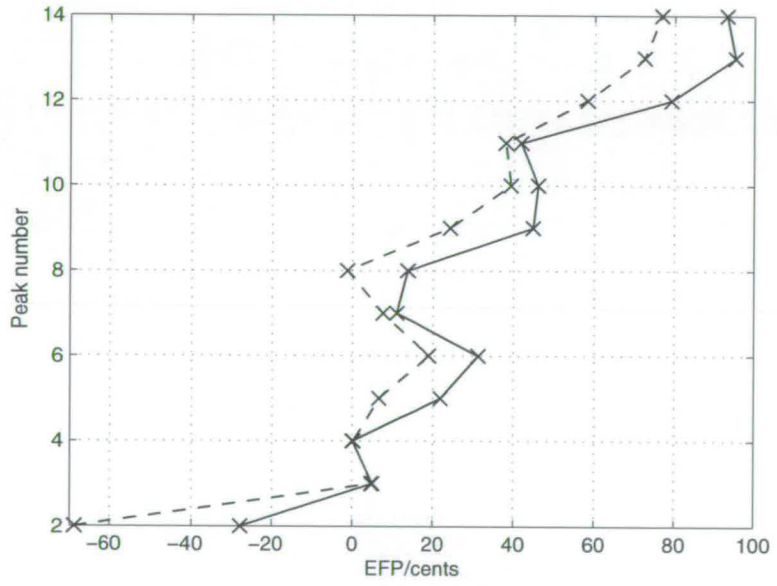
It was established in Chapter 7 through experimental comparison of a tenor and bass trombone that the tuning of peak 2 relative to the higher even-numbered peaks (particularly peak 4) has a significant effect on the quality of the lower playing register of the instrument. In this experiment, we take an existing tenor trombone design and attempt to improve its lower register by shifting this peak into a closer alignment without compromising any other resonance properties.

Experiment VII takes again the template for the 88H (a large-bore tenor). The impedance of the instrument was calculated with the spherical-wave model and the peak frequencies set as an optimisation target, with the following modifications: the target frequency of peak 2 was shifted from 112.0Hz to 115.0Hz (a shift of +45c), and peak 5 shifted from 292.5Hz to 295.0Hz (a shift of +15c). It was found that allowing freedom in the location of peak 5 granted greater flexibility in the location of peak 2. Note also that it often proved advantageous for target shifts to be greater than that actually desired, allowing the optimiser space in which to compromise. The 'correct' geometry for the 88H was set as the starting point. Objective function $\{0, 1, 0\}$ was used (i.e. the peak frequencies were optimised without magnitude data), with tuning parameters $\mu_\phi = 10$, $\nu_\phi = 10$Hz. Only peaks 2, 3, 4 and 5 were optimised over – i.e. no targets were set for the other peak locations. The optimisation took 1516 evaluations and 14 minutes.

As we can see in figure 9.7, the optimiser has converged to a solution which closely matches the target set. Peaks 2 and 5 are now located at 114.7Hz and 295.1Hz respectively. Peaks 3 and 4 have not been significantly altered. We see from the EFP plot (figure 9.6b) that the desired re-alignment of peak 2 has occurred. As a compromise, the higher peaks have all been shifted by between +5c and +20c; these are all rather smaller than the +41c shift of peak 2. The impedance magnitudes of the peaks have been affected, notably peaks 3-6, which have been shortened, and peaks 8 & 9, which are taller. Bearing in mind the discussion in Chapter 7, it is reasonable to suggest that this optimised trombone would have a significantly improved lower range (due to the alignment shift of peak 2) and an improved upper range (due to the taller peaks). The effect on the middle range is more difficult to predict; the sharper upper-range peaks will tend to reduce the harmonic content in the radiated sound

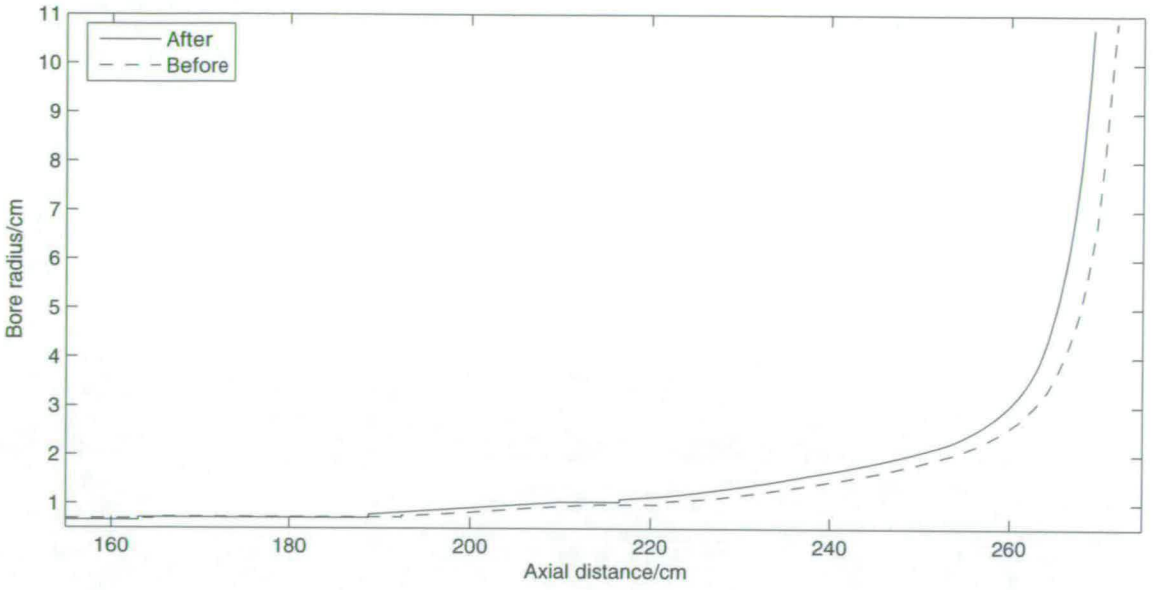


(a) Input impedance



(b) EFP (F=58.28Hz)

Figure 9.6: Results of tuning peak 2 of 88H before (dashed) and after (solid) optimisation (expt VII).



(a) Bell section

Figure 9.7: Bore profile from tuning peak 2 of 88H before (dashed) and after (solid) optimisation (expt VII).

and make it more difficult to play, but the increased peak height would tend towards the opposite effects. Which of these will dominate is unclear at this point.

We now turn our attention to the optimised bore profile figure 9.7a. Firstly, note that the bore design is smooth, entirely plausible, and could be built by a manufacturer with no additional difficulty; it can be described as a modified version of an existing design (indeed a very popular one). The bore of the cylindrical section (i.e. the main slide) has been reduced from 0.695cm (or, in the conventions of the instrument industry, a 0.547" diameter bore) to 0.663cm (0.522" diameter). The overall instrument length is some 2.6cm shorter than before, and the bell contour has been subtly altered, as has the taper of the tuning slide. Given the predicted performance in the previous paragraph, these results are surprising; we are attempting to replicate a feature of a bass trombone and so we would have expected the instrument to have been made generally larger in radius.

Clearly the trombone design space is complicated. Certain regions of high performance are well-established in the industry (for example the American large-bore paradigm to which the 88H and R4 belong [51]), and it may be that the optimiser has uncovered a region which might have otherwise been unexplored. Exactly how the optimised trombone will play, and

the desirability of same, will remain unknown until an example has been built and tested. The design given here is at least worthy of such experimentation, and a proof of the capability of this software to perform whole-instrument intonation optimisation.

9.3.2 Modifying Peak Magnitudes

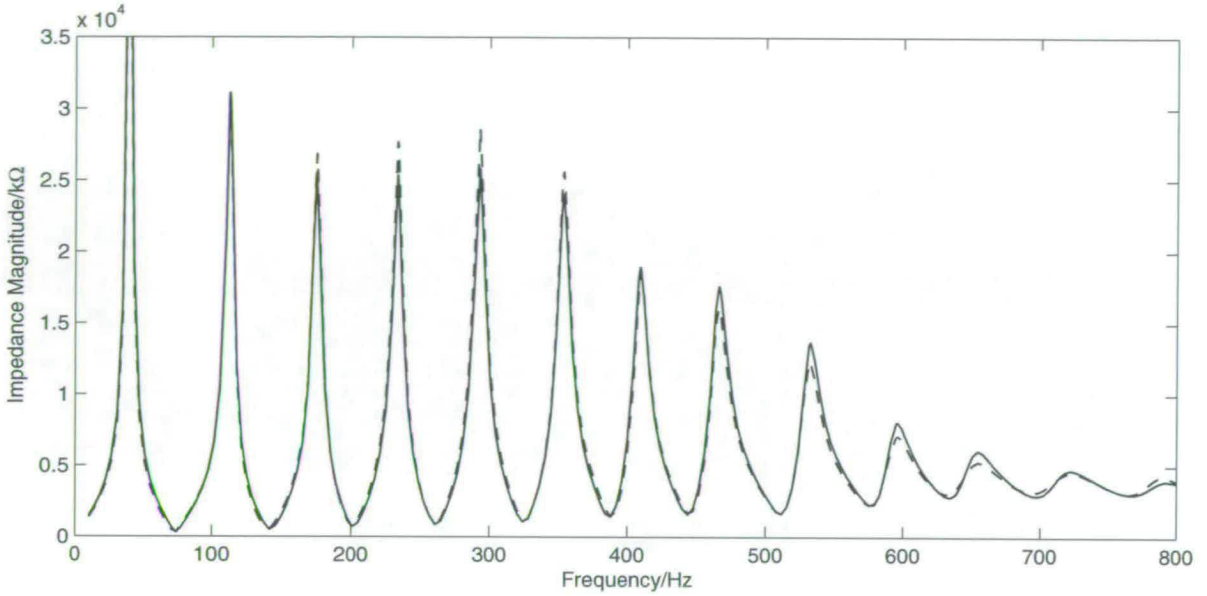
The 88H was again used, here in an attempt to modify the impedance envelope. Experiment VIII took the usual base target, this time modified such that peaks 8, 9 & 10 were targeted to increase in magnitude by 10%, and no target was set for the magnitude of peaks 1-7. The peak frequencies were left unchanged. Objective $\{0, 1, 1\}$ was used, with parameters $\nu_\phi = 10\text{Hz}$, $\mu_\phi = 20$, $\nu_Z = 0.1$, $\mu_Z = 10$, $N_{pk} = 10$. The optimisation required 1509 evaluations and 14 minutes.

