A Study of Automatic Transcription of Music Using a Standard PC

Victor Poon

2001

Abstract

This thesis describes using a Personal Computer to identify notes that are played by a musical instrument. Several groups have been doing this work with more sophisticated laboratories and equipment with only moderate success.

We have found the waves created by musical instruments vary, between instruments, a great deal in their stability and inherent vibration. It was more difficult to identify notes with very low frequencies than those with more central frequencies. We found it was very important to choose the correct starting point for the analysis with Fourier Transform otherwise we would not be analysing the stable stage of the wave.

We tried simple strategies to initially reduce the number of computer operations, and memory requirements, with marginal success. We followed with more complex subtraction strategies which were much more successful. The most useful technique involved creating a "calculated percentage multiple" which was almost 100% successful in identifying single notes. For multiple notes we were surprised to find that a group of five different instruments from MIDI were a better source of "known" notes to compare with the "unknown" notes than the MIDI equivalent of the real instrument playing the music.

These methods were developed using midi instruments but were verified using a real grand piano.

We suggest some further lines of enquiry that may make this technique more successful.

Acknowledgement

I would like to express my thanks to my supervisor, Dr. Peter Kay, for his help and advice during the completion of this thesis.

I am also very grateful for the help and the time given by Dr. Murray Sampson and Mrs. Lorraine Sampson who patiently proof read this thesis.

Finally I would like to thank my parents Yu Lin Poon and Tit Wing Poon for their financial and family support during my university study.

Contents

Contents i

List of Figures iv

List of Tables ix

Introduction 1

Literature Review 3

Chapter One: Background 10

1.1 Music Notes 10

1.2 Digital Audio 12

1.3 Wave File Format 12

1.4 MIDI 15

1.5 MATLAB 17

1.6 Fourier Transforms 18

1.6.1 The differences between DFT and FFT 19
1.6.1.1 Time efficiencies 19

1.6.2 Symmetry 20

Chapter Two: Using MIDI To Analyse Sound 23

2.1 Sample Rate, Sample Size, Frequency and Wavelength 25

2.1.1 The sample rate 25

2.1.2 The frequency, the wavelength and the sample size 26

2.2 The Effect of Using Different Sample Sizes 28

2.3 The Effect of Using Different Starting Points 31

2.4 Recording in MIDI 36

2.5 Playing Two Notes Simultaneously on a Single Instrument 41

2.6 Playing Three Notes at Once 45

2.7 Why is the Highest Peak of the Single Note not the Answer in Some Cases? 46

2.8 What is the Relationship between a Single Note and Two Notes or Three Notes? 48

2.9 The Reason for Different Instruments having Different Results from the Fourier Transform 50

2.10 Using the Fourier Transform of Recorded Notes to Identify Unknown Notes 58

2.10.1 The file size of the recording format 58

2.10.2 The choice of representative value 58

2.11 Possible Filtering Methods to Identify Unknown Notes 62

2.11.1 The first strategy 63

2.11.2 The second strategy 64

2.11.3 The third strategy 66

2.11.4 The fourth strategy 67

2.11.5 The fifth strategy 69

2.11.6 Summary 71

2.12 The Conditional or Tolerance Values 72

2.12.1 The first strategy 72

2.12.2 The second strategy 73

2.12.3 The third strategy 74

2.12.4 The fourth strategy 75

2.12.5 The fifth strategy 76

2.13 The Accuracy Rate for Different Strategies when Playing a Simple Single Note 77

2.13.1 Application of using strategies 1 to 4 77

2.13.2 Application of strategy 5 79

2.14 Application of Strategies 3-2-1 when Playing Two Notes 81

2.15 Application of Strategy 5 on Two Notes 83

2.15.1 Finding the first solution 83

2.15.2 Subtraction of notes to find the second solution 84
2.15.2.1 Subtraction excluding the first solution 85
2.15.2.1.1 Simple subtraction 85
2.15.2.1.2 Complex subtraction method 1 87
2.15.2.1.3 Complex subtraction method 2 87
2.15.2.2 Summary of subtraction excluding the first solution 90
2.15.2.3 Preliminary conclusions about strategy 5 91

2.16 Application of Strategy 5 on Three Notes Together 91

2.17 Sources of Error within Calculations 92

2.18 Finding the Number of Notes that are Playing Simultaneously? 93

Chapter Three: Real Instruments 98

3.1 Application of Strategy 5 on a Single Note 105

3.2 Application of Strategy 5 on Two Notes Played Together 105
3.2.1 Experimented results 106
3.2.2 Estimating the number of solutions 107

3.3 Application of Single Note Using MIDI as the Unknown Note 108

3.4 Application on Two Notes Using MIDI as the Unknown Note 109
3.4.1 Known notes from a single instrument 109
3.4.2 Known notes from 5 different instruments 110

3.5 Practical Experiment 111

3.6 Result of Method B 111

Chapter Four: Conclusion 116

Reference 119

List of Figures

Chapter One: Background

Figure 1.1 Example of a sound wave (Acoustic Grand Piano from MIDI, A4) 14

Figure 1.2 The number of operations of using DFT and FFT 20

Figure 1.3 Results of FT using MATLAB 20

Figure 1.4 Results of FT after rearrangement.. 21

Chapter Two: Using MIDI To Analyse Sound

Figure 2.1 The concept of translating text files to *.wav* files............................ 25

Figure 2.2 Wave format for A0 (Samples from 0 to 80,000) 26

Figure 2.3 Wave format for C8 (Samples from 0 to 3,150) 27

Figure 2.4 The results of Fourier Transform in 2D when using different sample sizes on an acoustic grand piano when playing the note C2 (65.41 Hz) 30

Figure 2.5 The results of Fourier Transform in 3D when using different sample sizes on an acoustic grand piano when playing the note C2 (65.41 Hz) 30

Figure 2.6 Simplified amplitude evolution of a music note 32

Figure 2.7 The results of Fourier Transform in 2D when using different starting points, on an acoustic grand piano when playing the note C2 (65.41 Hz).. 33

Figure 2.8 The results of Fourier Transform in 3D when using different starting points, on an acoustic grand piano when playing the note C2 (65.41 Hz).. 33

Figure 2.9 The results of Fourier Transform in 2D when using different starting points, on an acoustic grand piano when playing the note C2 (65.41 Hz) (in percentages)... 34

Figure 2.10 The results of Fourier Transform in 3D when using different starting points, on an acoustic grand piano when playing the note C2 (65.41 Hz) (in percentages)... 34

Figure 2.11 The results of Fourier Transform when using different starting points, on an acoustic grand piano when playing the note C2 (65.41 Hz) after totalling all the spectra 35

Figure 2.12 The results of Fourier Transform in 2D when using different starting points, on an acoustic grand piano when playing the note C2 (65.41 Hz) in percentages after totalling all the spectra 35

Figure 2.13 Example of recorder from MIDI (75) when playing the note A4 (440 Hz) (from Sample '0' to Sample 1500) 36

Figure 2.14 The sound wave of two different instruments when playing C2 (65.41 Hz) 47

Figure 2.15 The results of Fourier Transform on an acoustic grand piano when playing C2 (65.41 Hz) 47

Figure 2.16 The results of Fourier Transform on a recorder when playing C2 (65.41 Hz) 47

Figure 2.17 The sound wave of an acoustic grand piano when playing C4 48

Figure 2.18 The sound wave of an acoustic grand piano when playing A4 48

Figure 2.19 The sound wave adding up the value of C4 and A4 on an acoustic grand piano 49

Figure 2.20 The sound wave for an acoustic grand piano when playing C4 and A4 together 49

Figure 2.21 The sound wave of the note A4 on an acoustic grand piano (MIDI 1) 50

Figure 2.22 The sound wave of the note A4 on an acoustic guitar (steel) (MIDI 26) 50

Figure 2.23 The sound wave of the note A4 on an acoustic bass (MIDI 33) 50

Figure 2.24 The sound wave of the note A4 on a violin (MIDI 41) 51

Figure 2.25 The sound wave of the note A4 on a recorder (MIDI 75) 51

Figure 2.26 The sound wave of the note A4 when using different instruments .. 52

Figure 2.27 The results of the Fourier Transform of the note A4 on an acoustic grand piano (MIDI 1) 53

Figure 2.28 The results of the Fourier Transform of the note A4 on an acoustic guitar (steel) (MIDI 26) 53

Figure 2.29 The results of the Fourier Transform of the note A4 on an acoustic bass (MIDI 33) 53

Figure 2.30 The results of the Fourier Transform of the note A4 on a violin (MIDI 41) 54

Figure 2.31 The results of the Fourier Transform of the note A4 on a recorder (MIDI 75) 54

Figure 2.32 Example of tone spectra in FT for A4 (440 Hz) 57

Figure 2.33 The results of Fourier Transform when playing C2 (65.41Hz) on an acoustic grand piano 59

Figure 2.34 Figure 2.33 with note boundaries added 60

Figure 2.35 The "lowest" results of Fourier Transform from within each frequency range when playing C2 (acoustic grand piano) 61

Figure 2.36 The "highest" results of Fourier Transform from within each frequency range when playing C2 (acoustic grand piano) 61

Figure 2.37 The "average" results of Fourier Transform from within each frequency range when playing C2 (acoustic grand piano) 61

Figure 2.38 Summary of different methods of subtraction when using average frequency starting point 90

Figure 2.39 Summary of different methods of subtraction when using lower frequency starting point 90

Chapter Three: Real Instruments

Figure 3.1 An example of the process of detecting the starting point when playing with real instruments 102

Figure 3.2 A close up of the earlier parts of the FTs in Figure 3.1 103

Figure 3.3 The sound wave of C2 104

Figure 3.4 The maximum difference in each FT 104

Figure 3.5 The sum of difference in each FT 104

Appendix B

Figure B.1 The sample of acoustic grand piano when playing the note C2 (65.41 Hz).

Figure B.2 The results of Four Transform in 2D when using different sample sizes on an acoustic grand piano when playing the note C2 (65.41 Hz).

Figure B.3 The results of Four Transform in 3D when using different sample sizes on an acoustic grand piano when playing the note C2 (65.41 Hz).

Figure B.4 The results of Four Transform in 2D when using different starting point on an acoustic grand piano when playing the note C2 (65.41 Hz).

Figure B.5 The results of Four Transform in 3D when using different starting point on an acoustic grand piano when playing the note C2 (65.41 Hz).

Figure B.6 The results of Four Transform in 2D when using different starting point on an acoustic grand piano when playing the note C2 (65.41 Hz) (in percentages).

Figure B.7 The results of Four Transform in 3D when using different starting point on an acoustic grand piano when playing the note C2 (65.41 Hz) (in percentages).

Figure B.8 The results of Four Transform in 2D when using different starting point on an acoustic grand piano when playing the note C2 (65.41 Hz) after totalling all the spectra.

Figure B.9 The results of Four Transform in 2D when using different starting point on an acoustic grand piano when playing the note C2 (65.41 Hz) in percentages after totalling all the spectra.

Figure B.10 The sample of bright acoustic piano when playing the note C2 (65.41 Hz).

Figure B.11 The results of Four Transform in 2D when using different sample sizes on a bright acoustic piano when playing the note C2 (65.41 Hz).

Figure B.12 The results of Four Transform in 3D when using different sample sizes on a bright acoustic piano when playing the note C2 (65.41 Hz).

Figure B.13 The results of Four Transform in 2D when using different starting point on a bright acoustic piano when playing the note C2 (65.41 Hz).

Figure B.14 The results of Four Transform in 3D when using different starting point on a bright acoustic piano when playing the note C2 (65.41 Hz).

Figure B.15 The results of Four Transform in 2D when using different starting point on a bright acoustic piano when playing the note C2 (65.41 Hz) (in percentages).

Figure B.16 The results of Four Transform in 3D when using different starting point on a bright acoustic piano when playing the note C2 (65.41 Hz) (in percentages).

Figure B.17 The results of Four Transform in 2D when using different starting point on a bright acoustic piano when playing the note C2 (65.41 Hz) after totalling all the spectra.

Figure B.18 The results of Four Transform in 2D when using different starting point on a bright acoustic piano when playing the note C2 (65.41 Hz) in percentages after totalling all the spectra.

Figure B.19 The sample of an electronic acoustic piano when playing the note C2 (65.41 Hz).

Figure B.20 The results of Four Transform in 2D when using different sample sizes on an electronic acoustic piano when playing the note C2 (65.41 Hz).

Figure B.21 The results of Four Transform in 3D when using different sample sizes on an electronic acoustic piano when playing the note C2 (65.41 Hz).

Figure B.22 The results of Four Transform in 2D when using different starting point on an electronic acoustic piano when playing the note C2 (65.41 Hz).

Figure B.23 The results of Four Transform in 3D when using different starting point on an electronic acoustic piano when playing the note C2 (65.41 Hz).

Figure B.24 The results of Four Transform in 2D when using different starting point on an electronic acoustic piano when playing the note C2 (65.41 Hz) (in percentages).

Figure B.25 The results of Four Transform in 3D when using different starting point an electronic acoustic piano when playing the note C2 (65.41 Hz) (in percentages).

Figure B.26 The results of Four Transform in 2D when using different starting point on an electronic t acoustic piano when playing the note C2 (65.41 Hz) after totalling all the spectra.

Figure B.27 The results of Four Transform in 2D when using different starting point on an electronic acoustic piano when playing the note C2 (65.41 Hz) in percentages after totalling all the spectra.

List of Tables

Chapter One: Background

Table 1.1 The process of finding frequencies 11

Table 1.2 Table of frequencies displayed in octaves 11

Table 1.3 Table of the instruments in MIDI and the numbers to denote each instrument 16

Table 1.4 The number of operations using DFT and FFT 19

Chapter Two: Using MIDI To Analyse Sound

Table 2.1 The estimated duration of notes 27

Table 2.2 The relationship between frequency, the results for the Fourier Transform and the sample size 29

Table 2.3 The results of the highest peak of using Fourier Transform on recorder (MIDI 75) 37

Table 2.4 The accuracy rate when comparing the expected frequencies and the highest peak of the Fourier Transform grouped into types of instrument 40

Table 2.5 The results of playing two notes on MIDI by using recorder 41

Table 2.6 The results of playing two notes using recorder on MIDI 43

Table 2.7 The accuracy when using different ranges of expected note on the recorder 44

Table 2.8 The accuracy rate when using the two highest peaks to predict the results on different instruments 45

Table 2.9 The results of the Fourier Transform of the note A4 on different instruments 55

Table 2.10 The results of the Fourier Transform of the note C4 on different instruments 55

Table 2.11 A theoretical example of the highest peaks 63

Table 2.12 Applying the first strategy to Table 2.11 64

Table 2.13 Establishing the values accepted 65

Table 2.14 Applying the second strategy to Table 2.13 66

Table 2.15 A theoretical example applying strategy 3 using the information from Tables 2.13 and 2.14 67

Table 2.16 Table 2.13 with showing the difference between successive peaks.... 68

Table 2.17 Applying strategy 4 to Table 2.16 69

Table 2.18 Applying strategy 5 using the information from Tables 2.13, 2.14 and 2.15 70

Table 2.19 Summary of different minimum values using various starting points on 5 different instruments (in percentages) 73

Table 2.20 Minimum differences for 5 different instruments 74

Table 2.21 Minimum matches for 5 different instruments 75

Table 2.22 The number of zeros when using different conditional percentages... 75

Table 2.23 The minimum results of "known" minus "unknown" using different conditional percentages 76

Table 2.24 Results with notes from the same instrument 77

Table 2.25 Results with notes from 5 different instruments 78

Table 2.26 The summary of the number of rejected known notes using strategy 3 followed by strategies 1 and 2 in alternative order 78

Table 2.27 The summary of applying three different strategies in series showing the number of the correctly located notes 79

Table 2.28 The summary of the rank of the correct notes using strategy 5 80

Table 2.29 The results of choosing the highest rank in strategy 5 as the solution 81

Table 2.30 The average number of known notes in different instruments that have been filtered out when using average frequency as the starting point 82

Table 2.31 The average number of known notes in different instruments that have been filtered out when using lower frequency as the starting point.... 82

Table 2.32 The number of known notes that have been filtered out as we apply the strategies 83

Table 2.33 The number of cases which the highest rank is equal to the expected note 83

Table 2.34 The number of results is correct using strategy 5 (with repeated answer) 85

Table 2.35 The number of cases where the second solution is correct when preventing the first solution from being repeated 86

Table 2.36 The number of correct second solution using strategy 5 with complete FT results 87

Table 2.37 The number of correct second solution using strategy 5 using different subtraction methods 89

Table 2.38 Summary of the accuracy rate using different methods to find the solution when three notes are played together 91

Table 2.39 Summary of the averages of the highest peak of the FT results using different methods when two notes are played together....................... 94

Table 2.40 Summary of the averages of the highest peak of the FT results using different methods when three notes are played together..................... 95

Chapter Three: Real Instruments

Table 3.1 Summary of the accuracy rate when playing single notes 105

Table 3.2 Summary of the accuracy rate in different experiments when two notes are played together 106

Table 3.3 Summary of the averages of the maximum height of FT in different calculation method in the three experiments in section 3.2.1 107

Table 3.4 Summary of the accuracy rate when using different sets of known notes including results from the previous section 108

Table 3.5 Summary of the accuracy rate using single MIDI instrument (AGP) as the known note 109

Table 3.6 Summary of accuracy rate of using 5 different instruments as the known note 110

Table 3.7 Summary of the accuracy rate playing a simple tune 111

Table 3.8 The accuracy rate when playing two notes using Method A (same as Table 3.2) 112

Table 3.9 The accuracy rate when playing two notes using the lowest of the maximum difference between the previous FT and the current FT as the starting point.. 112

Table 3.10 The difference between the results in Tables 3.8 and 3.9 113

Table 3.11 The accuracy rate when playing two notes using the lowest of the average difference between the previous FT and the current FT as the starting point 113

Table 3.12 The difference between the results in Tables 3.8 and 3.11 114

Table 3.13 A summary of the differences in the accuracy rate between method A and the two versions of method B 114

Appendix B

Table B.1 The comparison between the highest peak from Fourier Transform and the expected results when using different instrument on MIDI.

Table B.2 The results of playing three notes together.

Table B.3 The relationship on an acoustic grand piano using different starting point and compared to the original curve (15 Periods of wavelength) in percentage.

Table B.4 The relationship on an acoustic guitar (steel) using different starting point and compared to the original curve (15 Periods of wavelength) in percentage.

Table B.5 The relationship on an acoustic bass using different starting point and compared to the original curve (15 Periods of wavelength) in percentage.

Table B.6 The relationship on a violin using different starting point and compared to the original curve (15 Periods of wavelength) in percentage.

Table B.7 The relationship on a recorder using different starting point and compared to the original curve (15 Periods of wavelength) in percentage.

Table B.8 Summary of the number of correct results when three notes are played together.

Table B.9 Number of Errors occurs using the absolute value method with complete FT results when two notes are played together.

Appendix C

Table C.1 The results using different types of known notes when playing a song on a real grand piano.

Here we provide a CD which contains the programs we wrote to enable us to perform the experiments and analyses.

Introduction

This thesis describes the work we have done to study the automatic transcription of music to printed notation using a standard PC. After the subject of this thesis was accepted we began a literature review. This review showed we would need to reduce the scope of the work to enable us to make a valid contribution to the field. We have therefore limited the study to identifying notes by comparing them to MIDI and also notes from the same instrument. Later in our work we realised that the selection of the starting point for the comparison was very important. However the time available did not allow us to more than touch on the issue. This complexity has also created difficulties for other work (Hamer 2001 [16]) done by larger groups using more sophisticated equipment.

We quote from another worker "Our scope and treatise is limited by several factors, but especially by the limited amount of resources compared to the wide range of topics that are related with music transcription. Moreover, engaging in a research that is quite new to our laboratory, analysis of musical signals, called for paying the required attention to just finding the right points of emphasis and avoiding wrong assumptions in an early phase." (Klapuri 2002 [21]) The quoted author is one member of a group six. We quote further discussion on commercial products "Polyphonic transcribers – sad to say – work very poorly. Monophonic systems are more robust, of course."

Our research has shown similar difficulties to those of other workers and taking into account the limitations, our accuracy has been acceptable.

Chapter 1 gives the background to the basic techniques used in this thesis. In sections one to four, we will explain the music and file systems we used. In sections five and six, we will look at the advantages of using MATLAB and the use of Fourier Transform.

Chapter 2 explains and reports on the analysis of MIDI generated files. It covers the analysis on single and a small number of notes played together.

Chapter 3 will perform the similar analysis on a real piano.

We then report our conclusions.

Literature Review

When one examines the literature, it becomes clear that the mathematical study of music is a difficult field studied over many years and from many different aspects. Such studies are the basis for the ability to transcribe music using a PC. In this review we give a general overview with reference examples rather than extensive lists. Our other approach is to cite reviews, textbooks and homepages to enable the reader to home in on any particular interest they may have.

When the unaware see sheet music they would be inclined to believe that what they see tells them what they will hear, if it is played by a competent musician. Unfortunately the simple indication of the note to be played does not reveal the way the music is created by instrument or voice. The sound produced varies a great deal depending on the instrument. The pitch is the same but the timbre is different.

Pitched instruments may make their sound through mechanisms that depend on creating a vibration in a physical object and modifying it before it transmits through air to the audience. A "string" maybe struck as in a piano or "rubbed' as in a violin. The length of the string, either in the instrument, or modified by "fingering", will set the frequency of the vibration, hence the sound. There will usually be more than one frequency created as part of one "note". Further modification of sound is via the construction material of the instrument, through hollow spaces such as in the body of a violin, or sounding boards in a piano.

Wind instruments make sound through the use of air forced through by either special oral techniques or mechanisms such as reeds to cause vibration. Their sound is affected by the length of the tube the air passes through. Most Brass instruments shorten the length of the tube by using valves to close off part of the tube. Other wind instruments such as the flute, depend on shortening the amount of tube that can allow air vibration, by covering the holes in the tube. Shortening the tube length alters the frequency of the note. see Suggs (1966) [35]. The physics of musical instruments is very complicated and is explained in Fletcher and Rossing (1998) [10]. This reference may be the prior edition of Rossing, Wheeler and Moore (2002) [34] but is sufficiently different to be worth referring to separately.

These differences, and others, are the source of the complications found in the automatic transcription of music. Notes may "overlap" or be played together.

The science of sound is explained in Rossing, Wheeler, Moore (2002) [34].

A recent review of mathematics and music is Benson (2002) [3]. This is an expanded and augmented version of notes given for an undergraduate course at the University of Georgia. Water waves are waves where the local movement is at right angles to the direction of propagation (transverse waves) whereas sound waves are longitudinal waves with the local movements in the same direction as the propagation. Sound waves have four main attributes:

1. Amplitude is the size of the vibration and is perceived as loudness. The amplitude of typical everyday sound is only a small fraction of a millimetre.
2. Pitch corresponds to the frequency of the vibration.
3. Timbre corresponds to the shape of the frequency spectrum of the sound, and
4. Duration which is the length of time for which the note lasts.

However most vibrations do not consist of a single frequency and naming the "defining" frequency can be difficult. Music should be defined more in terms of the perception of sound rather than in terms of the sound itself. The perceived pitch of a sound may represent a frequency not actually present in the waveform -- a "missing fundamental". This phenomenon is a part of "Psychoacoustics".

Benson describes the functioning of the human ear and its limitations. It "hears within a range of about 20Hz (Hertz) to 20,000 Hz although it may "feel" the sound below 20Hz.

The relevance of sine waves in relation to sound is explained as lying in the differential equation for simple harmonic motion which can be taken as a close approximation to the equation of motion of a particular point on the basilar membrane of the human ear or anywhere else on the chain from the outside air to the cochlea. The limitations of the approximation are discussed. Harmonic motion is explained in mathematical terms. This leads to damped harmonic motion and resonance.

Refer to Benson (2002) [3] where he then discusses Fourier theory and a useful mathematical system, which is often used for the analysis of sound, being Fourier Transform.

The article usually regarded as the original, announcing the fast Fourier Transform as a practical algorithm, is Cooley and Tukey (1965) [6]. An early algorithm is given by Gold and Rabiner (1969) [12]. This algorithm has been used much later in New Zealand to develop an initial system for the identification of melodies. McNab, Smith, Bainbridge and Witten (1997) [25]. An earlier description of signal processing, which involves measurement of spectra using Fourier transform is given in the text, Rabiner and Gold (1975) [31].

The creation of music by instrument is one way of starting the chain. The other is the voice. Some studies are involved in relating sound produced to the characteristics of the larynx. An example is Bagshaw, Hiller and Jack (1993) [1]. The next link is the transmission of sound through air and its subsequent interaction with objects within the space or enclosing the space. These matters of acoustics are of importance in the transcription of music because they indicate the possibility of the sound being changed before it is identified. (Kinsler, L.E, Frey, Coppens, A.B. , and Sanders 1982 [19]). The shape of the room in which it is played, and the nature of the floor and its foundations, are two examples of the factors which affect the acoustics.

The human perception of music is another aspect which is also very complicated (Roederer 1995 [33]). The functioning of the human ear is an integral part of this perception. (Hudspeth 1985 [17], 1989 [18]) is another reference for this aspect. Further references are continually updated on the homepage (CSTR Publications 2001 [7]). Neuropsychology is the discipline that studies the neural system processes and functions linking the input received from the environment and body with the full behavioural and mental output. A text covering this is McAdams and Bigand (1993) [24]. A homepage that covers this field is Neuropsychology Central (2001) [29].

The transcription of music involves artificial perception and music recognition rather than human perception. Tanguiane (1993) [36] gives us a good overview of this topic. Work dates back to the early 1970's with the first experiments on automatic notation of monophonic music and a program for transcribing polyphonic music performed on a computer-wired keyboard. The limitations experienced at that time included only admitting pitched sounds, avoiding some pitch combinations, and requiring simple rhythmic structure and a constant tempo. Artificial intelligence methods began to be used in the early to the mid 1980's.

Artificial perception is the usual approach to the pattern recognition which is necessary to develop a computer system for the automatic notation of the performed music. It attempts to follow nature. A difficulty is that if a system of artificial perception is developed for one type of music, it may not be as successful with different styles or musical cultures. Another problem is that compared to human perception, a much longer passage of music is required for artificial perception to establish a reliable result. A singing voice, being less stable than an instrumental sound, is more difficult to analyse. By 1985 the studies branch into either chord/note recognition or rhythm/tempo recognition.

An early method for rhythm recognition was to develop an hypothesis concerning the rhythmic structure from the very first events which is then continuously confronted with current data and being modified if necessary.

We paraphrase a passage that states that a particular difficulty is that there are no explicit definitions of notes, chords, rhythm, and tempo, and so their recognition is complicated. Goto and Muraoka (1997) [14] discuss the issue of the evaluation of beat tracking systems and the need to rely on human intervention as an evaluation tool since the beat is a perceptual concept a person feels in music which is difficult to define in an objective way. In designing their measure for evaluation they considered subjective hand-labeled beat positions to be the correct beat times. The positions can be finely adjusted by playing back the audio with click tones at beat times and the user also defines a hierarchical rhythmic structure. They later (Goto and Muraoka, 1998 [15]) report on a real-time beat tracking system that recognises a rhythmic structure in real-world audio signals sampled from popular-music compact discs. A homepage as an entry into the literature on rhythm/tempo recognition is Goto (2002) [13].

A recent method (Klapuri, Virtanen and Holm 2000 [22]) for the estimation of the multiple pitches of concurrent musical sounds comprised of sung vowels and the whole pitch range of 26 musical instruments had error rates for mixtures ranging from one to six simultaneous sounds were 2.1%, 2.4%, 3.8%, 8.1%, 12% and 18% respectively. In musical interval and chord identification tasks, the algorithm outperformed the average of ten trained musicians. This gives an indication of the current accuracy being achieved.

It is interesting to realise that music recognition and voice recognition use the same methods and the studies overlap. A problem in common is the separation of several sounds being heard by the recognition system some of which are "interference" outside the voice or music being studied. A recent review is De Cheveigne (1993) [8]. Results from an implementation of this approach illustrated its ability to analyse complex, ambient sound scenes that would confound previous systems.

Homepages for groups working in this field include MIT Media Lab (2002) [27] and Klapuri (2002) [21]. Because these are updated continually, they are ongoing sources of recent work also giving references to other workers.

A system of artificial perception, which has some similarity to artificial intelligence is the use of neural network. Examples of this concept are given in Roberto (1993) [32]. Other work with neural networks can be accessed at CSTR publications (2001) [7].

The review to this stage showed that it was going to be very difficult for us to create a research project within the limit of time and facilities. Klapuri (2002) [21] mentions a similar problem. It is interesting to note that Klapuri made this statement after completing the thesis (Klapuri 1998 [20]) within Tampere University of Technology, Finland, which has an Audio Research group. Klapuri (1998) [20] gives a review of systems of artificial perception. The phrase onset time is defined as the instant of time when the sound starts playing.

An important distinction made by Tanguiane (1993) [36] is quoted.

"The difference between artificial perception and artificial intelligence in pattern recognition is understood as follows. Artificial perception is used for discovering structure in visual and audio images by self-organisation of data and segregation of patterns. Artificial intelligence is used for pattern identification by their matching to known concepts. Usually, the identification of already segregated patterns is much simpler than their recognition in data flows: thus artificial perception and artificial intelligence are complementary."

This distinction by Tanguiane (1993) [36] is not necessarily clear in the use of the phrase, artificial intelligence, by other workers, but it leads to another approach which is used in our studies.

In August 1983, music manufacturers agreed on a protocol that is called the "MIDI 1.0 Specification". General MIDI (Musical Instrument Digital Interface) is a standard that defines specific locations for different instrument sounds in the present memory of synthesisers (MIDI Manufacturers Association Incorporated 2001 [26]).

Therefore MIDI could be used as the known concept, referred to by Tanguiane (1993) [36] to develop a system of musical note identification. Benson (2002) [3] gives a recent review of concepts in digital music.

Work which has some similarities to ours has used matching methods to compare a hummed tune to a recorded database of songs. We refer to Ghias and Logan (1995) [11] and Blackburn and DeRoure (1998) [4]. They have used a pitch tracking method represented as a sequence of relative pitch changes (i.e a melodic pitch contour). We note that when a person hums a song it is monophonic.

The New Zealand Digital Library MELody inDEX (McNab, Smith. Bainbridge and Witten (1997) [25]) is a system that accepts acoustic input from the user created by a few notes sung into the microphone. It transcribes the melodies automatically from the microphone input then searches a database for tunes that contain the same or similar sung patterns. The tunes retrieved are ranked according to the closeness of the match. Different search criteria were used such as melodic contour, musical intervals and rhythm, and tests were performed using both exact and approximate string matching. They also performed tests on how people remember tunes. They concluded from these experiments that people needed a choice of several matching procedures and should be able to explore the results interactively in their search for a particular melody. This recent effort shows the recurring need for human interaction in these processes. The voice range was limited to from F2 to G5. This reference is dealing with monophonic sound and uses the algorithm on Gold and Rabiner (1969) [12]. The system depends on the user separating each note by singing da or ta to create a note boundary. The solutions to the other problems and the techniques used are described. On going access to the work of this group can be made via the homepage of Bainbridge (2002) [2].

Chapter 1
Background

1.1 Music Notes

There are two systems of notation used today. The British and American system is based on A B C D E F G. The French system (sol-fa) uses a system of syllables to represent the pitch. These syllables are doh ray me fah so la te. Doh always equals C. Both systems of notation use a clef sign and the placement of the notes on a stave, to denote the pitch of the note to be played or sung. Middle C which is common to both staves (the note which used to be on the line removed from the Great Stave) is equal to C4. C4 being C in the fourth octave on the piano. As the piano is the most commonly played instrument, and also has the greatest number of notes able to be played, my thesis will use the term C4 as opposed to Middle C.

A normal piano has a range from A1 to C8. The difference between each note is in their frequencies. They go from a lower frequency where A1 is 55 Hz to a high frequency where C8 is 4186Hz. Different musical instruments may differ slightly from the exact frequency standard that is used for each note. However in concept, all notes follow the same rules.

When a note of the same name is pitched an octave higher, the higher note is doubled in frequency. For example, A4 is 440Hz then A5 is 880Hz and A6 is 1760 Hz. Each semitone is equally spaced on a logarithmic scale, and because there are 11 semitones in each octave, the frequencies are as in Table 1.1 and Table 1.2. From now on we will use the word "notes" to replace "semitones".

Note		Process	Frequency (Hz)	Note	Process		Frequency (Hz)
A4	0		440	A5	$440*2^{12/12}$	$880*2^{0/12}$	880
A#4	1	$440*2^{1/12}$	466.16	A#5	$440*2^{13/12}$	$880*2^{1/12}$	932.33
B4	2	$440*2^{2/12}$	493.88	B5	$440*2^{14/12}$	$880*2^{2/12}$	987.77
C5	3	$440*2^{3/12}$	523.25	C6	$440*2^{15/12}$	$880*2^{3/12}$	1046.50
C#5	4	$440*2^{4/12}$	554.37	C#6	$440*2^{16/12}$	$880*2^{4/12}$	1108.73
D5	5	$440*2^{5/12}$	587.33	D6	$440*2^{17/12}$	$880*2^{5/12}$	1174.66
D#5	6	$440*2^{6/12}$	622.25	D#6	$440*2^{18/12}$	$880*2^{6/12}$	1244.51
E5	7	$440*2^{7/12}$	659.26	E6	$440*2^{19/12}$	$880*2^{7/12}$	1318.51
F5	8	$440*2^{8/12}$	698.46	F6	$440*2^{20/12}$	$880*2^{8/12}$	1396.91
F#5	9	$440*2^{9/12}$	739.99	F#6	$440*2^{21/12}$	$880*2^{9/12}$	1479.98
G5	10	$440*2^{10/12}$	783.99	G6	$440*2^{22/12}$	$880*2^{10/12}$	1567.98
G#5	11	$440*2^{11/12}$	830.61	G#6	$440*2^{23/12}$	$880*2^{11/12}$	1661.22
A5	12	$440*2^{12/12}$	880	A6	$440*2^{24/12}$	$880*2^{12/12}$	1760

Table 1.1 The process of finding frequencies.

Start from	C1	C2	C3	C4	C5	C6	C7	C8
C	32.70	65.41	130.81	261.63	523.25	1046.50	2093.00	4186.01
C#/Db	34.65	69.30	138.59	277.18	554.37	1108.73	2217.46	4434.92
D	36.71	73.42	146.83	293.66	587.33	1174.66	2349.32	4698.64
D#/Eb	38.89	77.78	155.56	311.13	622.25	1244.51	2489.02	4978.03
E	41.20	82.41	164.81	329.63	659.26	1318.51	2637.02	5274.04
F	43.65	87.31	174.61	349.23	698.46	1396.91	2793.83	5587.65
F#/Gb	46.25	92.50	185.00	369.99	739.99	1479.98	2959.96	5919.91
G	49.00	98.00	196.00	392.00	783.99	1567.98	3135.96	6271.93
G#/Ab	51.91	103.83	207.65	415.30	830.61	1661.22	3322.44	6644.88
A	55	110	220	440	880	1760	3520	7040
A#/Bb	58.27	116.54	233.08	466.16	932.33	1864.66	3729.31	7458.62
B	61.74	123.47	246.94	493.88	987.77	1975.53	3951.07	7902.13

Table 1.2 Table of frequencies displayed in octaves.

1.2 Digital Audio

The Compact Disc (CD) is the most commonly used recording medium. It uses Pulse Code Modulation (PCM) for recording digital audio. This technique is also used to create the wave files (with the file extension name *.wav*), which is a popular format for computer music files.

In PCM recording hardware, the recording device converts the variation of air pressure into voltages. An analog-to-digital converter then translates the voltages back to data (samples) and it is saved. Wave files are put into different formats, different *Sample Rates*, different *Byte Rates* and either one (mono), two (stereo) or more channels for the sound.

In Compact Disc (CD) recording, the standard sample rate is 44,100Hz with 16 bits for each sample (44,100Hz means 44,100 samples per second). There are two sets of 16-bits long data for each sample. Each set represents one channel. One channel represents the left output device and the second channel represents the right output device. This is commonly known as stereo.

1.3 Wave File Format

Microsoft has published a standard called "RIFF" to store multimedia data. Microsoft is the largest computer software company in the world, so the standard is used very frequently. The type of data that a RIFF file contains is indicated by the file extension.

Examples of data that could be stored by RIFF files are:

- Audio/visual interleaved data (.AVI)
- Waveform data (.WAV)
- Bitmapped data (.RDI)
- MIDI information (.RMI)
- Colour Palette (.PAL)

- Multimedia movie (.RMN)
- Animated cursor (.ANI)
- A bundle of other RIFF files (.BND)

The wave (*.wav*) file format is a subset of the RIFF specification. The function of wave files is to store audio data.

The RIFF file starts with a file header followed by a sequence of data *chunks*.

Offset (hex)	Contents
0000	'R', 'I', 'F', 'F'
0004	Length of the entire file
0008	Form type (4 characters)
000B	First chunk type (4 characters)
0010	First chunk length
0014	First chunk's data
…	

In *.wav* files more than one format can be used. We have used the most basic. The description of this basic format and examples are shown in Appendix A.

Figure 1.1 is a graphical example of a wave file.

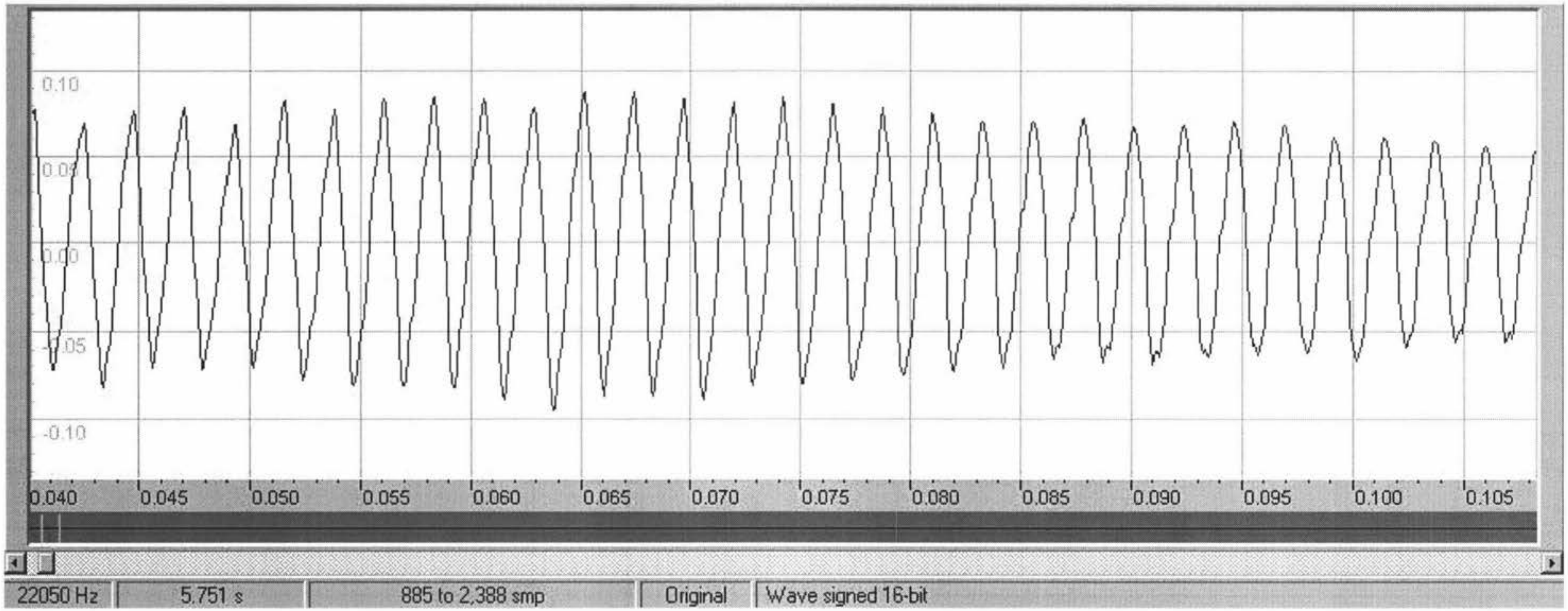

Figure 1.1 Example of a sound wave (Acoustic Grand Piano from MIDI, A4).

1.4 MIDI

MIDI stands for "Musical Instrument Digital Interface" and is a digital communication language that allows instruments to communicate with each other, even if they are made by other manufacturers, by use of a simple cable. The special feature of MIDI is that it transmits commands instead of audio signals.

The sounds in a General MIDI synthesizer or sound module are organized into 16 groups, with 8 variations of the instrument within each group. For example, all piano sounds are from 1 to 8, with an acoustic piano sound starting at 1. This makes it easier to re-create the right instrument for a specific part of a MIDI sequence. Table 1.3 displays the table of instruments used by MIDI.

In MIDI, numbers are also used to represent notes. e.g. B3 -> 59, C4 -> 60, C#4 -> 61, D4 -> 62, etc.

With the advent of MIDI, a whole new world for musicians was opened, allowing them to be more creative and to be able to compose music in a way never previously possible.

For further details about MIDI see "basic MIDI" (White, 1999 [38])

Piano		Bass		Reed	
1	Acoustic Grand Piano	33	Acoustic Bass	65	Soprano Sax
2	Bright Acoustic Piano	34	Electric Bass (Finger)	66	Alto Sax
3	Electric Grand Piano	35	Electric Bass (Pick)	67	Tenor Sax
4	Honky-Tonk Piano	36	Fretless Bass	68	Baritone Sax
5	Electric Piano 1	37	Slap Bass 1	69	Oboe
6	Electric Piano 2	38	Slap Bass 2	70	English Horn
7	Harpsichord	39	Synth Bass 1	71	Bassoon
8	Clavi	40	Synth Bass 2	72	Clarinet
Chromatic Percussion		**Strings**		**Pipe**	
9	Celesta	41	Violin	73	Piccolo
10	Glockenspiel	42	Viola	74	Flute
11	Music Box	43	Cello	75	Recorder
12	Vibraphone	44	Contrabass	76	Pan Flute
13	Marimba	45	Tremolo Strings	77	Blown Bottle
14	Xylophone	46	Pizzicato Strings	78	Skakuhachi
15	Tubular Bells	47	Orchestral Harp	79	Whistle
16	Dulcimer	48	Timpani	80	Ocarina
Organ		**Ensemble**		**Synth Lead**	
17	Drawbar Organ	49	String Ensemble 1	81	Lead 1 (Square)
18	Percussive Organ	50	String Ensemble 2	82	Lead 2 (Sawtooth)
19	Rock Organ	51	SynthStrings 1	83	Lead 3 (Calliope)
20	Church Organ	52	SynthStrings 2	84	Lead 4 (Chiff)
21	Reed Organ	53	Choir Aahs	85	Lead 5 (Charang)
22	Accordion	54	Voice Oohs	86	Lead 6 (Voice)
23	Harmonica	55	Synth Voice	87	Lead 7 (Fifths)
24	Tango Accordion	56	Orchestra Hit	88	Lead 8 (Bass & Lead)
Guitar		**Brass**		**Synth Pad**	
25	Acoustic Guitar (nylon)	57	Trumpet	89	Pad 1 (New Age)
26	Acoustic Guitar (steel)	58	Trombone	90	Pad 2 (Warm)
27	Electric Guitar (jazz)	59	Tuba	91	Pad 3 (Polysynth)
28	Electric Guitar (clean)	60	Muted Trumpet	92	Pad 4 (Choir)
29	Electric Guitar (muted)	61	French Horn	93	Pad 5 (Bowed)
30	Overdriven Guitar	62	Brass Section	94	Pad 6 (Metallic)
31	Distortion Guitar	63	SynthBrass 1	95	Pad 7 (Halo)
32	Guitar Harmonics	64	SynthBrass 2	96	Pad 8 (Sweep)

Table 1.3 Table of the instruments in MIDI and the numbers to denote each instrument (Continued on next page).

Table 1.3 (cont.)

Synth Effect		Percussive		
97	FX 1 (Rain)	113	Tinkle Bell	
98	FX 2 (Soundtrack)	114	Agogo	
99	FX 3 (Crystal)	115	Steel Drums	
100	FX 4 (Atmosphere)	116	Woodblock	
101	FX 5 (Brightness)	117	Taiko Drum	
102	FX 6 (Goblins)	118	Melodic Tom	
103	FX 7 (Echoes)	119	Synth Drum	
104	FX 8 (Sci-Fi)	120	Reverse Cymbal	
Ethnic		**Sound Effect**		
105	Sitar	121	Guitar Fret Noise	
106	Banjo	122	Breath Noise	
107	Shamisen	123	Seashore	
108	Koto	124	Bird Tweet	
109	Kalimba	125	Telephone Ring	
110	Bagpipe	126	Helicopter	
111	Fiddle	127	Applause	
112	Shanai	128	Gunshot	

1.5 MATLAB

MATLAB is a computer program for numerical analysis well known to mathematicians. It allows us to produce highly accurate results. It is very easy to produce graphs and solutions by using a wide range of built-in functions (include DFT and FFT).

There are other advantages for C programmers using MATLAB. It allows users to create their own script files, like C, and also provides some functions which are exactly like C, e.g. sprintf, fscanf, fwrite, etc. C programmers can easily pick up the program and perform all kinds of functions without the work of a C compiler. MATLAB is a post-compiler program, so it will run a compiler every time you start the program. It may spend extra time at the beginning, but as it moves on to construct and look at graphs, it is quicker and easier for users.

In this thesis, we will use MATLAB 5.1 (student version) to run all programs. There are some differences between the student and the professional version. One is that the professional version can calculate and store with a larger memory. The precision for calculation is also better.

1.6 Fourier Transforms

Figure 1.1 (shown earlier) is an example of a sound curve (A4 on recorder). It was created using MIDI converted into wave format. In this example a 22050Hz sample rate is used, and the sample size is 1504 (from sample 885 to 2388). By counting the number of periods, we will get the result 30, and consequently the wavelength is 50.1. (i.e. 1504 / 30 = 50.1)

$\Rightarrow$ 50.1 / 22050 = 0.002274 seconds (wavelength in time)

$\Rightarrow$ 440Hz which is A4 [as noted earlier].

Counting the periods and calculating the wavelength, or the frequency, is tedious. To alleviate this problem we use the Fourier Transform (FT). It is named after a French mathematician, Joseph Fourier (1768-1830), who developed a mathematical way of expressing the time-domain functions as frequency spectra. It converts any periodic function (waves) into the form of sine and cosine functions.

There are three types of Fourier Transform. The basic form of Fourier Transform is called Continuous Fourier Transform, but the most commonly used forms are the Discrete Fourier Transform (DFT) and the Fast Fourier Transform (FFT). They are conceptually similar (all based on the Fourier Series), but the DFT and FFT use summation of fixed intervals instead of using integration over the whole function as does CFT.

$$a_k = \int_{-\infty}^{\infty} x(t)\exp(-j2\pi kt)\, dt \qquad \text{............ (Continuous Fourier Transform)}$$

$$a_k = \frac{1}{N}\sum_{n=0}^{N-1} x(n)\exp(-j2\pi kn/N) \qquad \text{............ (Discrete Fourier Transform)}$$

$$x(t) = a_0 + \sum_{i=-\infty}^{\infty} a_i \cos(2\pi it/T) + b_i \sin(2\pi it/T) \quad \text{............ (Fourier Series)}$$

1.6.1 The differences between DFT and FFT

The sample size for FFT is always fixed at a power of 2 while the sample size for DFT can be any number. FFT is numerically more efficient when solving problems. It requires approximately N*$\log_2$(N) complex multiplication and addition operations which are less operations then for DFT which requires approximately N^2. For the details of the mathematical aspects of Fourier Transform the reader can refer to "*Digital Signal Processing By Lynn*" (Lynn P.A., Fuerst W., 1994 [23]) and "*C++ Algorithms for Digital Signal Processing*" (Embree P.M., Danieli D., 1999 [9]).

1.6.1.1 Time efficiencies

The time efficiencies of DFT and FFT are shown in Table 1.4 and Figure 1.2. We can see FFT is more efficient than DFT when the sample size is large and fixed.

Transformation Length (N)	DFT Operations (N^2)	FFT Operations (N * $\log_2$(N))
8	64	24
16	256	64
32	1024	160
64	4096	384
128	16384	896
256	65536	2048
512	262144	4608
1024	1048576	10240

Table 1.4 The number of operations using DFT and FFT.

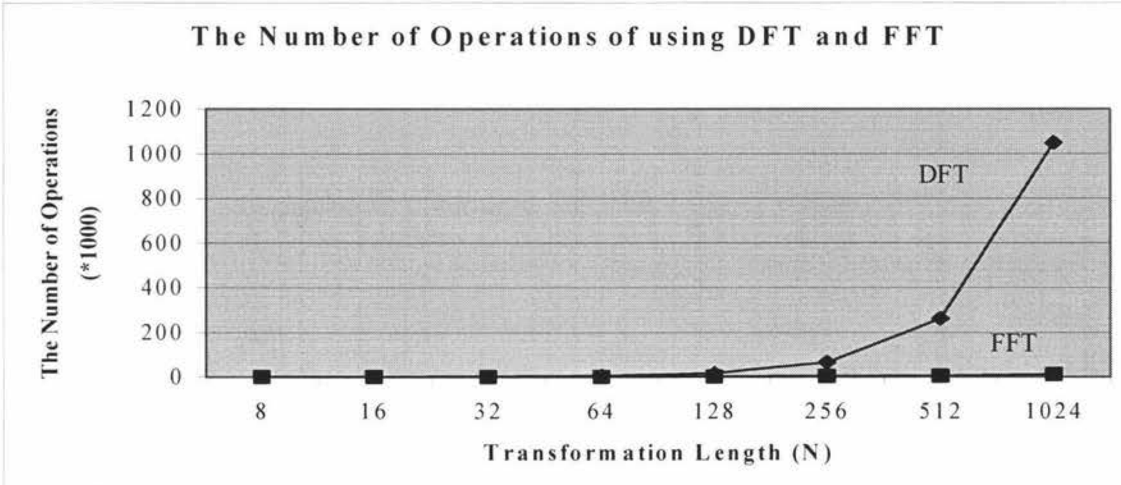

Figure 1.2 The number of operations of using DFT and FFT.

1.6.2 Symmetry

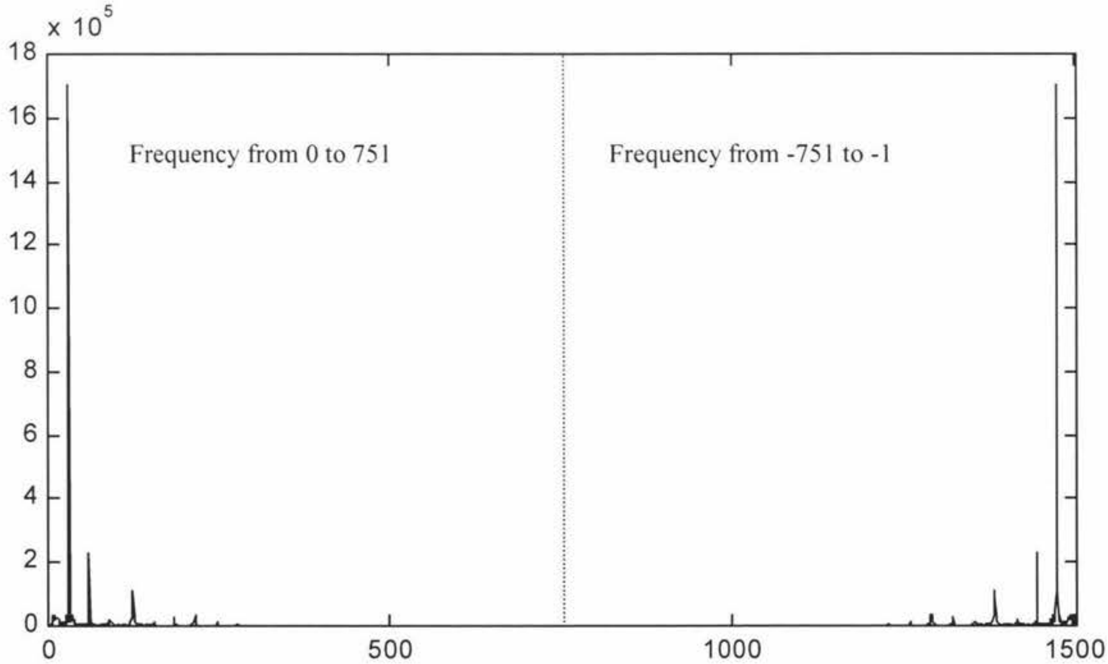

Figure 1.3 Results of FT using MATLAB.

The results of FT are in complex value. The result of the FT is a combination of a sine and a cosine function. i.e. an image and a real value of the FT. This value can be combined to produce an absolute value which is the square root of the sum of the squares of the real value and the image value.

Figure 1.3 shows the results of FT in MATLAB with a sample size of 1504. MATLAB uses the "fft" program. If the sample size is a power of 2, FFT will be calculated, otherwise it will calculate using DFT. The value on the x-axis describes the number of periods that can be found in the samples. e.g. one of the peaks in Figure 1.3 is at 30 (on x-axis) which means there are 30 periods in the samples. The results from MATLAB start from 0 and go to 1503. The results of the FT (which is frequency) should be displayed from – Sample Size/2 to Sample Size/2.

Figure 1.3. only plots part of the full transformation. The plot continues in concept before and after the figure. Using this concept, we recast the results to Figure 1.4. We have removed 751 samples from the end of Figure 1.3 and begun Figure 1.4 751 samples earlier than Figure 1.3

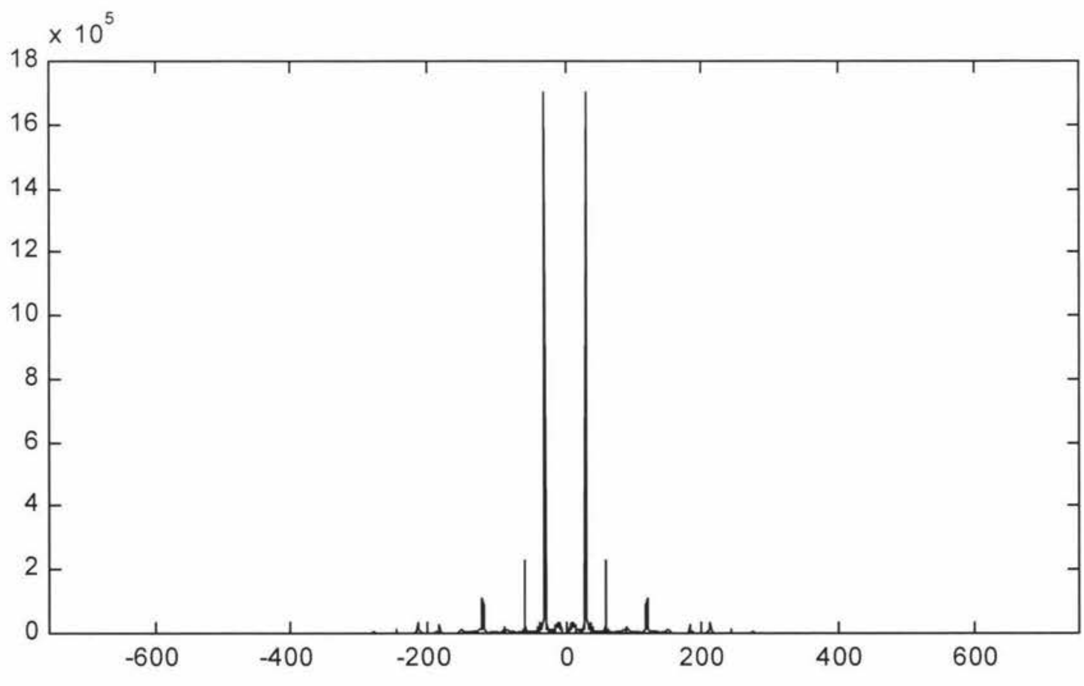

Figure 1.4 Results of FT after rearrangement.

i.e.

1503	in Figure 1.3	becomes	-1	in Figure 1.4	(i.e. 1503-1504)
1502		$\rightarrow$	-2		
1501		$\rightarrow$	-3		

...

1474 (another peak in Figure 1.3)

$\rightarrow$ -30 (i.e. 1474-1504)

...

In Figure 1.4 there are two peaks in the graph which are at 30 and –30 (30 and 1473 in Figure 1.3). 30 is the number of cycles or periods in the sample. This is the result we found when counting the number of periods in Figure 1.1. We quote from Bourke (1993) [5]. *"The negative frequency samples are also the inverse of the positive frequency samples."*

Because the FT consists of two parts one being the mirror of the other for simplicity the remainder of this thesis will only display the results of the first part.

For single notes, any error created by the FT starting and finishing at different points in the phase of the wave is not significant. Figures 2.7 and 2.8 show in 2D and 3D the resulting FT when using different starting point with the same instrument and the same note. It can be seen from these figures that any variations in the FT is indiscernible. The effects of phase on FT with multiple notes is recognised later.

Chapter 2
Using MIDI To Analyse Sound

We used MIDI because it can produce better results without interference by noise from the outside world. In the first section we will explain how we generated the MIDI files. The second to the fifth sections will explain the meaning of, and the differences between, the main parameters. In the remaining sections we will look at the analysis using different methods for finding the correct solution for the note.

We have explained that MIDI is a communication language that allows instruments to communicate with each other. A MIDI file on a computer (which is the file standard to store MIDI commands) uses a *.mid* as its extension. The MIDI language is written in binary form which is difficult for humans to modify. There are many MIDI based programs that contain graphical user interfaces such as "*Anvil Studio*" or "*Cakewalk Home Studio*".

We used a program called *t2mf* (mf2t, 1997 [29]) to convert a special text file into a *.mid* file. The following is an example of a simple text file that can be converted into *.mid* file using *t2mf*.

```
MFile 1 1 240
MTrk
0 Tempo 750000
0 PrCh ch=1 p=0
0 On ch=1 n=60 v=64
240 Off ch=1 n=60 v=0
TrkEnd
```

Explanation:

```
MFile 1 1 240
[Mfile <format, 0 or 1> <No of tracks> <division>]

[<division> is either a positive number (giving the
time resolution in clicks per quarter note) or a
negative number followed by a positive number (giving
SMPTE timing).]

MTrk                    [Start of track]

0 Tempo 750000
0 PrCh ch=1 p=0
[0 is the time, ch is the channel number and p is the
number of the instrument (0 = acoustic grand piano),
which is 1 less than the value in Table 1.3 on page
16.  That is because the MIDI file starts counting
from zero.]

0 On ch=1 n=60 v=64
["On" means turn on the volume/note.  n is the number
to represent the note, 60 is C4.  v is the volume.]

240 Off ch=1 n=60 v=0
[Again 240 is the time, which is equal to a crotchet
as the number for division is 240. Off means turn off
the volume/note.  v = 0 means no sound is produced.]

TrkEnd    [End of the track]
```

Although MIDI produces sounds, the computer (using a program such as *Media Player*) cannot display it as wave format unless we record the sound prior to the analysis and this is too time consuming. Therefore we used a program called *TiMidity* (TiMidity, 2001 [37]). It converts *.mid* files to *.wav* files. In Figure 2.1 we show the concept of using *t2mf* and *TiMidity*.

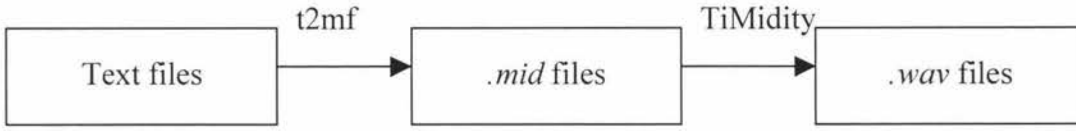

Figure 2.1 The concept of translating text files to *.wav* files.

There are two versions of *TiMidity*. "*timidity-gui*" is for the Windows environment, and *timidity-con* is for DOS. We will use the DOS version of TiMidity because MATLAB can make use of the DOS commands and runs it automatically without any human interaction.

The following is an example of TiMidity commands in the DOS environment.

```
timidity-con -OwM -s 22050 -o test.wav test.mid
```

`-OwM`	-Ow means Output format is RIFF WAVE, and M means monophonic
`-s 22050`	means Sample Rate is 22050Hz.
`-o test.wav`	means the output file name is 001.wav.
`test.mid`	is the input file name.

2.1 Sample Rate, Sample Size, Frequency and Wavelength

In this section we discuss the meanings and the relationships of the sample rate, the sample size, the frequency and the wavelength.

2.1.1 The sample rate

A Sample rate stands for the number of samples per second. The unit is Hz. There must be at least 2 samples to create a periodic wave (otherwise it is only a straight line). Because of this the FT can only detect a frequency range from 0 to half of the sample size of the data. This thesis will report on experiments analysing a frequency range below 10,000 Hz. Hence the sample rate needs to be at least 20,000 Hz. We used 22050 Hz as the sample rate because it is half of the sample rate for most CD's (44100 Hz), and it is also commonly the standard used in the studies of sound.

2.1.2 The frequency, the wavelength and the sample size

For most pianos, the range of frequencies starts at 55 Hz (A1) and goes to 4086Hz (C8) i.e. from 882 samples per cycle [22050/55] (wavelength = 1 / frequency) to 5 samples per cycle [22050/4086], using the 22050Hz sample rate. The duration of the note is also important, especially for piano type musical instruments. In the case of a relatively short duration note, the volume may reduce to zero in the middle of the sample data if the sample size is too large. This will invalidate the result of FT. As a general rule, duration is inversely proportional to the frequency. Figure 2.2 and 2.3 show the sample data for A1 and C8. We can see from the figures, the duration of A1 with a low frequency and a long wavelength is longer than that of C8 with a high frequency and a short wavelength. This is because the higher the frequency, the more friction there is with the air and in the instrument itself, which reduces the duration of the note. Table 2.1 shows an estimation of the duration of several notes on Acoustic Grand Piano (AGP) (MIDI 1 in Table 1.3) where all the notes are equal to a crochet (generally this applies to all the experiments in chapter 2). In the example of C8, the sample size needs to be as low as 2000 to get an accurate FT result.

All the notes we played in MIDI are equal to a crotchet (in common time).

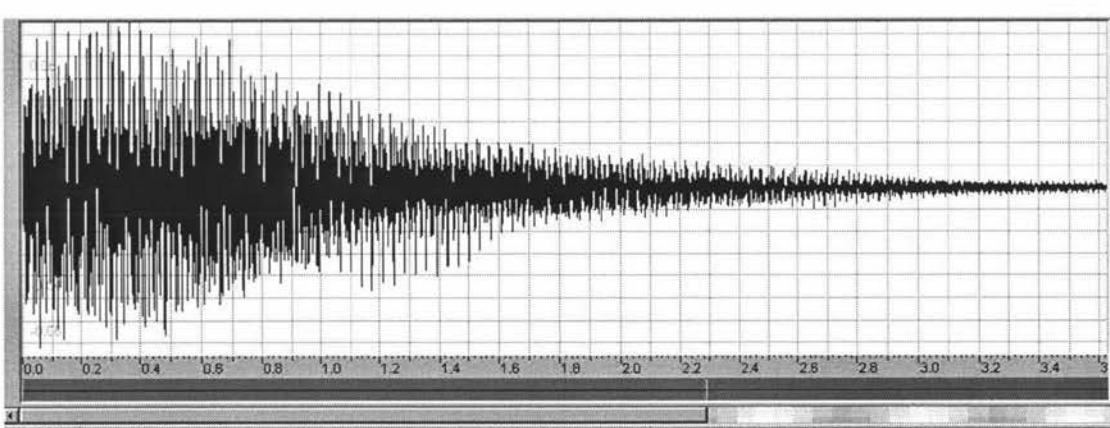

Figure 2.2 Wave format for A0 (Samples from 0 to 80,000).

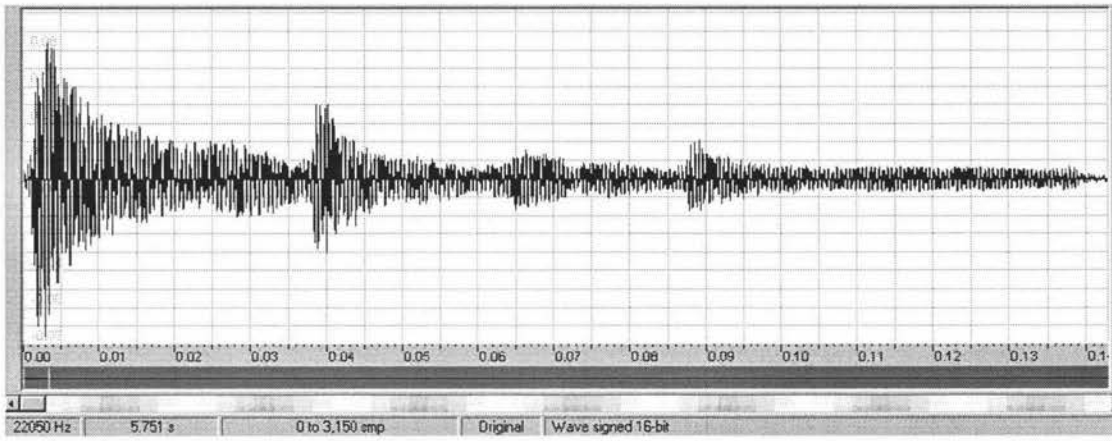

Figure 2.3 Wave format for C8 (Samples from 0 to 3,150).

MIDI Value	Note	Duration in Samples 22050	Duration Time (secs)
21	A0	60,271	2.7334
31	G1	44,414	2.0142
41	F2	23,003	1.0432
51	D#3	28,971	1.3139
61	C#4	25,998	1.1790
71	B4	12,277	0.5568
81	A5	6,639	0.3011
91	G6	5,357	0.2429
101	F7	4,171	0.1892
108	C8	2,119	0.0961

Table 2.1 The estimated duration of notes.

However a problem arises from the fact that DFT or FFT are only notated in natural numbers on the domain, instead of real numbers. The results of FT for each note are shown in Table 2.2. We can see there are several notes using the same expected period which do not appear in the table. The empty spaces on the table signify the result is the same as that in the closest lowest frequency note on that row. For example, with a sample size of 2000, there are notes like B0 to D1 which would be 3 and E1 to G1 which would be 4. 4096 samples appear to be a good choice when using FFT but it only covers the range of frequency approximately between D#2 to G7. We used 5000 as the sample size and so we were forced to use DFT because 5000 is not a power of 2. It covered the frequency from A#1 to B6. With a sample size of 5000 certain notes with short duration may die out as small values. This could affect the results of the FT, but these differences are assumed to be minor as it is only a small part within the whole sample. From now on we will study the frequency range between C2 to B6 because this problem is not significant above C2.

2.2 The Effect of Using Different Sample Sizes

Figure 2.4 and 2.5 show the effect of using different sample sizes (From 4000 to 7000) using 100 as the step size, on an AGP when playing C2, Figure 2.4 displays in two dimensions and Figure 2.5 displays in three dimensions. The two graphs show that C2 (65.41 Hz) is not the highest peak but C3 (130.81 Hz) or F5 (698.46 Hz) is. We will explain this later. In this section we will concentrate on the effect of using different sample sizes. Figure 2.4 (same as Figure B.2 in Appendix B) contains all the results using different sample sizes. The thickness of the line in Figure 2.4 is due to the changes of the sample size. The purpose of Figure 2.4 is to show that there are no sudden discrepancies as the sample size increases. Figure 2.5 (same as Figure B.3 in Appendix B) is the three-dimensional result which shows that the results using different sample sizes are very similar. As we compare the line for each sample size, we find the changes in spectrum are consistent. The conclusion is that the effect is small if the changes in the sample size are small, but grows as the sample size increases.

Figure B.11 and B.12 (in Appendix B) are the results of FT of a "Bright Acoustic Piano" (MIDI 2 in Table 1.3) when playing note G#3 (207.65 Hz), and Figure B.20 and B.21 are the results of FT of "Electric Acoustic Piano" (MIDI 3) when playing note E5 (659.26 Hz). The highest peaks in these curves are as expected.

Sample Size	**Frequency (Hz)**												
	A0	**A#**	**B**	**C1**	**C#**	**D**	**D#**	**E**	**F**	**F#**	**G**	**G#**	**A**
	27.50	**29.14**	**30.87**	**32.70**	**34.65**	**36.71**	**38.89**	**41.20**	**43.65**	**46.25**	**49.00**	**51.91**	**55.00**
2000	2	3					4					5	
3000	4				5			6			7		
4000	5		6			7			8		9		10
5000	6	7			8		9		10		11	12	
	A1	A#	B	C2	C#	D	D#	E	F	F#	G	G#	A
	55.00	58.27	61.74	65.41	69.30	73.42	77.78	82.41	87.31	92.50	98.00	103.83	110.00
2000			6			7			8		9		10
3000		8		9		10	11		12	13		14	15
4000		11		12	13		14	15	16	17	18	19	20
5000		13	14	15	16	17	18	19	20	21	22	24	25
	A2	A#	B	C3	C#	D	D#	E	F	F#	G	G#	A
	110.00	116.54	123.47	130.81	138.59	146.83	155.56	164.81	174.61	185.00	196.00	207.65	220.00
2000		11		12	13		14	15	16	17	18	19	20
3000		16	17	18	19	20	21	22	24	25	27	28	30
4000	20	21	22	24	25	27	28	30	32	34	36	38	40
5000	25	26	28	30	31	33	35	37	40	42	44	47	50

Table 2.2 The relationship between frequency, the results for the Fourier Transform and the sample size.

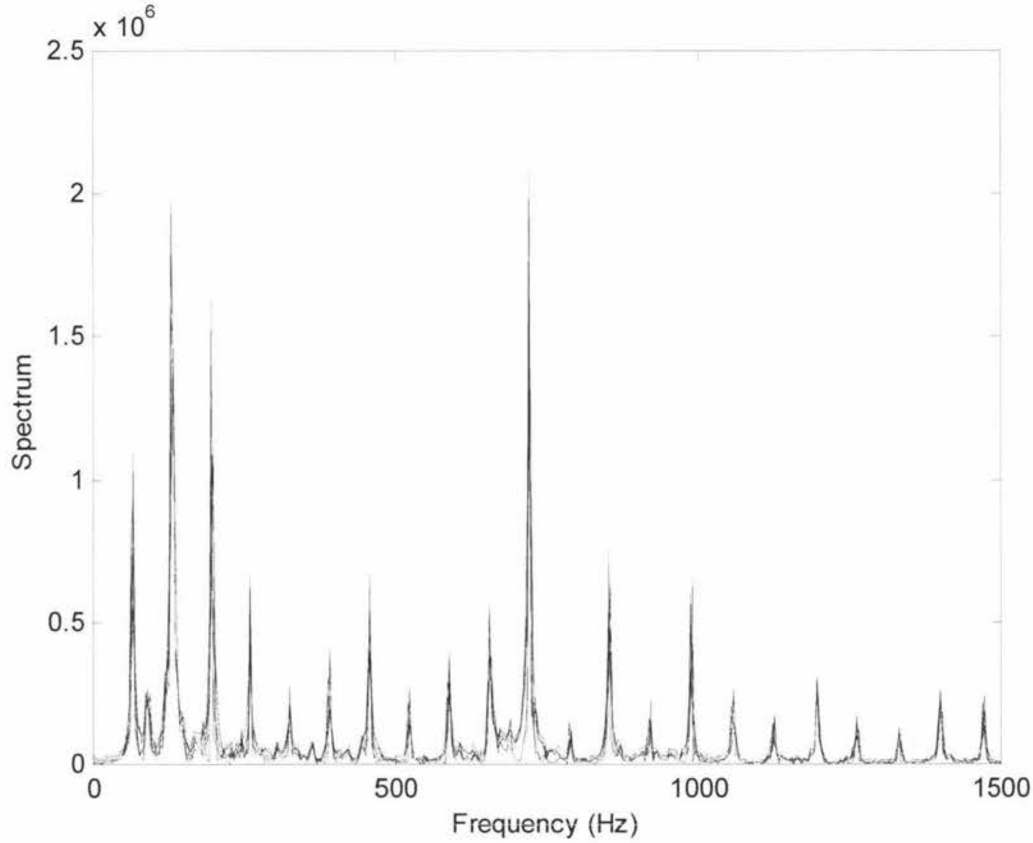

Figure 2.4 The results of Fourier Transform in 2D when using different sample sizes on an acoustic grand piano when playing the note C2 (65.41 Hz).

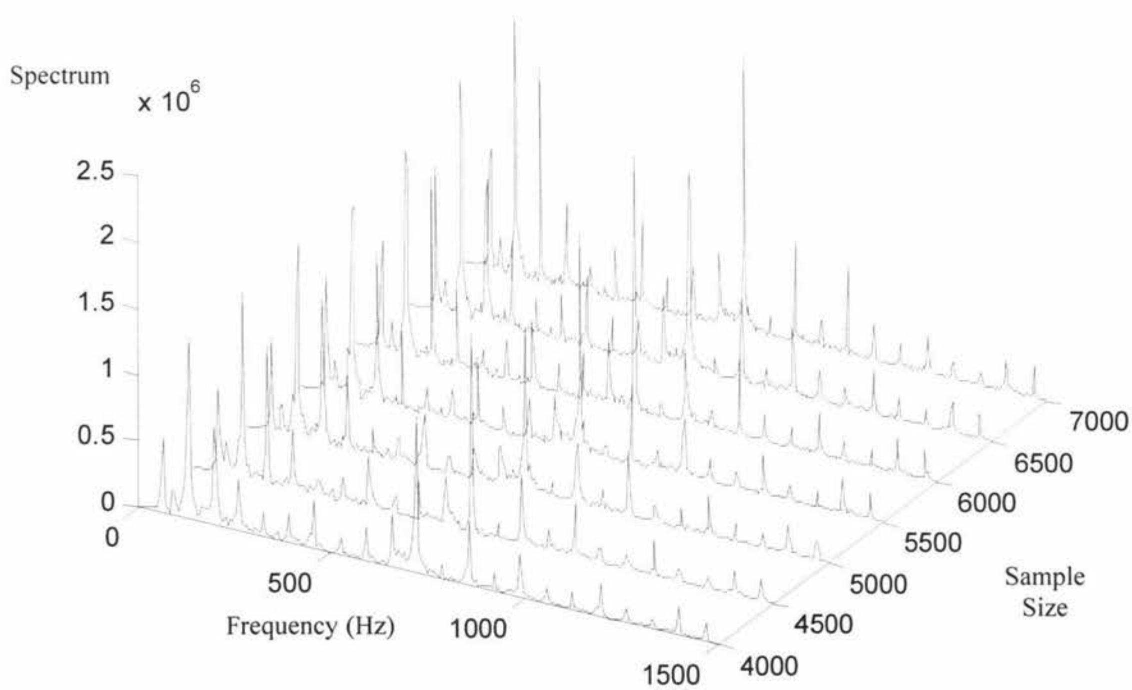

Figure 2.5 The results of Fourier Transform in 3D when using different sample sizes on an acoustic grand piano when playing the note C2 (65.41 Hz).

2.3 The Effect of Using Different Starting Points

Figures 2.7 to 2.12 (more details in Appendix B) are the FT results of using different Starting Points (from 1 to 5000 with a step size of 100) on an AGP when playing C2.

Figure 2.7 is the results of FT displayed in two dimensions and Figure 2.8 is three dimensional. Another way of graphing the results are displayed in Figure 2.9 and 2.10. The results of each FT is redrawn with each frequency being shown as a percentage of the highest peak. Figure 2.9 shows that C3 (130.81 Hz) and F5 (718.89 Hz) can both be the highest peak for at least one of the starting points. Figures 2.8 and 2.10 both show the highest peak varies between C3 or F5 depending on the starting point. This is because the frequencies in each FT vary unevenly as you change the starting point

A third way to display the results (taking this variation into account), is to add all the heights from each FT for each note. This display is in Figure 2.11. In Figure 2.12, we display Figure 2.11 as percentages with the highest peak as 100 (%). This shows that on average, F5 is the highest peak and C3 is approximately 92% of it.

Figure B.13 to B.18 and B.22 to B.27 (in Appendix B) show the results of FTs using different notes with different sample sizes. We found that the results for the Bright Acoustic Piano (MIDI 2) (G#3) and the Electric Grand Piano (MIDI 3) (E5) are stable, but we cannot assume that is the case for all notes (such as in Figure 2.12).

The choice of a starting point is very important. When we analysed from the beginning to the end of the note, we found that the wave and the FT results of the note will only become stable after a short period of time. Because a note usually combines several frequencies, the various frequencies within a note decay at different rates. Therefore it is better to exclude both these unwanted instabilities from the analysis to avoid an inaccurate reading.

In general the waveform of a musical note has three different stages. These stages are attack, steady-state, and decay. They may be described differently in different books. The duration of each stage is different and can be affected by the duration of the note.

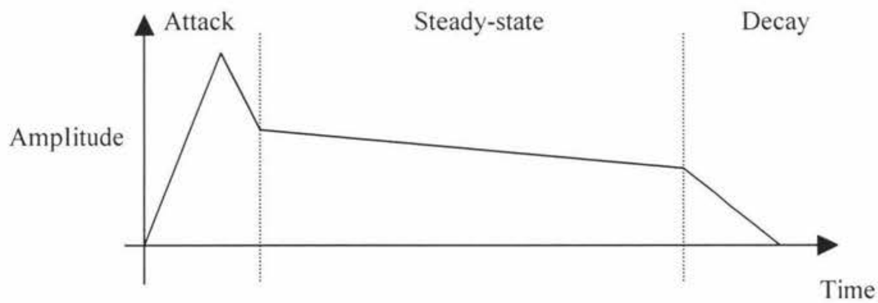

Figure 2.6 Simplified amplitude evolution of a musical note.