We see from figure 9.8a that the optimisation was successful, increasing peak 8 from 16,110kΩ to 17,602kΩ, peak 9 from 12,358kΩ to 13,668kΩ, and peak 10 from 7,094kΩ to 8,064kΩ, respectively increases of 9.2%, 10.6% and 13.6%. Certain lower-index peaks were shortened, most notably peak 5 from 28,671kΩ to 25,820kΩ, a decrease of 9.9%. The frequency (and therefore EFP) of peaks 2-10 was not significantly changed; higher peaks were each made roughly 5 cents sharper. The bore-profile (figure 9.8b) again shows a reduction in the radius of the main slide from 0.695cm (0.547" diameter) to 0.665cm (0.524"), and some subtle changes (a maximum of 0.5mm) to the bell contour. The experiment was repeated but with 20% increases specified; the optimiser was unable to converge to a geometrically-satisfying solution.

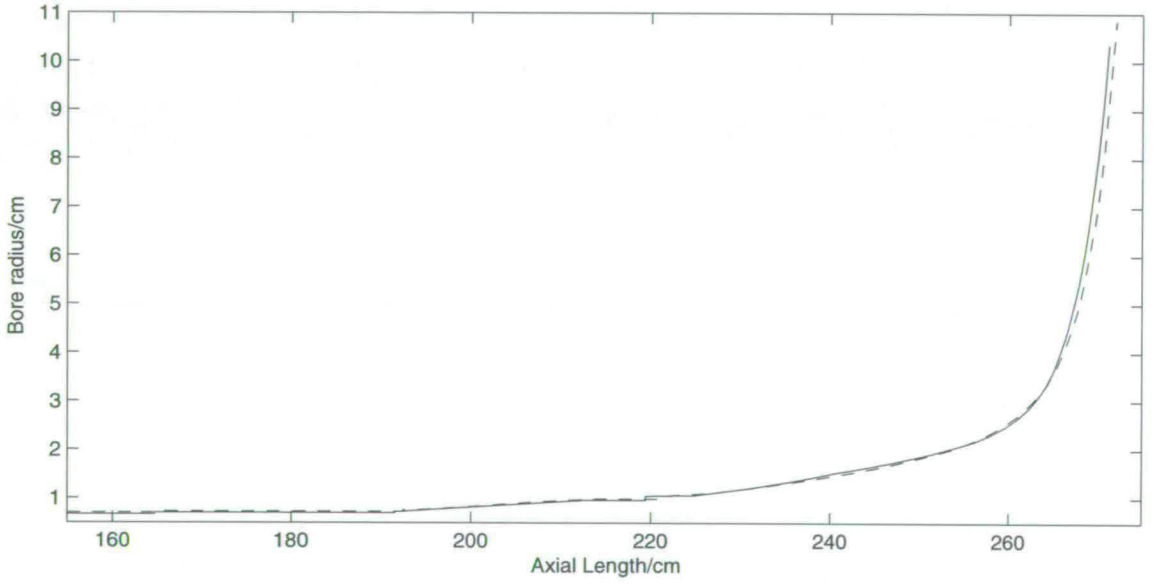
These results demonstrate that it is possible to modify the impedance envelope without modifying the peak intonation. The changes made here to the envelope are of similar proportion to those found in experiment VII, but had only a small effect on intonation. As a result of these changes, we would expect the optimised trombone to produce a somewhat brighter timbre than its predecessor, and possess a higher register which is somewhat easier to play.

9.3.3 Tuning F-attachments

It was shown in section 7.2.3 that trombones with the F-valve engaged have very flat (~150c) second peaks. We here attempt to modify the additional cylindrical tubing to rectify this



(a) Input impedance



(b) Bell section

Figure 9.8: Results of increasing the magnitudes of peaks 8-10 of 88H before (dashed) and after (solid) optimisation (expt VIII).

discrepancy without compromising other impedance features (expt IX). A design which is simple to manufacture is also a priority.

As a starting point, a virtual 88H had an additional 1m of cylindrical tubing added in the neckpipe area (the valve itself was assumed to be of constant bore, and bends were neglected). The impedance was calculated with the spherical-wave model and was used, with its second peak modified, as the optimisation target. We are interested in the low-frequency behaviour, so higher modes are not necessary and the transmission-line model will be sufficiently accurate for our purposes. Given that players will seldom, if ever, use the fifth and seventh resonances (note the prime index), the target location of these peaks was left free. Objective $\{0, 1, 0\}$ was used, with parameters $\nu_\phi = 30\text{Hz}$, $\mu_\phi = 20$, $N_{pk} = 20$.

A variety of different geometrical templates were tried. The radius of the extra tubing at each end was fixed, but all other parameters left free. The most successful template was a series of stepped tubes of equal length and varying radii (this would be relatively easy to manufacture when compared with other potential designs). The results, compared with the cylindrical tube, are shown in figures 9.9 & 9.10.

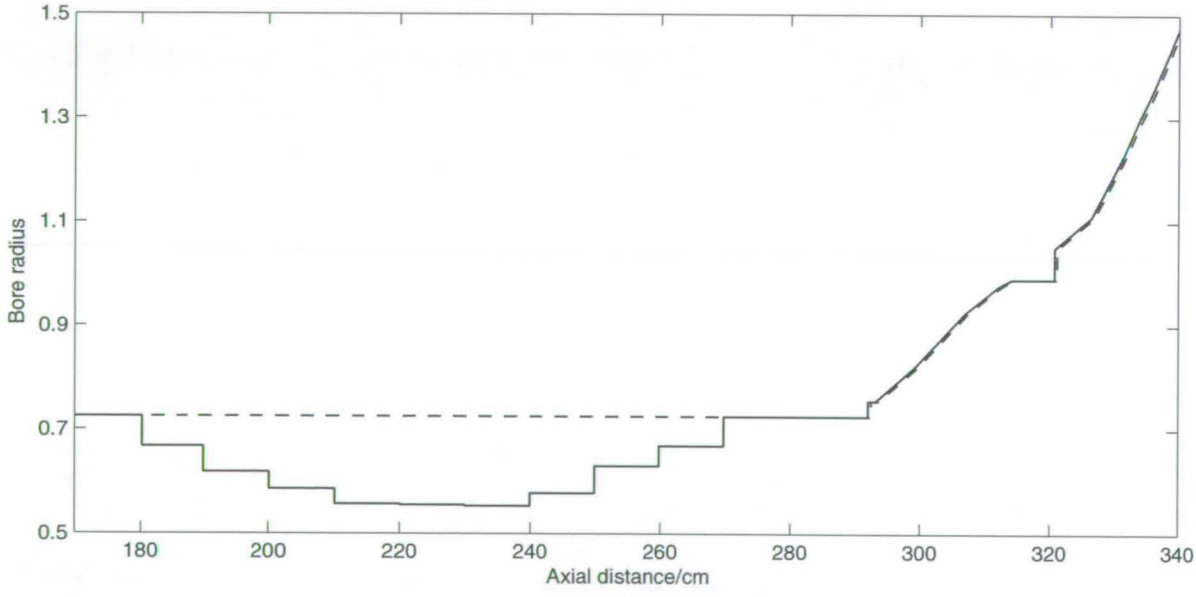


Figure 9.9: Bore profile of F-attachment tubing before (dashed) and after (solid) optimisation (expt IX).

We see in figure 9.10 that the second peak was successfully shifted by 125c from 77.1Hz to 80.0Hz; peak 4 was shifted by -13c, giving a net change in misalignment (i.e. the difference in

cents between the two equivalent fundamentals) of -139c. Peaks 5 and 7 were of no interest, and the peaks from 6 upwards were largely unaffected. However, the compromise needed was for peak 3 to be shifted by -24c, and substantially reduced in magnitude. One might speculate that this new geometry would make the low register (peak 2) rather easier to play, and the next register (peak 3) somewhat more difficult. In the absence of a more complete understanding of the impedance-performance relationship, a prototype must be built in order to examine the true playing characteristics of this revised instrument.

9.4 Difficulties Encountered

When designing an optimisation technique, certain features unique to a given problem will arise and must be considered on a case-by-case basis.

9.4.1 General Observations

As a general rule, it is almost impossible to modify a single feature of an impedance curve without affecting other features. The design problems solved above represent good compromises of a variety of features; they do not modify a single impedance feature, but rather modify all of them in such a way as to make certain changes negligible. This is a subtle distinction; it is not a question of finding the single change in bore to effect the correct impedance change, but rather making many changes in the bore so that the undesirable impedance changes largely cancel out, leaving only the desired change. There is an inherent lack of freedom, certainly within the continuum of realistic bore shapes, in the brass instrument optimisation problem, but this can be circumvented, to a certain extent, by careful use of the optimiser.

There are certain tasks to which the optimiser is not suited. For example, if we were attempting to alter a given trombone such that the intonation of a single peak is changed while keeping every other peak unmoved, the optimiser will generally not move away from the starting point. This is partly due to the aforesaid lack of freedom in altering the impedance curve, and partly due to the optimiser starting in a very strong local maximum and therefore having difficulty navigating away from it. This can be counteracted, at least in part, in several ways. Firstly, certain less important features can be left unspecified in the target; for

example, with trombones, the location of peak 1 is almost entirely irrelevant to the player, and can be left free without a loss of performance. Secondly, the number of design targets N_{pk} can be reduced (in principle this is not very different from the previous point, but in practice it is implemented differently). Thirdly, the optimiser may be modified in future to assign weights to each individual peak; at present they all score on an equal weighting, but in cases of this kind it would allow a greater 'incentive' for the optimiser to move away from the local maximum on which it starts. In general, such weighting is not necessary, but may prove a useful addition to the range of tools available.

The optimiser may be given a target which is outside of its capabilities. Most generally, this is a target which is not only unattainable within the constraints of the specified design space (this has already been described as 'perfect-infeasible'), but for which no reasonable approximation exists either. This is again a subtle distinction, but to the optimiser they are quite different, as the two problems will differ greatly in ease of convergence. If a reasonable match exists, the optimiser is likely to converge to it relatively easily; if no such approximate match exists, it is likely to explore the more extreme regions of the space in an attempt to find a solution (as any good optimiser might) and, as a consequence, will sometimes return unacceptable geometries (e.g. figure 9.2). If such results are returned, it can be assumed quite safely that the target is not attainable with the given template (and therefore constraints) and one or both must be modified to gain a more satisfactory result.

9.4.2 Displaced Peaks

If only the least-squares objective $\{1, 0, 0\}$ is used, there is a tendency for it to give disproportionately high scores to poorly-located peaks if the candidate and target impedance curves overlap.

In figure 9.11 are shown two peaks which demonstrate this point. Both are equally poor matches of peak location, but peak (b) will give a much higher least-squares score than peak (a), because the right-hand tail of the peak is close to the target; the peak is the wrong location and the wrong shape, but gains a good score, thus providing a spurious local maximum. In the hybrid objectives $\{1, 1, 0\}$ and $\{1, 1, 1\}$, this effect is generally avoided by the peak-frequency optimisation.

9.4.3 Non-Uniqueness and Insensitivity

Given the limited amount of information describing the target impedance, particularly if the least-squares whole-curve fitting is not used, it is to be expected that solutions may not be unique, and additionally that there may be local maxima describing close, 'incorrect' matches. This makes the optimiser's task more difficult, as it must either avoid or traverse away from these local maxima in search of the global optimum. Likewise, if there are two regions in the search space that provide two or more matches of nearly equal quality, it is difficult for the optimiser to choose between them without prior knowledge of the 'correct' solution (which would violate the purpose of the optimisation). A complete study of the uniqueness problem is beyond the scope of this work, but we can demonstrate the problem with numerical results.

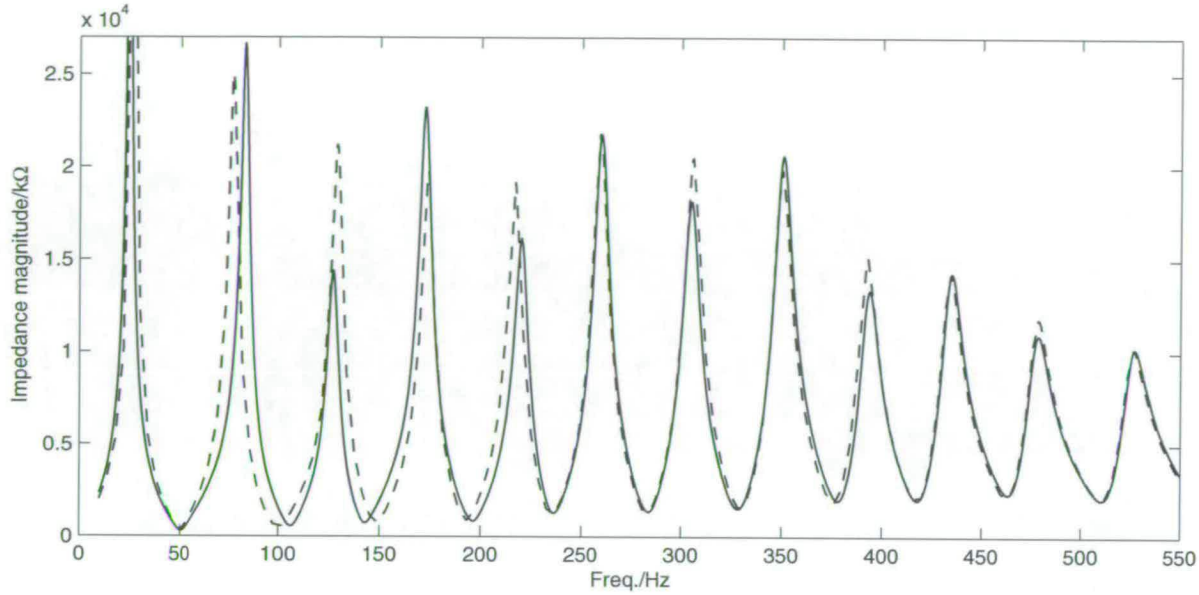
The impedance of an approximation of the 88H was calculated. This instrument has several jump discontinuities in its bore, resulting from the various tuning slides and fixings required of practical instruments; here we focus on the discontinuity between the main slide and the neckpipe, which is a join between two long cylindrical sections of slightly different radius. The virtual instrument was then modified such that location of this jump was altered (fig. 9.12), but the total length and the radii of the cylinders were unchanged. The jump was moved, by respective shortening and lengthening of the cylinders, by 1cm in either direction. This is a large change in the scale of the problem, but resulted in only a minimal change in impedance of ~ 0.1Hz and ~ 20kΩ – the impedance is comparatively insensitive to this change. The optimiser may then choose 'incorrect' values for the location of this jump, and make much smaller changes in a more sensitive area to compensate. This is compounded by the known tendency of the Rosenbrock algorithm to converge poorly when very close to an optimum.

In considering the results of reconstructions featuring bore discontinuities in otherwise fairly straight sections, it must be accepted that a small degree of error results from the inability of the optimiser to resolve such discrepancies.

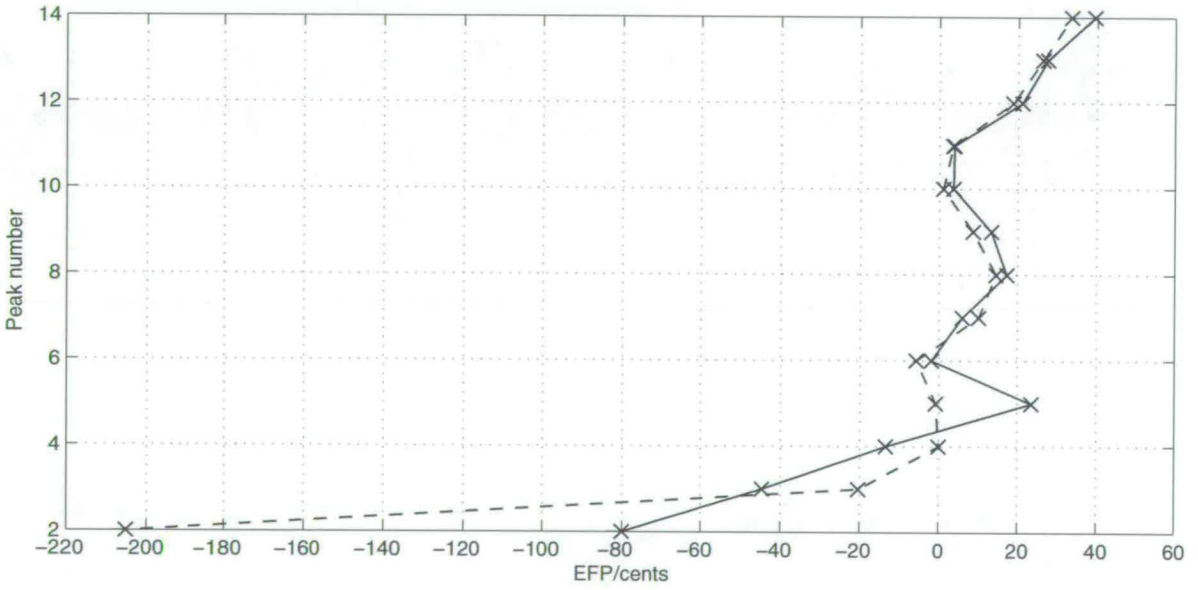
9.4.4 Multi-Bessel Bells

One disadvantage of describing a bell with multiple Bessel-horns is the non-uniqueness that results. For a given horn shape, there is a continuum of piece-wise-Bessel horns which are

close interpolations. To the optimiser, this will appear as a series of very similar local maxima which may differ only fractionally in quality. In practice, we are not interested in how the bore-profile of a result is described, but rather its shape, and therefore a user will not consider this problem with any great interest.

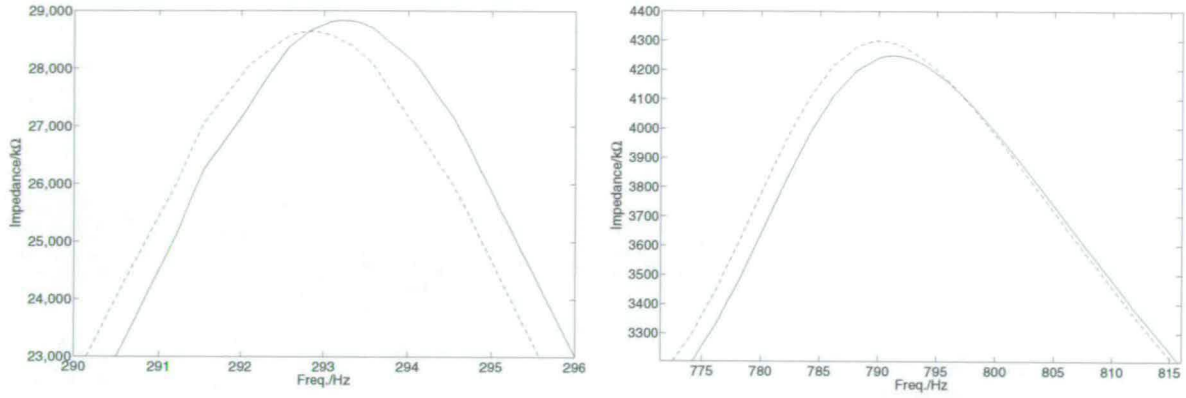


(a) Input impedance



(b) EFP (F=43.4Hz)

Figure 9.10: Plots of 88H with F-attachment before (dashed) and after (solid) optimisation (expt IX).



(a) Poor match

(b) Good match

Figure 9.11: Plots of 2 equally-poor candidate peaks (solid) compared with targets (dashed) gaining quite different objective scores, demonstrating the 'displaced peaks' problem.

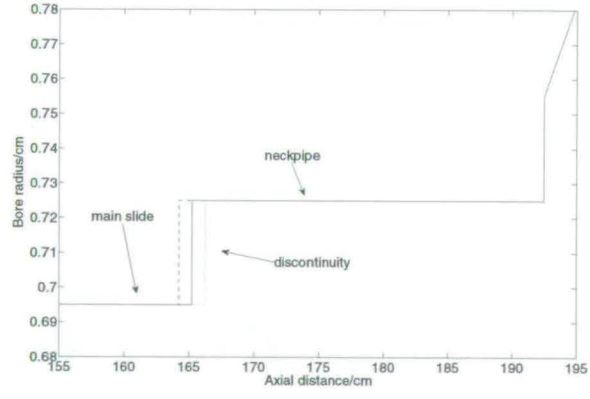


Figure 9.12: Part of bore profile showing modifications to location of jump discontinuity.

Chapter 10

Conclusions

10.1 Fulfilment of Aims

10.1.1 Aim 1

To review existing methods of input impedance calculation, and to derive appropriate methods for bent tubing.

A full review of the derivation of the plane-wave transmission-line model was given, stating wall effects and radiation impedance. New transmission matrices for Bessel and exponential horns were derived. An alternative derivation of the multi-modal method was given, and used as a basis for the derivation of a model for bent waveguides, which extended previous work to produce input impedance matrices compatible with the existing multi-modal method. Results for bore discontinuities and radiation impedance were extended to include the nodal-diameter modes. The multi-modal wall loss mechanism was briefly reviewed for the purposes of highlighting the contradictory boundary conditions on which the model is based.

A comparison of the numerical results for the different models was conducted, using a complete instrument including a mouthpiece. In agreement with previous studies, the addition of higher modes has a significant effect, diminishing progressively with each additional mode. The spherical- and plane-wave assumptions were compared with the multi-modal method; the spherical wave is shown to be an overestimate of the the effects of higher modes, but with a greatly smaller computational cost.