The position of the starting point and the sample period that we use are ideally inside the steady-state. In the fixed sample size (5000 samples) certain notes with high frequency may not fit inside the steady-state. However we shall ignore this situation and assume the result is not seriously affected because we have limited the frequency range to below C7. In any note below C7 the portion of the note in the unstable states is small.

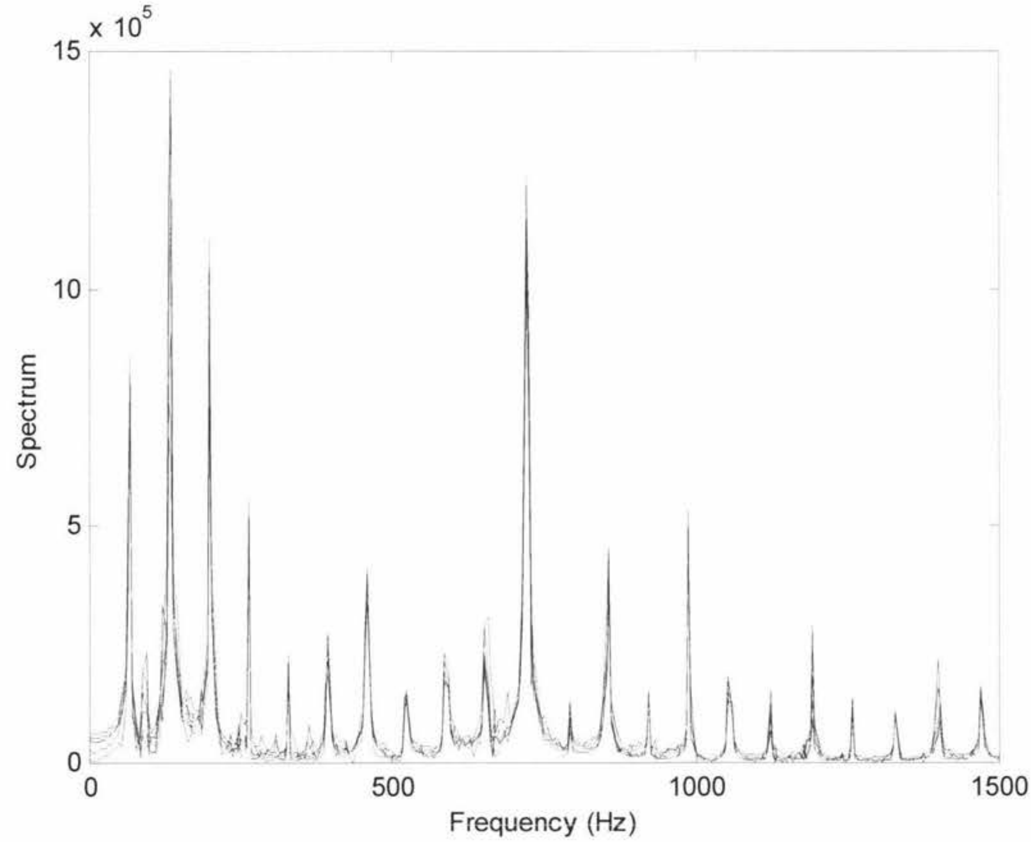

Figure 2.7 The results of Fourier Transform in 2D when using different starting points, on an acoustic grand piano when playing the note C2 (65.41 Hz).

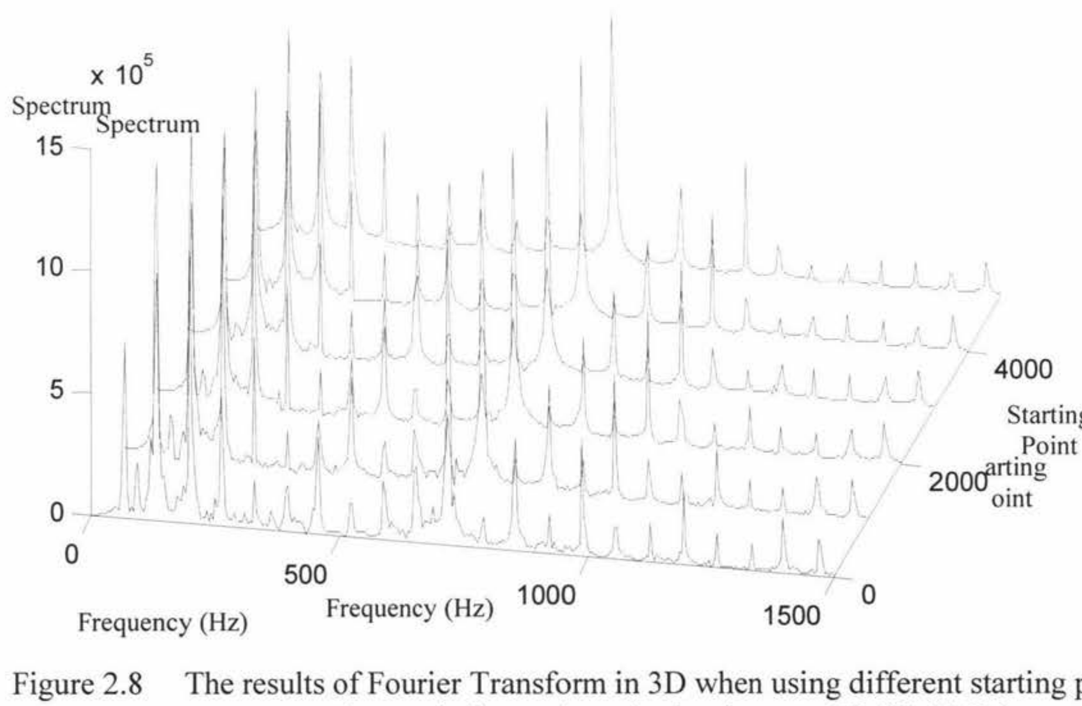

Figure 2.8 The results of Fourier Transform in 3D when using different starting points, on an acoustic grand piano when playing the note C2 (65.41 Hz).

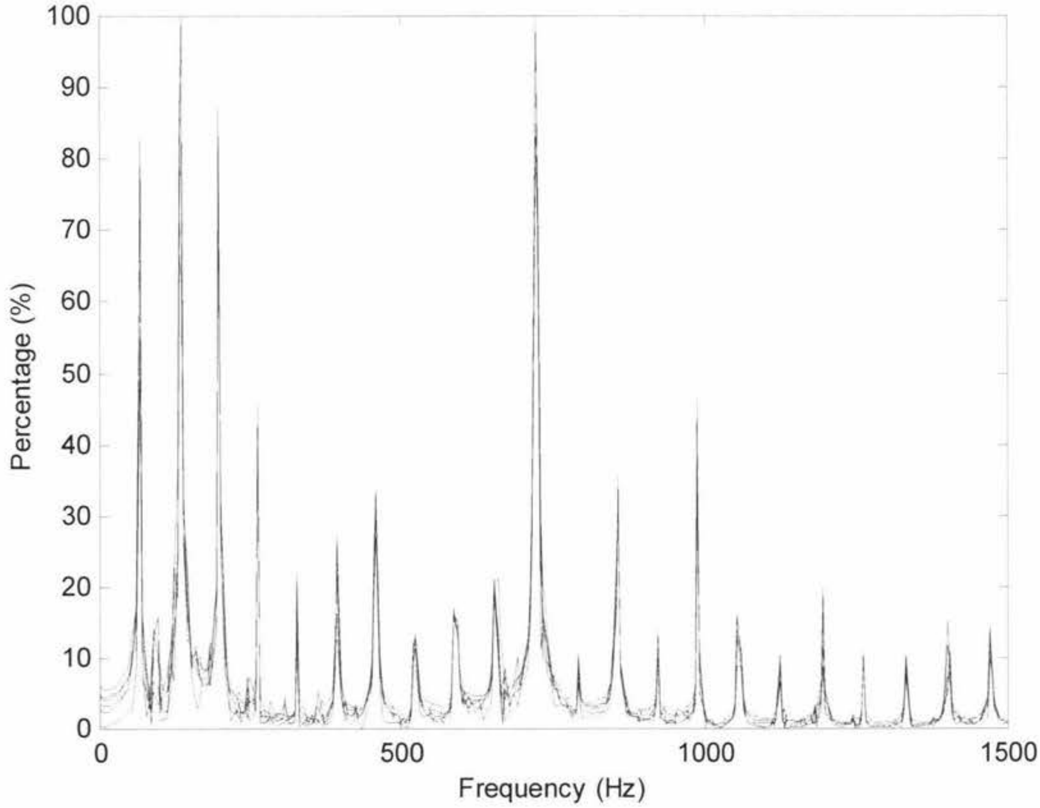

Figure 2.9 The results of Fourier Transform in 2D when using different starting points, on an acoustic grand piano when playing the note C2 (65.41 Hz) (in percentages).

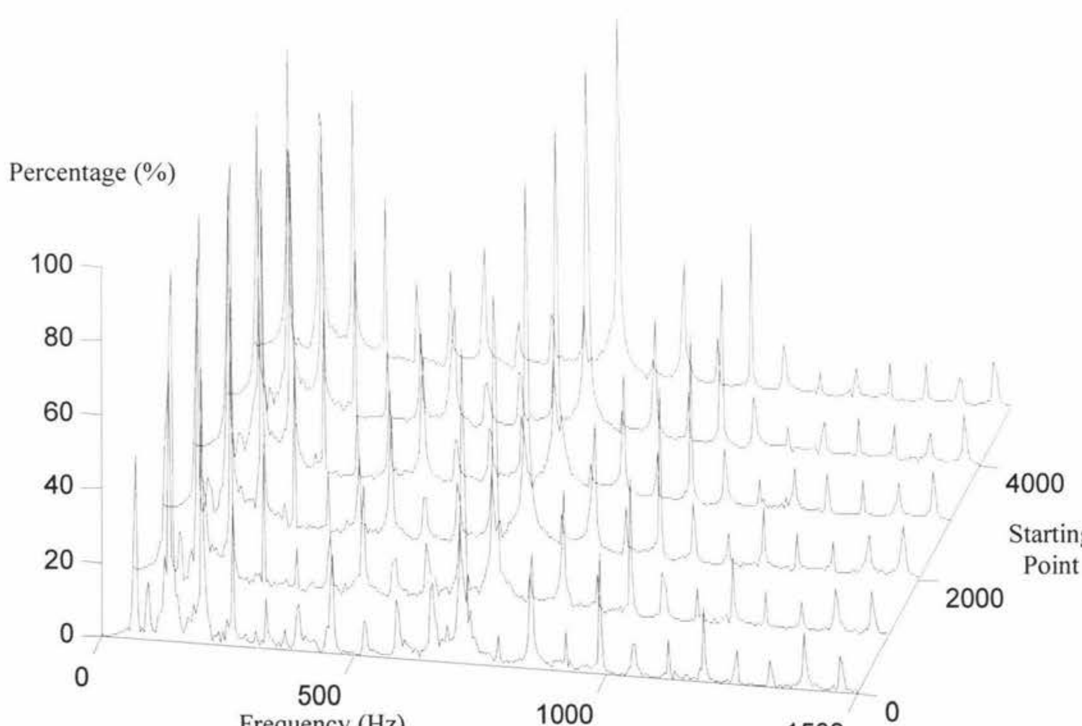

Figure 2.10 The results of Fourier Transform in 3D when using different starting points, on an acoustic grand piano when playing the note C2 (65.41 Hz) (in percentages).

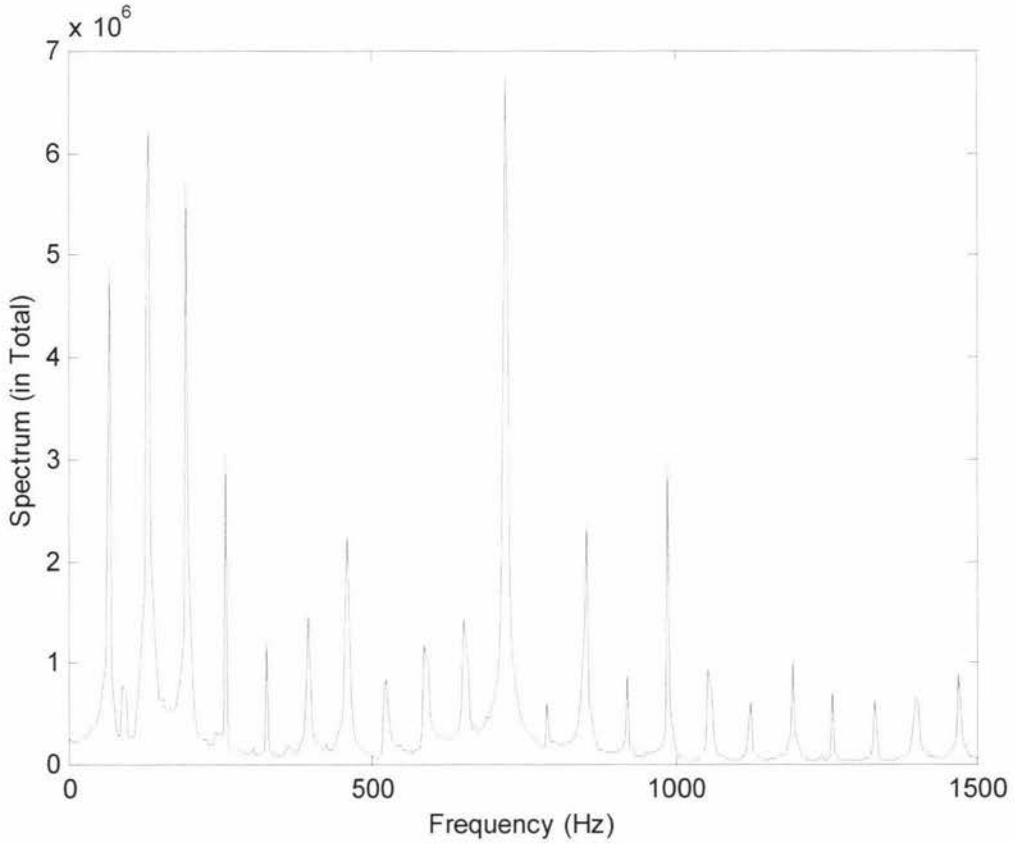

Figure 2.11 The results of Fourier Transform when using different starting points, on an acoustic grand piano when playing the note C2 (65.41 Hz) after totalling all the spectra.

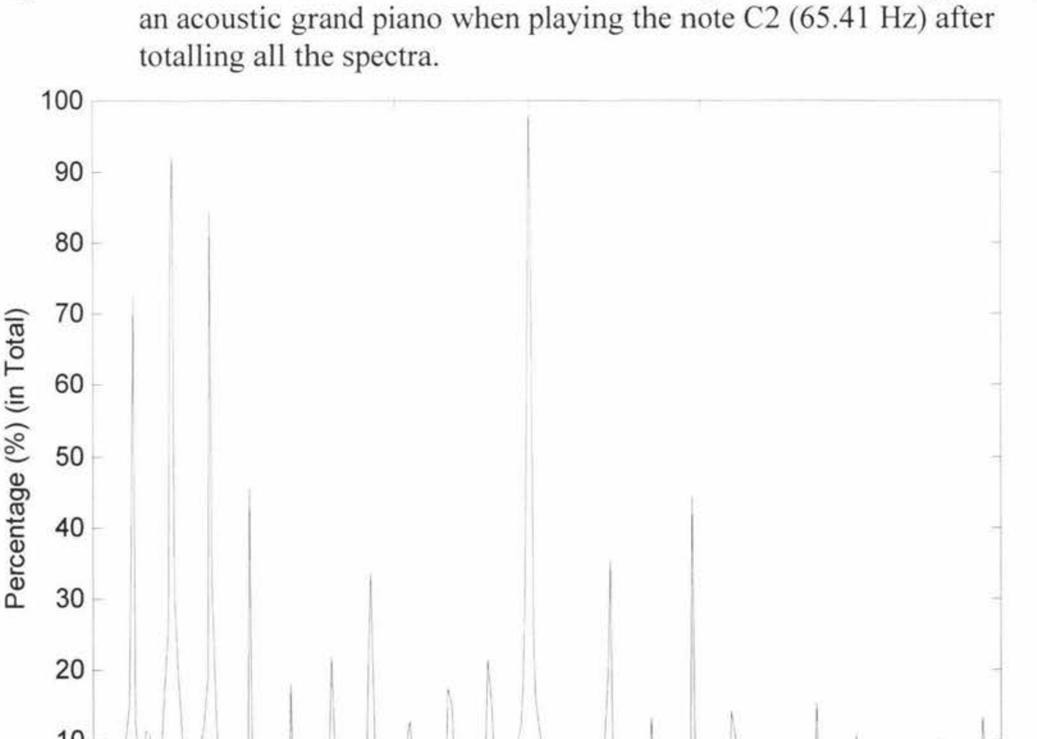

Figure 2.12 The results of Fourier Transform when using different starting points, on an acoustic grand piano when playing the note C2 (65.41 Hz) in percentages after totalling all the spectra.

2.4 Recording in MIDI

We started this section using a recorder (MIDI 75 in Table 1.3) to perform our experiments. A recorder is a simple instrument which will create smooth and stable waves that generally are only composed of sine waves. The sample rate is 22050Hz and all the notes are equal to a crochet.

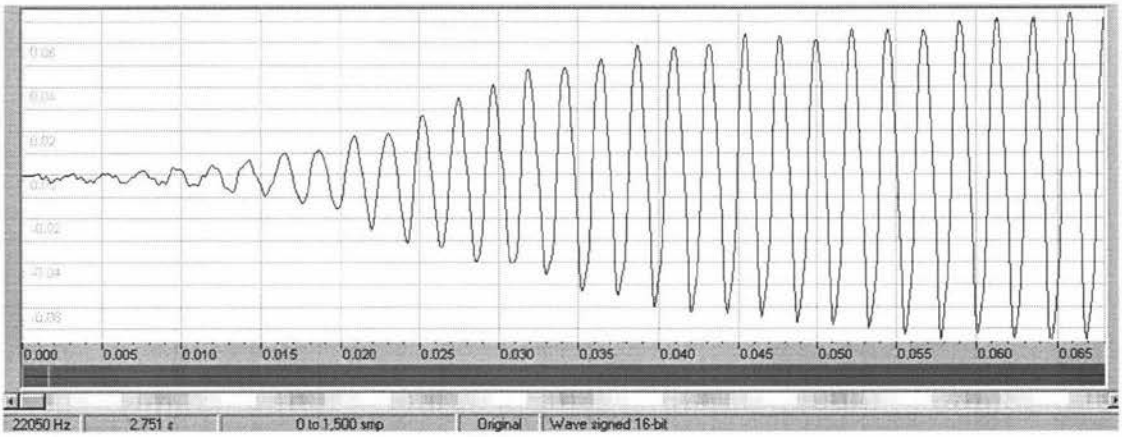

Figure 2.13 Example of recorder from MIDI (75) when playing the note A4 (440 Hz) (from Sample '0' to Sample 1500).

Figure 2.13 shows the initial part of the wave. The examples are played from sample 'zero'. We find that after approximately 10 periods of the note, the wave becomes stable. i.e. for A4 (440 Hz), with a wavelength of 50.1 samples and a sample rate of 22050Hz, the graph shows the wave becomes stable after 501 Samples (1/3 of the graph above). To avoid the attack stage in the following experiments, we chose the starting point of 15 periods after the beginning of the sample data (e.g. for A4 when using 22050 sample rate we start at sample 751). To be safe we have allowed 5 more periods.

Table 2.3 shows the results comparing the highest peak in the FT and comparing it with the expected frequency. Some results at lower frequencies are incorrect. There are two possible reasons. A single result in the FT of lower frequencies may represent several frequencies depending on the sample size. Later we showed that at some frequencies, some instruments can produce several frequencies when producing one note, so we must consider that this possibility may also occur with a recorder.

Table 2.3 The results of the highest peak of using Fourier Transform on recorder (MIDI 75).

Note Name	Frequency (Hz)	Max. Peak (Hz)	Note Name	Correct ?
A0	27.5	26.46	G#0	No
A#0	29.14	30.87	B0	No
B0	30.87	30.87	B0	Yes
C1	32.7	30.87	B0	No
C#1	34.65	35.28	C#1	Yes
D1	36.71	35.28	C#1	No
D#1	38.89	39.69	D#1	Yes
E1	41.2	39.69	D#1	No
F1	43.65	44.1	F1	Yes
F#1	46.25	48.51	G1	No
G1	49	48.51	G1	Yes
G#1	51.91	52.92	G#1	Yes
A1	55	57.33	A#1	No
A#1	58.27	57.33	A#1	Yes
B1	61.74	61.74	B1	Yes
C2	65.41	66.15	C2	Yes
C#2	69.3	70.56	C#2	Yes
D2	73.42	74.97	D2	Yes
D#2	77.78	79.38	D#2	Yes
E2	82.41	83.79	E2	Yes
F2	87.31	88.2	F2	Yes
F#2	92.5	92.61	F#2	Yes
G2	98	97.02	G2	Yes
G#2	103.83	105.84	G#2	Yes
A2	110	110.25	A2	Yes
A#2	116.54	119.07	A#2	Yes
B2	123.47	123.48	B2	Yes
C3	130.81	132.3	C3	Yes
C#3	138.59	141.12	C#3	Yes
D3	146.83	145.53	D3	Yes
D#3	155.56	154.35	D#3	Yes
E3	164.81	167.58	E3	Yes
F3	174.61	176.4	F3	Yes
F#3	185	185.22	F#3	Yes
G3	196	198.45	G3	Yes
G#3	207.65	207.27	G#3	Yes
A3	220	220.5	A3	Yes
A#3	233.08	233.73	A#3	Yes
B3	246.94	246.96	B3	Yes

(Continued on next page)

Table 2.3 (cont.)

Note Name	Frequency (Hz)	Max. Peak (Hz)	Note Name	Correct ?
C4	261.63	264.6	C4	Yes
C#4	277.18	277.83	C#4	Yes
D4	293.66	295.47	D4	Yes
D#4	311.13	313.11	D#4	Yes
E4	329.63	330.75	E4	Yes
F4	349.23	348.39	F4	Yes
F#4	369.99	370.44	F#4	Yes
G4	392	392.49	G4	Yes
G#4	415.3	414.54	G#4	Yes
A4	440	441	A4	Yes
A#4	466.16	467.46	A#4	Yes
B4	493.88	493.92	B4	Yes
C5	523.25	524.79	C5	Yes
C#5	554.37	555.66	C#5	Yes
D5	587.33	586.53	D5	Yes
D#5	622.25	621.81	D#5	Yes
E5	659.26	661.5	E5	Yes
F5	698.46	696.78	F5	Yes
F#5	739.99	740.88	F#5	Yes
G5	783.99	789.39	G5	Yes
G#5	830.61	829.08	G#5	Yes
A5	880	877.59	A5	Yes
A#5	932.33	930.51	A#5	Yes
B5	987.77	987.84	B5	Yes
C6	1046.5	1045.17	C6	Yes
C#6	1108.73	1115.73	C#6	Yes
D6	1174.66	1173.06	D6	Yes
D#6	1244.51	1239.21	D#6	Yes
E6	1318.51	1314.18	E6	Yes
F6	1396.91	1389.15	F6	Yes
F#6	1479.98	1472.94	F#6	Yes
G6	1567.98	1561.14	G6	Yes
G#6	1661.22	1662.57	G#6	Yes
A6	1760	1755.18	A6	Yes
A#6	1864.66	1856.61	A#6	Yes
B6	1975.53	1966.86	B6	Yes

(Continued on next page)

Table 2.3 (cont.)

Note Name	Frequency (Hz)	Max. Peak (Hz)	Note Name	Correct ?
C7	2093	2094.75	C7	Yes
C#7	2217.46	2209.41	C#7	Yes
D7	2349.32	2350.53	D7	Yes
D#7	2489.02	2478.42	D#7	Yes
E7	2637.02	2637.18	E7	Yes
F7	2793.83	2795.94	F7	Yes
F#7	2959.96	2963.52	F#7	Yes
G7	3135.96	3135.51	G7	Yes
G#7	3322.44	3325.14	G#7	Yes
A7	3520	3519.18	A7	Yes
A#7	3729.31	3730.86	A#7	Yes
B7	3951.07	3951.36	B7	Yes
C8	4186.01	4189.5	C8	Yes

The column "*Correct ?*" displays the results of comparing the highest peak of the FT with the highest peak of the expected note and shows the accuracy rate of matching them is 81 out of 88 ~ 92% - which is quite accurate. As we have mentioned earlier, our experiments will concentrate on the range between C2 to B6. The accuracy rate of this range is 100%.

Table B.1 in the Appendix B shows the results of the same methods applied to all the other instruments. We see that about 50%~60% of the notes are located by choosing the highest peak of the FT. However the accuracy rate for music will be higher because MIDI includes a large number of "instruments" that are not musical. e.g. Gunshot (128), Helicopter (127), etc. This is shown in Table 2.4.

Instrument Groups	Midi instrument range		All Octaves
Pipe	73	80	84%
Synth Pad	89	96	79%
Synth Effect	97	104	73%
Synth Lead	81	88	72%
Chromatic Percussion	9	16	68%
Strings	41	48	66%
Bass	33	40	63%
Piano	1	8	63%
Ensemble	49	56	57%
Brass	57	64	43%
Guitar	25	32	40%
Reed	65	72	39%
Organ	17	24	31%
Ethnic	105	112	22%
Sound Effect	121	128	12%
Percussive	113	120	10%

Table 2.4 The accuracy rate when comparing the expected frequencies and the highest peak of the Fourier Transform grouped into types of instrument.

We can make two observations.

i) There are large differences between instruments, so the accuracy of the identification of the note is directly related to the choice of instrument.

ii) Most instruments with higher frequency notes have a higher accuracy rate than instruments with the lower frequency notes.

2.5 Playing Two Notes Simultaneously on a Single Instrument

We used a recorder to produce two notes at the same time. We then chose several different instruments to produce the same output, and compared the results. The starting points were 15 periods of the wavelength of the lowest frequency note which produces a better "stable" stage.

First Note		Second Note		Highest Peak		Second Highest		First Note?	Second Note?	Number of Correct Answer	Both Correct?
C4	262	C#4	277	278	C#4	265	C4	Yes	Yes	2	Yes
C4	262	D4	294	296	D4	265	C4	Yes	Yes	2	Yes
C4	262	D#4	311	313	D#4	265	C4	Yes	Yes	2	Yes
C4	262	E4	330	331	E4	265	C4	Yes	Yes	2	Yes
C4	262	F4	349	348	F4	265	C4	Yes	Yes	2	Yes
C4	262	F#4	370	370	F#4	265	C4	Yes	Yes	2	Yes
C4	262	G4	392	393	G4	265	C4	Yes	Yes	2	Yes
C4	262	G#4	415	415	G#4	265	C4	Yes	Yes	2	Yes
C4	262	A4	440	441	A4	265	C4	Yes	Yes	2	Yes
C4	262	A#4	466	468	A#4	265	C4	Yes	Yes	2	Yes
C4	262	B4	494	494	B4	265	C4	Yes	Yes	2	Yes

Table 2.5 The results of playing two notes on MIDI by using recorder.

The first four columns are the properties of the expected notes. The 5th to 8th columns ("Highest Peak" & "Second Highest") are the results of the Fourier Transform. The 9th and 10th columns ("First Note?" & "Second Note?") check whether the highest or the second highest peaks are the expected notes. The 11th column shows the number of correct answers and the 12th column displays whether both of the notes appear as the highest or the second highest in the Fourier Transform.

The table shows that this initial experiment produced accurate results. We repeated this experiment with "note one" ranging from A0 to C8 and "note two" being all the combinations within the same octave range as "note one"

We excluded the cases where "note one" and "note two" are in reverse order and where they are same. Without the exclusions, there would be 88*88 = 7744 (FT) analyses. The exclusions reduce the number to 3828 (FT).

The results are in Table 2.6. The results of the reciprocal cases are counted twice because one result belongs to note one and another result belongs to note two. The first two columns list the note and its number. The third column shows the number of cases which have at least one result correct, and the fourth column shows the number of cases which have both results correct.

At the bottom we show the average of columns 3 and 4 expressed as a percentage. Table 2.3 shows that the accuracy rate for a single note is approximately 92%. Using this we can calculate the expected average of column 3 and 4.

For column 3 the chance that the first result or the second result is correct is 92% + 8% * 92% of cases i.e. 99.36%. This percentage is very close to the experimental result of 98.3%.

For Column 4 the FT was correct for 92% of the first notes and also correct for 92% of the second notes. i.e. 92% * 92% = 84.64%. Again this percentage is very close to the experimental result of 84.33%.

Note No. & Note Name		One Correct	Two Correct	Note No. & Note Name		One Correct	Two Correct
21	A0	80	0	67	G4	87	80
22	A#0	79	0	68	G#4	87	80
23	B0	87	80	69	A4	87	80
24	C1	69	0	70	A#4	87	80
25	C#1	87	80	71	B4	87	80
26	D1	70	0	72	C5	87	80
27	D#1	87	79	73	C#5	87	80
28	E1	79	0	74	D5	87	80
29	F1	87	80	75	D#5	87	80
30	F#1	73	0	76	E5	87	80
31	G1	87	80	77	F5	87	80
32	G#1	87	80	78	F#5	87	80
33	A1	73	0	79	G5	83	80
34	A#1	87	79	80	G#5	87	80
35	B1	87	80	81	A5	87	80
36	C2	87	80	82	A#5	87	80
37	C#2	87	80	83	B5	87	80
38	D2	87	78	84	C6	87	80
39	D#2	87	78	85	C#6	86	80
40	E2	87	79	86	D6	87	80
41	F2	87	80	87	D#6	87	80
42	F#2	87	80	88	E6	80	79
43	G2	87	78	89	F6	87	80
44	G#2	87	79	90	F#6	83	80
45	A2	87	80	91	G6	83	80
46	A#2	87	78	92	G#6	87	80
47	B2	87	80	93	A6	83	75
48	C3	87	80	94	A#6	83	80
49	C#3	87	80	95	B6	85	80
50	D3	87	80	96	C7	87	80
51	D#3	87	80	97	C#7	83	80
52	E3	87	80	98	D7	87	80
53	F3	87	80	99	D#7	81	74
54	F#3	87	80	100	E7	87	80
55	G3	87	80	101	F7	87	80
56	G#3	87	80	102	F#7	87	80
57	A3	87	80	103	G7	87	80
58	A#3	87	80	104	G#7	87	80
59	B3	87	80	105	A7	87	80
60	C4	87	80	106	A#7	87	80
61	C#4	87	80	107	B7	85	80
62	D4	87	80	108	C8	87	80
63	D#4	87	80	**Averages**		85.52	73.36
64	E4	87	80	**out of**		87	87
65	F4	87	80	**Percentages**		98.3	84.33
66	F#4	85	80				

Table 2.6 The results of playing two notes using recorder on MIDI.

This predictability is because these sounds are synthesised by computer (MIDI), therefore the results for more notes can also calculated. However, with different instruments (even though they are MIDI generated), the difference between the calculated and the actual results are larger. The important question is that of the actual accuracy rate of finding the correct solutions.

We rearrange the results in Table 2.6 into octave groups and display them in Table 2.7. The accuracy is high for notes C2 and above.

Range Started		**Range Ended**		**Number of Cases**	**One Correct**		**Two Correct**	
21	A0	108	C8	7656	7528	98.33%	6456	84.33%
36	C2	95	B6	3540	3540	100%	3532	99.77%
48	C3	95	B6	2256	2256	100%	2256	100%

Table 2.7 The accuracy when using different ranges of expected note on the recorder.

Table 2.8 shows a similar analysis of results using different instruments. The tables are organised in octave order and once again the accuracy outside the investigated range is lower.

In these experiments we chose to use instruments with a high accuracy when playing a single note, to analyse the practicality of this method. It is clear from these results that the accuracy using different instruments varied greatly. In some instruments higher accuracy occurs with higher frequencies. We will explain later why such instruments have lower accuracy and why different instruments had different results.

Range Start		Range Ended		Octave	Number of Cases	One Correct		Two Correct	
Acoustic Grand Piano (MIDI 1)									
21	A0	108	C8	All	7656	5440	71.06%	2206	28.81%
36	C2	95	B6	2-6	3540	3104	87.68%	1552	43.84%
48	C3	95	B6	3-6	2256	2216	98.23%	1426	63.21%
60	C4	95	B6	4-6	1260	1260	100%	968	76.83%
72	C5	95	B6	5-6	552	552	100%	502	90.94%
84	C6	95	B6	6	132	132	100%	132	100%
Acoustic Guitar (steel) (MIDI 26)									
21	A0	108	C8	All	7656	4940	64.52%	1086	14.18%
36	C2	95	B6	2-6	3540	2830	79.94%	794	22.43%
48	C3	95	B6	3-6	2256	1840	81.56%	544	24.11%
60	C4	95	B6	4-6	1260	1048	83.17%	304	24.13%
72	C5	95	B6	5-6	552	478	86.59%	150	27.17%
84	C6	95	B6	6	132	118	89.39%	30	22.73%
Acoustic Bass (MIDI 33)									
21	A0	108	C8	All	7656	7210	94.17%	5446	71.13%
36	C2	95	B6	2-6	3540	3540	100%	3076	86.89%
48	C3	95	B6	3-6	2256	2256	100%	1992	88.3%
60	C4	95	B6	4-6	1260	1260	100%	1070	84.92%
72	C5	95	B6	5-6	552	552	100%	450	81.52%
84	C6	95	B6	6	132	132	100%	106	80.3%
Violin (MIDI 41)									
21	A0	108	C8	All	7656	6436	84.06%	1398	18.26%
36	C2	95	B6	2-6	3540	3306	93.39%	1006	28.42%
48	C3	95	B6	3-6	2256	2076	92.02%	612	27.13%
60	C4	95	B6	4-6	1260	1088	86.35%	232	18.41%
72	C5	95	B6	5-6	552	386	69.93%	56	10.14%
84	C6	95	B6	6	132	112	84.85%	30	22.73%

Table 2.8 The accuracy rate when using the two highest peaks to predict the results on different instruments.

2.6 Playing Three Notes at Once

The design of this experiment is the same as the previous experiments with the same exclusions. We concentrated on the range from C2 (36) to B6 (95). Because there are many combinations for three notes, we only tried the combinations which involve even MIDI numbers (every second note). e.g. "36, 38 & 40" and "36, 70 & 94". The number of cases is 4060. Table B.2 (in Appendix B) shows the results.

This experiment shows the same pattern as for two notes. Different instruments show different levels of accuracy. The accuracy at different frequencies on different instruments varies.

2.7 Why is the Highest Peak of the Single Note Not the Answer in Some Cases?

There are two possibilities. Perhaps other frequencies created by the instrument at the same time "confuses" the Fourier Transform. The other possibility is the highest peak is an overtone of the note[1] (This will be demonstrated in section 2.9 on page 50).

Figure 2.14 demonstrates the first possibility comparing the wave from a recorder to that from an Acoustic Grand Piano (AGP) when playing C2. The complicated curve created by AGP has a similar shape to the simple curve of the recorder. Figures 2.15 and 2.16 are the results of FT. Although the peak of the C2 frequency (65.41 Hz) on AGP (Figure 2.15) is higher than most of the frequencies there is one frequency higher.

This is different to the results from Figure 2.4 on page 30 or 2.7 on page 33 because the starting point is later. We may be able to see these from the "back" end of Figure 2.8 (in 3 dimensions) (on page 33).

In Figures 2.15 and 2.16 show the spectra of the notes from the simple instrument (recorder) are much higher than those of the instrument that produces more frequencies. This could be because the energy is divided over more frequencies. The highest peak for the recorder is at least 5 times higher than any peak for the acoustic grand piano (on the y-axis, the recorder is $5.2*10^6$ and the acoustic grand piano is $8.2*10^5$) (more examples will be shown in section 2.9 on page 50).

---

[1] Overtone: *Acoustics & Mus.* An upper partial tone; a harmonic. *Physics*. An analogous component of any kind of oscillation, having a frequency that is an integral multiple of the fundamental frequency [Excerpted from *Oxford Talking Dictionary* [30]]. We will used the physical meaning in this thesis.

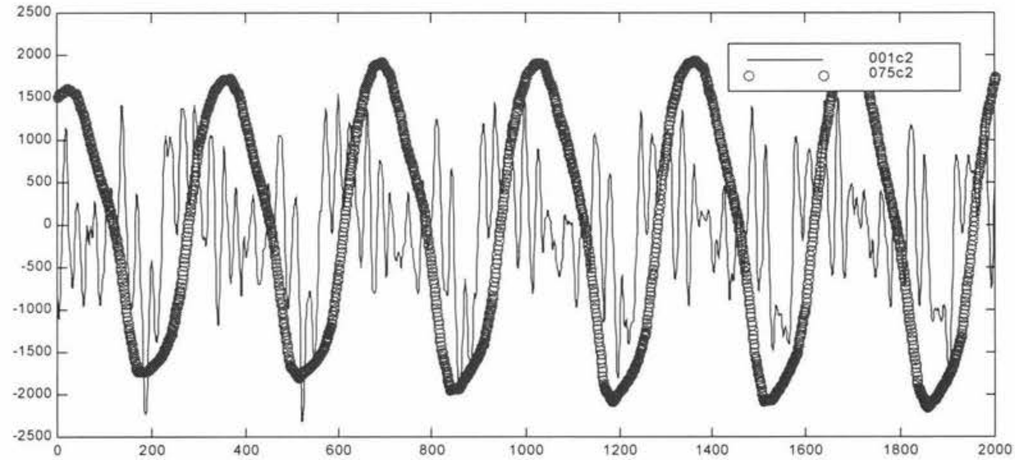

Figure 2.14 The sound wave of two different instruments when playing C2 (65.41 Hz).

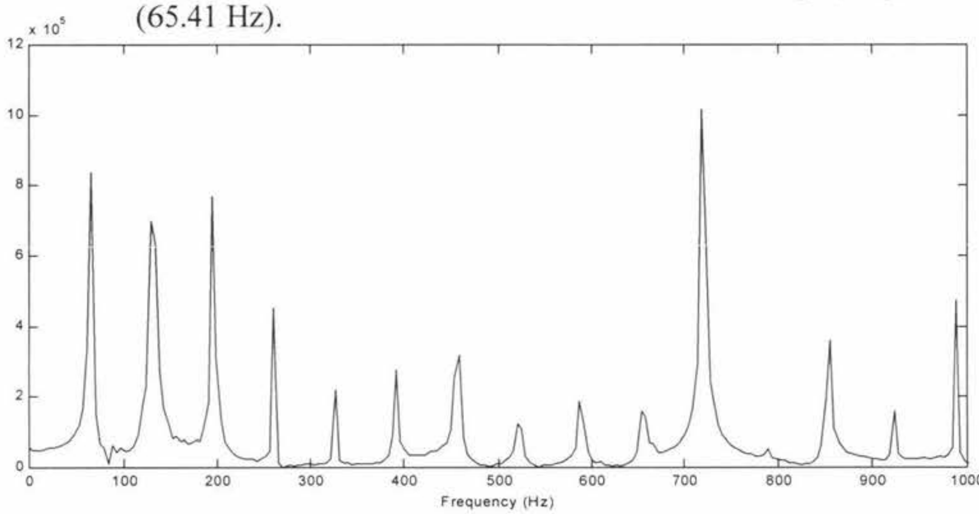

Figure 2.15 The results of Fourier Transform on an acoustic grand piano when playing C2 (65.41 Hz).