The newly-derived bends model was used to investigate the effects of curvature on tubing

resembling that in brass instruments. It was shown analytically that plane-wave propagation is not affected by tubing bends. Numerically, increasing the curvature has the effect of raising the resonance frequency and lowering the impedance magnitude. The model was included with the higher-modes model to simulate a complete trombone, including bends. The effect of the curvature was found to be clear but negligibly small; if a similar result is found once losses are included, we will be able to say definitively that bends have no direct effect on the impedance of a trombone. Further work is required to derive such a loss model.

10.1.2 Aim 2

To develop the underlying techniques used in brass instrument optimisation, particularly the instrument representation.

A number of innovations in brass instrument optimisation technique are presented. New objective functions are presented, as are methods for finding impedance peaks with minimal computational overhead. A novel template-based instrument representation is proposed and implemented, including models of instrument bells with a series of short Bessel-horns, providing close approximations to real instruments with very few design variables.

The template representation allows the geometry of an instrument to be defined without knowledge of its exact dimensions; if one knows the class of instrument (e.g. trombone), then a template can be defined based on prior knowledge. This can be used to constrain the search space to include only reasonable 'trombone shapes' without omitting any useful potential solutions; by thus reducing the search space, the optimisation speed is substantially improved. The use of multi-Bessel bell approximations, and therefore fewer design variables, further improves the optimisation speed. Use of the templates requires some *a priori* knowledge of the instrument, and reconstructions therefore lack the 'black-box' assumption; in practice this is seldom necessary, and the generality which is traded off is not generally required. Fully general problems can still be defined if needed.

Regardless of how the geometry is represented (e.g. point-wise, Bessel-horns), the template implementation has proved to be an elegant and very flexible solution. Both conceptually and in practice, it is compatible with any geometrical representation and any optimisation technique, and can therefore form a sound and useful basis for any future work in instrument

optimisation.

10.1.3 Aim 3

To investigate the comparative efficacy of the Rosenbrock and genetic algorithms for brass instrument optimisation.

Reconstruction experiments were performed with both optimisation algorithms; the genetic algorithm was significantly outperformed by Rosenbrock, which gave better results, converged faster and has deterministic repeatability. Given that problems with few design variables are better suited to genetic methods, and considering the results in [19], we conclude that further use of genetic algorithms for brass instrument optimisation is not worthwhile.

10.1.4 Aim 4

To develop integrated and easy-to-use software, written in C++, to perform optimisation and other related tasks.

A significant time investment has produced software to perform a variety of useful tasks related to instrument optimisation. The software is written in C++, is fully object-oriented and uses a variety of advanced programming techniques. It was developed in Linux, but will be easily ported to other operating systems. It is run from the command-line, using convenient plain-text input files, and outputs to either MATLAB or Maple for easy plotting.

The software can perform optimisation using either the Rosenbrock or genetic algorithms. Targets are defined in terms of a saved impedance curve and/or peak data. The optimisation template (along with its constraints), target data and tuning parameters are specified for each problem in a user-defined file. It can evaluate and save the impedance of a given instrument, calculated with a user-specified number of modes. Experimental data from BIAS [43] can be imported and set as targets. Benchmark tests can be performed, evaluating a simple instrument repeatedly, to compare the computation times of various methods. Each of these tasks is included in a single application, and is easily available through various command-line parameters.

During the course of this work much effort has been spent to ensure that the software is as flexible, powerful, expandable and easy-to-use as possible. The author has used this

software extensively, and is convinced by this experience that the complex programming task of implementing the sophisticated template representation has been amply justified by the end result. Due to the choice of language and the carefully-designed source code, future additions and modifications to the software will present little difficulty, ranging from the addition of other (more accurate) physical models to new element types representing other geometries.

10.1.5 Aim 5

To conduct experimental measurements of real trombones to gain insight into the impedance properties of good instruments.

A series of BIAS impedance measurements were performed on a variety of modern trombones and mouthpieces. While the BIAS equipment is very portable and easy to use, the model available gives results which suffer from some noise, and imperfect repeatability (though subsequent results using a newer model in factory condition are markedly better). Nevertheless, useful results were gained. Equivalent Fundamental Pitch plots were used to assist analysis of the harmonicity of the resonances.

Results demonstrate that exactly harmonic alignment of resonances is not always desirable; trombones, particularly the larger models, show a trend of higher peaks having gradually sharper equivalent fundamentals, whereas French Horns do not [56]; clearly there is some variation between instrument classes. It appears that a degree of inharmonicity is essential to give the characteristic timbral qualities required of certain instruments.

Smaller trombones generally demonstrate taller impedance peaks, particularly in the upper frequency range. Varying the mouthpiece likewise affects the peak heights, and has a small effect on the peak frequencies. When taken in combination, the results suggest that taller and more closely aligned peaks contribute to the ease of playing and the brightness of the tone produced. The measured impedance of a trombone is significantly dependent on the bore of both the mouthpiece and the trombone.

The experimental measurement of a trombone was compared to a theoretical plot based on geometrical measurements of the same instrument; while there is broad agreement, significant discrepancies are present. Possible causes include experimental error, the lack of an infinite

rigid baffle in the measurements, inaccuracies in the measurement of the instrument geometry and shortcomings in the theoretical modelling. As far as the latter is concerned, the boundary conditions in the multimodal wall loss model are proposed as the most likely cause.

10.1.6 Aim 6

To use the optimiser to reconstruct an instrument from its impedance.

The optimiser was successfully used to reconstruct, to within a fine tolerance, a complete virtual trombone from its calculated impedance, both using the spherical-wave assumption and with the multi-modal method. This task required no *a priori* knowledge of the specific dimensions of the instrument, and was successful both with targets based on simplified geometry and those based on fully-detailed measurements of a real instrument. These results demonstrate the effectiveness of the optimiser to solve bore-reconstruction problems in the absence of modelling discrepancies.

A more stern test was the reconstruction of an instrument from its experimental impedance as measured by BIAS. A good reconstruction was found, with the problem of being inaccurate in the bell region. This can largely be ascribed to the discrepancy between the theoretical and experimental impedance plots found for this instrument outwith the optimiser. The magnitude of the discrepancy made the reconstruction problem very difficult, and the optimiser was still able to produce a reasonable approximation. Given that the optimisation method itself has separately proven to be sound, we conclude that, once the discrepancy between theory and experiment has been resolved, the optimiser will be capable of such experimental bore-reconstruction.

10.1.7 Aim 7

To use the optimiser to modify existing instruments to alter certain impedance properties.

The optimiser was successfully used to provide solutions to design problems relevant to modern instrument design. The experiments were based on a Conn 88H, a popular large-bore tenor trombone model, and attempted to modify the geometry of the instrument to improve its performance. Firstly, the alignment of peak 2, which has a strong effect on the lower register performance of the instrument, was successfully shifted to be more closely harmonic

with the other peaks. It is expected that the modified design, were it built, would have a significantly improved lower register as a result of the changes. Secondly, the height of several peaks in the upper playing range were made taller, giving the expectation of an improved upper register. In both cases some compromises in other aspects of the impedance curve were needed, but were rather smaller than the desired change. Both experiments produced smooth designs which could be built with no additional difficulty. Thirdly, the extra tubing of the F-attachment was optimised to make the badly misaligned peak 2 more closely harmonic; at present the author is aware of no previous attempts by any manufacturer at having a non-cylindrical bore in such an attachment. This novel optimisation successfully brought the peak into much closer alignment, with an associated expectation of a large improvement in performance in that register, but required a compromise in the location of peak 3, which is likely to have an adverse effect on the next highest register. The construction and testing of a prototype will show whether this trade-off is favourable, and further work may reveal better compromises.

These results demonstrate the effectiveness of the optimiser to solve realistic problems in trombone design. Due to the above discrepancy between the experiment and the multimodal theory, and the further discrepancy between the multimodal model and the spherical-wave model used for the optimisation, we cannot make any firm conclusions about how the proposed designs would perform if they were built; furthermore, musicians would be required to provide subjective judgements as to the quality of the performance. However, most of the improvements focus on the lower-frequency performance, in which region the spherical-wave model assumption holds, and can therefore be considered a good indication of performance in this range. Excepting improvements in the modelling, the next stage of the process is to build prototypes of the designs, and perform tests to investigate the real performance of these proposed designs.

10.2 Comparison with Previous Studies

10.2.1 Kausel

Kausel's pioneering work provided the first genuine attempt at an automated instrument optimiser. Using a point-wise representation, a least-squares objective function, and the Rosen-

brock algorithm (a genetic algorithm was also unsuccessfully used), it successfully reconstructs entire instruments, albeit with some 'ripples' in the bell section and using "admittedly countless" evaluations. Intonation optimisation can be performed over all 7 valve combinations, a capability which remains unique. This software has a graphical user-interface, and has been successfully used commercially. The present work may be viewed as a continuation of Kausel's approach.

10.2.2 Noreland

This approach was based around use of the gradient-based Levenberg-Marquadt algorithm, and reformulates the transmission-line model to give the gradient of the least-squares objective function with respect to the geometric parameters (generally point-wise, but parametrisations were also attempted). Under the assumption that the imaginary part of the input impedance vanishes at the magnitude maxima, it is evaluated at the target frequencies and used directly as a measure of fitness of peak frequency. The second derivative of the (point-wise) segment radius with respect to length is used to give smooth, if sometimes wiggly, bell geometries. The method was generally used to reconstruct bell sections for which the length, input and output radii were known and fixed; although good results are found with very rapid convergence, it is a rather less difficult reconstruction problem than that attempted here.

10.3 Future Work

10.3.1 Theory-Experiment Discrepancy

The most obvious avenue for future work is that of the theory-experiment discrepancy. Work is elsewhere in progress to improve input impedance measurement techniques [58]; use of such equipment in detailed experiments with a large rigid baffle in place may yield improved experimental results. Likewise, improvements in the modelling can yield dividends.

Consider first the plane-mode models. Noreland proposed a hybrid spherical-wave model with finite-element modelling of the bell; while this yielded improved accuracy, it was computationally too expensive for optimisation purposes. Another possibility is that of combining theory with experimental results. Under the plane or spherical wave assumptions, it can be shown that the transmission matrix of an arbitrary bore can be deduced if it is measured ex-

perimentally with four different known termination impedances. If, for example, a trombone without a bell were terminated in turn by four different stopped tubes and measured, then its transmission-matrix can be deduced. If the bell were then added and another measurement taken, then the impedance of the bell and its radiating end can also be deduced. This can then be used as the termination impedance in theoretical calculations of the instrument, and may confer an increase in accuracy as the spherical-wave assumption is good in the rest of the instrument. The same technique may also be employed to produce a model of the radiation impedance of a French Horn including the player's hand in the bell, allowing horns to be optimised in a similar way. Such an experimental technique has not, to the author's knowledge, been attempted.

Moving onto the multi-mode model, the wall-loss mechanism merits further attention. It is based on two contradictory boundary conditions, and neglects the effect of the boundary layer on the transverse behaviour. It is not clear to what extent this assumption is valid, but given the wide variation in radius found in brass instrument bells it is not obvious that such transverse effects are negligible when compared to the axial effects. Further study in this area, possibly including experimental data, may prove valuable.

Continuing in this vein, we come to the bends model. The current experiments have indicated that the curvature has a negligible effect on the impedance of a complete trombone. However, this model lacks any account of the wall losses. A complete formulation of such losses would give a definitive answer to the question of the effect of the bends. Such a formulation is likely to be very involved. Similarly, an experimental investigation into the possible indirect effect of bends would be valuable – for example, the precise geometry of a nominally toroidal bend in an instrument. It is possible that the process by which such bends are manufactured causes some small perturbations to the geometry, which may in turn have an effect on the impedance. Nederveen [6] suggests that such perturbations exist, though his destructive method is unsuitable for precious instruments. The conventional rod-and-ruler based method of measuring bore-profiles suffers from rather limited precision, and is useless for bent ducts. One might suggest some small laser-based probe attached to a flexible rod capable of navigating the entire instrument as being suitable, though this may prove difficult in practice.

Further work with multi-modal methods may be facilitated by rederiving the models in tensor form; given the multiple indices (which are currently lumped together) and non-trivial coordinate spaces, a tensor analysis may provide a good deal more flexibility than the existing model.

Once the discrepancy has been resolved, the optimiser will become a viable alternative to acoustic pulse reflectometry [5, 59] for the task of bore reconstruction.

10.3.2 Software & Optimisation

The ongoing ARTSIM project [40] intends to provide a comprehensive library of computational tools to model acoustic systems. Much of the code written for this thesis could be included in such a library with relatively little modification.

The template model and implementation are applicable to any other optimisation technique, and will prove useful to any future attempts at similar problems. Combining the template representation with the gradient-based optimisation technique in [21] may yield a further improvement in convergence speed.

At present, the design space is normalised to be of uniform size. This normalisation is a simple linear scaling. Some advantage may be gained by investigating non-linear scaling, chosen so as to bias the optimiser toward being near the centre of the space without reducing the design space.

The computation speed of the optimiser may be improved by the retention of previously calculated impedance data. Not every section is altered in each step of the optimiser, so those sections unaltered may use existing data. For this to be worthwhile, the computational overhead pertaining to the storage and subsequent access of a database of impedance matrices must be less than the time saved by such retention; it is not clear that this condition will necessarily be met in all cases, but the matter is certainly worthy of investigation. The computation speed of the spherical-wave model may be improved by the derivation of a lossy transmission matrix for a Bessel-horn; at present only a lossless model exists.

The addition of a graphical user interface to the software would significantly increase its ease of use, particularly for those less familiar with its inner workings. Any commercial product must have such an interface.

The templates developed for this work have proven to be effective, but in order to be useful in general bore reconstruction problems, they must be tested carefully on other instruments different to those for which the templates were developed. A broad database of templates for a variety of different instruments and instrument classes can be constructed.

10.3.3 Instrument Performance

The construction of a large database of impedance measurements of a wide variety of instruments would be valuable, particularly if it can be combined with the evaluations of professional players. Ideally, one will be able to obtain a reliable method of predicting how a musician will judge an instrument based on its impedance curve. Some work to this end appears in this thesis, but it is neither comprehensive nor rigorous enough to give a sufficiently solid grounding. Experiments can be devised whereby a component of a given instrument is chosen, and various alternative designs of that component are designed with the optimiser to investigate certain impedance properties. Such components can then be built and tested by musicians, and correlations sought between their evaluations and the impedance properties.

Current plotting methods are unable to display satisfactorily on a single plot the harmonicity and magnitude of a given impedance curve. Such a plot would assist interpretation and prediction of performance. For example, a broad, tall but somewhat inharmonic higher peak may provide a greater contribution to a regime than a narrow, short harmonic peak; the EFP plot makes no account of this. One possible method is to use an existing EFP plot, and overlay horizontally around each peak a row of narrow coloured vertical bands, the colour of which represents the magnitude of the impedance a given number of cents from that peak. Such a plot would resemble a rectangle filled with rows of coloured bands, and would allow one to see at a glance not just the harmonicity, but also the associated impedance magnitude both at and around the peaks.

The optimiser has proven capable of making feasible modifications to complete instruments and meeting specified parameters. Prototypes of such designs could be built, their impedance measured and compared with that predicted by the modelling, and played by musicians to determine if they perform as predicted. Kausel's optimiser has already been used to produce commercial products. Given the improvements made in this work, it is reasonable

to expect that this software will be capable of similar usage. Once its effectiveness has been proven in the experimental domain, it may prove invaluable to an instrument designer. A detailed and comprehensive exploration of the capabilities of the optimiser as applied to a wide range of instruments is likely to prove most interesting.

The effect of a player moving between two regimes with quite different harmonicity is unclear. It is proposed in section 7.2.3 that the player may find greater ease in slurring between two regimes of similar harmonicity. An experimental investigation into such effects would be valuable.

Although the effect, in the absence of losses, of the bends in a trombone was examined and found to be negligible, this does not make any statement about other instruments. The trumpet features some sharp bends in its valve tubing, as does the French Horn which also features bent tubing almost exclusively throughout its bore. Similar numerical investigations on these instruments may yield interesting results.

Appendix A

Conical Plane-Wave Derivation

Equation 2.27 gives the impedance at the mouth of a conical horn. Using, for convenience, the impedance per unit area $\widehat{z} = Sz$, we have

$$\widehat{z_1} = \frac{\rho\omega[(A-B)\sin(kx_1)) + j(A+B)\cos(kx_1)]}{\frac{1}{x_1}[(A+B)\cos(kx_1) - j(A-B)\sin(kx_1)] + k[(A+B)\sin(kx_1) + j(A-B)\cos(kx_1)]},$$

we have

$$\frac{A+B}{A-B} = \frac{[\frac{1}{x_1}j\sin(kx_1) - jk\cos(kx_1)]\widehat{z_1} + \rho\omega\sin(kx_1)}{[\frac{1}{x_1}\cos(kx_1) + k\sin(kx_1)]\widehat{z_1} - j\rho\omega\cos(kx_1)}. \tag{A.1}$$

Equating this fraction at the mouth and the throat x_0 gives

$$\frac{[\frac{1}{x_0}j\sin(kx_0) - jk\cos(kx_0)]\widehat{z_0} + \rho\omega\sin(kx_0)}{[\frac{1}{x_0}\cos(kx_0) + k\sin(kx_0)]\widehat{z_0} - j\rho\omega\cos(kx_0)} = \frac{[\frac{1}{x_1}j\sin(kx_1) - jk\cos(kx_1)]\widehat{z_1} + \rho\omega\sin(kx_1)}{[\frac{1}{x_1}\cos(kx_1) + k\sin(kx_1)]\widehat{z_1} - j\rho\omega\cos(kx_1)}. \tag{A.2}$$

For ease of notation we define a series of dummy variables $\gamma_{1\ldots8}$ such that

$$\frac{\gamma_1\widehat{z_0} + \gamma_2}{\gamma_3\widehat{z_0} + \gamma_4} = \frac{\gamma_5\widehat{z_1} + \gamma_6}{\gamma_7\widehat{z_1} + \gamma_8}, \tag{A.3}$$

giving

$$\widehat{z_0} = \frac{(\gamma_4\gamma_5 - \gamma_2\gamma_7)\widehat{z_1} + (\gamma_4\gamma_6 - \gamma_2\gamma_8)}{(\gamma_1\gamma_7 - \gamma_3\gamma_5)\widehat{z_1} + (\gamma_1\gamma_8 - \gamma_3\gamma_6)}. \tag{A.4}$$

Then, noting the the length of the cone section $d = x_1 - x_0$, we have

$$\cos(kd) = \cos(kx_1)\cos(kx_0) + \sin(kx_1)\sin(kx_0)$$
$$\sin(kd) = \sin(kx_1)\cos(kx_0) - \cos(kx_1)\sin(kx_0). \tag{A.5}$$

We then simplify (A.4) term-by-term:

$$\begin{aligned}
\gamma_4\gamma_5 - \gamma_2\gamma_7 &= [-j\rho\omega\cos(kx_0)]\left[\frac{1}{x_1}j\sin(kx_1) - jk\cos(kx_1)\right] \\
&\quad -[\rho\omega\sin(kx_0)]\left[\frac{1}{x_1}\cos(kx_1) + k\sin(kx_1)\right] \\
&= \rho\omega\left(\frac{\sin(kd)}{x_1} - k\cos(kd)\right), \qquad \text{(A.6)} \\
\gamma_4\gamma_6 - \gamma_2\gamma_8 &= [-j\rho\omega\cos(kx_0)][\rho\omega\sin(kx_1)] - [\rho\omega\sin(kx_0)][-j\rho\omega\cos(kx_1)] \\
&= -j\rho^2\omega^2\sin(kd), \qquad \text{(A.7)} \\
\gamma_1\gamma_7 - \gamma_3\gamma_5 &= \left[\frac{1}{x_0}j\sin(kx_0) - jk\cos(kx_0)\right]\left[\frac{1}{x_1}\cos(kx_1) + k\sin(kx_1)\right] \\
&\quad -\left[\frac{1}{x_0}\cos(kx_0) + k\sin(kx_0)\right]\left[\frac{1}{x_1}j\sin(kx_1) - jk\cos(kx_1)\right] \\
&= jk^2\left[-\left(\frac{1}{k^2x_0x_1} + 1\right)\sin(kd) + \left(\frac{d}{kx_0x_1}\right)\cos(kd)\right], \qquad \text{(A.8)} \\
\gamma_1\gamma_8 - \gamma_3\gamma_6 &= \left[\frac{1}{x_0}j\sin(kx_0) - jk\cos(kx_0)\right][-j\rho\omega\cos(kx_1)] \\
&\quad -\left[\frac{1}{x_0}\cos(kx_0) + k\sin(kx_0)\right][\rho\omega\sin(kx_1)] \\
&= -\rho\omega\left(\frac{\sin(kd)}{x_0} + k\cos(kd)\right). \qquad \text{(A.9)}
\end{aligned}$$