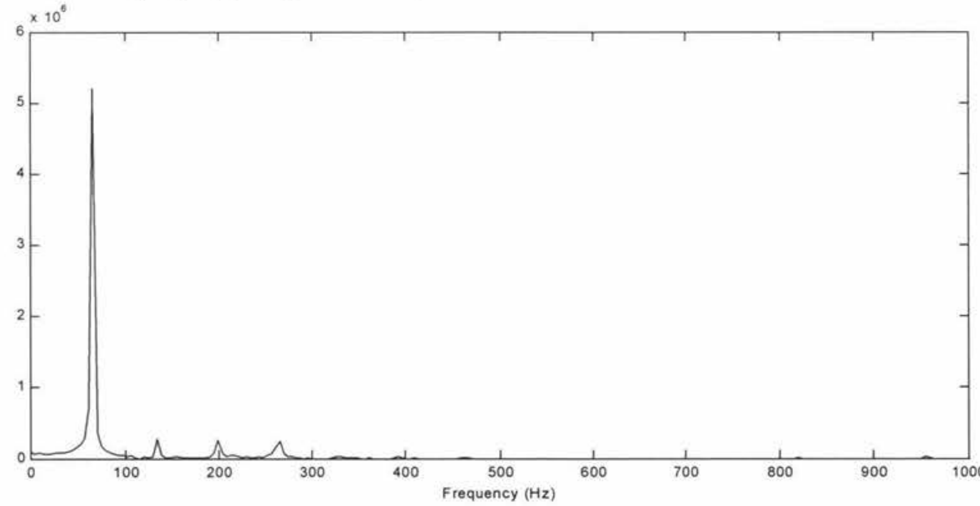

Figure 2.16 The results of Fourier Transform on a recorder when playing C2 (65.41 Hz).

2.8 What is the Relationship between a Single Note and Two Notes or Three Notes?

There is a simple relationship between a single note and two notes or three notes. Figure 2.17 and 2.18 show a part of the sound wave of AGP when playing C4 and A4 separately. Figure 2.19 is the result of adding the values from Figure 2.17 and 2.18. Figure 2.20 is the actual sound wave created by an acoustic grand piano when playing C4 and A4 together in MIDI.

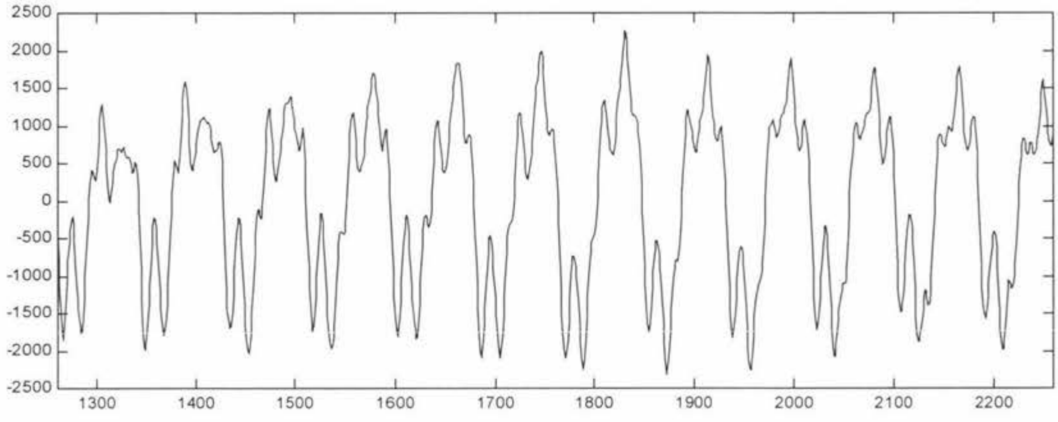

Figure 2.17 The sound wave of an acoustic grand piano when playing C4.

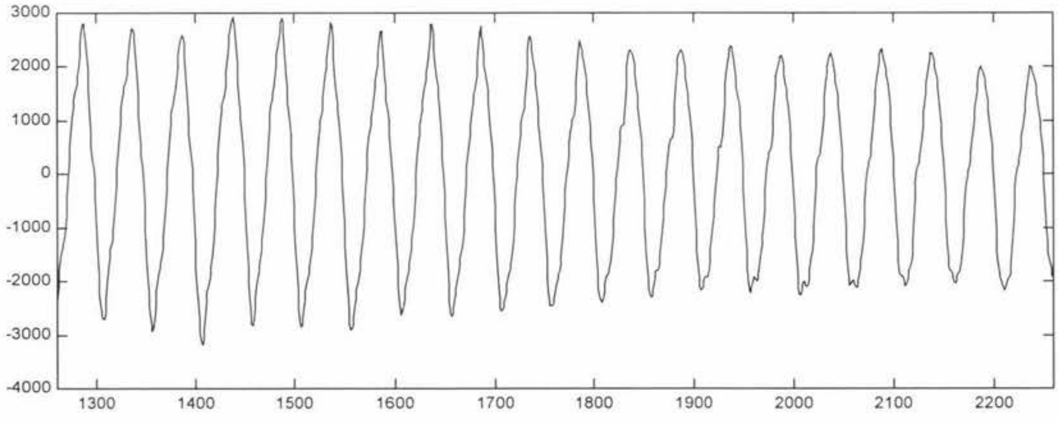

Figure 2.18 The sound wave of an acoustic grand piano when playing A4.

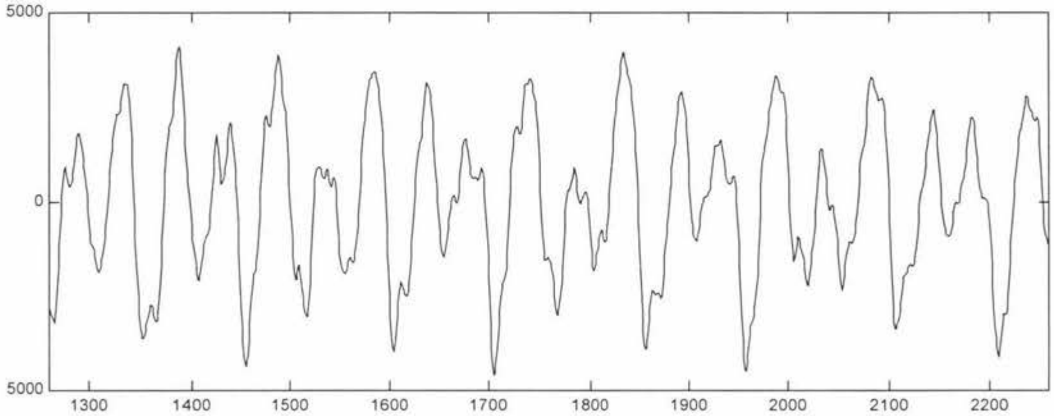

Figure 2.19 The sound wave adding up the value of C4 and A4 on an acoustic grand piano.

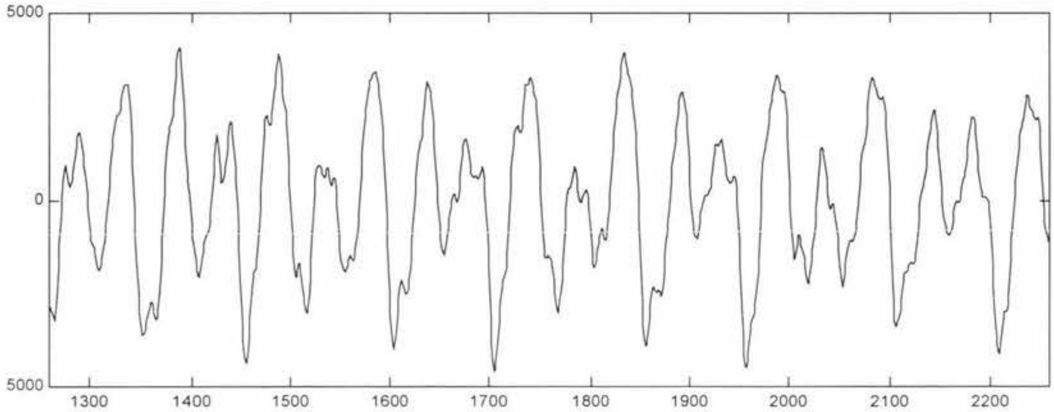

Figure 2.20 The sound wave for acoustic grand piano when playing C4 and A4 together.

We can see that the waves in Figure 2.19 and 2.20 are almost identical. This happens with different notes on the same instrument, or different notes each from different instruments. The differences between actual and calculated are minor. The sound wave for two notes together is the same as the sound wave that is created by adding the single sound waves from each note. This also applies to three or more notes. The same addition can be done with FT.

This leads to the concept of the reverse. If we subtract from the FT of two notes the FT of one of them, the remainder will be the FT of the other.

2.9 The Reason for Different Instruments having Different Results from the Fourier Transform

In section 2.5 and 2.6 we have seen that the accuracy of the expected notes being identified, by using the highest peaks of the FT, varies greatly between different instruments and different frequencies. This can be explained by observing the following examples for the note A4 on different instruments.

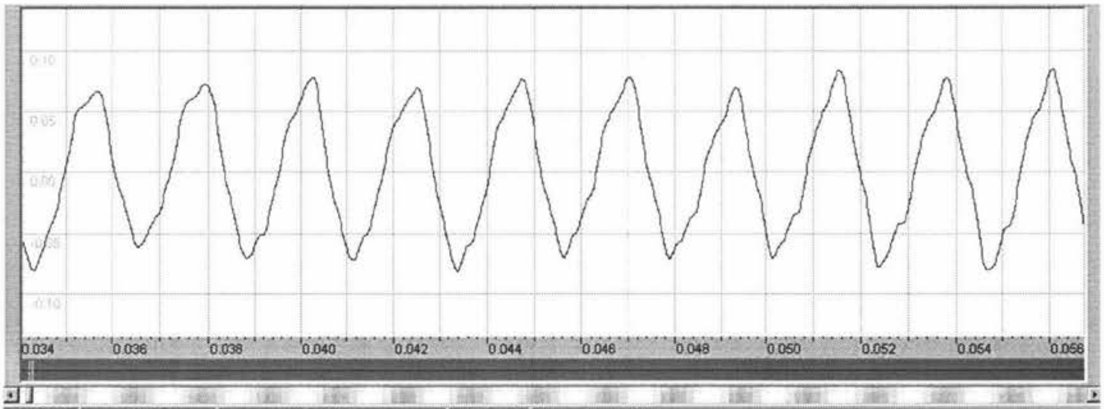

Figure 2.21 The sound wave of the note A4 on an acoustic grand piano (MIDI 1).

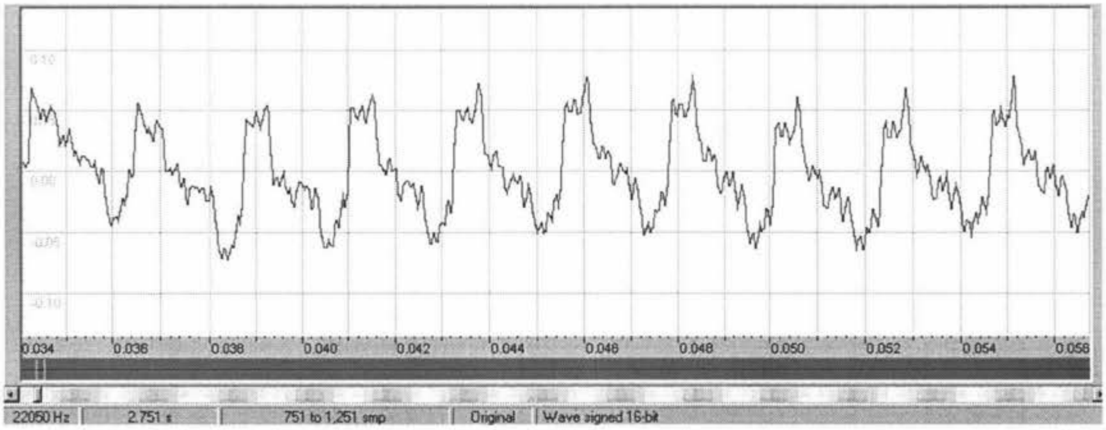

Figure 2.22 The sound wave of the note A4 on an acoustic guitar (steel) (MIDI 26).

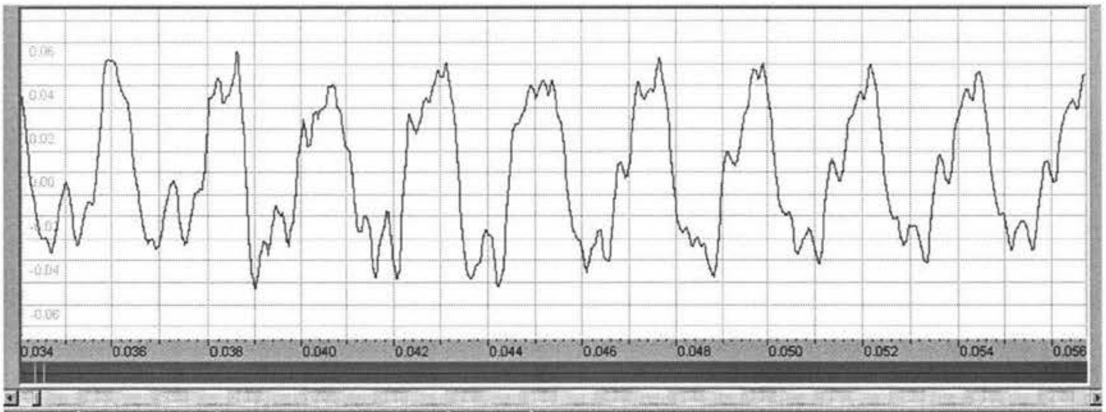

Figure 2.23 The sound wave of the note A4 on an acoustic bass (MIDI 33).

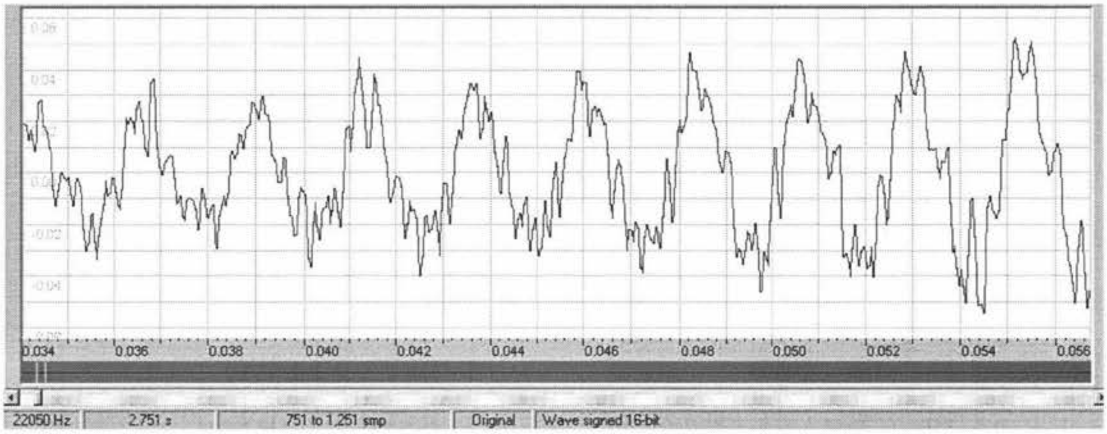

Figure 2.24 The sound wave of the note A4 on a violin (MIDI 41).

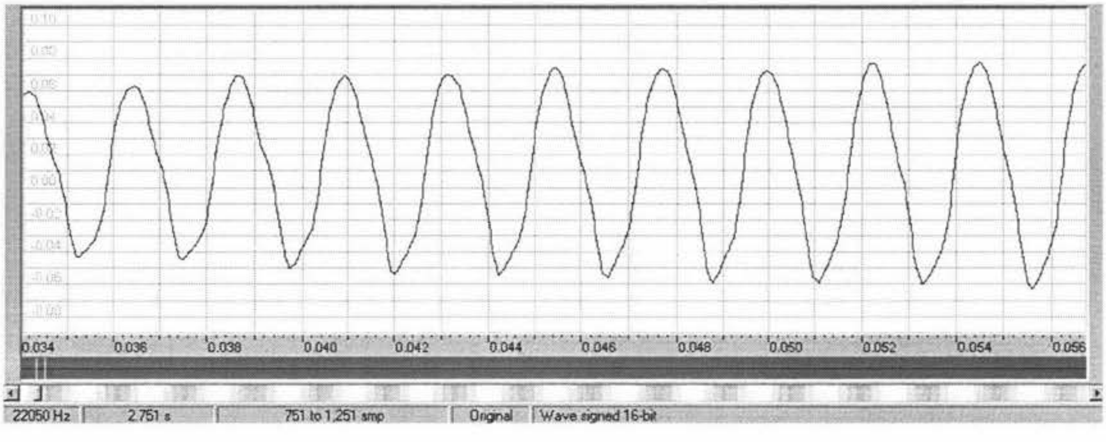

Figure 2.25 The sound wave of the note A4 on a recorder (MIDI 75).

Though they play the same note, the sound waves are different to a greater or lesser extent. Ideally they are the same frequency (the expected frequency). Figure 2.26 shows the sound waves of all the instruments, after rearrangement to the same phase, displayed on the same graph.

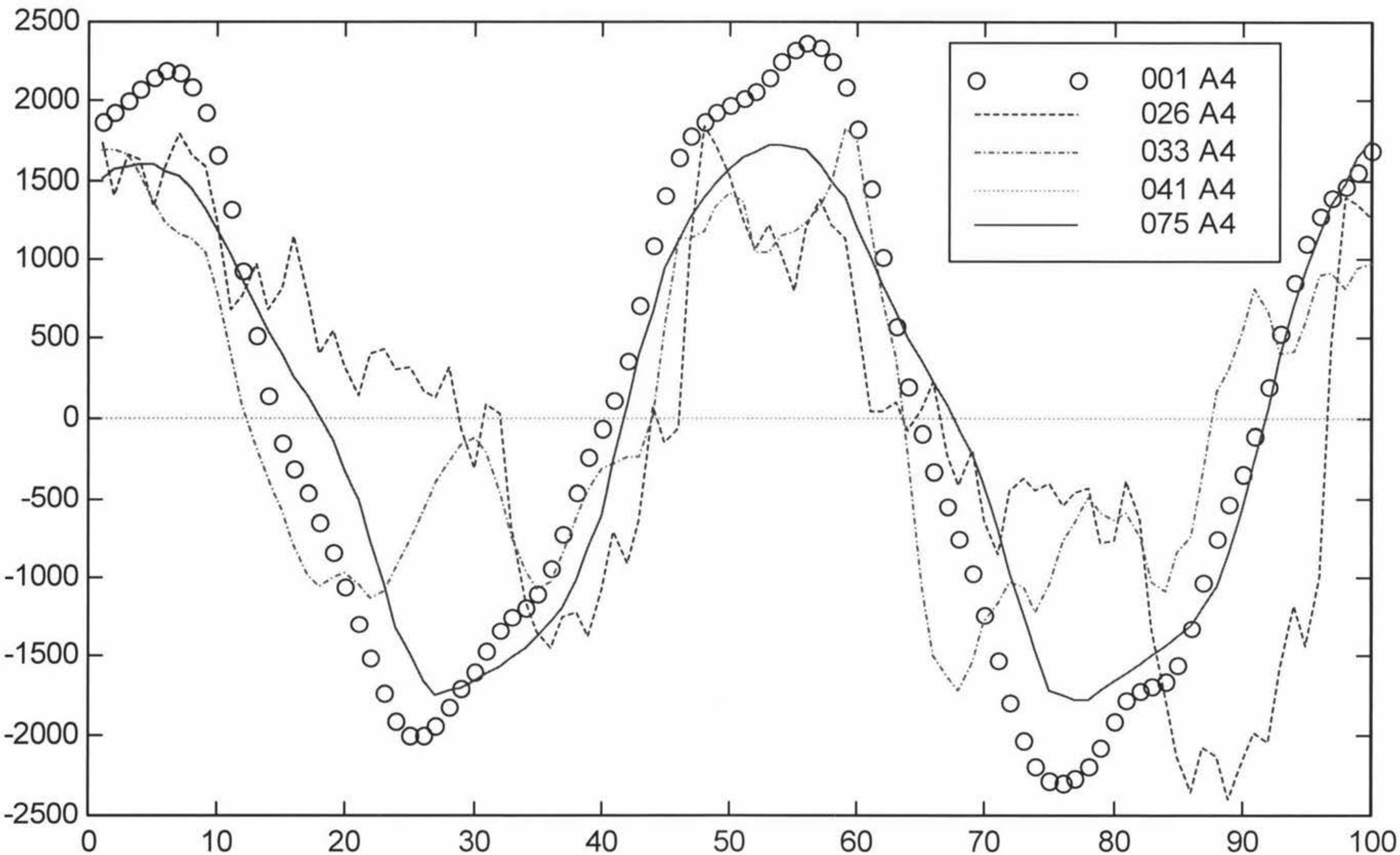

Figure 2.26 The sound wave of the note A4 when using different instruments.

We then looked at the Fourier Transforms of these notes (Figure 2.27 to 2.31).

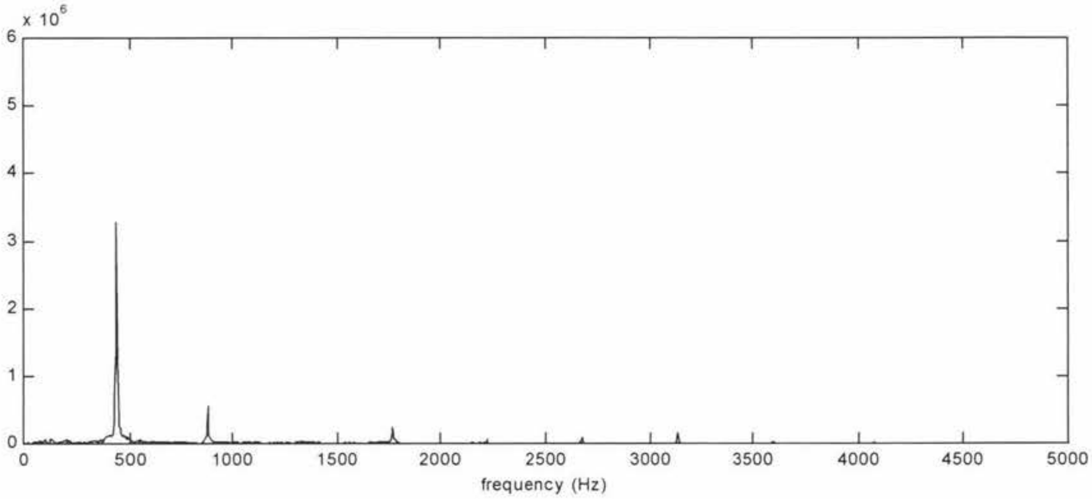

Figure 2.27 The results of the Fourier Transform of the note A4 on an acoustic grand piano (MIDI 1).

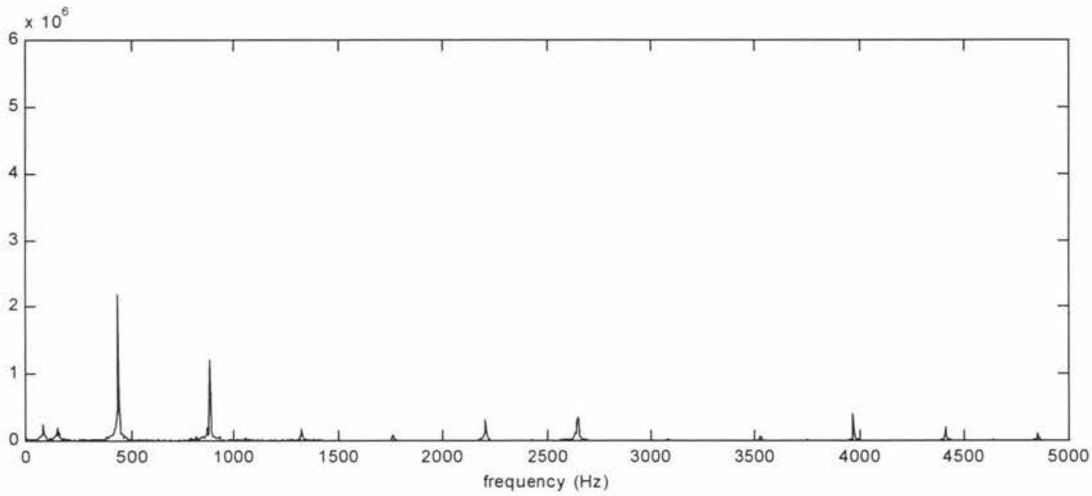

Figure 2.28 The results of the Fourier Transform of the note A4 on an acoustic guitar (steel) (MIDI 26).

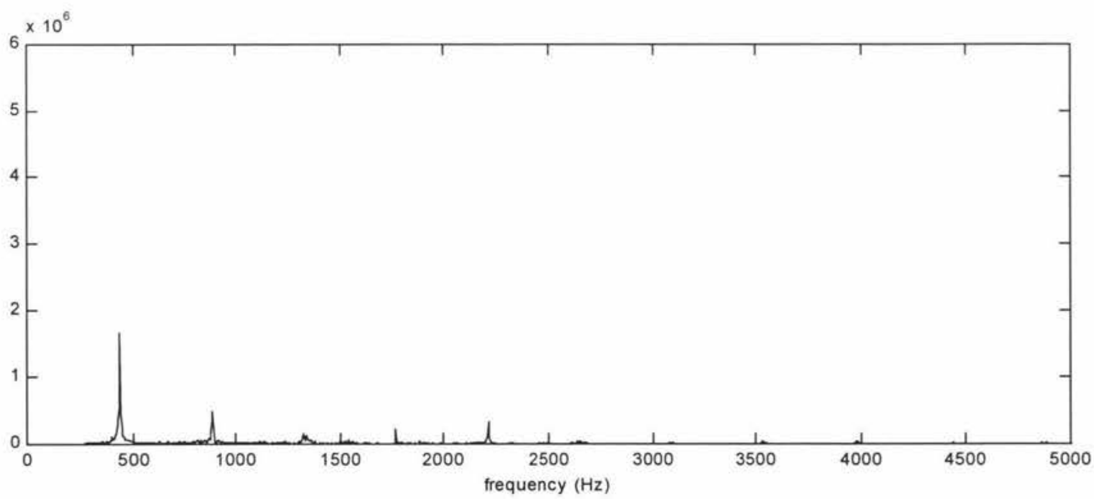

Figure 2.29 The results of the Fourier Transform of the note A4 on an acoustic bass (MIDI 33).

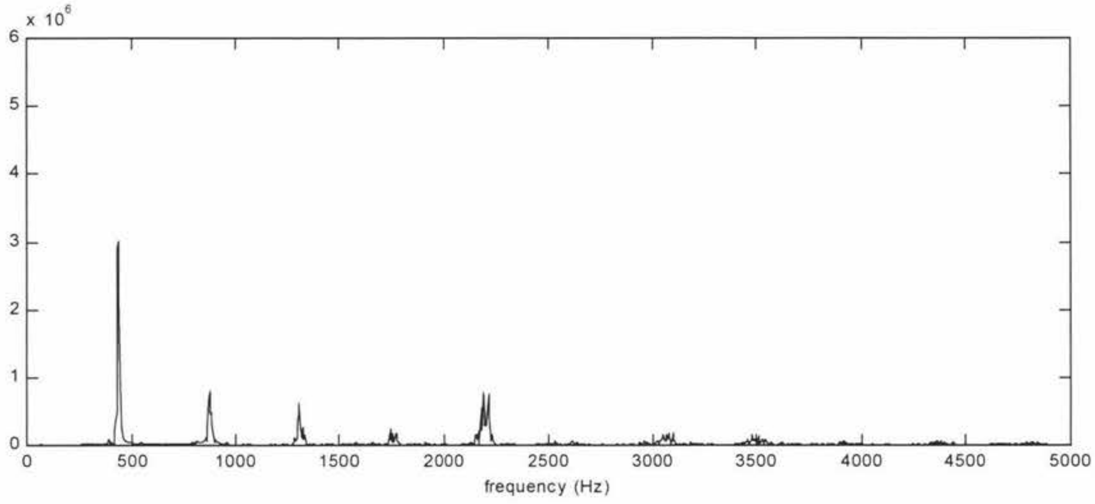

Figure 2.30 The results of the Fourier Transform of the note A4 on a violin (MIDI 41).

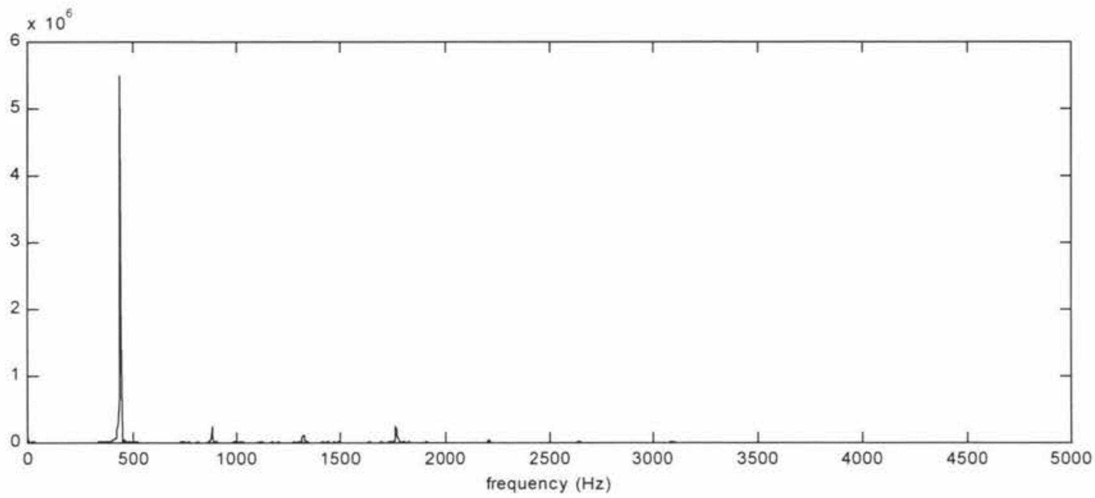

Figure 2.31 The results of the Fourier Transform of the note A4 on a recorder (MIDI 75).

The results from the Fourier Transform followed the same pattern as for the waves. Their highest peaks are very close to the expected frequency. Their highest peaks are at the same frequency but their spectra and the values for the other peaks differed. These variations are different for different notes and different instruments. The highest peak may not be the same as the expected frequency for some of the notes (as Figure 2.4 on page 30, the expected frequency is C2, and the highest peak is C3/F5).

Table 2.9 shows the highest three peaks when playing A4 on different instruments and compares the amplitude of them within each instrument, and over other instruments. The values in Column 5 are displayed in scientific format, so 3.28E+06 is equal $3.28*10^6$, etc. Table 2.10 shows the same operation on a different note, C4.

Instru. No.	Note No.	Frequency of the 3 highest peaks (Hz)	Note Name	Spectrum	Peak as percentage related to the highest peak	Peak as percentage of the highest peak within instrument	Peak as percentage of the highest peak within 5 instruments
1	50	440	A4	3.28E+06	100%	81%	48%
1	62	880	A5	5.43E+05	17%	13%	8%
1	51	466.16	A#4	2.60E+05	8%	6%	4%
26	50	440	A4	2.19E+06	100%	77%	32%
26	62	880	A5	1.20E+06	55%	42%	18%
26	88	3951.07	B7	3.99E+05	18%	14%	6%
33	50	440	A4	1.65E+06	100%	48%	24%
33	62	880	A5	4.89E+05	30%	14%	7%
33	78	2217.46	C#7	3.34E+05	20%	10%	5%
41	50	440	A4	3.02E+06	100%	68%	45%
41	62	880	A5	7.90E+05	26%	18%	12%
41	78	2217.46	C#7	7.76E+05	26%	17%	11%
75	50	440	A4	5.49E+06	100%	81%	81%
75	62	880	A5	2.54E+05	5%	4%	4%
75	74	1760	A6	2.52E+05	5%	4%	4%

Table 2.9 The results of the Fourier Transform of the note A4 on different instruments.

Instru. No.	Note No.	Frequency of the 3 highest peaks (Hz)	Note Name	Spectrum	Peak as percentage related to the highest peak	Peak as percentage of the highest peak within instrument	Peak as percentage of the highest peak within 5 instruments
1	41	261.63	C4	3.20E+06	100%	79%	47%
1	60	783.99	G5	7.78E+05	24%	19%	11%
1	69	1318.51	E6	7.73E+05	24%	19%	11%
26	41	261.63	C4	1.14E+06	100%	40%	17%
26	53	523.25	C5	9.41E+05	83%	33%	14%
26	75	1864.66	A#6	4.03E+05	35%	14%	6%
33	41	261.63	C4	1.22E+06	100%	35%	18%
33	53	523.25	C5	8.52E+05	70%	25%	13%
33	69	1318.51	E6	5.45E+05	45%	16%	8%
41	41	261.63	C4	2.06E+06	100%	46%	30%
41	53	523.25	C5	2.03E+06	99%	46%	30%
41	72	1567.98	G6	9.39E+05	46%	21%	14%
75	41	261.63	C4	3.67E+06	100%	54%	54%
75	40	246.94	B3	6.64E+05	18%	10%	10%
75	42	277.18	C#4	6.32E+05	17%	9%	9%

Table 2.10 The results of the Fourier Transform of the note C4 on different instruments.

We made three observations from Tables 2.9 and 2.10.

i) There is no standard pattern for the peaks produced by the same note on different instruments. e.g. For A4 when using an acoustic grand piano (MIDI 1) the three highest peaks are A4, A5, A#5 and when using the recorder (MIDI 75) they are A4, A5 and A6.

ii) On the same instrument there is no pattern for each note. e.g. On AGP, there are 12 semitones difference between the highest and the second highest peaks for A4 (A4 to A5), and 19 for C4 (C4 to G5).

iii) The relative height of the highest peak of the Fourier Transform is different for different instruments with different notes. e.g. The ratios between recorder and AGP are:

$5.49*10^6 : 3.28*10^6$ (col. 5) ~ 100 : 60 on A4 and

$3.67*10^6 : 3.20*10^6$ ~ 100 : 87 on C4.

The tables also show other common phenomena. Firstly the second highest peak is often the overtone of the highest peak.

Example 1. From Table 2.9 playing A4. A4 is always the highest peak, 100%, A5 is always the second highest. A6 (also the overtone of A4) appears in one of the results.

Example 2. From Table 2.10 playing C4. C4 is always the highest peak. In three out of five cases C5 is the second highest.

This phenomenon occurs frequently in other situations and should be noted when it does. In some cases such extra peaks are almost as high as the expected note and, rarely, may be higher than the expected note.

Most musical instruments produce harmonic tones.

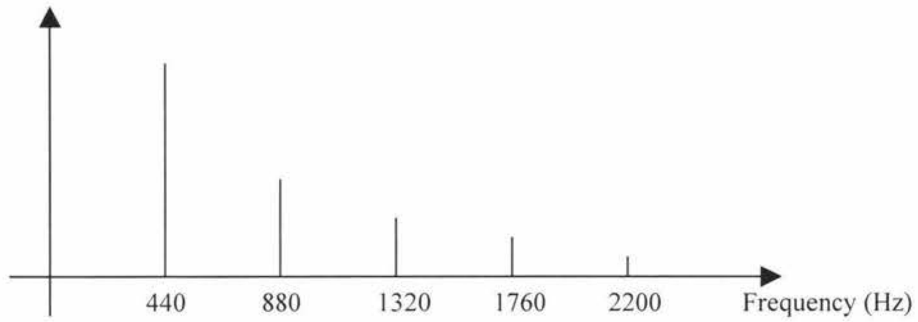

Fig 2.32 Example of tone spectra in FT for A4 (440 Hz).

From the figure, A5 (880 Hz) is one octave higher than A4, and it is double the frequency of A4. A6 (1320 Hz) is two octaves higher than A4, and it is three times the frequency of A4, and so on.

The amplitude on these harmonic tones usually shrinks as we move further from the original tone through the octaves, however there is no standard rate of decay. This may not happen in some instruments but should be considered.

The second phenomenon is that some simple instruments producing a simple wave (e.g. recorder) have a major difference between the highest and the second highest peaks. These instruments have a high accuracy of prediction using FT. We have seen this earlier comparing the recorder with the AGP.

When we revisit sections 2.5 and 2.6 (on page 41 and 45), we find using the highest peaks as the predictor is inaccurate with several instruments. There are several high peaks generated by a "single" note, which confuses the result. This is another reason different instruments produce different results.

Because of these difficulties, in the next section we used the actual results from the FT of the single note, rather than the highest peak, as the basis for comparison. Then we took the FT of an unknown note and use saved results to identify it. It may be possible to use the same operation applied on a different but similar instrument.

2.10 Using the Fourier Transform of Recorded Notes to Identify Unknown Notes

2.10.1 The file size of the recording format

We know that for each note, or for each set of FT, there are 5000 results. Saving all the notes in the file without compression or special file format will be time and memory consuming (5000 * 16 Bytes (for complex value) ~ 78 Kbytes). If we save all the notes from A0 to C8 for all 128 instruments it will use approximately 859 Mbytes of space. Therefore we will only save one result in each frequency range for each note in each set of FT. The boundaries, which encompass this range, are the midpoints between the frequencies of adjacent notes, on an arithmetic scale.

For example, for A4 (440 Hz), the boundaries are 427.65 Hz and 453.08 Hz. There are 6 results within this range and we only chose one result to represent A4.

We chose the boundary between two adjacent frequencies in the arithmetic scale instead of logarithmic scale because this gave more chance of a clearer signature for each tone. Usually the result was closer to the expected frequency, rather than to the boundary, so the choice of scale had only a minor effect compared to the size of the difference between two boundaries.

2.10.2 The choice of representative value

What is the best value inside the frequency range that we choose to best represent the original curve. The lowest, the highest or the average? Figure 2.33 to 2.37 are examples of C2 by AGP. Recall that the distance between two tones increases as the frequency increases. This also increases the number of results from the FT between the two tones. Figure 2.34 shows the original curve with the boundary drawn. These boundaries separate the results for each frequency range for notes.