Substituting back into (A.4) gives

$$\widehat{z_0} = \frac{\rho\omega\left(\frac{1}{x_1}\sin(kd) - k\cos(kd)\right)\widehat{z_1} - j\rho^2\omega^2\sin(kd)}{jk^2\left[-\left(\frac{1}{k^2x_0x_1} + 1\right)\sin(kd) + \left(\frac{d}{kx_0x_1}\right)\cos(kd)\right]\widehat{z_1} - \rho\omega\left(\frac{1}{x_0}\sin(kd) + k\cos(kd)\right)}. \tag{A.10}$$

Then, using $\omega = kc$, mouth and throat impedances $\widehat{z_0} = S_0 z_0$ and $\widehat{z_1} = S_1 z_1$, where S_0 and S_1 are the cross-sectional areas at the throat and mouth, gives

$$z_0 = \frac{1}{S_0}\frac{\rho ck\left(\frac{1}{x_1}\sin(kd) - k\cos(kd)\right)z_1 - \frac{\rho c}{S_1}j\rho ck^2\sin(kd)}{jk^2\left[-\left(\frac{1}{k^2x_0x_1} + 1\right)\sin(kd) + \left(\frac{d}{kx_0x_1}\right)\cos(kd)\right]z_1 - \frac{\rho c}{S_1}k\left(\frac{1}{x_0}\sin(kd) + k\cos(kd)\right)}. \tag{A.11}$$

Defining the characteristic impedance $z_c = \rho c/S_0$ and observing that $1/S_1 = x_0^2/x_1^2 S_0$ yields

$$z_0 = z_c \frac{k\left(\frac{1}{x_1}\sin(kd) - k\cos(kd)\right) z_1 - z_c \frac{x_0^2}{x_1^2} jk^2 \sin(kd)}{jk^2\left[-\left(\frac{1}{k^2 x_0 x_1} + 1\right)\sin(kd) + \left(\frac{d}{k x_0 x_1}\right)\cos(kd)\right] z_1 - z_c \frac{x_0^2}{x_1^2} k\left(\frac{1}{x_0}\sin(kd) + k\cos(kd)\right)}. \tag{A.12}$$

Multiply by $-x_1/k^2 x_0$ and cancel over z_c to find

$$z_0 = \frac{\left(\frac{x_1}{x_0}\cos(kd) - \frac{1}{kx_0}\sin(kd)\right) z_1 + jz_c \frac{x_0}{x_1}\sin(kd)}{jz_c^{-1}\left[\left(\frac{1}{k^2 x_0^2} + \frac{x_1}{x_0}\right)\sin(kd) - \left(\frac{d}{k x_0^2}\right)\cos(kd)\right] z_1 + \frac{x_0}{x_1}\left(\frac{1}{kx_0}\sin(kd) + \cos(kd)\right)}. \tag{A.13}$$

Appendix B

Bessel & Exponential Horns

We consider in this appendix two horn shapes which offer analytical solutions to the Webster equation (2.9).

B.1 Bessel Horn

A waveguide of the following shape is defined to be a Bessel-horn:

$$\mathcal{R}(x) = b(-x)^{-\gamma}, \tag{B.1}$$

where $\mathcal{R}$ is the bore radius a distance x along the axis of instrument, γ is a flare constant and b a constant. The horn lies to the left of the origin, with the mouth facing right. The bore function is valid only in the domain $(-\infty, 0)$ and becomes unbounded at the origin.

Now consider a Bessel-horn, flare γ, length d and radius ranging from $\mathcal{R}_0$ at the throat to $\mathcal{R}_1$ at the mouth. From (B.1), we can choose b such that the horn expands (in (x, r) coordinates) from $(x_0, \mathcal{R}_0)$ to $(x_0 + d, \mathcal{R}_1)$ as x increases.

$$\begin{aligned} \mathcal{R}_0 &= b(-x_0)^{-\gamma} \\ \mathcal{R}_1 &= b(-x_0 - d)^{-\gamma}. \end{aligned} \tag{B.2}$$

Solving for b, x_0 gives

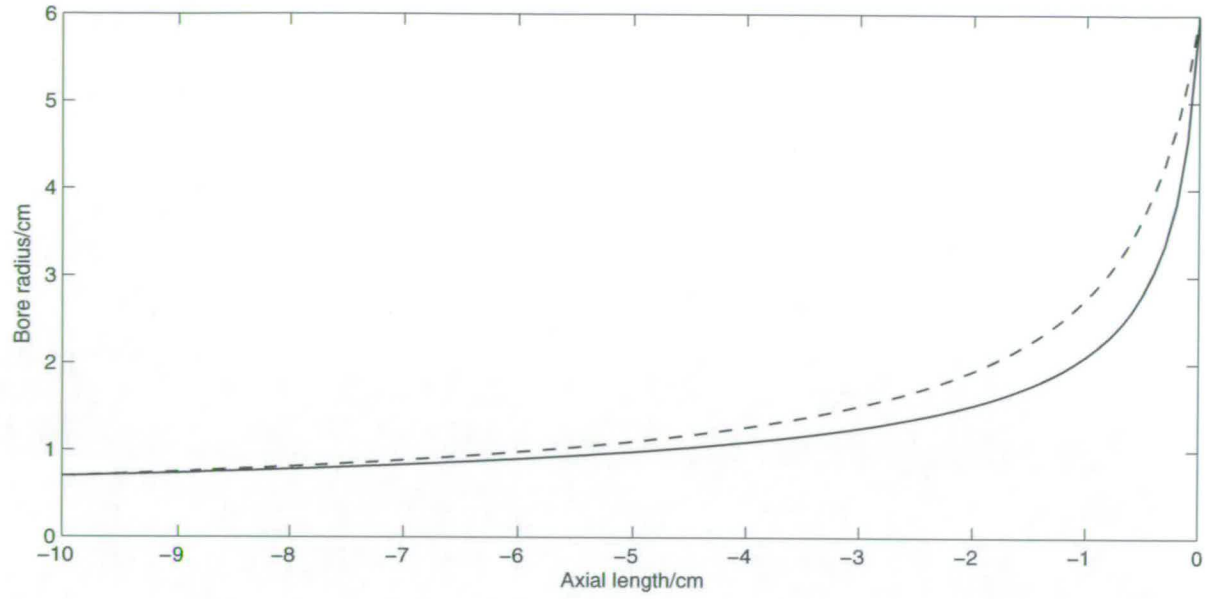


Figure B.1: Two Bessel-horns with same length and radius properties, with flare constants 0.5 (solid) and 0.7 (dashed). Dimensions approximately match a trombone bell.

$$\begin{aligned} b &= \left(\frac{d}{\mathcal{R}_0^{-1/\gamma} - \mathcal{R}_1^{-1/\gamma}} \right)^{\gamma} \\ x_0 &= -\left(\frac{\mathcal{R}_0}{b} \right)^{-1/\gamma}. \end{aligned} \tag{B.3}$$

We can use this result to specify a continuum of Bessel-horns with given length and input/output radii by varying the choice of flare γ (figure B.1).

B.2 Exponential Horn

A waveguide of the following shape is defined to be an exponential horn

$$\mathcal{R}(x) = \mathcal{R}_0 e^{\frac{\gamma x}{2}}, \tag{B.4}$$

where γ is a flare constant and $\mathcal{R}_0$ an arbitrary parameter. Consider an exponential horn of length d, with throat and mouth radii $\mathcal{R}_0$ and $\mathcal{R}_1$ arbitrarily located on the x-axis at x_0 and x_1 respectively; we seek a set of such horns matching these given criteria. We attempt a coordinate transform $\xi + \xi_0 = x$, defined such that the throat of the horn is at $\xi = 0$ and the mouth at $\xi = d$. However,

$$\begin{aligned}\mathcal{R}(\xi) &= \mathcal{R}_0 e^{\frac{\gamma}{2}(\xi+\xi_0)} \\ &= \left(\mathcal{R}_0 e^{\frac{\gamma}{2}\xi_0}\right) e^{\frac{\gamma}{2}\xi},\end{aligned} \tag{B.5}$$

so a translation ξ_0 in the x-axis can be absorbed into the multiplicative transform $\mathcal{R}_0$; we may therefore discard the translation by setting $\xi_0 = 0$ and safely assume, without loss of generality, that the throat is located at $x_0 = 0$ and the mouth at $x_1 = d$. Then

$$\begin{aligned}\mathcal{R}(0) &= \mathcal{R}_0 \\ \mathcal{R}(d) &= \mathcal{R}_0 e^{\frac{\gamma}{2}d},\end{aligned} \tag{B.6}$$

giving

$$\gamma = \frac{2}{d} \ln\left(\frac{\mathcal{R}_1}{\mathcal{R}_0}\right), \tag{B.7}$$

therefore given any two points there is only one exponential horn, with flare γ as shown, that passes through both these points.

Appendix C

Properties of Bessel Functions

C.1 Solutions of Bessel's Equation

Bowman [27] transforms the Bessel differential equation such that

$$x^2\frac{\partial^2 y}{\partial x^2} + (2p+1)x\frac{\partial y}{\partial x} + (\alpha^2 x^{2r} + \beta^2)y = 0 \tag{C.1}$$

has general solution

$$y(x) = x^{-p}\left[C_1 J_{q/r}\left(\frac{\alpha}{r}x^r\right) + C_2 Y_{q/r}\left(\frac{\alpha}{r}x^r\right)\right], \tag{C.2}$$

valid for all real values of the order q/r, where $q = \sqrt{p^2 - \beta^2}$, and J and Y are respectively Bessel-functions of the first and second kinds. Bowman gives this solution to supplement, in cases of integer order, a similar one in which a $J_{-q/r}$ term replaces that of $Y_{q/r}$, but does not mention that the latter solution can be used in all cases; [26] (p.76) makes it clear that $J_\nu(x)$ and $Y_\nu(x)$ always form a fundamental system of solutions, and therefore a cylinder function $\mathcal{C}(x) = a_\nu J_\nu(x) + b_\nu Y_\nu(x)$ is a general solution of the Bessel equation or its transforms. We recall the definition [26] (p.64)

$$Y_\nu(x) = \frac{J_\nu(x)\cos(\nu\pi) - J_{-\nu}(x)}{\sin(\nu\pi)} \tag{C.3}$$

such that the function $Y_\nu(x)$ is continuous in ν (except where $x = 0$), because the limit of the above expression is taken as $\nu \to n$ where n is an integer. This definition may also be used, with retained values of J_ν and $J_{-\nu}$, to improve computation speed.

C.2 Some Identities

The following results hold [26] (p.54-55):

$$\begin{aligned} J_{-1/2}(x) &= \left(\frac{2}{\pi x}\right)^{1/2} \cos(x), \\ J_{-3/2}(x) &= \left(\frac{2}{\pi x}\right)^{1/2} \left(\frac{-\cos(x)}{x} - \sin(x)\right), \\ J_{1/2}(x) &= \left(\frac{2}{\pi x}\right)^{1/2} \sin(x), \\ J_{3/2}(x) &= \left(\frac{2}{\pi x}\right)^{1/2} \left(\frac{\sin(x)}{x} - \cos(x)\right). \end{aligned} \tag{C.4}$$

Then, using (C.3), we have

$$\begin{aligned} Y_{-1/2}(x) &= J_{1/2}(x), \\ Y_{-3/2}(x) &= J_{3/2}(x). \end{aligned} \tag{C.5}$$

From [28] p.146,

$$\int x J_m(ax) J_m(bx) dx = \frac{bx J_m(ax) J_{m-1}(bx) - ax J_{m-1}(ax) J_m(bx)}{a^2 - b^2}. \tag{C.6}$$

From [27] equation 6.4

$$J_{-q}(x) = (-1)^q J_q(x). \tag{C.7}$$

C.3 Zeros

From equation 9.5.5 on p.370 of [60]

$$J_m(\gamma_{mn}) = \frac{\gamma_{mn}}{m} J_{m-1}(\gamma_{mn}) = \frac{\gamma_{mn}}{m} J_{m+1}(\gamma_{mn}), \tag{C.8}$$

where γ_{mn} is the $(n+1)^{\text{th}}$ zero of derivative J'_m.

Table of Zeros

For the Bessel-function J_m of the first kind, we tabulate γ_{mn} [60] (p.411)

		n							
	γ_{mn}	0	1	2	3	4	5	6	7
	0	0.0000	3.8317	7.0156	10.1735	13.3237	16.4706	19.6159	22.7601
	1	1.8411	5.3314	8.5363	11.7060	14.8636	18.0155	21.1643	24.3113
	2	3.0542	6.7061	9.9695	13.1704	16.3475	19.5129	22.6716	25.8260
	3	4.2012	8.0152	11.3459	14.5859	17.7888	20.9725	24.1449	27.3101
m	4	5.3178	9.2824	12.6819	15.9641	19.1960	22.4010	25.5898	28.7678
	5	6.4156	10.5199	13.9872	17.3128	20.5755	23.8036	27.0103	30.2029
	6	7.5013	11.7349	15.2682	18.6374	21.9317	25.1839	28.4098	31.6179
	7	8.5778	12.9324	16.5293	19.9419	23.2681	26.5450	29.7908	33.0152

Table C.1: Zeros γ_{mn} of Bessel function derivative J'_m.

Appendix D

Physical Constants

Keefe [13] gives values for the thermodynamic constants of air at room temperatures. For air at atmospheric pressure at temperature T ℃, using the value $\Delta T = T - 26.85$ ℃, we have

$$\begin{aligned} \rho &= 1.1769 \times 10^{-3}(1 - 0.00335\Delta T)\text{g cm}^{-3} \\ \eta &= 1.846 \times 10^{-4}(1 + 0.0025\Delta T)\text{g s}^{-1}\,\text{cm}^{-1} \\ c &= 3.4723 \times 10^{4}(1 + 0.001166\Delta T)\text{cm s}^{-1}, \end{aligned} \tag{D.1}$$

accurate within $\Delta T = \pm 10$ ℃.

Appendix E

Repeat Experimental Measurements

E.1 Experimental Measurements

Since the initial submission of this work, a new BIAS head has been made available. This is the latest version, offering improved performance over its predecessor, and is factory condition. This new equipment was used to repeat certain of the measurements above in an attempt to produce results with smaller measurement errors. The exact instruments and mouthpieces were not available, so the closest available match was substituted.

It has already been commented that directly comparing measurements with the same equipment on the same instruments taken during different sessions in different conditions is not valid. It is therefore certainly not valid to compare directly measurements taken with different equipment on different instruments during sessions separated by over two years. Nevertheless, given the existence of presumably improved experimental equipment it is necessary to employ it to validate the analysis of Chapter 7. Most of this analysis is based on EFP, which appears to be much less sensitive to the above inconsistencies, and can be normalised in frequency; direct comparisons between the two measurements, keeping in mind differences in the instruments, is sufficient to test whether the qualitative EFP features observed are unaffected.

We see in Figure E.2a the EFP of a Conn 8H with a Denis Wick 5AL mouthpiece measured

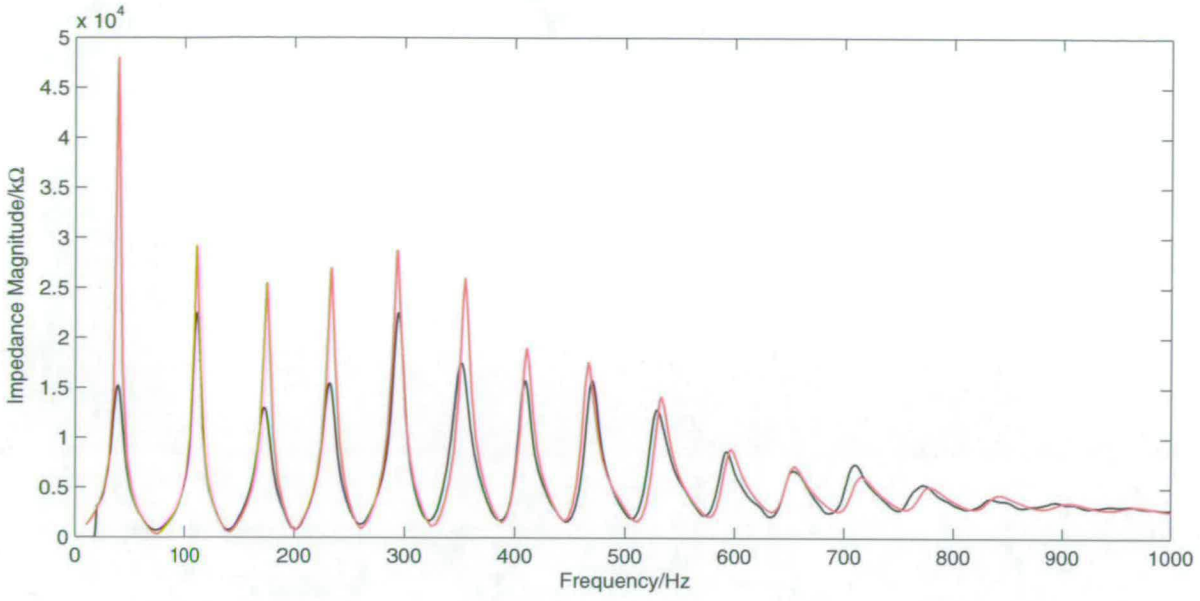
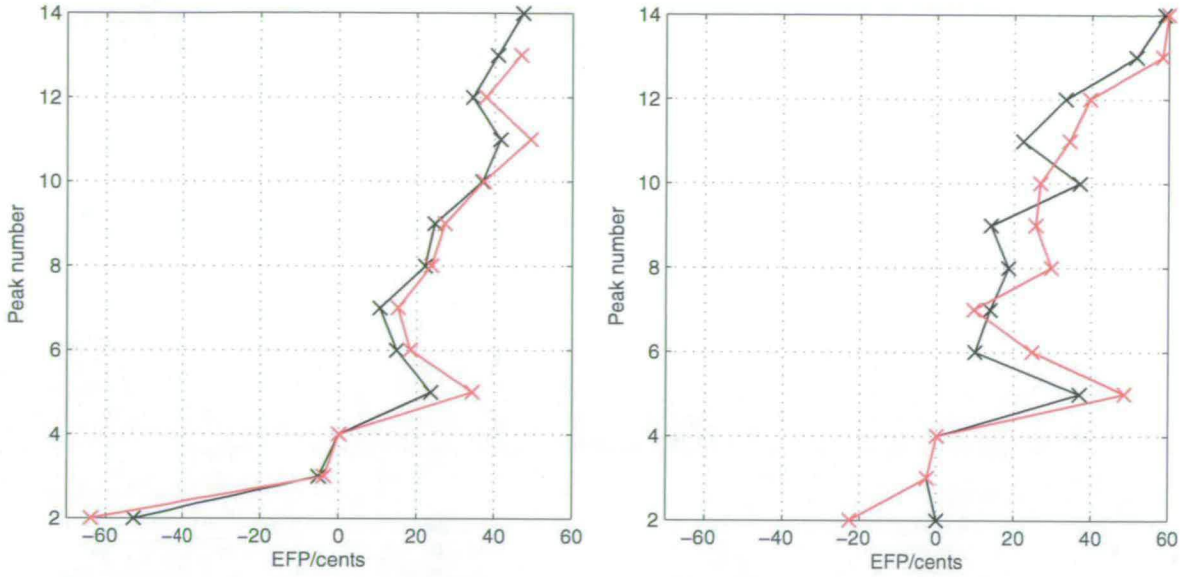


Figure E.1: Impedance of 8H, experimental measurement (black), and calculated with 11 modes (red).

with the new BIAS compared with that of a Conn 88H with another 5AL measured with the old BIAS. According to the manufacturers, the two trombones are nominally identical in bore-profile save for a rotary valve mid-way through the length of the 88H. The mouthpieces are likewise nominally identical. A musician would expect the two instruments to be virtually indistinguishable during playing. We see no significant difference in EFP between the two measurements.

In Figure E.2b the EFP of a Rath R9 with a Denis WIck 3AL mouthpiece measured with the new BIAS compared with that of another R9 with B1.5 mouthpiece measured with the old BIAS. The trombones are very similar, but not identical in design, and the mouthpieces are rather different in design. Nevertheless, the results are still qualitatively comparable; the only significant difference being the alignment of peak 2; the new measurement is, if anything, more supportive of the preceding analysis than was the old measurement.

We conclude that there is nothing in the new measurements which would cause us to reconsider the above discussion.



(a) 8H with new BIAS (black) and 88H with old BIAS (red). Each measured with 5AL mouthpieces

(b) An R9 with new BIAS (black) and 3AL mouthpiece, another R9 with old BIAS (red) and B1.5 mouthpiece.

Figure E.2: EFP plots for trombones comparing old and new BIAS equipment, normalised to 4^{th} resonance.

E.2 Optimisation

Experiment VI, being the reconstruction of a trombone from an experimental BIAS measurement, was repeated using results gained from the new equipment. The result was not a significant improvement over that already found; given the continued presence of the theory-experiment discrepancy this is not surprising.

Publications

A.C.P. Braden. Optimisation techniques for solving design problems in modern trombones. In *Proceedings of Forum Acusticum, Budapest*, pages 557–562, 2005.

A.C.P. Braden. Higher modes and the input impedance of trombones. In *Proceedings of Vienna Talk on The Future of Wind Instruments*, 2005.

A.C.P. Braden. Input impedance of bent waveguides. In *Proceedings of the Institute of Acoustics Spring Conference, Southampton*, volume 28(1), 2006.