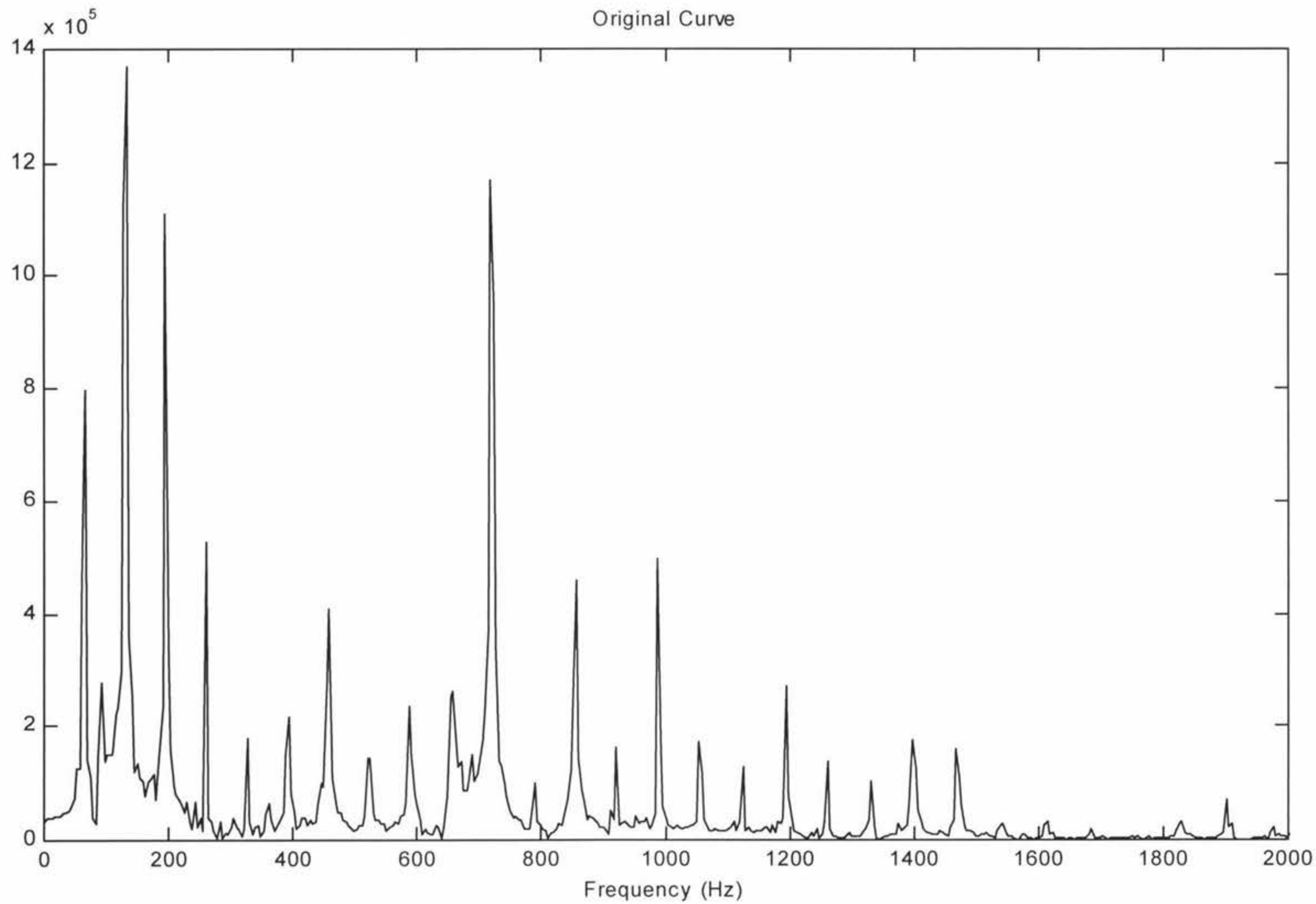

Fig 2.33 The results of Fourier Transform when playing C2 (65.41 Hz) on an acoustic grand piano.

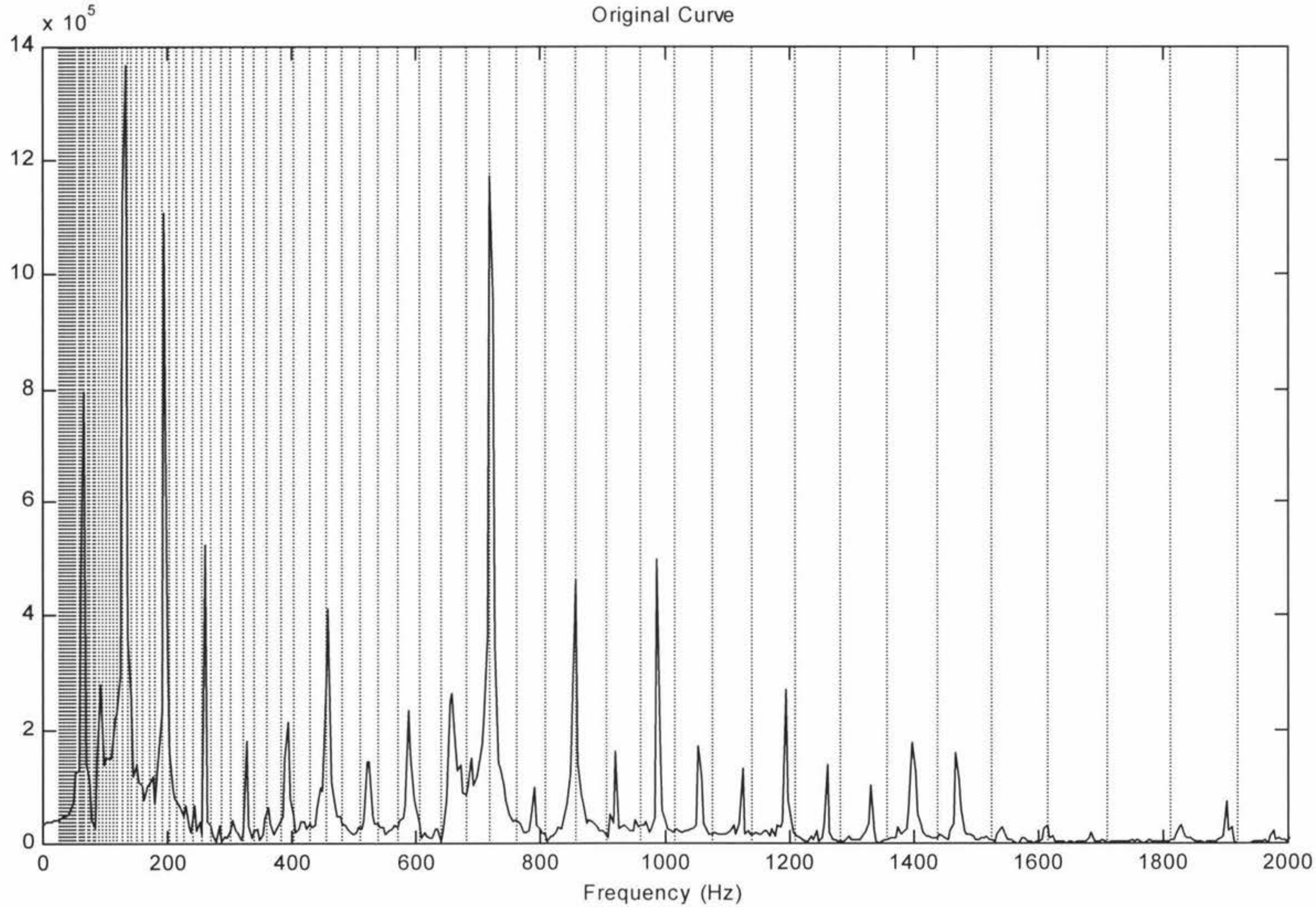

Figure 2.34 Figure 2.33 with note boundaries added.

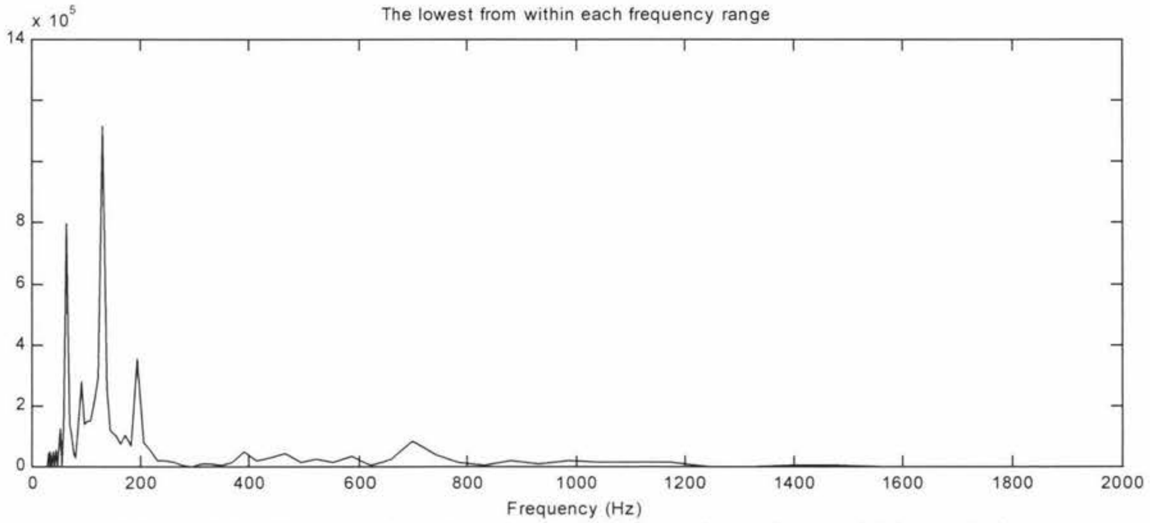

Figure 2.35 The "lowest" results of Fourier Transform from within each frequency range when playing C2 (acoustic grand piano).

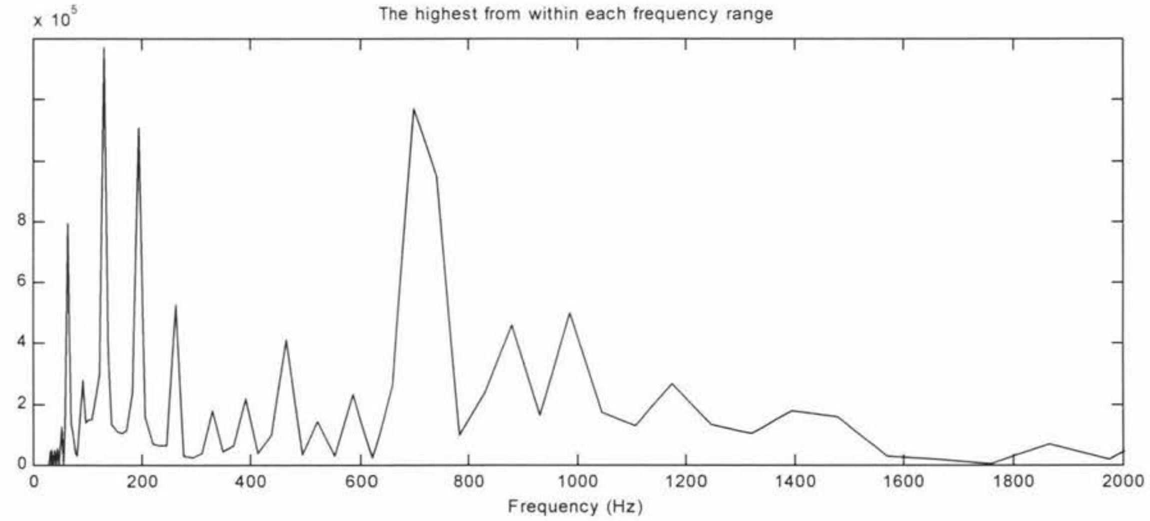

Figure 2.36 The "highest" results of Fourier Transform from within each frequency range when playing C2 (acoustic grand piano).

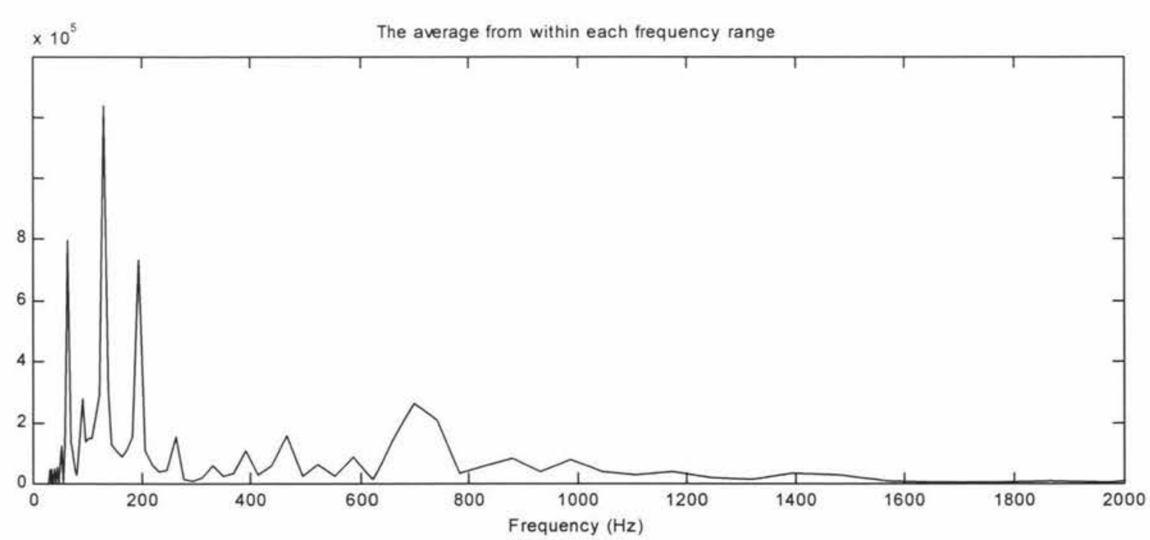

Figure 2.37 The "average" results of Fourier Transform from within each frequency range when playing C2 (acoustic grand piano).

The graphs for the lowest and the average values have a different shape to the original but the graph for the highest values retains the original shape.

Since our expected frequencies can be up to C8 we saved all the maximum results for each tone from G#0 to G#9. The size of the file for each note reduces by this method to 109 samples (from G#0 to G#9) * 16 Bytes ~ 1.7 Kbytes. The file size for all notes from A0 to C8 (88 notes) and all 128 instruments is approximately 18.7 Mbytes (1.7 * 88 * 128) which is about 45 times smaller than using the results for the whole period.

2.11 Possible Filtering Methods to Identify Unknown Notes

In this section (2.11.1 to 2.11.6), we will describe the strategies that we tested to remove extraneous results. The purpose of the strategies was to reduce to the smallest possible size, the group of potential solutions, without removing the correct solution. We were willing to accept that it may be necessary to use more than one strategy in succession to finalise the correct solution. The strategies are methods of artificial perception which in various ways attempt to characterise the "shape" of the Fourier Transform.

For ease of description, the words following have the listed definitions:

Known note	are the results of a single note. In the following pages, the words *known note* imply the results of the FT of the note.
Unknown note	is the note which we are going to compare with all known notes to identify it. In the following pages, *unknown note* means the results of the FT of the unknown note.

Remember in the following sections the results of FT are only one result for each tone or for each frequency range.

2.11.1 The first strategy

A known note is rejected if the highest peak of the unknown note is less than half of its highest peak. This strategy is especially useful when the unknown note contains many values that disperse the energy and such a note is compared to known notes containing only a few high values. Most of these situations occur when the known and the unknown note are from different instruments. e.g. most of the notes from a recorder contain only one or two higher values. Therefore when they are compared to the unknown note of other instruments, they are very easy to filter out.

We set the tolerance at a half to make allowance for the variation in the FT created by the choice of starting point and this will be proven later (see section 2.12.1 on page 71). This also reduces the usefulness of this strategy. Commonly this strategy cut out about 10%-20% of the known notes. When we moved on to multi-notes the amplitude increased, and that made the strategy less usable. This strategy may only be acceptable when detecting single notes.

The strategy written in mathematical style follows,

If *the highest peak of the known note * (Acceptance Percentage) <= the highest peak of the unknown note*, the known notes will be accepted.

All other known notes will be rejected.

An example using made up values:

Note		Highest Result		Second		Third		Fourth	
Unknown Note	(?) (piano)	18000	C3	16200	C2	4500	A4	900	F2
Known Notes	C2 (piano)	18000	C3	16200	C2	4500	A4	900	F2
	B3 (piano)	30000	B3	15000	B4	1500	B5	300	B2
	C4 (piano)	35000	C4	3150	C5	700	D4	30	E2
	A4 (piano)	40000	A4	9000	C4	3000	D4	2000	E4

Table 2.11 A theoretical example of the highest peaks.

Note		Highest Result	The relation between the known note and the unknown note	The application of strategy No. 1.
Unknown Note	(?) (piano)	18000		
Known Note	C2 (piano)	18000	18000 * 0.5 < 18000	Accepted
	B3 (piano)	30000	30000 * 0.5 < 18000	Accepted
	C4 (piano)	35000	35000 * 0.5 < 18000	Accepted
	A4 (piano)	40000	40000 * 0.5 > 18000	Rejected

Table 2.12 Applying the first strategy to Table 2.11.

Table 2.11 shows the FT values for each note. The highest peak from the unknown note is C3 which is 18000 (no unit for FT results) and the lowest, F2, which is 900 (lower values are ignored). For the known note, the highest value for C2 is 18000. For B3 is 30000, C4 is 35000 and A4 is 40000.

In Table 2.12 the first three columns show the note and the highest value for the note. The next column compares the known note to the unknown note. Because half of the highest result from A4 (20000) is higher than the highest result for the unknown note (18000), A4 will be rejected. By applying the same strategy, the others will be accepted.

2.11.2 The second strategy

We first calculate each of the values for a known or unknown note as a percentage with respect to the highest peak in each FT. We then count the number of values that are more than 10% of the highest peak and compare the total number of such values of each known note, to that of each unknown note. If, for any known note, the number of values is greater than the number for the unknown note (plus a tolerance) then that known note is rejected. The tolerance is to compensate for the variation which is created by the selection of starting point.

It is useful if the unknown note is a single note, for this strategy removed many known notes on the same or different instruments. However when the number of unknown notes was greater than two the number of "10% or higher" values increased, and this decreased the number of known notes rejected.

We chose 10% because there are many small values that need to be ignored in order to reduce the work required but other percentages could be investigated in the future.

Example:

Assume there is an unknown note which is actually two notes together and the number of tones which are greater than 10% of the highest peak is 3. Therefore one of the known notes must have the count of 3 or less. It is possible for the other known note to have any value where it contains peaks which are low when compared to the first known note.

If *the number of peaks in the known note – (the number of the peaks in the unknown note + tolerance) <= 0* then these known notes will be accepted, otherwise the known notes will be rejected.

A theoretical example:

Made up result for unknown note, (?) (piano)			
Note	Value of FT	Percentage	
A4	40000	100%	Accepted
C4	9000	22.5%	Accepted
D4	3000	7.5%	Ignored
E4	2000	5%	Ignored
Number of values greater than or equal to 10% of the highest value		2	

Table 2.13 Establishing the values accepted.

Table 2.13 shows the FT results of an unknown note with each value shown as a percentage of the highest value. The highest value is 40000, so the percentage of "tone" C4 is 9000/40000 ~ 22.5%. Since we only count the number of values which are 10% or more of the highest peak, only A4 and C4 are accepted.

Made up result for known note											
C2 (piano)			B3 (piano)			C4 (piano)			A4 (piano)		
C3	18000	100%	B3	30000	100%	C4	35000	100%	A4	40000	100%
C2	16200	90%	B4	15000	50%	C5	3150	9%	C4	9000	22.5%
A4	4500	25%	B5	1500	5%	D4	700	2%	D4	3000	7.5%
F2	900	5%	B2	300	1%	E2	30	0%	E4	2000	5%
No of values >= 10%		3			2			1			2
3 - (2+0) = 1	> 0		2 - (2+0) = 0	= 0		1 - (2+0) = -1	< 0		2 - (2+0) = 0	= 0	
*Rejected known note			*Accepted known note			*Accepted known note			*Accepted known note		

*Assume the tolerance is 0, so if the number of tones is greater than 2+(0) = 2, that known note is ignored.

Table 2.14 Applying the second strategy to Table 2.13.

Table 2.14 shows there are 4 known notes if the tolerance is zero. If there are notes which have counts greater than the count from the unknown note (plus a tolerance) then that note is rejected. e.g. C2 has 3 tones above 10% which is more than the unknown note (2), so C2 is rejected.

2.11.3 The third strategy

This uses the accepted percentages (10% or higher) from the second strategy. Without counting the number, we check each result from the unknown note against the results of all known notes. If the accepted tone of the unknown note is not present in the accepted tones for the known note then that known note will be rejected. Again, this strategy is very useful on single notes. This strategy removed many known notes on the same or different instruments. When the unknown note consisted of two or more notes the number of "10% or higher" values increased and this rapidly decreased the number of rejected known notes.

Filter out the unwanted known note using strategy No. 3								
Musical Note		Results from Fourier Transform						
		C2	C3	B3	C4	A4	B4	
Unknown Note	(?) (piano)				23%	100%		
Known Note	C2 (piano)	90%	100%			25%		Accepted
	B3 (piano)			100%			50%	Rejected
	C4 (piano)				100%			Accepted
	A4 (piano)				23%	100%		Accepted

Table 2.15 A theoretical example applying strategy 3 using the information from Tables 2.13 and 2.14.

In this example, the known notes of C2, C4 and A4 have at least one tone the same as the unknown note and because of that they are accepted. The known note of B3, because it has no matches to the unknown note, is rejected.

If the FT results between the known note and the unknown note are well matched, we may increase the number of matches to exclude more unwanted notes. e.g. if we only accept the notes which have 2 matches or more, A4 is accepted and the others will be rejected. This modification will not work if notes (e.g. recorder) only have one result above 10%.

2.11.4 The fourth strategy

The fourth possible strategy is applied after listing the values of the FT results in descending order. The purpose is to identify the change in the shape of the FT where the amplitude drops sharply. We list the percentages (as in Strategy 2 and 3) in descending order for both the known note and the unknown note. We then locate the position in the series where the difference in amplitude between two adjacent values is more than a set percentage of the highest value. Counting the number of higher values above that position allocates a number to represent the different shapes of the FT. If the count for the known note is larger than the count for the unknown note (plus a tolerance for errors due to the boundary) then that known note is rejected.

In some cases there are no differences more than the conditional percentage, therefore rank could not be determined and we will assume the result is zero. For example, if the biggest difference is 77.5% and the conditional percentage is set at 80%, the result will be recognised as zero.

If the conditional percentage is too small, there are only a small number of notes filtered out. If the percentage is too high, it may create a lot of "zero" results, and that also decreases the number of notes filtered out. Therefore the choice of the conditional percentage is important.

Ideally if the known note fits any of the following requirements, then that note will be accepted.

i) the rank number of the known note – (the rank number of the unknown note + tolerance) <= 0

ii) the rank number of the known note is "zero"

iii) the rank number of the unknown note is "zero"

otherwise, the known note will be rejected.

When the rank is "0" there are too many possibilities, so we will accept the note and let other strategies filter it out.

A theoretical example:

Result for unknown note, (?) (piano)			
Note	Value of FT	Different in Percentages	
A4	100%	77.5%	Accepted
C4	22.5%	12.5%	Accepted
Boundary	10%	-	-
D4	7.5%	2.5%	Ignored
E4	5%	0%	Ignored
The number of differences between each tone greater than or equal to 60% of the highest value.		1	

Table 2.16 Table 2.13 with showing the difference between successive peaks.

In Table 2.16 we count the number of tones for the unknown note which are above the difference (60% or more). The difference between A4 and C4 is 77.5%, and because it is after the first peak, the rank is 1. If the conditional percentage is less than 45% there may be two more differences between peaks which are more than the conditional percentage. This would means there are two ranks indicating the position of the change of shape. An alternative could be either choosing the first rank as the solution or making the conditional percentage higher.

Result for known notes											
C2 (piano)			B3 (piano)			C4 (piano)			A4 (piano)		
C3	100%	10%	B3	100%	50%	C4	100%	90%	A4	100%	77.5%
C2	90%	65%	B4	50%	40%	-	10%	-	C4	22.5%	12.5%
A4	25%	15%	-	10%	-				-	10%	-
-	10%	-									
The differences is >= 60%		2			0			1			1
1 – (2+0) = -1		< 0	1 – (0+0) = 1		> 0	1 – (1+0) = 0		= 0	1 – (1+0) = 0		= 0
*Accepted known note			*Rejected known note			*Accepted known note			*Accepted known note		

*By ranking with a difference in percentages greater than or equal to 60% with a tolerance of 0. If the difference between each tone is less than 1+(0) = 1, that known note is rejected.

Table 2.17 Applying strategy 4 to Table 2.16.

The top part of Table 2.17 shows the differences between the peaks which then leads to the rank. If we set the tolerance as zero the limit will be 1+0 = 1. That means at least one of the notes from the unknown "wave" could contain the rank of either 1 or 0. Others like C2 with the rank of 2 will be rejected.

2.11.5 The fifth strategy

This strategy follows the second and third strategies. It may be possible to identify the unknown note by choosing the correct known note which has the most results from the strategy 2 and 3 in common with the unknown note. There is a possibility that this simple method is flawed. In an example when the unknown note has 3 values matching the known note and 2 values matching with another known note, another consideration may come into play. If the unknown note has 4 values above 10%, the first known note has 8 values above 10%, and the second known note has 5 values above 10%, is it likely that the second known note is the correct answer even though it has 1 less matched value?

We have developed a method which may solve the problem. We calculate the number of matching tones as a percentage of the possible number of tones for each note (either known or unknown), then multiply those percentages. We will then use the highest multiple to indicate the solution.

Example:

1/2 *(Number of matches / Number of tones for the unknown note)*

1/3 *(Number of matches / Number of tones for the known note)*

then multiply them together to get the final percentage,

1/2 * 1/3 = 1/6 ~ 17%.

Expected Note	Number of matches between the known and the unknown notes	Percentage of matches with the result from unknown note (out of 2).	Number of tones for the known note which is greater than or equal to 10%.	Percentage of matches with respect to each known note.	Calculation	Percentage with respect to both unknown and known note.
C2 (piano)	1	50%	3	33.33%	= 50% * 33.33%	16.67%
B3 (piano)	0	0%	2	0%	= 0% * 0%	0%
C4 (piano)	1	50%	1	100%	= 50% * 100%	50%
A4 (piano)	2	100%	2	100%	= 100% * 100%	100%

Table 2.18 Applying strategy 5 using the information from Tables 2.13, 2.14 and 2.15.

In the above table we see the expected note (A4) has the highest final percentage which suggests that this method may be successful. This strategy will be tested in later sections.

2.11.6 Summary

Strategies 1 to 4 are used to reduce extraneous results.

Strategy One: Using the highest peak of both notes

i) If *the highest peak of the known note * (Acceptance Percentage) <= the highest peak of the unknown note*, it is accepted.

Strategy Two: Counting the number of percentages which are over 10% of the highest peak.

i) If *the number of peaks in the known note – (the number of peaks in the known note + tolerance) <= 0*, it is accepted.

Strategy Three: Counting the number of matching tones (which are over 10% of the highest) in both notes.

i) If *"Number of matches" >= (minimum accepted matches)*, it is accepted

Strategy Four: Ranking the shape of FT. If the known note passes any of the following conditions then it is accepted.

i) the rank number of the known note – (the rank number of the unknown note + tolerance) <= 0

ii) *the rank number of the known note is "zero"*

iii) *the rank number of the unknown note is "zero"*

Strategy Five: To find the known note that most closely parallels the unknown note with respect to matching the peaks. The method of calculation of the percentage multiple is

*(Number of matches / Number of tones for the unknown note) * (Number of matches / Number of tones for the known note)*

2.12 Finding the Conditional or Tolerance Values

In this section we give the results for each single note from C2 to B6, when using 6 different starting points and compare the results to the result of the original starting point (15 periods of wavelength). The starting points of the first three experiments are 0, 30, 45 periods of the wavelength. The other three starting points are 15 periods of wavelength plus 500, 1000 or 2000 samples. We do this to show the effect of changing the starting point and the results will also help us find the various tolerances needed.

The selection of starting point in this chapter is, to some extent, artificial because accurate selection is a difficulty which is beyond the scope of this thesis. The accurate selection of starting point will require more work by others however this does not affect the validity of our studies on the methods of identification of notes. Starting point is also discussed later in the appropriate sections.

We note in passing Goto and Muraoka (1997) in their discussion on the evaluation of beat tracking system have mentioned the finding of the onset times as one of the three main sets of problems for beat tracking researchers.

In Chapter 3 the selection of starting point for real instruments was done by visual examination of a number of FT's for each note to select the FT which best covered the stable stage. This visual selection avoided the need to know the frequency of the notes before the frequency is detected. We also did initial testing on a program to automatically find the best starting point.

2.12.1 The first strategy

We use the highest peak of the original FT as the base for comparison. Then we compare the highest peaks from other starting points as percentages of it.

The results of the experiment measuring the percentage difference between the highest peaks of the known and the unknown note on AGP is displayed in Table B.3 in Appendix B. We summarise the results at the bottom of table. Because the unknown note could be more than one note the maximum is meaningless. The minimum is the percentage of the peak compared to the peak from the original starting point. The minimum is important because it shows either the likely level of the acceptance percentage will be too high and reject a valid solution, or if it is too low, we will not reject enough extraneous results.

The table (and also Table 2.19 below) shows the greatest reduction is the minimum of 22% of the original starting point with AGP. This percentage is affected most by later starting points because some of the FT is pushed out of the stable stage. If we choose a starting point reasonably close to the original starting point, the acceptance percentage can be more stringent. Taking this into account, we can conclude from Table B.3 that using "50%" as the acceptance percentage is acceptable for the cases where the starting points are close the original (less than 1000 samples or 15 periods of wavelength away from the original starting point).

Table 2.19 (details see Tables B.3 to B.7 in Appendix B) summarises the same information for five instruments and leads to the same conclusion. Although for some instruments a smaller acceptance percentage is possible, because the instrument of the unknown note is also unknown, making a smaller tolerance was not considered.

Minimum value on different instrument						
Instrument	**Period**			**15 Periods plus number of samples**		
	0	**30**	**45**	**500**	**1000**	**2000**
1	90%	61%	48%	54%	37%	22%
26	80%	83%	68%	77%	66%	49%
33	97%	53%	41%	81%	71%	59%
41	50%	72%	63%	87%	80%	64%
75	26%	87%	44%	78%	76%	59%
Min.	26%	53%	41%	54%	37%	22%

Table 2.19 Summary of different minimum values using various starting points on 5 different instruments (in percentages).

2.12.2 The second strategy

We count the number of tones of the unknown note which are 10% of the original highest peak, or higher, for each different starting point. We then subtract that number from the number of the original 15 periods starting point FT. i.e. known data count –

unknown data count. e.g. if the original is 7, and the new result is 8, then the value will be –1. If we find the minimum differences of this calculation we will be able to approximate a tolerance.

Minimum difference on different instruments						
Note value in MIDI	Starting Point					
	Periods			15 Periods plus number of samples		
	0	30	45	500	1000	2000
1	-3	-6	-5	-7	-3	-10
26	-8	-4	-5	-2	-4	-4
33	-7	-3	-3	-3	-4	-2
41	-18	-8	-10	-3	-4	-4
75	-8	-2	-5	-1	-2	-2
Min.	-18	-8	-10	-7	-4	-10

Table 2.20 Minimum differences for 5 different instruments.

As we can see that the minimum differences varies for different instruments and starting points. The difference due to the starting point has already been discussed. We can set the tolerance about 9 by ensuring the starting point is close to the original. This tolerance is a large part of the range (which for AGP is about 30) therefore this strategy is not efficient.

2.12.3 The third strategy

The third strategy is to match the peaks which are 10% or higher of the highest peak of the known note to those of the unknown note. Table 2.21 shows the minimum matches on 5 different instruments.

The table shows the minimum number of matches is 1. This means that if there is no match for any known note to the unknown note, then that known note will be rejected. One possible reason for only one match on some instruments, is that those instruments may only produce a small number of higher peaks (e.g. recorder). Another possibility is some of the tones may die off during the note. Therefore the minimum number of matches that is acceptable cannot be more than 1.

Minimum matches for different instruments						
Instrument	Starting Point					
	Period			15 Period plus number of samples		
	0	30	45	500	1000	2000
1	1	2	2	2	2	2
26	3	4	4	4	3	2
33	2	2	2	2	2	2
41	4	3	3	4	3	3
75	1	1	1	1	1	1
Min	1	1	1	1	1	1

Table 2.21 Minimum matches for 5 different instruments.

2.12.4 The fourth strategy

This strategy is designed to find the shape of the FT. We need to find the value of the conditional percentage and the tolerance (due to the boundary errors of the notes). Tables 2.22 and 2.23 show the results of the analyses using different conditional percentages. We rank the number of peaks for each starting point and compare these to the number of peaks in the FT of the 15 periods starting point. We did this on 5 different instruments (acoustic grand piano, acoustic guitar (steel), acoustic bass, violin, and recorder).

Starting Point		**The number of zeros (using different conditional percentages)**						
		30%	**40%**	**50%**	**60%**	**70%**	**80%**	**90%**
Periods	**15**	21	63	102	164	204	245	268
	0	34	81	130	167	214	249	271
	30	15	58	106	152	205	247	263
	45	18	68	118	161	198	249	265
15 periods plus number of samples	**500**	17	60	110	153	204	247	265
	1000	19	64	116	154	201	245	267
	2000	15	59	106	147	198	241	268
Average		19.86	64.71	112.57	156.86	203.43	246.14	266.71
No. of cases		300	300	300	300	300	300	300
Average (in percentage)		6.62%	21.57%	37.52%	52.29%	67.81%	82.05%	88.90%

Table 2.22 The number of zeros when using different conditional percentages.

As we mentioned previously if the rank of the known note is zero, then that known note will be accepted. This means the number of zeros is inversely related to the number of notes that will be rejected. Table 2.22 shows the number of zeros for different notes. From the table, we can see that as the conditional percentage increases, the number of zeros also increases. Therefore there will be a smaller chance of rejecting the known note. This is because the vertical spaces between the "higher" peaks and the "lower" peaks are small, hence a small conditional percentage is needed.

Table 2.23 shows the smallest differences between the rank of the known note and the rank of the unknown note. i.e. "known" – "unknown". These values will give us an idea of the tolerance required. The table shows the smaller conditional percentages requires a larger tolerance (up to 6). This is because the changes due to different starting points have become larger than the conditional percentages. This tolerance of 6 when added to the rank of the unknown note causes much fewer notes to be rejected, therefore a small tolerance is needed.

The balance of options discussed above leads to the better choices being "40%" or "50%" and their tolerances are 2 for "40%" and 1 for "50%". We will do some more analysis on these parameters in the next section.

Starting Point		**The minimum results (using different percentages)**						
		30%	**40%**	**50%**	**60%**	**70%**	**80%**	**90%**
Periods	**0**	-2	-1	-1	0	0	0	0
	30	-2	-1	0	0	0	0	0
	45	-4	-2	0	0	0	0	0
15 periods plus number of samples	**500**	-2	-2	0	0	0	0	0
	1000	-6	-1	-1	-1	0	0	0
	2000	-2	-1	-1	-1	0	0	0
Minimum		-6	-2	-1	-1	0	0	0

Table 2.23 The minimum results of "known" minus "unknown" using different conditional percentages.

2.12.5 The fifth strategy

The calculated percentage multiple varies greatly, therefore there are no conditional or tolerance values required.

2.13 The Accuracy Rate for Different Strategies when Playing a Simple Single Note

A simple single note is defined as one note per file without any distortion from the previous note.

There are two cases. First we compare each unknown note to each known note from the same instrument (60 possible solutions), and then we compare each unknown note to each known note from 5 different instruments (60*5 = 300 possible solutions). All the unknown notes start 500 samples (in real data) after the original starting point.

2.13.1 Application of using strategies 1 to 4

Table 2.24 shows the results using the known note which comes from the same instrument. In the table, there are two columns for strategy 4. One column represents 40% as the conditional percentage with tolerance value 2, and the other column represents 50% with tolerance value 1.

	Instrument Number	Strategy				
		1	2	3	4	
					(50%)	(60%)
Average number of known notes rejected	1	15.37	8.82	21.75	0	0
	26	8.75	6.8	24.1	0	0
	33	2.85	0.55	42.8	0	0
	41	7.95	2.02	28.82	0	1.2
	75	2.22	0	57.13	0	0
Average		7.43	3.64	34.92	0	0.24
No.of Cases		60	60	60	60	60
Percentage Rejected		12.38%	6.06%	58.20%	0%	0.40%

Table 2.24 Results with notes from the same instrument.

From the table we see the largest rejection rate is by strategy 3. Strategies 1, 2 and 4 are progressively less selective. Strategy 1 rejects less than a quarter of the number of notes rejected by strategy 3. This is because the use of the tolerance from strategy 1 reduces the number of rejected notes (especially on some instruments, e.g.

MIDI 33 & 75). The results of strategy 4 are close to zeros, so this means strategy 4 is not efficient.

	Instrument Number	Strategy				
		1	2	3	4	
					(50%)	(60%)
Average number of known notes rejected	1	103.65	21.73	155.47	0	2.67
	26	110.77	16.10	165.08	0	2.13
	33	92.73	26.85	201.75	0	2.87
	41	37.97	18.67	171.52	0	1.2
	75	2.23	53.62	260.97	0	4
Average		69.47	27.39	190.96	0	2.57
No. of Cases		300	300	300	300	300
Percentage Ignored		23.16%	9.13%	63.65%	0%	0.86%

Table 2.25 Results with notes from 5 different instruments.

Table 2.25 shows similar results to Table 2.24. Strategy 3 is still has the highest removal rate (~ 64%) and strategy 4 has the lowest at 1%. Since the both tables show strategy 4 is inefficient in rejecting the known note, we will discard this strategy.

Since we know strategy 3 is the best we will use it first, but we need to find which strategy of 1 or 2 to use next.

	Instrument Number	Known notes come from the same instrument			Known notes come from 5 different instruments		
		Strategy 3	Strategy 3 then		Strategy 3	Strategy 3 then	
			Strategy 1	Strategy 2		Strategy 1	Strategy 2
Average number of known notes rejected	1	21.75	31.20	29.25	155.47	201	170.93
	26	24.1	28.98	29.12	165.08	203.63	176.85
	33	42.8	43.25	42.83	201.75	225.85	218.35
	41	28.82	31.28	29.80	171.52	176.37	184.12
	75	57.13	57.13	57.13	260.97	260.97	276.55
Average		34.92	38.37	37.63	190.96	213.56	205.36
No. of Cases		60	60	60	300	300	300
Percentage Ignored		58.2%	63.95%	62.71%	63.65%	71.19%	68.45%

Table 2.26 The summary of the number of rejected known notes using strategy 3 followed by strategies 1 and 2 in alternative order.