Bibliography

[1] D. Mapes-Riordan. Horn modeling with conical and cylindrical transmission-line elements. *J.Audio Eng. Soc.*, 41:471–483, 1993.

[2] N.H. Fletcher and T.D. Rossing. *The Physics of Musical Instruments.* Springer, 1991.

[3] V. Pagneux, N. Amir, and J. Kergomard. A study of wave propagation in varying cross-section waveguides by modal decomposition. Part I. Theory and validation. *J. Acoust. Soc. Am.*, 100(4):2034–2048, 1996.

[4] N. Amir, V. Pagneux, and J. Kergomard. A study of wave propagation in varying cross-section waveguides by modal decomposition. Part II. Results. *J. Acoust. Soc. Am.*, 101(5):2504–2517, 1996.

[5] J.A. Kemp. *Theoretical and experimental study of wave propagation in brass musical instruments.* PhD thesis, University of Edinburgh, 2002.

[6] C.J. Nederveen. Influence of a toroidal bend on wind instrument tuning. *J. Acoust. Soc. Am.*, 104:1616–1626, 1998.

[7] D.H. Keefe and A.H. Benade. Wave propagation in strongly curved ducts. *J. Acoust. Soc. Am.*, 74:320–332, 1983.

[8] S. Félix and V. Pagneux. Sound propagation in rigid bends: A multimodal approach. *J. Acoust. Soc. Am.*, 110(3):1329–1337, 2001.

[9] S. Félix and V. Pagneux. Multimodal analysis of acoustic propagation in three-dimensional bends. *Wave Motion*, 36:157–168, 2002.

[10] S. Félix and V. Pagneux. Sound attenuation in lined bends. *J. Acoust. Soc. Am.*, 116(4):1921–1931, 2004.

[11] S. Félix and V. Pagneux. Ray-wave correspondance in bent waveguides. *Wave Motion*, 41:339–355, 2005.

[12] W.P. Bi, V. Pagneux, D. Lafarge, and Y. Aurégan. Modelling of sound propagation in a non-uniform lined duct using a multi-modal propagation method. *J. Sound. Vib.*, 289:1091–1111, 2006.

[13] D.H. Keefe. Acoustical wave propagation in cylindrical ducts: Transmission line parameter approximations for isothermal and nonisothermal boundary conditions. *J. Acoust. Soc. Am.*, 75:58–62, 1984.

[14] A.M. Bruneau, M. Bruneau, Ph. Herzog, and J. Kergomard. Boundary layer attenuation of higher order modes in waveguides. *J. Sound. Vib.*, 119(1):15–27, 1987.

[15] A.D. Pierce. *Acoustics: an Introduction to its Physical Principles and Applications*, pages 523–531. New York: McGraw-Hill, 1981.

[16] H. Levine and J. Schwinger. On the radiation of the sound from an unflanged circular pipe. *Physical review*, 73(4):383–406, 1948.

[17] W.E. Zorumski. Generalised radiation impedances and reflection coeffiecients of circular and annular ducts. *J. Acoust. Soc. Am.*, 54(6):1667–1673, 1973.

[18] D.H. Noreland. A numerical method for acoustic waves in horns. *ACUSTICA*, 88:576–586, 2002.

[19] W. Kausel. Computer optimization of brass wind instruments. In *Diderot Forum on Mathematics and Music*, page 227, 1999.

[20] W. Kausel and R. Egger. The brass-wind-instrument-optimizer as a tool for instrument makers: A case study. In *Proceedings of Vienna Talk on The Future of Wind Instruments*, 2005.

[21] D. Noreland. *A gradient based optimisation algorithm for the design of brass-wind instruments.* PhD thesis, Uppsala University, 2003.

[22] N.W. McLachlan. *Loud Speakers: Theory, Performance, Testing and Design*, pages 177–197. Dover Publications inc., New York, 1954.

[23] M van Walstijn. *Discrete-Time Modelling of Brass and Reed Woodwind Instruments with Application to Musical Sound Synthesis.* PhD thesis, University of Edinburgh, 2002.

[24] D. Zwilliger. *Handbook of Differential Equations*, pages 354–359. Academic Press, Third edition, 1998.

[25] R.W.R. Darling. Converting matrix Riccati equations to second-order linear ode. *SIAM Rev.*, 39(3):508–510, 1997.

[26] G.N. Watson. *A treatise on the theory of Bessel functions.* Cambridge University Press, Second edition, 1962.

[27] F. Bowman. *Introduction to Bessel functions.* Longmans Green & Co., First edition, 1938.

[28] E. Jahnke and F. Emde. *Tables of functions with formulae and curves.* Dover Publications, New York, 1945.

[29] H.H. Rosenbrock. An automatic method for finding the greatest or least value of a function. *Comp. J.*, 3:175–184, 1960.

[30] J. Kowalik and M.R. Osborne. *Methods for Unconstrained Optimization Problems.* American Elsevier Publishing Company, Inc., 1968.

[31] J.R. Palmer. An improved procedure for orthoganilising the search vectors in Rosenbrock's and Swann's direct search optimisation methods. *Comp. J.*, 12(1):69–71, 1968.

[32] M. Wall. Galib, C/C++ genetic algorithm library. `http://lancet.mit.edu/ga`.

[33] D.H. Noreland. Gradient-based optimisation of brass instruments. In *Proceedings of the Stockholm Music Acoustics Conference SMAC03, Sweden*, 2003.

[34] J Petrolito and K.A. Legge. Designing musical structures using a constrained optimization approach. *J. Acoust. Soc. Am.*, 117(1):384–390, 2005.

[35] L. Henrique, O. Inácio, J. Paulino, and J. Antunes. Optimization of vibratory and acoustical components of percussion instruments: Theoretical and experimental results. In *Proceedings of Forum Acusticum, Budapest*, pages 485–490, 2005.

[36] V. Chilekwa and D. Sharp. Detection, location and sizing of multiple leaks in a duct. In *Proceedings of Forum Acusticum, Budapest*, pages 2569–2572, 2005.

[37] A.C.P Braden. Numerical evaluation and optimisation of brass instrument impedance curves. Master's thesis, University of Bristol, 2002.

[38] A.C.P. Braden. Optimisation techniques for solving design problems in modern trombones. In *Proceedings of Forum Acusticum, Budapest*, pages 557–562, 2005.

[39] Michael Rath Trombones, correspondance. `http://www.rathtrombones.com`, 2004–2006.

[40] EAA Musical Acoustics Workgroup. Acoustic Real-Time Simulation Package. `http://www.eaa-fenestra.org/TCs/MA/Workgroups/WG2/Public/ARTSIM/`.

[41] GCC, GNU Compiler. `http://gcc.gnu.org`.

[42] W.H. Press. *Numerical Recipes in C++*, pages 406–407. Cambridge University Press, 2002.

[43] W. Kausel. Brass Instrument Analysis System. `http://www.bias.at`.

[44] M. Bertsch. Bridging instrument control aspects of brass instruments with physics-based parameters. In *Proceedings of the Stockholm Music Acoustics Conference*, 2003.

[45] M. Bertsch, K. Waldherr, and W. Kausel. Sensory evaluation testing of trumpets and correlation with acoustic measurements. In *Proceedings of Forum Acusticum, Budapest*, pages 575–580, 2005.

[46] M. Bertsch, K. Waldherr, M. Horvat, and K. Jambrošic. Reliability of trumpet players' judgements during instrument blind-test. In *Proceedings of Second Congress of Alps-Adria Acoustics Association and First Congress of Acoustical Society of Croatia. Zagreb, Croatia*, 2005.

[47] G.R. Plitnik and B.A. Lawson. An investigation of correlations between geometry, acoustic variables, and psychoacoustic parameters for French horns. *J. Acoust. Soc. Am.*, 106(2), 1999.

[48] E. Poirson, J.F. Petiot, and J. Gilbert. Study of the brightness of trumpet tones. *J. Acoust. Soc. Am.*, 118(4), 2005.

[49] W. Kausel. *A Musical Acoustician's Guide to Computational Physics.* Institut für Wiener Klangstil, 2003.

[50] M.W. Thompson and W.J. Strong. Inclusion of wave steepening in a frequency-domain model of trombone sound production. *J. Acoust. Soc. Am.*, 110(1), 2001.

[51] Online Trombone Journal, discussion forum. `http://forum.trombone.org`, 2002–2006.

[52] J.W. Whitehouse. *A study of the wall vibrations excited during the playing of lip-reed instruments.* PhD thesis, Open University, 2003.

[53] T.R. Moore, E.T. Shirley, I.E.W. Codrey, and A.E. Daniels. The effects of bell vibrations on the sound of the modern trumpet. *ACUSTICA*, 91:578–589, 2005.

[54] S. Dorn, J. Whitehouse, and D.B. Sharp. The effect of wall material on the structural vibrations excited when lip-reed instruments are blown. In *Proceedings of the Institute of Acoustics Spring Conference, Southampton*, volume 28(1), 2006.

[55] A.H. Benade. *Fundamentals of Musical Acoustics*, pages 402–405. Oxford University Press, 1976.

[56] J.P. Chick, C. Lumb, and D.M. Campbell. Passive acoustic characteristics and intonation problems of modern orchestral horns. In *Proceedings of ISMA2004, Nara, Japan*, 2004.

[57] A.C.P. Braden. Higher modes and the input impedance of trombones. In *Proceedings of Vienna Talk on The Future of Wind Instruments*, 2005.

[58] D.A. Hendrie and D.M. Campbell. Musical wind instrument analysis. In *Proceedings of Forum Acusticum, Budapest*, pages 671–676, 2005.

[59] D.B. Sharpe. *Acoustic pulse reflectometry for the measurement of musical wind instruments.* PhD thesis, University of Edinburgh, 1996.

[60] M. Abramowitz and I.A. Stegun. *Handbook of Mathematical Functions.* Dover Publications, New York, 1965.

[61] A.C.P. Braden. Input impedance of bent waveguides. In *Proceedings of the Institute of Acoustics Spring Conference, Southampton*, volume 28(1), 2006.

[62] R.W. Pyle Jr. Effective length of horns. *J. Acoust. Soc. Am.*, 57(6), 1975.

[63] R.E. Beatty Jr. Boundary layer attenuation of higher order modes in rectangular and circular tubes. *J. Acoust. Soc. Am.*, 27(6), 1950.

[64] E. Bängtsson, D. Noreland, and M. Berggren. Shape optimization of an acoustic horn. *Comput. Methods Appl. Mech. Engrg.*, 192:1533–1571, 2003.

[65] W. Rostafinski. Propagation of long waves in acoustically treated, curved ducts. *J. Acoust. Soc. Am.*, 71(1):36–41, 1982.