Table 2.26 shows the difference between these alternatives is insignificant (less than 3% of the total in both cases). Since strategy 1 will become less useful than strategy 2 as the number of notes played increases, we use strategy 2 as the second step and strategy 1 as the third step.

	Instrument Number	Known notes come from the same instrument			Known notes come from 5 different instruments		
		Strategy 3	Strategy 2	Strategy 1	Strategy 3	Strategy 2	Strategy 1
Average number of known notes rejected	1	21.75	29.25	37.62	155.47	170.93	214.27
	26	24.10	29.12	33.97	165.08	176.85	215.08
	33	42.80	42.83	43.28	201.75	218.35	242.06
	41	28.82	29.80	32.27	171.52	184.12	188.97
	75	57.13	57.13	57.13	260.97	276.55	276.55
Average		34.92	37.63	40.85	190.96	205.36	227.38
No. of Cases		60	60	60	300	300	300
Percentage Ignored		58.20%	62.71%	68.09%	63.65%	68.45%	75.79%

Table 2.27 The summary of applying three different strategies in series showing the number of the correctly located notes.

Table 2.27 shows there are still 20% to 30% of the known notes remaining as potential solutions. Although we have not yet located the solution for the unknown note, the number of the potential solutions is a lot smaller.

We could use strategies "3, 2 and 1" to reduce the number of the potential solutions, while avoiding the possibility of incorrect results, when using strategy 5. For faster processing and to reduce the memory required, instead of using this three stage process we only use strategy 5, allowing less complicated programs. This reduces the possibility of getting incorrect programs and the outcomes of the analyses are more reliable.

2.13.2 Application of strategy 5

We calculate the percentage multiple (for the method see section 2.11.5 on page 69) and find the rank of the expected frequency.

Table 2.28 summarises these results. Columns 2 & 3 show the average rank in the cases of the known note which comes either from the same instrument, or from 5 different instruments. Columns 4 & 5 show the lowest rank the expected results could be.

Instrument Number	**Average Rank**		**Lowest Rank**	
	Known notes come from		**Known notes come from**	
	the same instrument	**5 different instruments**	**the same instrument**	**5 different instruments**
1	1	1	1	1
26	1	1.03	1	3
33	1	1.07	1	3
41	1	1	1	1
75	1.02	1.03	2	2
Average \ Lowest Rank	1	1.03	2	3

Table 2.28 The summary of the rank of the correct notes.

We make two observations from Table 2.28.

i) Because the "Average Rank" columns (the average rank of strategy 5) show the position of the correct solution is very close to 1.00, the highest rank is very likely to give a correct answer.

ii) The "Lowest Rank" columns (the lowest rank using strategy 5) show the correct solution (expected frequency) will be among the upper ranks but not necessarily be the highest. This observation can help us to limit the number of the potential solutions.

This is by far the best result from any of the strategies tested.

Instrument Value	Known notes come from	
	the same instrument	5 different instruments
1	60	60
26	60	59
33	60	57
41	60	60
75	59	58
Average	59.8	58.8
No. of Cases	60	60
Percentage	99.67%	98.00%

Table 2.29 The results of choosing the highest rank in strategy 5 as the solution.

Table 2.29 shows the average accuracy rate for detecting a single note, when the starting point of the unknown note is 500 samples away from the original starting point, is about 99%. This is much more accurate than choosing the highest peak in the FT results as the solution.

2.14 Application of Strategies 3-2 -1 when Playing Two Notes

We chose two different starting points for this experiment. The first is 15 periods of the average frequency of the two notes. The second is 15 periods of the lower frequency. For each starting point, we examined the cases where the known notes are from the same instrument and when they are from 5 different instruments.

Table 2.30 and 2.31 show the number of known notes that have been filtered out when using strategies 3, 2 and 1 as in Table 2.27 in previous section.

We have excluded the repeated and reciprocal cases. The results weren't doubled in contrast to section 2.5 (on page 41).

Average frequency starting point						
Instrument Number	Known notes come from the same instrument			Known notes come from 5 different instruments		
	Strategy 3	Strategy 2	Strategy 1	Strategy 3	Strategy 2	Strategy 1
1	14.75	20.33	22.65	122.86	132.68	149.68
26	15.51	17.32	18.25	122.99	129.08	151.16
33	33.20	33.20	33.24	158.97	169.75	182.48
41	15.54	15.70	16.95	112.03	116.35	121.26
75	54.67	54.67	54.71	234.18	254.40	254.45
Average	26.73	28.24	29.16	150.21	160.45	171.81
No. of Cases	60	60	60	300	300	300
Percentage Rejected	44.56%	47.07%	48.60%	50.07%	53.48%	57.27%

Table 2.30 The average number of known notes in different instruments that have been filtered out when using average frequency as the starting point.

Lower frequency starting point						
Instrument Number	Known notes come from the same instrument			Known notes come from 5 different instruments		
	Strategy 3	Strategy 2	Strategy 1	Strategy 3	Strategy 2	Strategy 1
1	14.39	19.72	24.11	120.33	129.87	158.18
26	15.42	17.07	19.58	119.34	125.19	159.51
33	33.78	33.78	33.85	161.19	172.12	193.12
41	15.81	15.96	16.97	114.09	118.69	121.14
75	54.64	54.64	54.64	233.27	253.41	253.41
Average	26.81	28.23	29.83	149.64	159.86	177.07
No. of Cases	60	60	60	300	300	300
Percentage Rejected	44.68%	47.06%	49.71%	49.88%	53.29%	59.02%

Table 2.31 The average number of known notes in different instruments that have been filtered out when using lower frequency as the starting point.

Approximately 40%-50% of the known notes are filtered out. These results appear to show the strategies are quite efficient. However Table 2.32 shows the number of unknown notes for which the expected solutions are erroneously filtered out, by strategies 1 and 2 because of the tolerance. Therefore these strategies, rather than being quite efficient, are flawed. After this, our investigations were limited to strategy 5 and its improvement.

Instrument Number	Average frequency starting point			Lower frequency starting point		
	Strategy 3	Strategy 2	Strategy 1	Strategy 3	Strategy 2	Strategy 1
1	0	0	0	0	0	0
26	0	3	4	0	0	1
33	0	0	0	0	0	0
41	0	0	0	0	0	0
75	0	0	0	0	0	0
All Total	0	3	4	0	0	1

Table 2.32 The number of known notes that have been filtered out as we apply the strategies.

2.15 Application of Strategy 5 on Two Notes

2.15.1 Finding the first solution

First we will look at the accuracy of choosing the highest rank through the calculations of strategy 5.

Instrument Number	Known notes come from the same instrument		Known notes come from 5 different instruments	
	Average freq. starting point	Lower freq. starting point	Average freq. starting point	Lower freq. starting point
1	1759	1767	1623	1694
26	1674	1767	1554	1729
33	1752	1763	1501	1648
41	1770	1769	1514	1698
75	1304	1622	1147	1502
Average	1651.8	1737.6	1467.8	1654.2
No. of Cases	1770	1770	1770	1770
Percentage Rejected	93.32%	98.17%	82.93%	93.46%

Table 2.33 The number of cases which the highest rank is equal to the expected note.

Table 2.33 leads to four observations.

i) The unknown notes which come from different instruments have different levels of accuracy rate.

ii) The accuracy rates using the known notes from the same instrument are higher than when the known notes come from 5 different instruments. This is because the choices for the known notes are less therefore there will be less chance of having a note which is similar, but not the same, as the unknown note.

iii) Choosing the lower frequency as the starting point is more accurate than choosing the average frequency as the starting point. When using the lower frequency starting point, the unknown note and the known note are in the same phase of the waveform, therefore the differences in the FT results are smaller.

iv) Using Strategy 5 is efficient in finding the solution. The accuracy rate for different cases varies between 83% and 98%

2.15.2 Subtraction of notes to find the second solution

We simply subtract the filtered (complex) value of the FT of the known note from the FT of the unknown note, then apply strategy 5 as used for the first solution.

In Table 2.34, the "A1" column shows the results for the first solution and the "A2" column is the results of the second solution. Column "A2" includes duplications which appear in both "A1" and "A2". Column 4 is the number of duplicated cases where they both matched one of the expected results. Column 5 is the duplicated cases where neither matched the expected solutions.

Instrument Number	No. of Correct Answers		No. of Duplicated Answers	
	A1	A2	Correct Once	Both Wrong
Average Frequency (Same Instrument)				
1	1759	1727	1512	10
26	1674	1663	1341	83
33	1752	1702	1422	14
41	1770	1755	1453	0
75	1304	1251	1064	370
Average	1651.8	1619.6	1358.4	95.4
No. of Cases	1770	1770	1770	1770
Percentage	93.32%	91.50%	76.75%	5.39%
Lower Frequency (Same Instrument)				
1	1767	1691	103	3
26	1767	1757	190	3
33	1763	1746	141	7
41	1769	1759	72	1
75	1622	1456	100	119
Average	1737.6	1681.8	121.2	26.6
No. of Cases	1770	1770	1770	1770
Percentage	98.17%	95.02%	6.85%	1.50%

Table 2.34 The number of results is correct using strategy 5 (with repeated answer).

Table 2.34 shows the large number of duplications, especially when using the average frequency for the starting point. This should not be possible, since we know there are two different notes playing at the same moment not one. This happens because the remainder after the subtraction still contains some characteristics of the first note. Evidence for this can be seen in the fourth column if we compare using lower frequency instead of average frequency as the starting point. The number of duplications is much smaller because the lower frequency starting point enables a better match between part of the unknown note and the known note which is the solution.

2.15.2.1 Subtraction excluding the first solution

2.15.2.1.1 Simple subtraction

One way of removing duplications is to ensure the first solution is not amongst the possible second solutions. The summary of experiments doing these is shown in Table 2.35.

Instrument Number	Known notes come from the same instrument		Known notes come from 5 different instruments	
	Average freq. starting point	Lower freq. starting point	Average freq. starting point	Lower freq. starting point
1	798	1667	439	1342
26	1084	1685	579	1330
33	842	1705	400	1339
41	1278	1751	629	1546
75	955	1531	727	1374
Average	991.4	1667.8	554.8	1386.2
No. of Cases	1770	1770	1770	1770
Percentage	56.01%	94.23%	31.34%	78.32%

Table 2.35 The number of cases where the second solution is correct when preventing the first solution from being repeated.

These results leads to the same observations as in Table 2.33.

The accuracy rates of the second solution for the average frequency starting point are very low (56% or 31%). This is not acceptable.

2.15.2.1.2 Complex subtraction method 1

We checked whether the results from the simple subtraction would be improved by reversing the process i.e. doing all the subtraction before doing the filtering. This change of order may have lessened the loss of correct solutions.

Instrument Value	Known notes come from the same instrument		Known notes come from 5 different instruments	
	Average freq. starting point	Lower freq. starting point	Average freq. starting point	Lower freq. starting point
1	802	1719	449	1485
26	1110	1687	612	1396
33	851	1707	408	1365
41	1217	1757	591	1619
75	959	1568	739	1421
Average	987.8	1687.6	559.8	1457.2
No. of Cases	1770	1770	1770	1770
Percentage	55.81%	95.34%	31.63%	82.33%

Table 2.36 The number of correct second solution using strategy 5 with complete FT results.

From Table 2.36, we can see the method produces only a slight improvement in the accuracy rate.

2.15.2.1.3 Complex subtraction method 2

We tried two other methods of subtractions. The result for the unknown note for a certain "tone" can be written as $(\pm)\ a + (\pm)\ b\ i$ and the result for the known note for a certain "tone" as $(\pm)\ ka + (\pm)\ kb\ i$. The $(\pm)$ means the results could be positive or negative.

In previous experiments, the methods used are based on a simple calculation.

i.e. $(\pm)\ a + (\pm)\ b\ i \quad - \quad (\pm)\ ka + (\pm)\ kb\ i$

Method I) make a, b, ka & kb always positive and do the subtraction. (We called this the "ignored sign" method) i.e. $(+)\,a + (+)\,b\,i \quad - \quad (+)\,ka + (+)\,kb\,i$

Method II) or find the absolute value on both results and do the subtraction. (We called this the "absolute value" method)

i.e. $abs((\pm)\,a + (\pm)\,b\,i\,) \quad - \quad abs((\pm)\,ka + (\pm)\,kb\,i\,)$

Table 2.37 shows the results from using these two methods when using different numbers of known notes and different types of the FT results (filtered and complete as in section 2.15.2.1.2 on page 87).

We made two new observations.

i) Under the same conditions, the accuracy rate using the complete FT results is higher than when using the filtered FT results.

ii) Under the same conditions, the accuracy rate using the absolute method is higher than when using the ignored sign method.

The table also shows the accuracy rate, especially using average frequency starting point, has increased markedly. The accuracy rate, having increased by 50%, has moved from being unacceptable to being acceptable.

Instrument Number	Filtered FT				Complete FT			
	Known notes come from the same instrument		Known notes come from 5 different instruments		Known notes come from the same instrument		Known notes come from 5 different instruments	
	Average freq. starting point	Lower freq. starting point	Average freq. starting point	Lower freq. starting point	Average freq. starting point	Lower freq. starting point	Average freq. starting point	Lower freq. starting point
Ignored Sign Method								
1	1370	1697	1041	1373	1400	1758	1073	1528
26	1654	1759	1422	1417	1663	1763	1512	1487
33	1501	1758	1100	1411	1503	1764	1122	1462
41	1735	1758	1462	1581	1702	1762	1287	1638
75	1443	1589	1319	1475	1439	1581	1315	1477
Average	1540.6	1712.2	1268.8	1451.4	1541.4	1725.6	1261.8	1518.4
No. of Cases	1770	1770	1770	1770	1770	1770	1770	1770
Percentage	87.04%	96.73%	71.68%	82.00%	87.08%	97.49%	71.29%	85.79%
Absolute Values Method								
1	1553	1692	1307	1350	1633	1757	1431	1515
26	1699	1765	1552	1406	1712	1768	1645	1492
33	1643	1761	1278	1402	1630	1763	1353	1457
41	1760	1761	1567	1562	1759	1764	1624	1641
75	1587	1601	1530	1510	1587	1591	1529	1504
Average	1648.4	1716	1446.8	1446	1664.2	1728.6	1516.4	1521.8
No. of Cases	1770	1770	1770	1770	1770	1770	1770	1770
Percentage	93.13%	96.95%	81.74%	81.69%	94.02%	97.66%	85.67%	85.98%

Table 2.37 The number of correct second solution using strategy 5 with different subtraction methods.

2.15.2.2 Summary of subtraction excluding the first solution

Figure 2.38 and 2.39 summarise the accuracy rate of the solution when ensuring the first solution is not repeated. The last row of boxes in each figure contains information about the second solution. The first row shows the method we used. The second row is the accuracy rate using known notes which come from the same instrument. The third row is the accuracy rate using known notes coming from five different instruments, and the last row is a ranking of each method's relative accuracy. The descriptions of the methods are abbreviated to two letters. These abbreviations are also used later.

SF = Simple Calculation Filtered FT	SC = Simple Calculation Complete FT
IF = Ignored Sign .... FT	IC = Ignored Sign .... FT
AF = Absolute Value .... FT	AC = Absolute Value .... FT

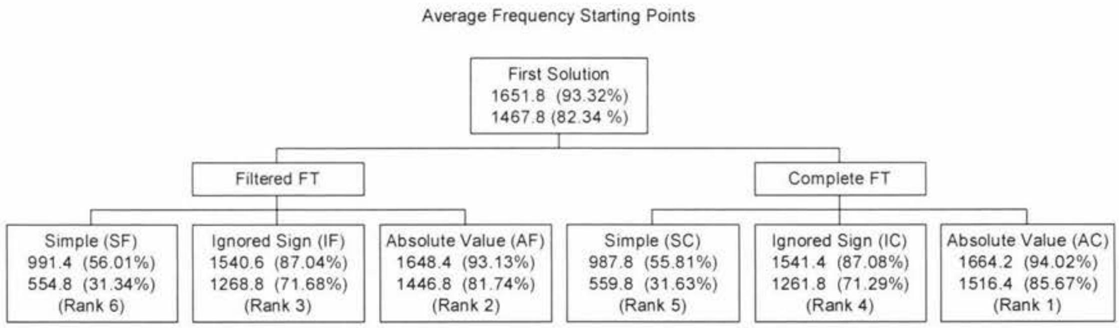

Figure 2.38 Summary of different methods when using average frequency starting point.

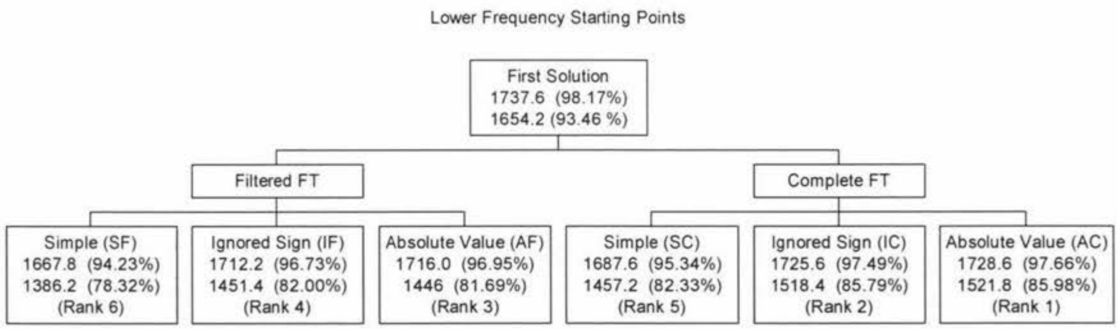

Figure 2.39 Summary of different methods when using lower frequency starting point.

2.15.2.3 Preliminary conclusions about strategy 5

i) The number of instruments for the known note, and the source of the unknown note will affect the accuracy rate.

ii) The accuracy rate using the simple method of calculation is unacceptable since the accuracy rates of the second solutions are as low as 32%.

iii) The difference between using the filtered FT results and the complete FT results is small, so the use of the filtered FT results can be considered when time is important.

iv) The absolute values method has a higher accuracy rate than the ignored sign method which in turn is more accurate than the simple method.

2.16 Application of Strategy 5 on Three Notes Together

We performed the experiment of IF, AF, IC and AC on three notes playing together. We then reverted to using the combinations of even MIDI numbers between C2 (36) to B6 (95). e.g. "36, 38 & 40" and "36, 70 & 94". As before, the experiment excludes the repeated notes and the reciprocal cases. The number of cases in each experiment is 4060. Details can be found in Table B.8 in Appendix B.

	Filtered FT			Complete FT		
	A1	A2	A3	A1	A2	A3
Ignored Sign Method						
Average Frequency (Same Instrument)	90%	89%	74%	90%	89%	76%
Lowest Frequency (Same Instrument)	96%	95%	83%	96%	96%	86%
Average Frequency (Five Instruments)	68%	67%	54%	68%	68%	59%
Lowest Frequency (Five Instruments)	86%	75%	58%	86%	78%	62%
Absolute Value Method						
Average Frequency (Same Instrument)	90%	92%	86%	90%	93%	88%
Lowest Frequency (Same Instrument)	96%	95%	91%	96%	96%	94%
Average Frequency (Five Instruments)	68%	73%	64%	68%	76%	75%
Lowest Frequency (Five Instruments)	86%	76%	62%	86%	79%	71%

*A1 = First Solution, A2 = Second Solution, A3 = Third Solution

Table 2.38 Summary of the accuracy rate using different methods to find the solution when three notes are played together.

Although there were small differences, we observed the same features we have discussed while describing our experiments when two notes are playing together. These differences are a general reduction in accuracy, because of the greater complexity of the wave when three notes are being played together.

2.17 Sources of Error within Calculations

The calculation method will not affect the accuracy of the first solution. However as we located more solutions, the accuracy rate changed depending on the method of subtraction.

Table B.9 in Appendix B are the summary of the number of errors occurring in method AC when two notes are played together. It is very hard to see any pattern to the errors especially when using the lower frequency starting point. Errors when using a lower frequency starting point are less important because the results are usually more accurate than when using the average frequency starting point. However when using the average frequency starting point, the accuracy rate for the notes from the central frequencies (e.g. octave 4) is higher than when the notes are from lower or higher frequencies. Similar results also can be found when using methods IF, IC and AF and also when three notes are played together.

Other sources of error could be related to the cases where the "second" or the "third" notes are the harmonic terms of the "first" note. This error does not appear strongly when we look at results when using method AC when two notes are played together. When three notes are played together it was very difficult to separate the combinations where two or three notes are "related", so no conclusion has been made. We will look at this again in the next chapter when playing two notes together on a real grand piano.

2.18 Finding the Number of Notes that are Playing Simultaneously?

It is very hard to tell by looking at the FT results, the number of notes present, because the amplitude of the FT could be a combination of several "tones". It would appear that the only way is by looking at the amplitudes of the contributing waves. However the wave of the unknown note and that of the known note could be in different phases, or at different sound levels, so it is also difficult to recognise the number of notes from the maximum amplitudes of the combined wave. We concentrated on the investigation of the amplitude of the FT results because we can do the subtraction after finding the successive solutions and the remnant of the FT will have an ever reducing amplitude. This phenomenon does not usually occur when doing subtraction using waves.

In the previous experiments (AC, AF, ...), we had recorded the maximum value in each of the FT results (in absolute value format). Since the FT for single notes are created from a same sample wave file, using different starting points, the amplitudes in the FT result are very similar. The high accuracy rate when finding the solution for single notes means the unknown notes are correctly removed.

Table 2.39 is a summary of the averages of the maximums of the highest peaks using different methods of calculation when playing two notes together and Table 2.40 is the summary when playing three notes together.

	Using Filtered FT Results			Using Complete FT Results		
	Before Subtraction	After First Subtraction	After Second Subtractions	Before Subtraction	After First Subtraction	After Second Subtractions
Simple Method						
Average Frequency (Same Instrument)	2.86E+06	3.72E+06	4.04E+06	2.86E+06	3.72E+06	4.05E+06
Lower Frequency (Same Instrument)	2.81E+06	2.23E+06	2.79E+06	2.81E+06	2.21E+06	2.80E+06
Average Frequency (Five Instruments)	2.86E+06	3.70E+06	4.17E+06	2.86E+06	3.69E+06	4.19E+06
Lower Frequency (Five Instruments)	2.81E+06	2.27E+06	2.86E+06	2.81E+06	2.25E+06	2.85E+06
Average	2.84E+06	2.98E+06	3.47E+06	2.84E+06	2.97E+06	3.47E+06
Ignored Sign Method						
Average Frequency (Same Instrument)	2.86E+06	2.24E+06	1.44E+06	2.86E+06	2.27E+06	1.51E+06
Lower Frequency (Same Instrument)	2.81E+06	2.01E+06	8.70E+05	2.81E+06	2.04E+06	1.01E+06
Average Frequency (Five Instruments)	2.86E+06	2.25E+06	1.48E+06	2.86E+06	2.28E+06	1.58E+06
Lower Frequency (Five Instruments)	2.81E+06	2.02E+06	9.27E+05	2.81E+06	2.06E+06	1.08E+06
Average	2.84E+06	2.13E+06	1.18E+06	2.84E+06	2.16E+06	1.30E+06
Absolute Value Method						
Average Frequency (Same Instrument)	2.86E+06	2.11E+06	4.83E+05	2.86E+06	2.15E+06	5.53E+05
Lower Frequency (Same Instrument)	2.81E+06	1.98E+06	2.33E+05	2.81E+06	2.01E+06	3.83E+05
Average Frequency (Five Instruments)	2.86E+06	2.11E+06	5.98E+05	2.86E+06	2.16E+06	6.92E+05
Lower Frequency (Five Instruments)	2.81E+06	1.99E+06	3.33E+05	2.81E+06	2.03E+06	4.99E+05
Average	2.84E+06	2.05E+06	4.12E+05	2.84E+06	2.09E+06	5.32E+05

Table 2.39 Summary of the averages of the highest peak of the FT results using different methods when two notes are played together.

	Using Filtered FT Results				Using Complete FT Results			
	Before Subtraction	**After First Subtraction**	**After Second Subtractions**	**After Third Subtractions**	**Before Subtraction**	**After First Subtraction**	**After Second Subtractions**	**After Third Subtractions**
Ignored Sign Method								
Average Freq. (Same Inst.)	3.15E+06	2.76E+06	2.24E+06	1.72E+06	3.15E+06	2.78E+06	2.28E+06	1.78E+06
Lowest Freq. (Same Inst.)	3.03E+06	2.55E+06	2.02E+06	1.33E+06	3.03E+06	2.58E+06	2.06E+06	1.47E+06
Average Freq. (Five Inst.)	3.15E+06	2.76E+06	2.27E+06	1.78E+06	3.15E+06	2.79E+06	2.35E+06	1.88E+06
Lowest Freq. (Five Inst.)	3.03E+06	2.56E+06	2.05E+06	1.43E+06	3.03E+06	2.59E+06	2.11E+06	1.60E+06
Average	3.09E+06	2.66E+06	2.15E+06	1.57E+06	3.09E+06	2.69E+06	2.20E+06	1.68E+06
Absolute Value Method								
Average Freq. (Same Inst.)	3.15E+06	2.69E+06	1.99E+06	7.94E+05	3.15E+06	2.72E+06	2.05E+06	8.93E+05
Lowest Freq. (Same Inst.)	3.03E+06	2.52E+06	1.85E+06	5.00E+05	3.03E+06	2.55E+06	1.91E+06	7.18E+05
Average Freq. (Five Inst.)	3.15E+06	2.68E+06	2.01E+06	1.02E+06	3.15E+06	2.72E+06	2.11E+06	1.17E+06
Lowest Freq. (Five Inst.)	3.03E+06	2.52E+06	1.87E+06	7.55E+05	3.03E+06	2.56E+06	1.96E+06	1.00E+06
Average	3.09E+06	2.60E+06	1.93E+06	7.67E+05	3.09E+06	2.64E+06	2.01E+06	9.45E+05

Table 2.40 Summary of the averages of the highest peak of the FT results using different methods when three notes are played together.

There are five observations from this experiment.

i) The results of the simple method of calculation are very erratic. The average maximum peaks on the FT remnant for the other methods decreases as we progressively remove solutions.

ii) Other than for the simple method, the tables show the final remnant using the lower frequency starting point, is lower than that left after using the average frequency starting point. This is because the known note and the unknown note are in the same phase, so subtracting the first solution is more "complete".

iii) Whether we use a known note from the same instrument, or one from 5 different instruments, there is relatively little difference compared to the amplitude itself, in the remnant.

iv) The remnant using the filtered FT results is lower than the one when using the complete FT results because the filtered FT results are always the highest within the range for each tone. Therefore the subtraction will always remove the largest possible amount for the known note.

v) The remnant when using the ignored sign method is always higher than the one using the absolute value method. The accuracy of finding the correct solution with the ignored sign method is lower than that using the absolute value method. Therefore when removing the incorrect solution the removal of amplitude of the FT will also be incorrect.

To be more specific IC and IF are less useful because their final remnant was still too high when compared to AC or AF. By looking at the maximum height of an unknown note we can estimate the number of notes from their amplitudes. However some notes may contain very high or very small peaks. Therefore we needed to progressively remove each note until we could remove no more. For AC and AF the final remnant when two notes played was about $2 \sim 7 * 10^5$ (see Table 2.39) , which was slightly higher than the lowest maximum height of all the known notes ($4.68 * 10^5$). It was much lower than the average of them ($2.39 * 10^6$). If we used any value below $2.39 * 10^6$ as the boundary, we could usually correctly estimate the number of solution.

Table 2.40 leads to the same observations for three notes as from Table 2.39. The only difference was the range of the final remnant was higher (about $5 * 10^5 \sim 1.17 * 10^6$). The range was still below the average of the highest peak for the known note ($2.39 * 10^6$).

For MIDI instruments we have usually correctly estimated the number of notes being played. When the volume level of the known notes and the unknown notes, are different, or the number of the unknown notes increases, the possibility of locating the correct number of solutions may be less.

Chapter Three

Real Instruments

Chapter two discusses experiments where the wave files have been generated by computer. These are very stable and without any noise or any other mechanical interference caused by outside sources. We investigated the application of strategy 5 to a real grand piano playing either single notes or two notes together. The cases of three or more notes playing together would take longer to compute and it is hard to specify a limited range of notes to represent the whole situation. The conditions for the wave files and the range of the notes are the same as the previous section.

The starting point was located using two different methods. In this section we did not use a starting point defined in terms of the frequency of the wave (e.g. 15 periods of wavelength). This would not be valid in the real world although it was used during the design phase of the development of the methodology.

Method A

We wrote a simple program to find the appropriate starting point on the graph of the wave. We then did a series of FT's. The first FT started at the point chosen by the program just mentioned. Each following FT started 1000 samples later i.e. 1st at original starting point, 2nd at original starting point plus 1000 samples, 3rd at original starting point plus 2000 samples, etc. The series continued for 11 FTs. To find the representative stable state we examined the FTs on the screen and saw that for the earlier starting points there were often extra peaks at the beginning of the FTs or the "shape" of the FT may be variable. On some occasions only one of these "signs of instability" was seen. The starting point chosen was the earliest stage when these "the signs of instability" had subsided.

These signs of instability are illustrated in the following two figures although the printouts do not show them as clearly as they appeared on the screen.

Figure 3.1 shows a single example of this process. It shows the original wave, the series of 11 FTs and the selected solution of FT. In this example the starting point was just before the seventh FT.

Figure 3.2 shows a close up of the earlier parts of the FTs in Figure 3.1.

Method B

The previous method could be open to criticism as it required some human intervention although it was a good method of testing the system of identification of real notes. Unfortunately the time available for this thesis only allowed initial testing of an automatic method for the identification of the starting point.

We used the same computer program used for method A to locate the approximate position of the onset time of the wave. Then we performed an experiment similar to method A. The steps followed are now outlined:

1. We do an FT using the approximate starting point identified by the onset time program. We then perform another FT starting 1000 samples from the original starting point. We continue with another starting point 1000 samples later and progressively do FT's until we have done 11 FTs (which is same range as for method A).

2. We perform subtraction on each FT against its next FT to find the absolute difference. For example, if the first FT is y0, the second FT is y1, ... and the 11 FT is y10, then the first subtraction is absolute(y1-y0), the second subtraction is absolute(y2-y1), ... and the 10th subtraction is absolute(y10-y9).

3. We then find the maximum difference and the sum of all the differences in each subtraction.

   i.e.

   maximum (absolute(y1-y0)) and sum (absolute(y1-y0))
   maximum (absolute(y2-y1)) and sum (absolute(y2-y1))
   ...
   maximum (absolute(y10-y9)) and sum (absolute(y10-y9))

4. We divided each result with the maximum height of each relative FT.

   i.e.

   maximum(absolute(y1-y0)) / maximum(absolute(y1) or absolute(y0))
   and sum(absolute(y1-y0)) / maximum(absolute(y1) or absolute(y0))
   maximum(absolute(y2-y1)) / maximum(absolute(y2) or absolute(y1))
   and sum(absolute(y2-y1)) / maximum(absolute(y2) or absolute(y1))
   ...
   maximum(absolute(y10-y9)) / maximum(absolute(y10) or absolute(y9))
   and sum(absolute(y10-y9)) / maximum(absolute(y10) or absolute(y9))

5. Finally we choose the position where the lowest of the maximum difference and the lowest of the sum of all the differences is located as the position of the starting point.

   i.e. Location(minimum (
   maximum(absolute(y1-y0)) /
   maximum(absolute(y1) or absolute(y0)),
   maximum(absolute(y2-y1)) /
   maximum(absolute(y2) or absolute(y1)),
   ...
   maximum(absolute(y10-y9)) /
   maximum(absolute(y10) or absolute(y9))
   ))

   and

```
Location (minimum(
        sum (absolute(y1-y0)) /
                maximum(absolute(y1) or absolute(y0)),
        sum (absolute(y2-y1)) /
                maximum(absolute(y2) or absolute(y1)),
        ...
        sum (absolute(y10-y9)) /
                maximum(absolute(y10) or absolute(y9))
))
```

We graphed the maximum difference in the absolute values when performing the subtraction using different starting points (with the same step size of 1000 samples) for the note C2 (Figure 3.4) and the sum of the differences in each FT subtraction (Figure 3.5).

The purpose of this experiment is to set the starting point at the point of the lowest changes (hopefully that is the stable position) on the wave. As we also know the amplitude of the wave will decrease as time passes, we have used the maximum height from the selected FT as the basis, so we can recognise the changes of FT as a percentage of the maximum height of the selected FT.

It is probable that the sudden change from reducing differences to increasing differences in absolute value does not capture all the "signs of instability". Therefore this method may not be as successful as method A. We leave that work to others as we were first checking whether the method would be even partially successful.

The results of this method are in sections 3.6.

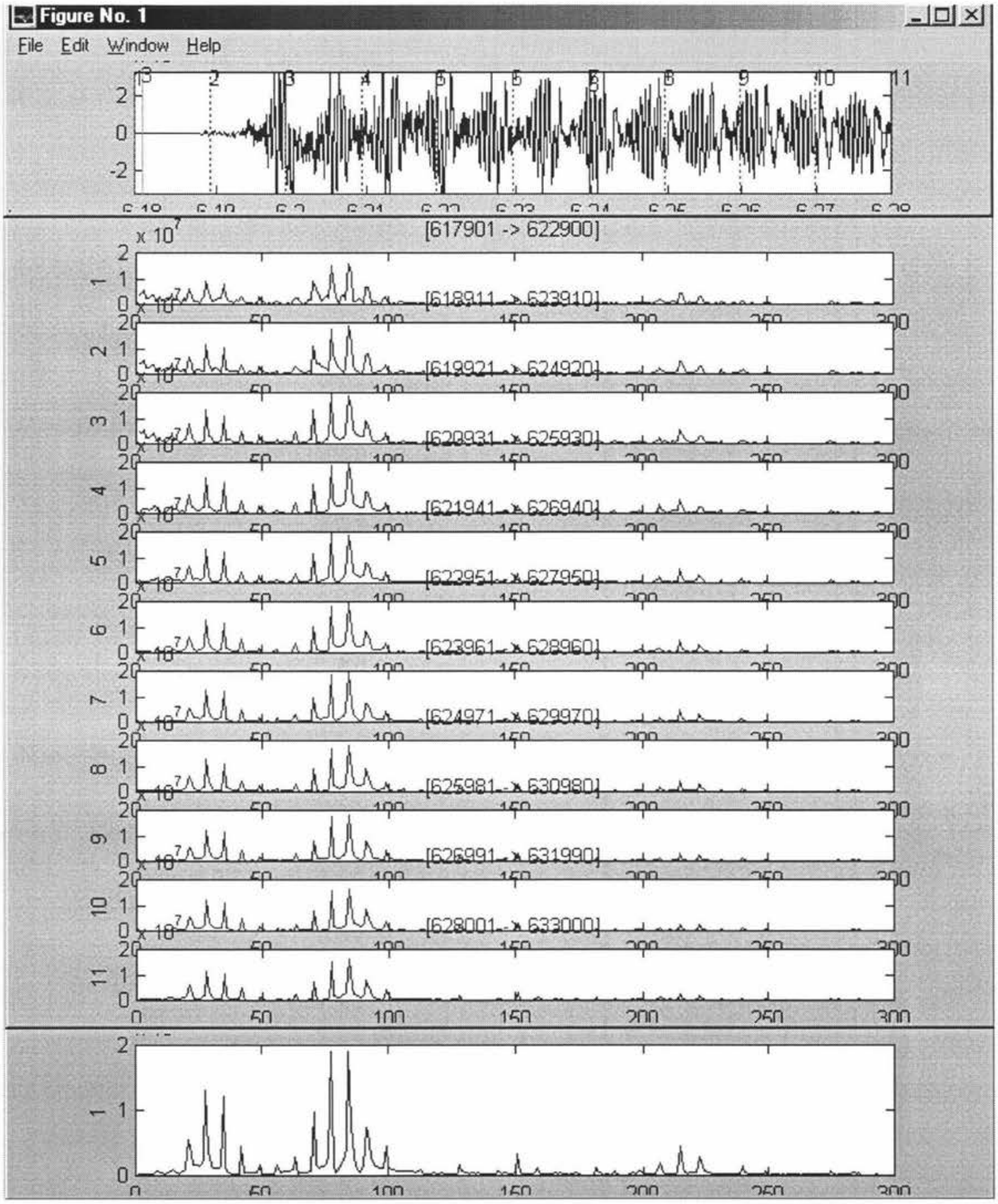

Figure 3.1 An example of the process of detecting the starting point when playing with real instruments.

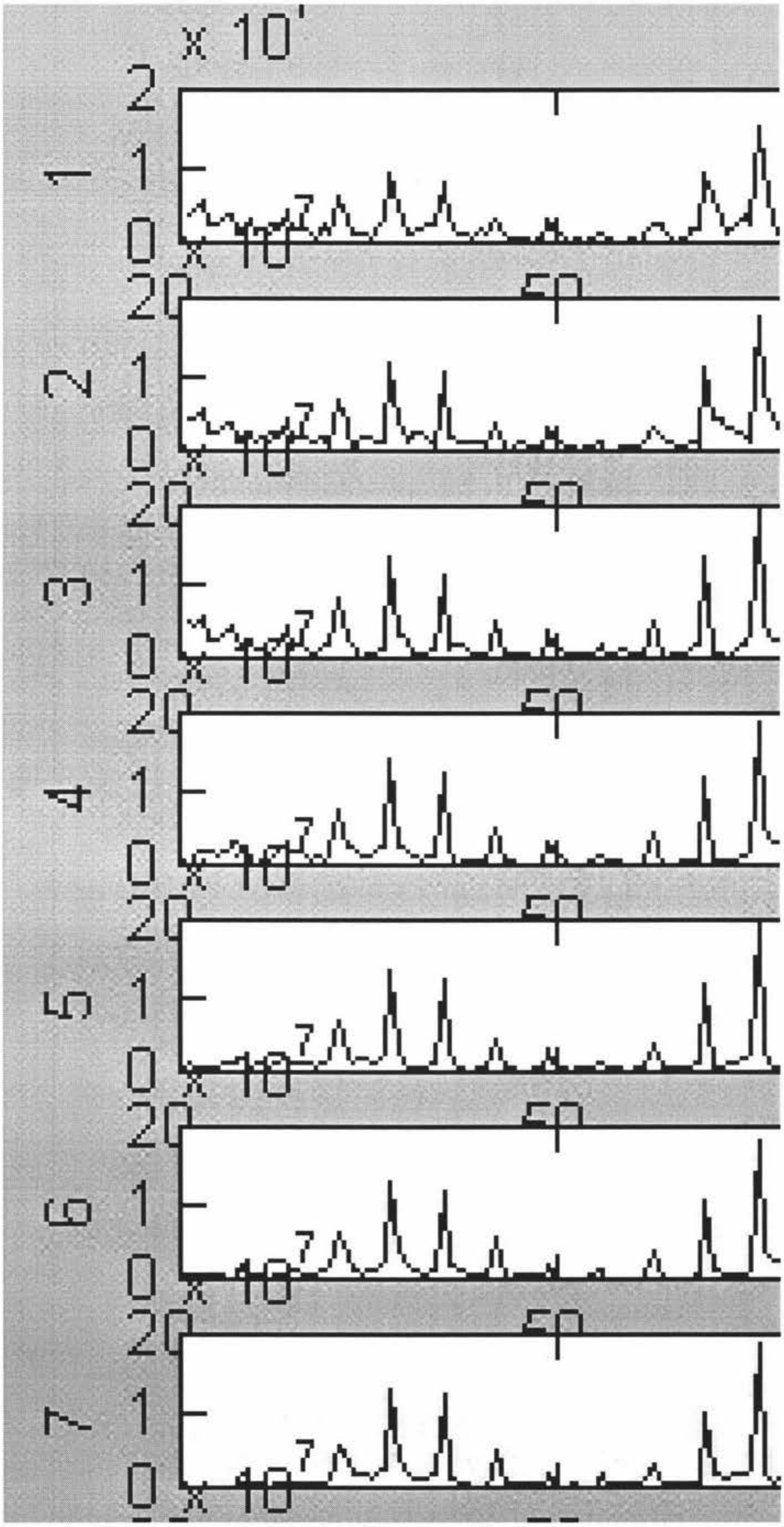

Figure 3.2 A close up of the earlier parts of the FTs in Figure 3.1.

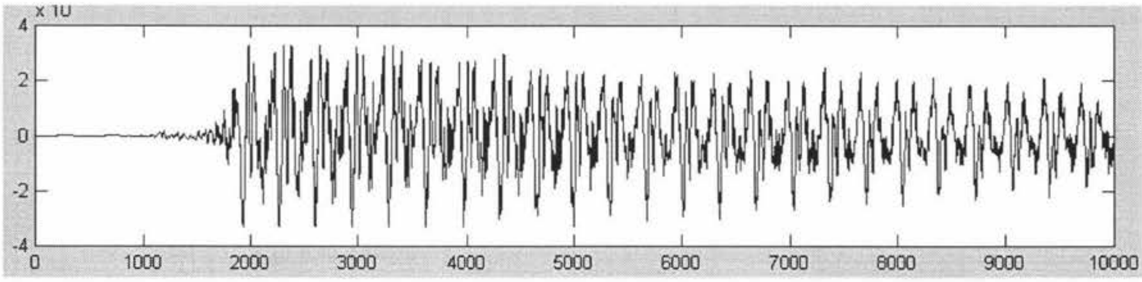

Figure 3.3 The sound wave of C2.

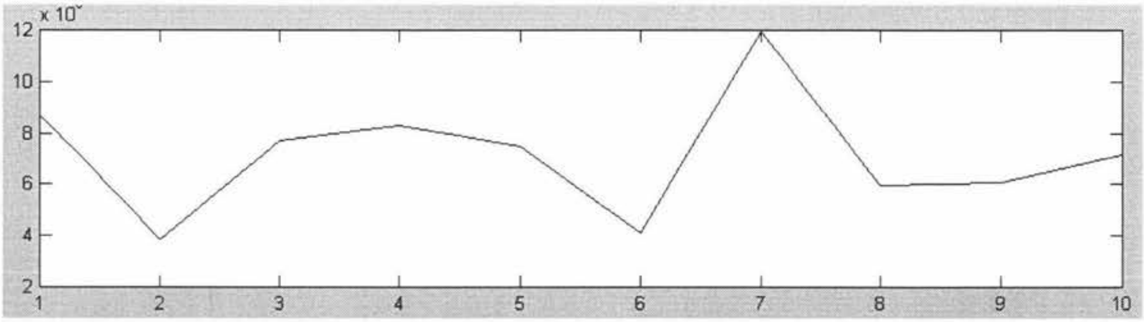

Figure 3.4 The maximum difference in each FT

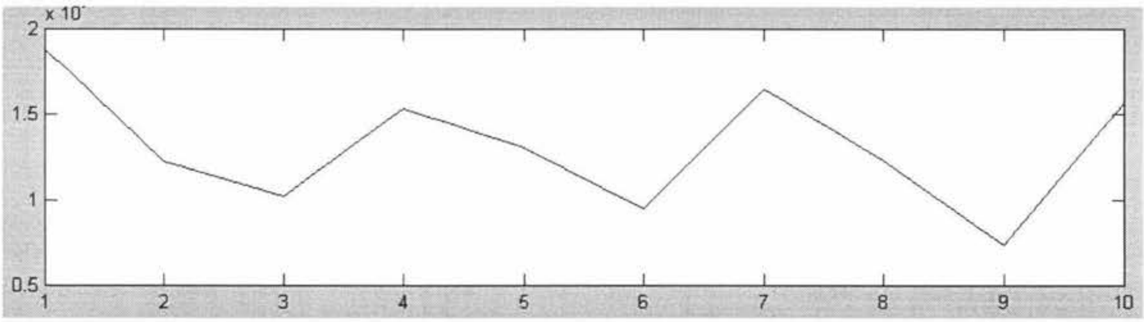

Figure 3.5 The sum of difference in each FT

Unfortunately the time available for this thesis precluded the solution to the problem of automatic identification of the starting point. This is not surprising when we remember the difficulties Goto (2002) [13] and his fellow workers have in selecting onset times when studying beat tracking.

3.1 Application of Strategy 5 on a Single Note

The starting point is 250 samples or 500 samples after the starting point of the known note. "250 samples" is about 0.74 periods of the wavelength using C2, and is up to 22.40 periods of wavelength when using B6. 500 samples is double that. The method of calculation (filtered FT or complete FT; and ignored sign method or absolute value method) is not relevant because there is no subtraction.

	Correct Results	No. Of Cases	Percentage
Original starting point plus 250 samples	60	60	100.00%
Original starting point plus 500 samples	59	60	98.33%

Table 3.1 Summary of the accuracy rate when playing single notes.

Table 3.1 shows the accuracy rate is very high and we can conclude that we will get a high accuracy rate if we correctly chose the starting point. This accuracy is in spite of the fact that the starting point is very hard to choose in some cases. For high frequency notes the difficulty is because the duration is very short, and the FT results can be very unstable. In other cases, vibration of the wave may also cause problems.

3.2 Application of Strategy 5 on Two Notes Played Together

We performed three different experiments using the different methods of subtraction (IF, IC, AF and AC).

I) We called the first experiment "skip one". The experiment did not include any "sharp" notes. The first note was chosen from between C2 to B6. The second note of the pair was two notes after the first. e.g. "C2 and E2", "D2 and F2", "E2 and G2", and so on until "G6 and B6". (33 cases for each method). This showed the average accuracy rate when using different notes.

II) We called the second experiment "single octave" which was all the combinations within the same octave from octave 2 through octave 6. (66*5 =

330 cases for each method). This shows the accuracy rate when performing within each octave.

III) We called the third experiment "different octave". We performed the experiment on the same note from different chosen octaves. These experiments cover "octave 2 & octave 3", "octave 2 & octave 4", "octave 3 & octave 4", "octave 3 & octave 5", "octave 4 & octave 5", "octave 4 & octave 6", and "octave 5 & octave 6". (12*7 = 84 cases for each method). This shows the accuracy rate when notes are played with one of its harmonic notes, one or two octaves higher.

3.2.1 Experimented results

Table 3.2 summarises the results of the experiments.

	No. Of Cases	Percentage				
		Accuracy of the First Solution	Accuracy of the Second Solution			
			IF	AF	IC	AC
Skip One	33	93.94%	63.64%	78.79%	63.64%	75.76%
Single Octave (oct. 2)	66	90.91%	81.82%	86.36%	81.82%	86.36%
Single Octave (oct. 3)	66	87.88%	65.15%	63.64%	65.15%	74.24%
Single Octave (oct. 4)	66	96.97%	80.30%	92.42%	83.33%	87.88%
Single Octave (oct. 5)	66	93.94%	74.24%	81.82%	71.21%	78.79%
Single Octave (oct. 6)	66	75.76%	51.52%	66.67%	63.64%	69.70%
Different Octave	84	96.43%	52.38%	50.00%	58.33%	59.52%

Table 3.2 Summary of the accuracy rate in different experiments when two notes are played together.

We made similar observations, from Table 3.2, to those we saw with MIDI, but there is some variation.

i) In most situations the accuracy rate finding the first solution was very high.

ii) In most situations the accuracy rate using the absolute value method was higher than when using the ignored sign method.

iii) In half of the experiments the accuracy rate using the filtered FT results was higher than when using the complete FT results. The rest showed the opposite. Therefore we cannot conclude which method is better.

We highlighted the results for the more accurate absolute value method.

For experiment "skip one", the accuracy rate for the first solution, when playing two notes together, using a real instrument, was 93.94%, which was very similar to the results on the MIDI instrument. The accuracy rate for the second solution was either 75.76% or 78.79%, which in both cases was twice the failure rate of the MIDI experiment.

The experiment, "single octave", gave similar results to "skip one". It showed there were some octaves (octaves 2, 4 and 5) where the accuracy rate was much higher than the others.

The experiment, "different octave", had a very low accuracy rate for the second solution. This was related to the harmonic nature of the paired notes created by the piano because when we removed the first solution some of the characteristics of the second solution might also be removed.

3.2.2 Estimating the number of solutions

Calculation Method	Before Subtraction	After First Subtraction	After Second Subtractions
IF	3.12E+07	2.02E+07	1.29E+07
AF	3.12E+07	1.76E+07	6.49E+06
IW	3.12E+07	2.06E+07	1.35E+07
AW	3.12E+07	1.82E+07	7.61E+06

Table 3.3 Summary of the averages of the maximum height of FT using different calculation method in the three experiments of previous section.

The table shows the average of the final remnant is about $1.3 * 10^7$ when using the ignored sign method, and $7 * 10^6$ when using the absolute value method. Although these two remnants are lower than the averages amplitude of the known note (2.52 *

10^7), the remnants after the first subtraction are already lower than that average. Possible reasons are, the volume level of the music, "incorrect" choice of the starting point, or an erroneous result. The lowest maximum peak of the known note is about $1.96*10^6$, which is very close to the final remnant, when using the absolute value method. If we continue subtracting FT results until the remnant is lower than "$1.96 * 10^6$", we may only gain one or two more solutions.

Further observations from this experiment on a real instrument are the same as for the MIDI experiments.

3.3 Application on Single Note Using MIDI as the Known Note

In the previous section the known note came from the real instrument. We used two sets of known notes from MIDI to identify the solutions for a single note. The first set of the known note is from the acoustic grand piano (MIDI 1). The second set of the known note is from 5 different instruments (MIDI 1, 26, 33, 41, 75). Table 3.4 shows the summary of this experiment.

	No. Of Cases	Accuracy rate of using different sets of known notes		
		Real Grand Piano	Single MIDI Instrument	5 Different MIDI Instruments
Original Starting Point plus 250 samples	60	100.00%	71.67%	85.00%
Original Starting Point plus 500 samples	60	98.33%	71.67%	85.00%

Table 3.4 Summary of the accuracy rate when using different sets of known notes including results from section 3.1.

The accuracy rate of the solution using MIDI instruments as the known note is worse than the real instrument (72% and 85% compared to 100% with one starting point and 98% with the other).

3.4 Application on Two Notes Using MIDI as the Known Note

We performed the same experiment as in the previous section, with two notes played together.

3.4.1 Known notes from a single instrument

	No. Of Cases	Percentage				
		Accuracy of the First Solution	Accuracy of the Second Solution			
			IF	AF	IC	AC
Skip One	33	66.67%	39.39%	36.36%	39.39%	39.39%
Single Octave (oct. 2)	66	83.33%	62.12%	62.12%	59.09%	56.06%
Single Octave (oct. 3)	66	75.76%	36.36%	36.36%	36.36%	36.36%
Single Octave (oct. 4)	66	74.24%	43.94%	43.94%	42.42%	42.42%
Single Octave (oct. 5)	66	68.18%	45.45%	46.97%	50.00%	50.00%
Single Octave (oct. 6)	66	45.45%	22.73%	27.27%	19.70%	21.21%
Different Octaves	84	83.33%	42.86%	44.05%	44.05%	44.05%

Table 3.5 Summary of the accuracy rate using single MIDI instrument (AGP) as the known note.

The table shows the accuracy rate is very low. The range of the accuracy rate for the first solution is between 45% and 83% and the second solution is between 22% and 62%. There are many reasons for such a low accuracy rate. e.g. the real piano may be tuned differently from the MIDI piano, or the volume levels of the known note and the unknown note are different. This causes the subtraction results to be very erratic.

3.4.2 Known notes from 5 different instruments

	No. Of Cases	Percentage				
		Accuracy of the First Solution	Accuracy of the Second Solution			
			IF	AF	IC	AC
Skip One	33	84.85%	57.58%	63.64%	66.67%	66.67%
Single Octave (oct. 2)	66	93.94%	75.76%	69.70%	77.27%	75.76%
Single Octave (oct. 3)	66	83.33%	71.21%	66.67%	77.27%	72.73%
Single Octave (oct. 4)	66	86.36%	75.76%	74.24%	74.24%	72.73%
Single Octave (oct. 5)	66	80.30%	75.76%	78.79%	74.24%	74.24%
Single Octave (oct. 6)	66	42.42%	36.36%	37.88%	33.33%	36.36%
Different Octaves	84	91.67%	85.71%	83.33%	84.52%	84.52%

Table 3.6 Summary of accuracy rate of using 5 different instruments as the known note.

The table shows some of the accuracy rates are still very low (especially in octave 6). However the accuracy rates have increased compared to using a single instrument in MIDI as the source of the known note. The range for the first solution is about 42%~94% and the second solution is about 34%~86%. The reason for getting a higher accuracy rate could be that one, or some, instruments are coincidentally better "matched" (similar FT results) when compared to our real instrument. The results of this experiment are much closer than those from the single MIDI instrument, to those when using the same real instrument, as the known note. This begs the question: Are a larger collection of synthetic MIDI instruments more likely to contain the match of a real instrument than the specific MIDI instrument with the same name? Whatever the reason, the results are unexpected.

These last two experiments confirm most of the observations made when using the real instrument as the known note. The other observed difference is that using the ignored sign method is often more accurate than using the absolute value method.

3.5 Practical Experiment

We recorded a short piece of music called "The Little Brown Jug" on a real grand piano. We then chose the starting points manually as described earlier, and compared the FT results of the known note from the real grand piano to the MIDI instruments (as in the previous sections). Although the time interval between each note is quite long, some of the notes may still overlap the previous note.

	No. Of Cases	Accuracy Rate of using different sets of known notes		
		Real Grand Piano	Single MIDI Instrument	5 Different MIDI Instruments
Recognitions on "Little Brown"	30	96.67%	63.33%	93.33%

Table 3.7 Summary of the accuracy rate playing a simple tune.

In spite of this potential problem, Table 3.7 shows the accuracy rate using the real instrument is very high (29/30 ~ 96.67%) and so is the rate using 5 different MIDI instruments as the known note (28/30 ~ 93.33%). More details can be found in Table C.1 in Appendix C. Using a single MIDI instrument (MIDI 1) as the source of the known note still shows a very low accuracy rate. This mirrors the observations we found earlier and further study of the reason for this, along the lines suggested, may well be a very fruitful area for the future.

3.6 Results of Method B

The experiment is similar to that in section 3.2 (on page 105) except the method of selection of starting point is different.

Because we wanted this initial test to be reasonably difficult we did not repeat a similar experiment to section 3.1 (on page 105). As section 3.1 is only testing the methods when one note is played we moved directly to repeating section 3.2 when two notes are played together.

The experiment for MIDI (as sections 3.3 and 3.4 on pages 108 and 109) would not help as MIDI is an artificial construct.

Experiment Results:

	No. Of Cases	Percentage				
		Accuracy of the First Solution	Accuracy of the Second Solution			
			IF	AF	IC	AC
Skip One	33	93.94%	63.64%	78.79%	63.64%	75.76%
Single Octave (oct. 2)	66	90.91%	81.82%	86.36%	81.82%	86.36%
Single Octave (oct. 3)	66	87.88%	65.15%	63.64%	65.15%	74.24%
Single Octave (oct. 4)	66	96.97%	80.30%	92.42%	83.33%	87.88%
Single Octave (oct. 5)	66	93.94%	74.24%	81.82%	71.21%	78.79%
Single Octave (oct. 6)	66	75.76%	51.52%	66.67%	63.64%	69.70%
Different Octave	84	96.43%	52.38%	50.00%	58.33%	59.52%

Table 3.8 The accuracy rate when playing two notes using Method A (same as Table 3.2)

	No. Of Cases	Percentage				
		Accuracy of the First Solution	Accuracy of the Second Solution			
			IF	AF	IC	AC
Skip One	33	78.79%	63.64%	72.73%	57.58%	69.70%
Single Octave (oct. 2)	66	78.79%	62.12%	78.79%	59.09%	71.21%
Single Octave (oct. 3)	66	87.88%	50.00%	56.06%	51.52%	59.09%
Single Octave (oct. 4)	66	86.36%	68.18%	74.24%	65.15%	77.27%
Single Octave (oct. 5)	66	84.85%	78.79%	83.33%	83.33%	84.85%
Single Octave (oct. 6)	66	84.85%	56.06%	51.52%	53.03%	56.06%
Different Octave	84	86.90%	54.76%	55.95%	52.38%	55.95%

Table 3.9 The accuracy rate when playing two notes using the lowest of the maximum difference between the previous FT and the current FT as the starting point.

	No. Of Cases	Percentage				
		Accuracy of the First Solution	Accuracy of the Second Solution			
			IF	AF	IC	AC
Skip One	33	15.15%	0.00%	6.06%	6.06%	6.06%
Single Octave (oct. 2)	66	12.12%	19.70%	7.57%	22.73%	15.15%
Single Octave (oct. 3)	66	0.00%	15.15%	7.58%	13.63%	15.15%
Single Octave (oct. 4)	66	10.61%	12.12%	18.18%	18.18%	10.61%
Single Octave (oct. 5)	66	9.09%	-4.55%	-1.51%	-12.12%	-6.06%
Single Octave (oct. 6)	66	-9.09%	-4.54%	15.15%	10.61%	13.64%
Different Octave	84	9.53%	-2.38%	-5.95%	5.95%	3.57%
Average All						
	447	6.77%	5.07%	6.73%	9.29%	8.30%

Table 3.10 The difference between the results in Tables 3.8 and 3.9.

	No. Of Cases	Percentage				
		Accuracy of the First Solution	Accuracy of the Second Solution			
			IF	AF	IC	AC
Skip One	33	78.79%	60.61%	69.70%	54.55%	57.58%
Single Octave (oct. 2)	66	68.18%	69.70%	68.18%	74.24%	74.24%
Single Octave (oct. 3)	66	84.85%	50.00%	63.64%	51.52%	65.15%
Single Octave (oct. 4)	66	87.88%	66.67%	72.73%	62.12%	74.24%
Single Octave (oct. 5)	66	83.33%	69.70%	83.33%	75.76%	83.33%
Single Octave (oct. 6)	66	78.79%	54.55%	51.52%	48.48%	53.03%
Different Octave	84	85.71%	52.38%	50.00%	48.81%	54.76%

Table 3.11 The accuracy rate when playing two notes using the lowest of the average difference between the previous FT and the current FT as the starting point.

	No. Of Cases	Percentage				
		Accuracy of the First Solution	Accuracy of the Second Solution			
			IF	AF	IC	AC
Skip One	33	15.15%	3.03%	9.09%	9.09%	18.18%
Single Octave (oct. 2)	66	22.73%	12.12%	18.18%	7.58%	12.12%
Single Octave (oct. 3)	66	3.03%	15.15%	0.00%	13.63%	9.09%
Single Octave (oct. 4)	66	9.09%	13.63%	19.69%	21.21%	13.64%
Single Octave (oct. 5)	66	10.61%	4.54%	-1.51%	-4.55%	-4.54%
Single Octave (oct. 6)	66	-3.03%	-3.03%	15.15%	15.16%	16.67%
Different Octave	84	10.72%	0.00%	0.00%	9.52%	4.76%
Average All						
	447	9.76%	6.49%	8.66%	10.23%	9.99%

Table 3.12 The difference between the results in Tables 3.8 and 3.11.

Difference using Different Method to Locate the Starting Point	No. Of Cases	Percentage						No. Of Cases Diff.
		Accuracy of the First Solution	Accuracy of the Second Solution				Average	
			IF	AF	IC	AC		
Using maximum difference in each FT	447	6.77%	5.07%	6.73%	9.29%	8.30%	7.23%	32.33
Using average difference in each FT	447	9.76%	6.49%	8.66%	10.23%	9.99%	9.03%	40.34

Table 3.13 A summary of the differences in the accuracy rate between method A and the two versions of method B.

Conclusion for Method B

Table 3.13 is the summary of this experiment. In both cases Method A has a higher accuracy rate (about 7%~9%) than Method B.

Using the lowest average difference in FT has a lower accuracy rate (about 2%) than using the lowest maximum difference in FT. Because of the small number of test cases, and the lack of testing on other single notes or two note groups, as well as on other instruments (e.g. on guitar or on recorder), the results are only indicative.

More experiments will need to be done on this field by a different researcher. Goto (2002) [13] is a good source of information about similar problems with beat tracking.

Chapter 4
Conclusion

The task we undertook is much more difficult than expected, however we were moderately successful. We give our conclusions with the successes and the difficulties.

These first conclusions are the results of studies in chapter 2, using MIDI instruments. Our efforts were based on the use of five strategies. The first three strategies we used, to find the note or notes played, were only partially successful. The first strategy used the highest peak of the Fourier Transform (FT) as the identifying feature. The feature of the second strategy was the number of peaks above 10% of the highest peak, and the feature of the third strategy was the number of matching peaks between the known and unknown note. The fourth strategy attempted to identify the shape of the FT but was unsuccessful. The fifth strategy which was the most successful strategy was a combination of strategy 2 and 3 (more fully described in section 2.13 on page 77). Because strategies 1 to 4 are less useful, we have concentrated on the development of strategy 5.

The method of strategy 5 to locate the solutions for the unknown note, has a higher accuracy rate than simply choosing the highest peaks (which was the method used for our initial attempt). When using strategy 5, the accuracy rate of solutions when a single note is played, and the first solution when several notes are played, is much higher than the accuracy rate of the other solutions for multi-notes.

The further development of strategy 5 involves comparing the "simple", the "ignored sign", and the "absolute value" methods of subtraction. The "absolute value" method, with the complete FT results, helped us improve the accuracy rate to an average of 87% when two notes are played and 76% when three notes are played, where the known notes are from 5 different instruments. The rate was 96% for two notes and 93% for three notes respectively when the known notes are from the same instrument.

Other experiments using filtered FT results lead to similar conclusions for the accuracy rate, but the time spent on processing is less. Therefore, if fast processing and memory savings are needed, filtered FT results can be used for subtraction.

These experiments are within the acoustic range from C2 to B6. Notes lower than C2 give less accurate solutions because some notes are missed during the FT analysis. Notes higher than B6 are less accurate because of the duration of the note.

The major sources of error are the choice of starting point and the method of subtraction. We know there are three stages for each note we play (attack, steady-state and decay), therefore the FT results are different if the starting point is different. It is important that the FT is within the steady-state, if at all possible, especially when multi-notes are played. We have also shown that the method of subtraction using absolute value is better than using other methods.

Other sources of error could be related to:

i) The sample size (which will affect the results of FT) (see section 2.2 on page 28).

ii) The number of instruments used for the known note.

iii) The source instrument, or instruments, of the unknown notes.

iv) The frequency of the unknown note (some octaves have a higher accuracy rate than others, which may also relate to the choice of starting point) (see section 2.17 on page 92).

v) When multi-notes are played on an acoustic instrument where one of the notes may be a harmonic term of the other (see section 2.17 on page 92).

We can estimate the number of notes in the unknown note more accurately if we progressively subtract the FT results for each note, found to be a part of the unknown note, until we reach a remnant that is too small to contain a note. We also see that using "absolute value" for subtraction leads to the remnant for the FT results being lower than other methods. This will give us a more accurate estimate of the number of notes.

In chapter 3, using a real instrument, we arrive at similar conclusions to those from chapter 2. The accuracy rate for a real instrument is lower and we were less able to find, correctly, the number of notes played. The subtraction methods showed less variation in their accuracy rates.

We have found an unexpected result in chapter 3 where using 5 different MIDI instruments as the source of the known note was more accurate than using the MIDI equivalent of the real instrument (Acoustic Grand Piano, MIDI 1). This may mean that it would be useful to repeat these experiments using a larger group of MIDI instruments as the source of the known note. It may also be useful to vary the instruments within those groups according to the type of music (e.g. classic, country, and rock'n'roll) being played.

Sources of error from real instruments include those already listed plus two more:

i) Notes may overlap the previous or following notes.

ii) The different volume levels of notes.

We tested two methods of determining the starting point in the real instrument. The first had human interaction and the second used a computer program. These results show that with further study, especially on the automatic identification of the starting point, automatic transcription of music to printed notation using a standard PC, may become a reality. The use of midi as a known concept to act as a standard against which we can identify unknown notes has been successfully begun. Using the phrases as defined by Tanguiane (1993) [36], this work shows that this artificial intelligence approach may be simpler than the more common artificial perception approach.

References

1. Bagshaw P.C., Hiller S.M. and Jack M.A. (1993) Enhanced Pitch Tracking and the Processing of F0 Contours for Computer Aided Intonation Teaching, in Proc. Eurospeech '93, Berlin, pp. 1003-1006.

2. Bainbridge D. http://www.cs.waikato.ac.nz/~davidb/, 2002.

3. Benson D. (2002) Mathematics and Music, http://www.mat.uga.edu/~djb/index.html, 2002.

4. Blackburn S. and DeRoure D. (1998) A Tool for Content Based Navigation of Music, in Proc. ACM Multimedia.

5. Bourke P., http://astronomy.swin.edu.au/~pbourke/analysis/dft, 1993.

6. Cooley J.W. and Tukey J.W. (1965) An algorithm for the machine calculation of complex Fourier series, Math. of Computation 19 (1965), 297-301.

7. CSTR Publications, http://www.cstr.ed.ac.uk/publications/, 2001.

8. De Cheveigne, A. (1993) Separation of concurrent harmonic sounds: Fundamental frequency estimation and a time-domain cancellation model of auditory processing, JASA 93(6), 3271-3290.

9. Embree P.M. and Danieli D. (1999) C++ Algorithms for Digital Signal Processing (2nd ed.). Upper Saddler River, NJ : Prentice Hall.

10. Fletcher N.H. and Rossing T.D. (1998) The Physics of Musical Instruments (2nd Ed.), Spriner-Verlag, ISBN: 0-387-98374-0.

11. Ghias A. and Logan D. (1995) Query By Humming – Musical Information Retrieval in an Audio Database, in Proc. ACM Multimedia.

12. Gold B. and Rabiner L.R. (1969) Parallel Processing Techniques of Estimating Pitch Periods of Speech in the Time Domain, K. Acoust. Soc. Am., 46, No. 2, 442-449, Aug 1969.

13. Goto M., http://staff.aist.go.jp/m.goto/publications.html, 2002.

14. Goto M. and Muraoka Y. (1997). Issues in Evaluating Beat Tracking Systems, Working Notes of the IJCAI-97 Workshop on Issues in AI and Music – Evaluation and Assessment, pp.9-16, August 1997.

15. Goto M. and Muraoka Y. (1998). An Audio-based Real-time Beat tracking System and its Applications. Proceedings ICMC98.

16. Hamer M. (Dec 2001). Pitch Perfect. New Scientist.

17. Hudspeth A.J. (1985) The cellular basis of hearing: The biophysics of hair cells. Science 230:745.

18. Hudspeth A.J. (1989) How the ear's works work. Nature 341:397.

19. Kinsler, L.E., Frey A.R., Coppens, A.B., and Sanders J.V. (1982) Fundamentals of Acoustics, 3 ed., Wiley, New York.

20. Klapuri A. (1998) Automatic Transcription of Music. Masterate thesis, Tampere University of Technology.

21. Klapuri A. http://www.cs.tut.fi/~klap/iiro/literature.html, 2002.

22. Klapuri, Virtanen and Holm (2000) Robust multipitch estimation for the analysis and manipulation of polyphonic musical signals". In Proc. COST-G6 Conference on Digital Audio Effects, DAFx-00, Verona, Italy.

23. Lynn P.A. and Fuerst W. (1994). Introductory digital signal processing with computer applications. Chichester, New York : Wiley.

24. McAdams S. and Bigand S. (1993) Thinking in Sound. Clarendon Press, ISBN: 0-19-852257-6.

25. McNab R.J., Smith L.A., Bainbridge D. and Witten I.H. (1997) The New Zealand Digital Library MELody inDEX, D-Lib Magazine, May 1997, ISSN: 1082-9873.

26. MIDI Manufacturers Association Incorporated (2001) Tutorial on MIDI and Music Synthesis http://www.midi.org/about-midi/tutorial/tutor.htm.

27. MIT Media Lab, http://sound.media.mit.edu/papers.html, 2002.

28. mf2t [Computer program] http://www.hitsquad.com/smm/programs/mf2t, 1997.

29. Neuropsychology Central http://www.neuropsychologycentral.com/, 2001.

30. Oxford Talking Dictionary (1998) The Learning Company, Inc.

31. Rabiner L.R. and Gold B. (1975) Theory and Applications of Digital Signal Processing. Englewood Cliffs, New Jersey: Prentice-Hall.

32. Roberto V. (1993) Intelligent Perceptual Systems. New Directions in Computational Perception, Springer-Verlag, ISBN: 3-540-57379-8.

33. Roederer J.G. (1995). The physics and psychophysics of music. An introduction, Springer-Verlag, ISBN: 0-387-94366-8.

34. Rossing T.D., Wheeler P. and Moore R. (2002) The Science of Sound (3rd Ed.), Addision Wesley. ISBN: 0-8053-8565-7.

35. Suggs W. (1966). Meet the Orchestra, The Macmillan Company, New York. Library of Congress catalogue No.66-11110.

36. Tanguiane A.S. (1993). Artificial Perception and Music Recognition. Springer-Verlag. ISBN: 3-540-57394-1.

37. TiMidity [Computer program] http://www.goice.co.jp/member/mo/timidity, 2001.

38. White P. (1999). basic MIDI. Sanctuary Publishing.

Appendix A

Format of wave file:

Endian	File Offset (bytes)	Field Size (bytes)	Field
Big	0	4	**Chunk ID**
Little	4	4	**Chunk Size**
Big	8	4	**Format**
Big	12	4	**SubChunk ID**
Little	16	4	**SubChunk Size**
Little	20	4	**Audio Format**
Little	22	2	**Number of Channels**
Little	24	2	**Sample Rate**
Little	28	4	**Byte Rate**
Little	32	2	**Block Align**
Little	34	2	**Bits Per Sample**
Big	36	4	**SubChunk ID**
Little	40	4	**SubChunk Size**
Little	44	SubChunk Size	**Sample Data**

The "RIFF" chunk descriptor

The Format of concern here is "WAVE", which requires two sub-chunks; "fmt"and "data".

The "fmt" sub-chunk

Describes the format of the sound information in the data sub-chunk.

The "data" sub-chunk

Indicates the size of the sound information and contains the new sound data.

Example of wave file (raw data):

52 49 46 46	D4 04 00 00	57 41 56 45	66 6D 74 20
10 00 00 00	01 00 01 00	11 2B 00 00	11 2B 00 00
01 00 08 00	64 61 74 61	B0 04 00 00	7D 7E 7A 7B

Translation:

The "RIFF" chunk descripter

52 49 46 46	can be translated into "RIFF" using ASCII (standard for *.wav*).
D4 04 00 00	Chunk size = 4 D4 (Hex) = 1236 (Dec).
57 41 56 45	can be translated into "WAVE" using ASCII (standard for *.wav*).

The "fmt" sub-chunk

66 6D 74 20	can be translated into "fmt " using ASCII.
10 00 00 00	Sub-Chunk size is 16 bytes.
01 00 01 00	
01 00	Audio format is 1 (PCM).
01 00	Number of channels is 1 (mono).
11 2B 00 00	Sample rate is 11025 samples per second.
11 2B 00 00	Byte rate is 11025 bytes per second.
01 00 08 00	
01 00	BlockAlign is 1 byte.
08 00	Bits per sample = 8 bits per Sample.
64 61 74 61	can be translated into "data" using ASCII.
B4 04 00 00	Sub-Chunk size = 4B0 (Hex) = 1200 (Dec).

The "data" sub-chunk

7D 7E 7A 7B	
7A	the first sample = 7D (Hex) = 125 (Dec) => -3 (Dec)
7E	the second sample = 7E (Hex) = 126 (Dec) => -2 (Dec)
7A	the third sample = 7A (Hex) = 122 (Dec) => -6 (Dec)
7B	the forth sample = 7B (Hex) = 123 (Dec) => -5 (Dec)

Appendix B

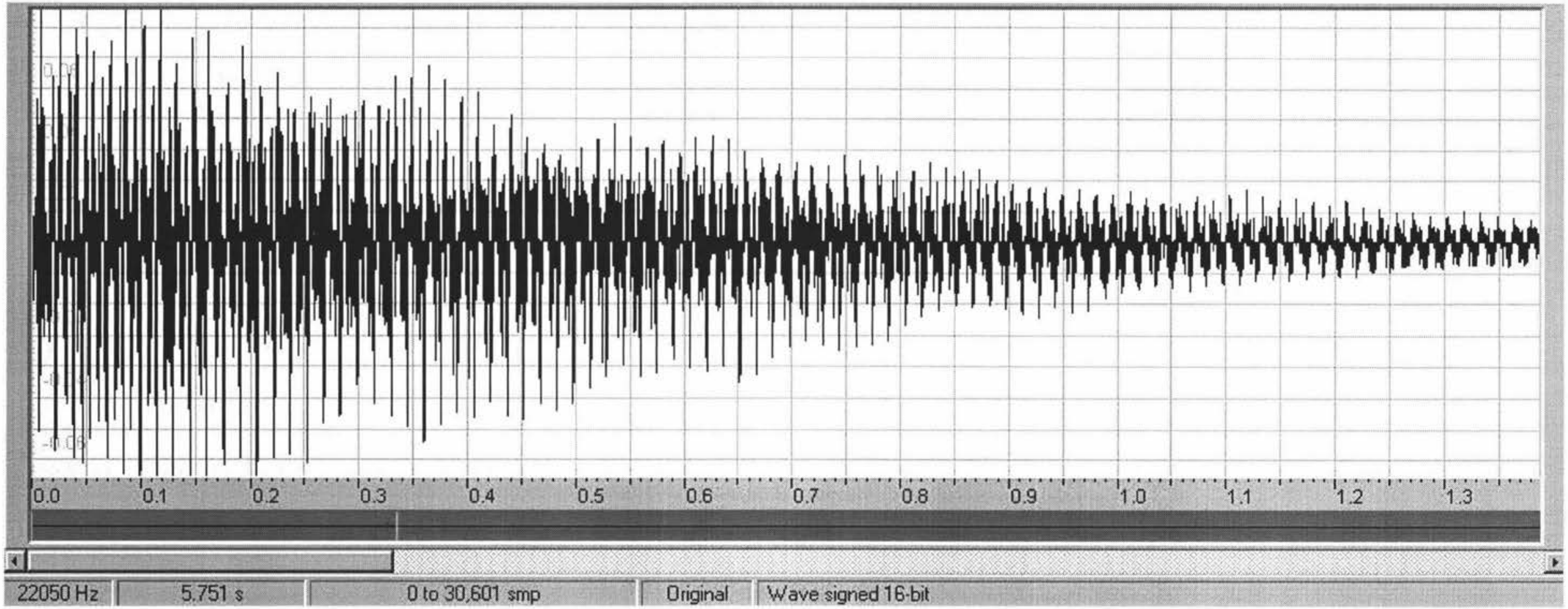

Figure B.1 The sample of an acoustic grand piano when playing the note C2 (65.41 Hz).

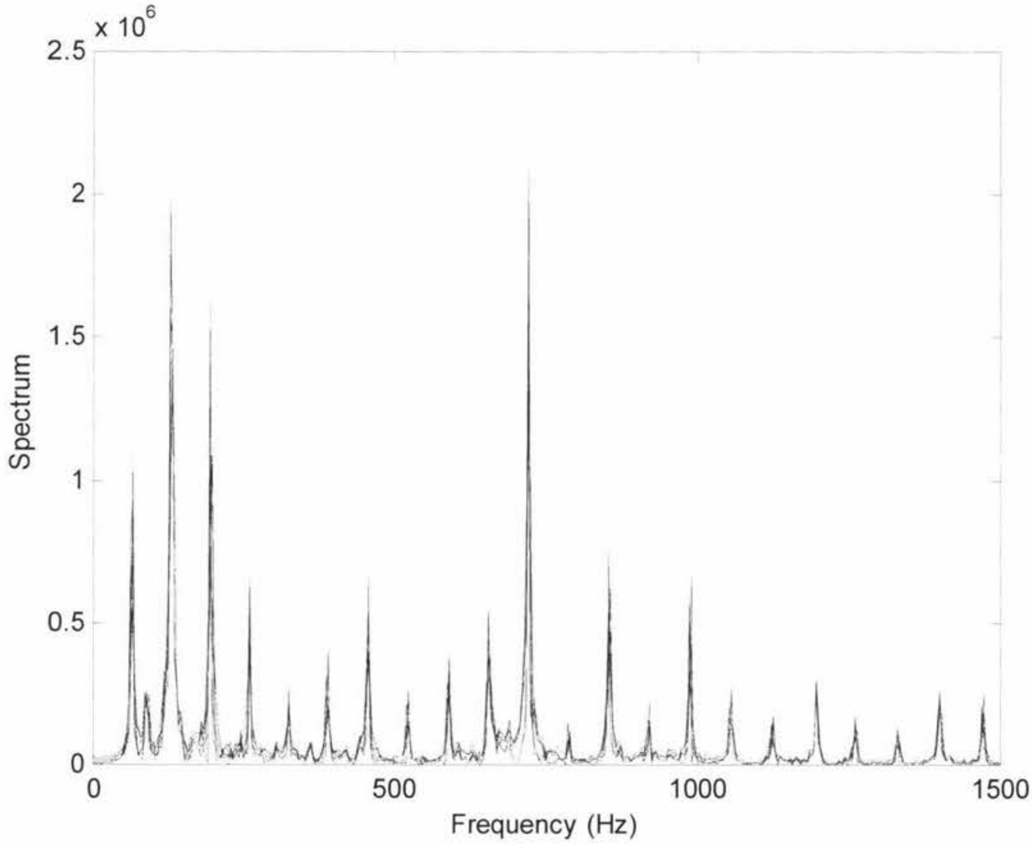

Figure B.2 The results of Fourier Transform in 2D when using different sample sizes on an acoustic grand piano when playing the note C2 (65.41 Hz).

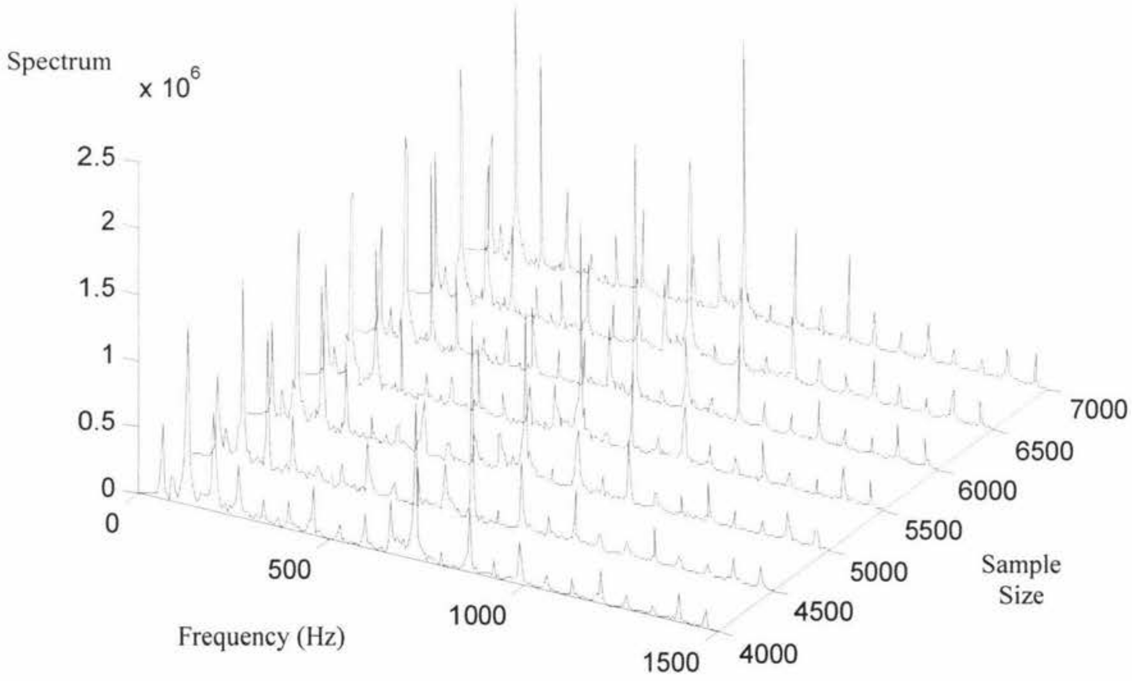

Figure B.3 The results of Fourier Transform in 3D when using different sample sizes on acoustic grand piano when playing the note C2 (65.41 Hz).

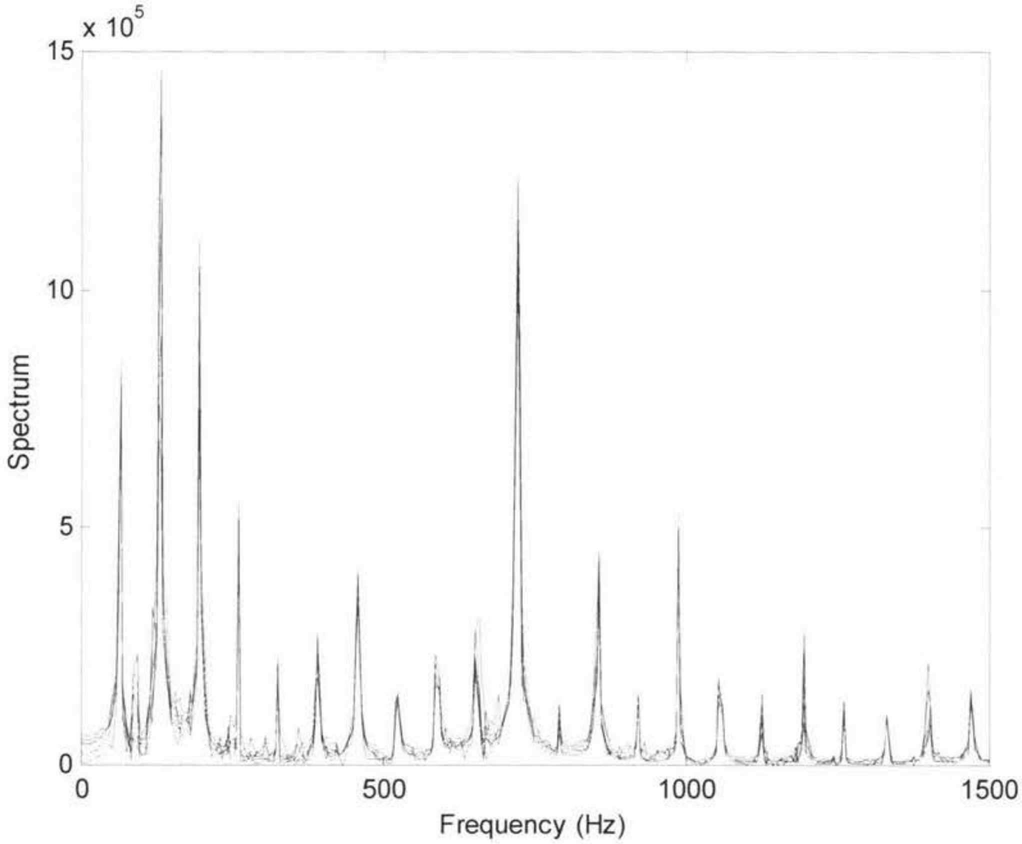

Figure B.4 The results of Fourier Transform in 2D when using different starting points on an acoustic grand piano when playing the note C2 (65.41 Hz).

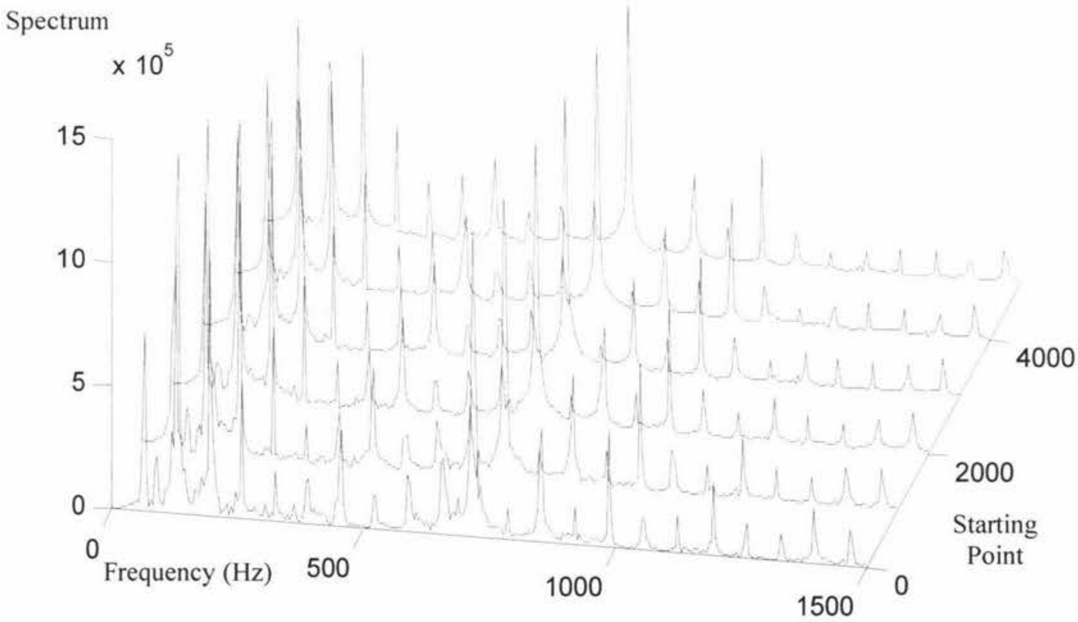

Figure B.5 The results of Fourier Transform in 3D when using different starting points on an acoustic grand piano when playing the note C2 (65.41 Hz).

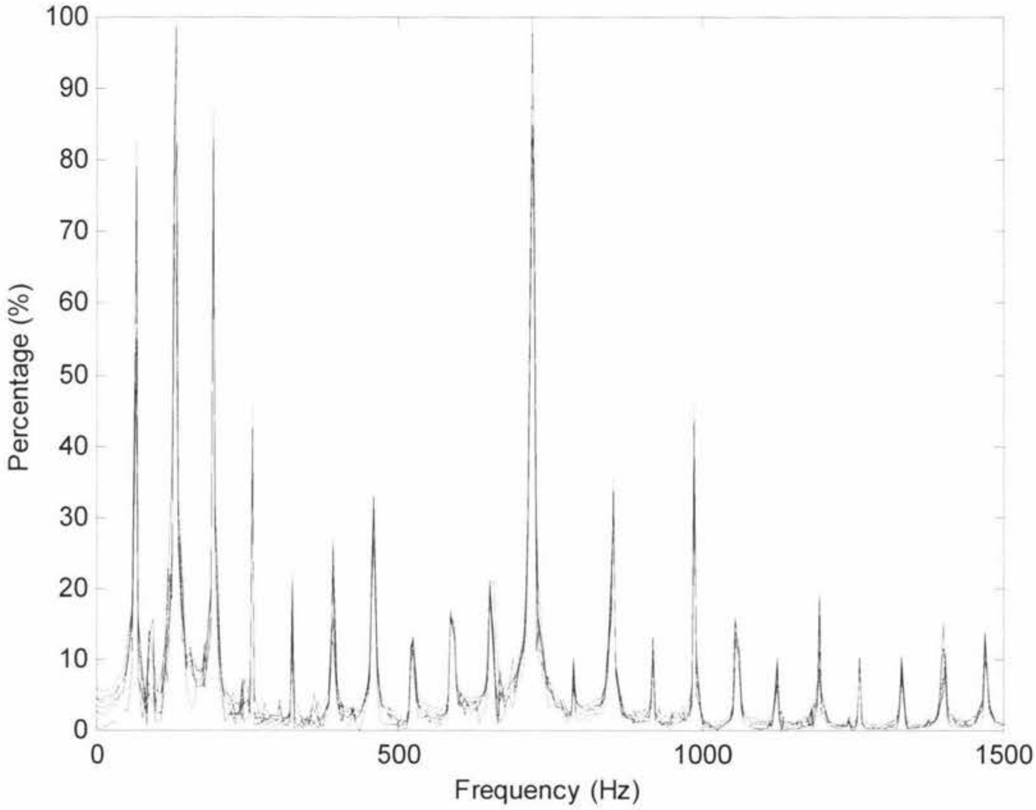

Figure B.6 The results of Fourier Transform in 2D when using different starting points on an acoustic grand piano when playing the note C2 (65.41 Hz) (in percentages).

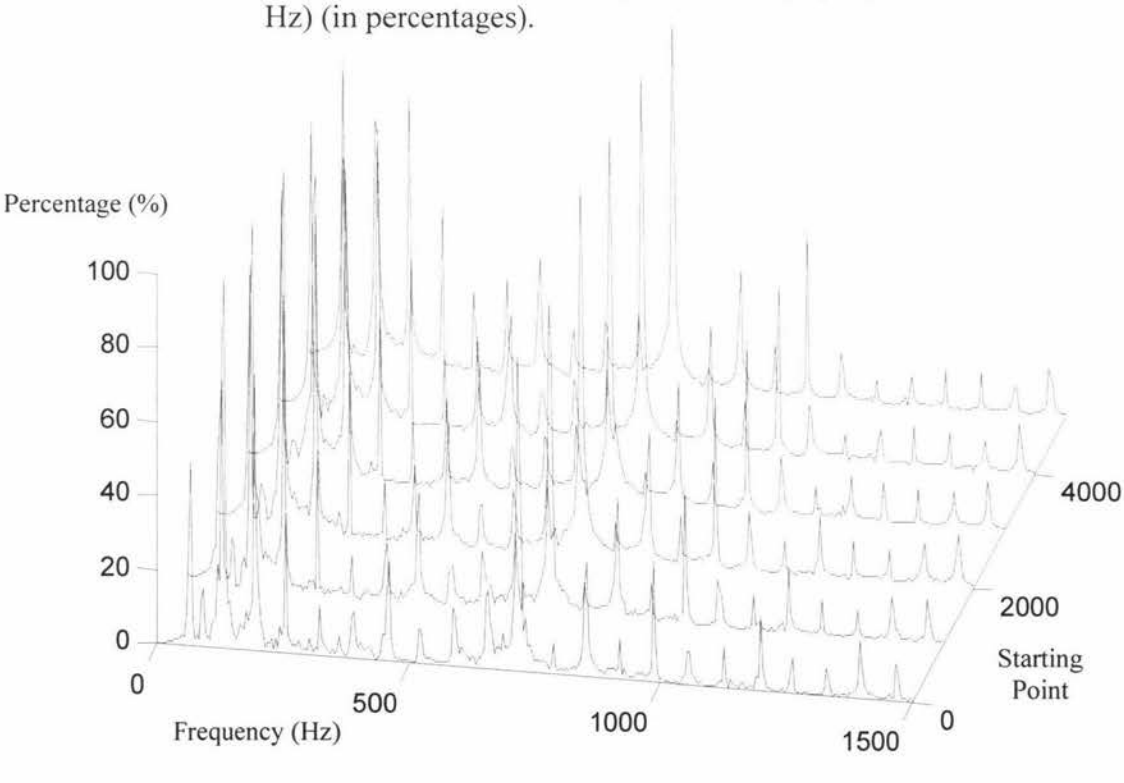

Figure B.7 The results of Fourier Transform in 3D when using different starting points on an acoustic grand piano when playing the note C2 (65.41 Hz) (in percentages).

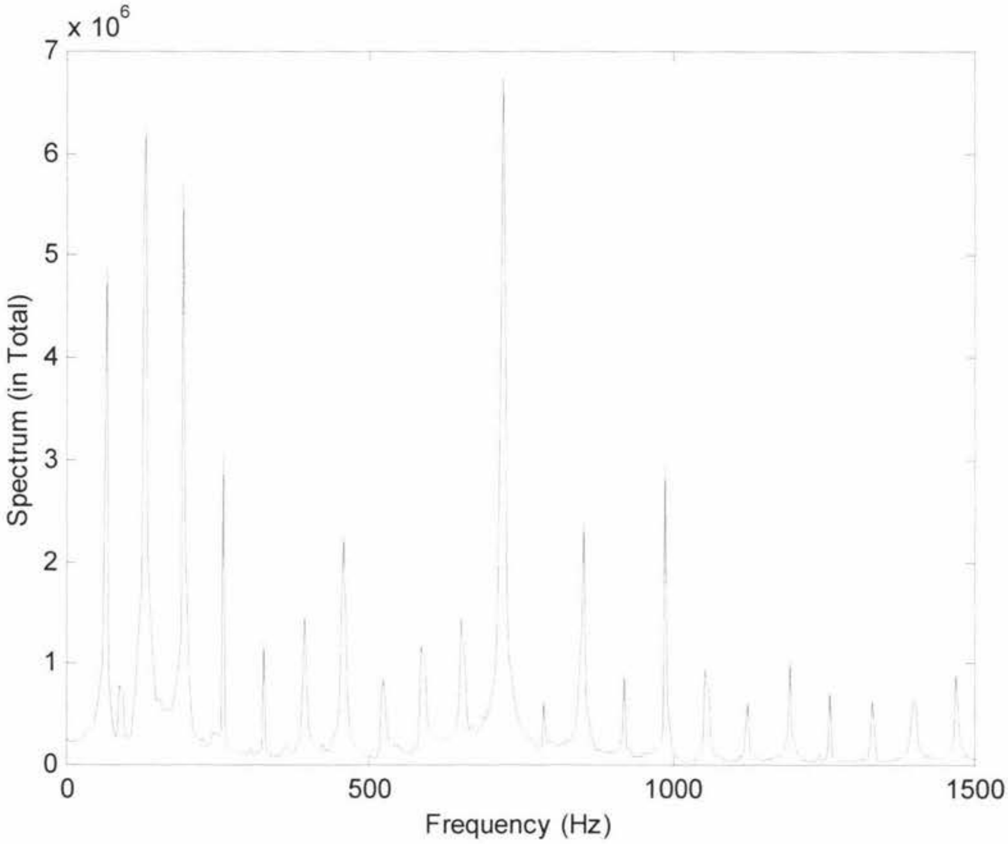

Figure B.8 The results of Fourier Transform when using different starting points on an acoustic grand piano when playing the note C2 (65.41 Hz) after totalling all the spectra.

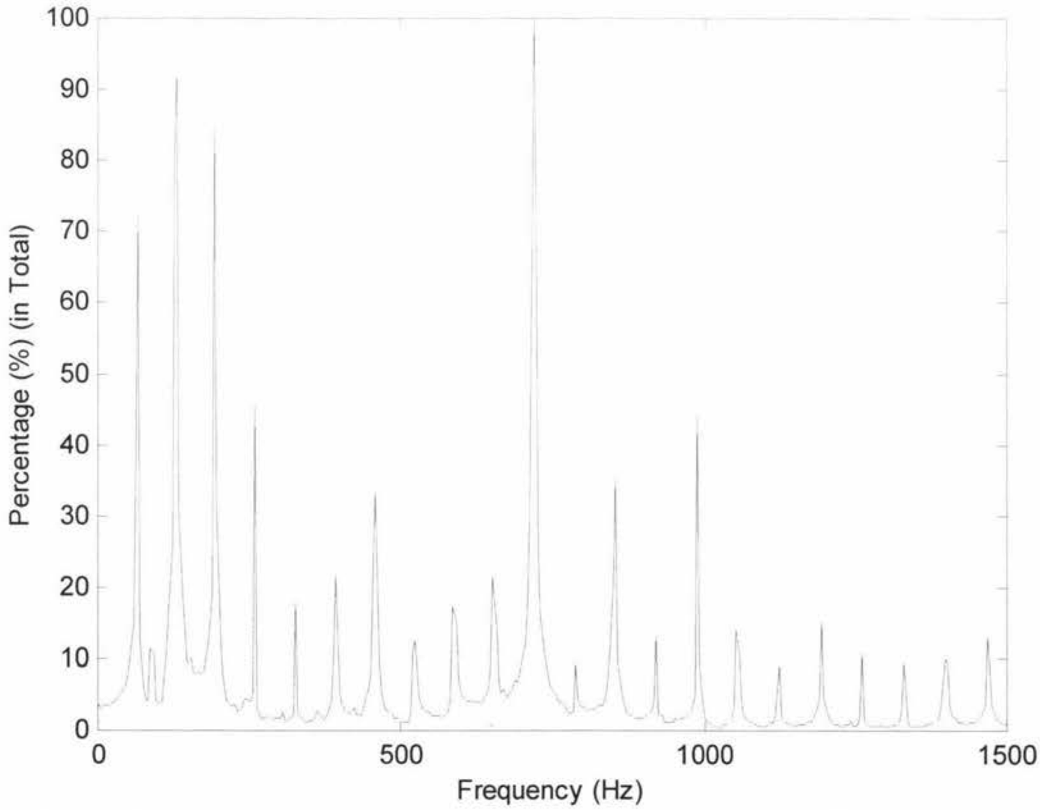

Figure B.9 The results of Fourier Transform when using different starting points on an acoustic grand piano when playing the note C2 (65.41 Hz) in percentages after totalling all the spectra.

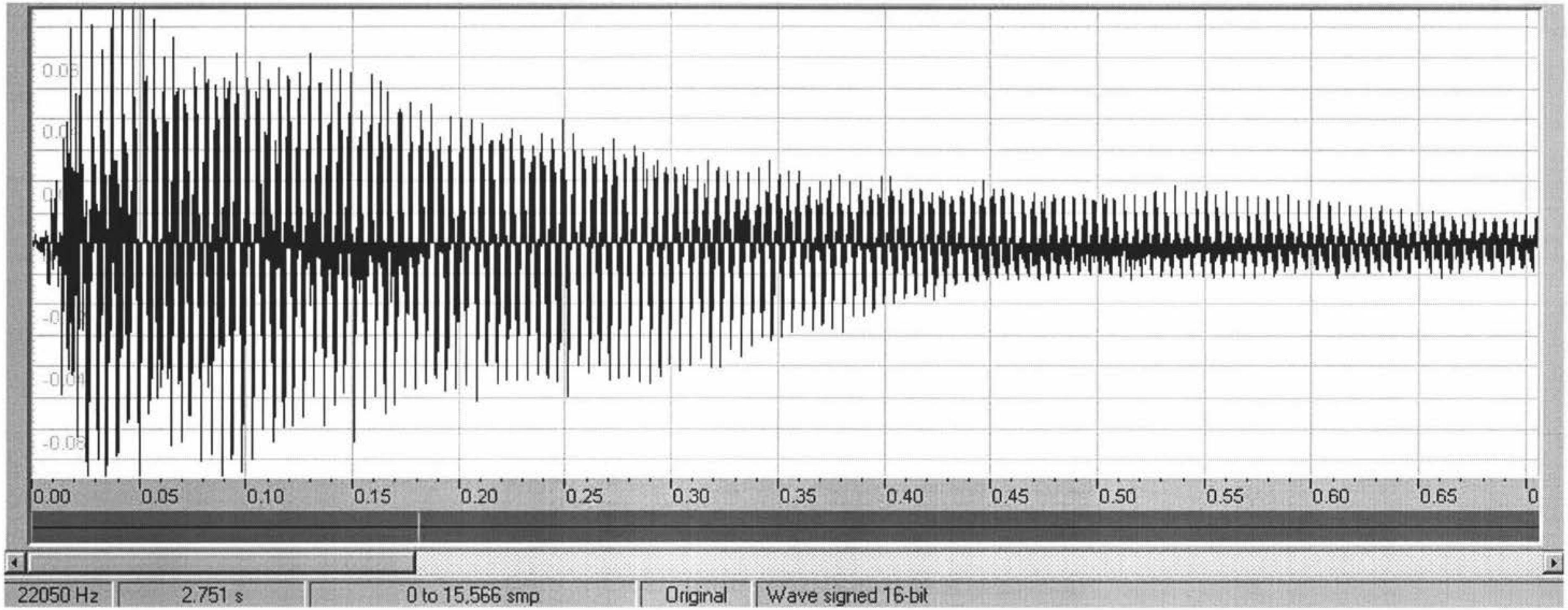

Figure B.10 The sample of a bright acoustic piano when playing the note G#3 (207.65 Hz).

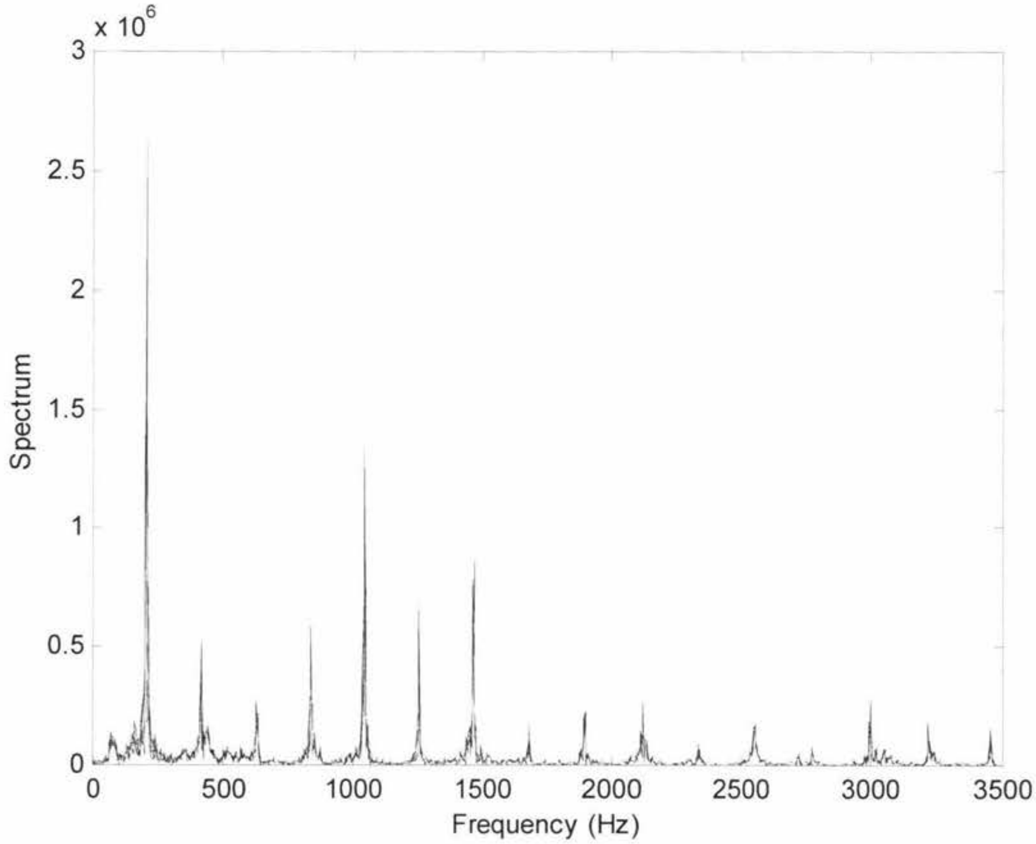

Figure B.11 The results of Fourier Transform in 2D when using different sample sizes on a bright acoustic piano when playing the note G#3 (207.65 Hz).

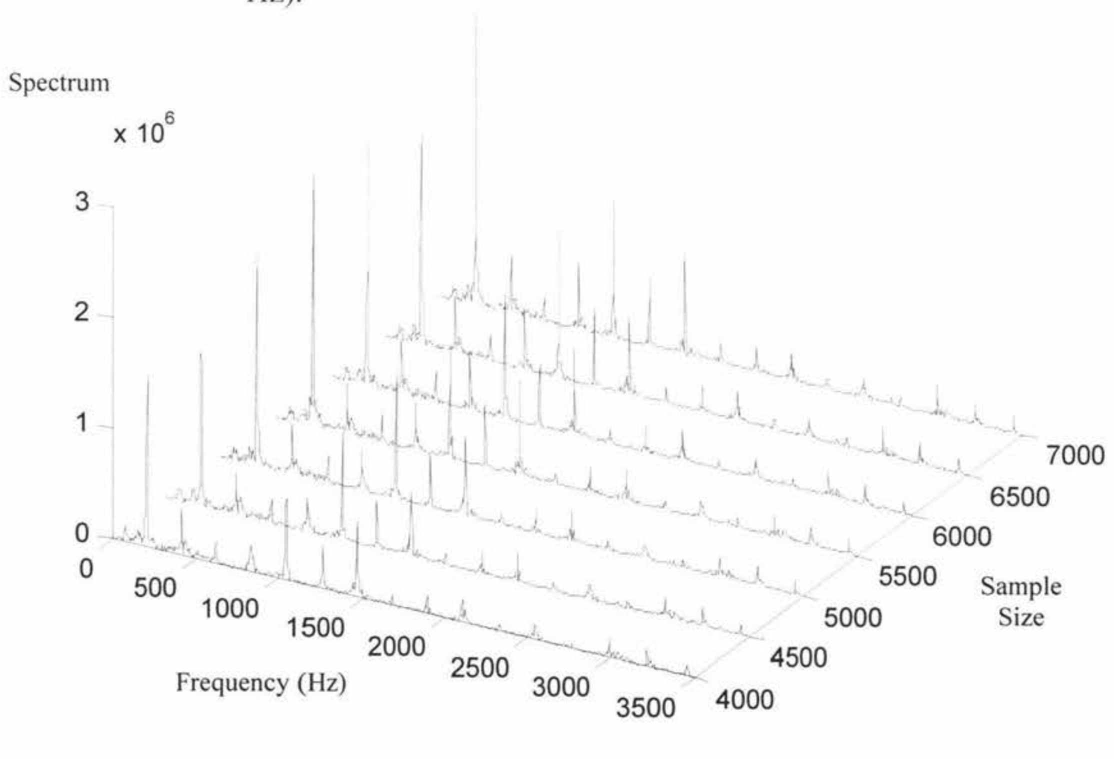

Figure B.12 The results of Fourier Transform in 3D when using different sample sizes on a bright acoustic piano when playing the note G#3 (207.65 Hz).

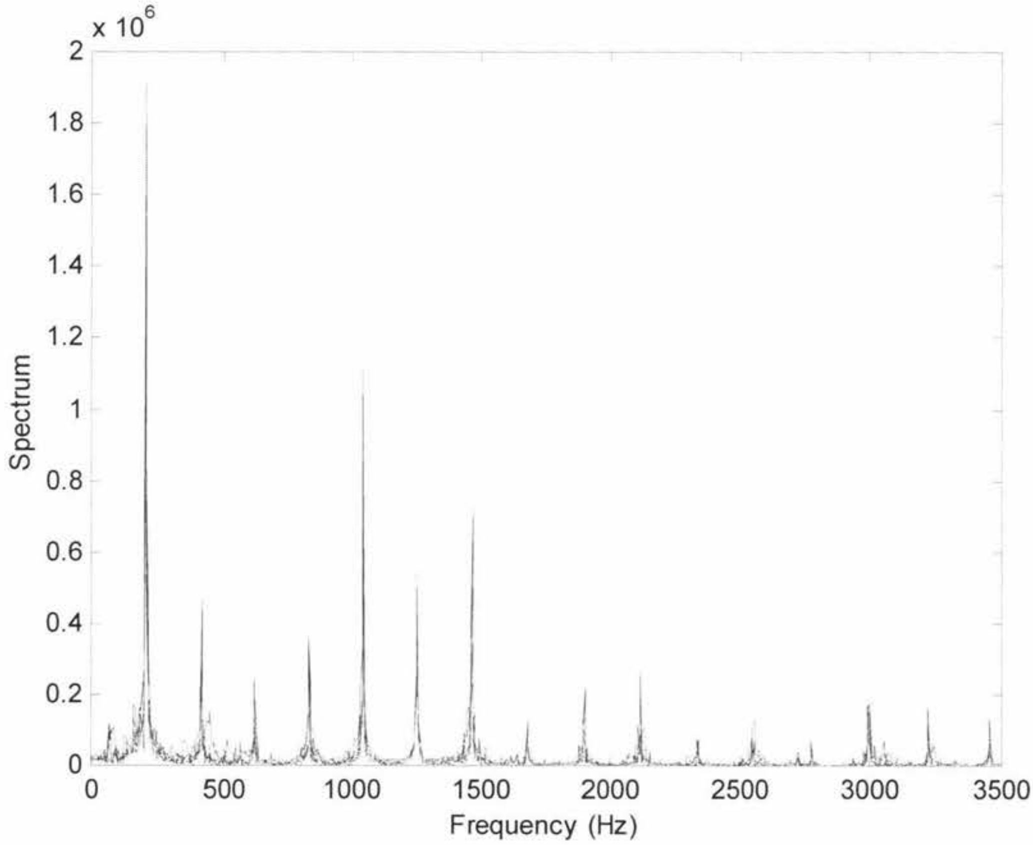

Figure B.13 The results of Fourier Transform in 2D when using different starting points on a bright acoustic piano when playing the note G#3 (207.65 Hz).

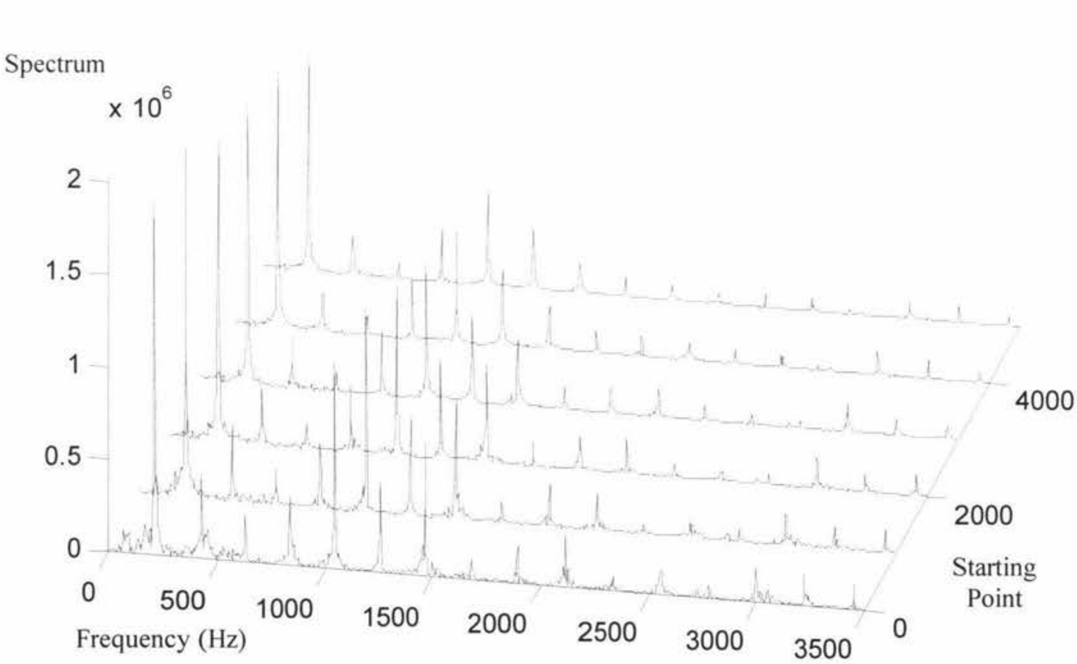

Figure B.14 The results of Fourier Transform in 3D when using different starting points on a bright acoustic piano when playing the note G#3 (207.65 Hz).

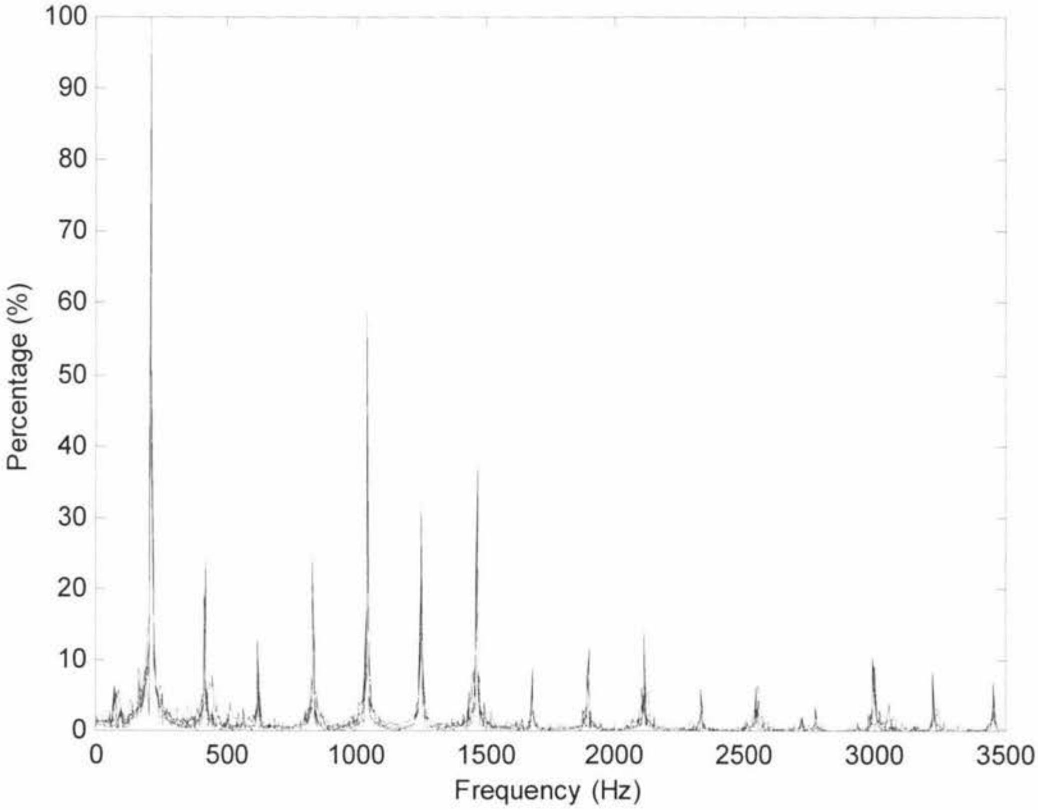

Figure B.15 The results of Fourier Transform in 2D when using different starting points on a bright acoustic piano when playing the note G#3 (207.65 Hz) (in percentages).

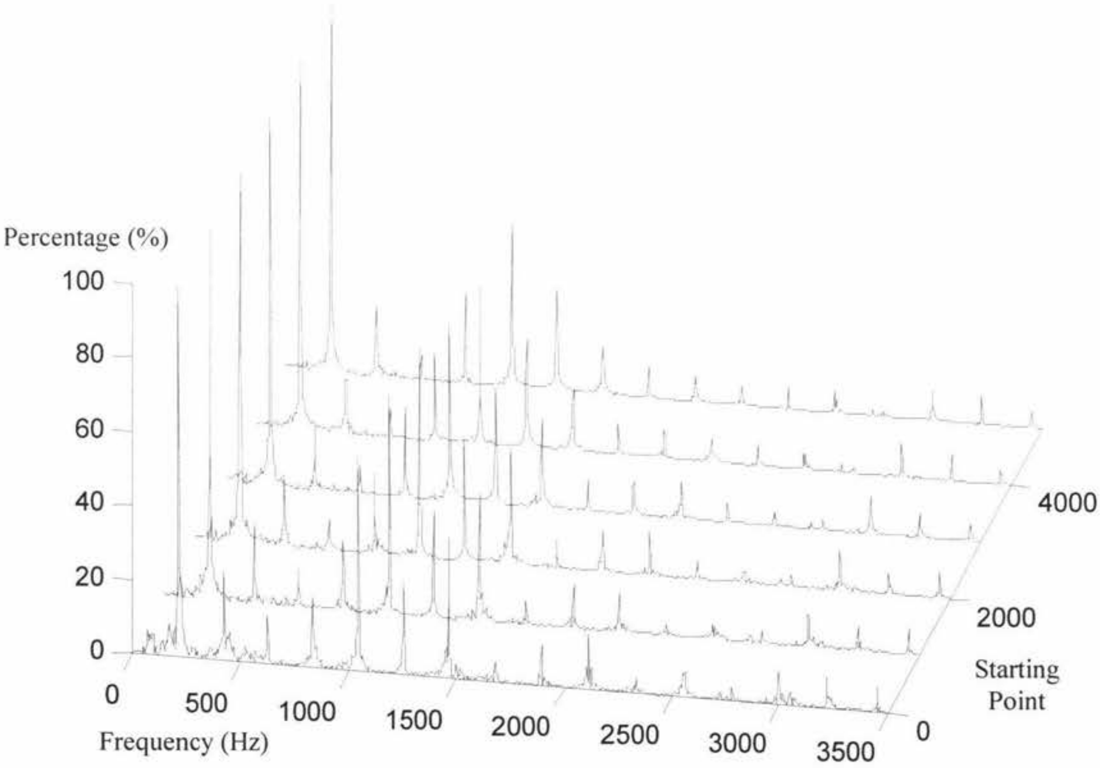

Figure B.16 The results of Fourier Transform in 3D when using different starting points on a bright acoustic piano when playing the note G#3 (207.65 Hz) (in percentages).

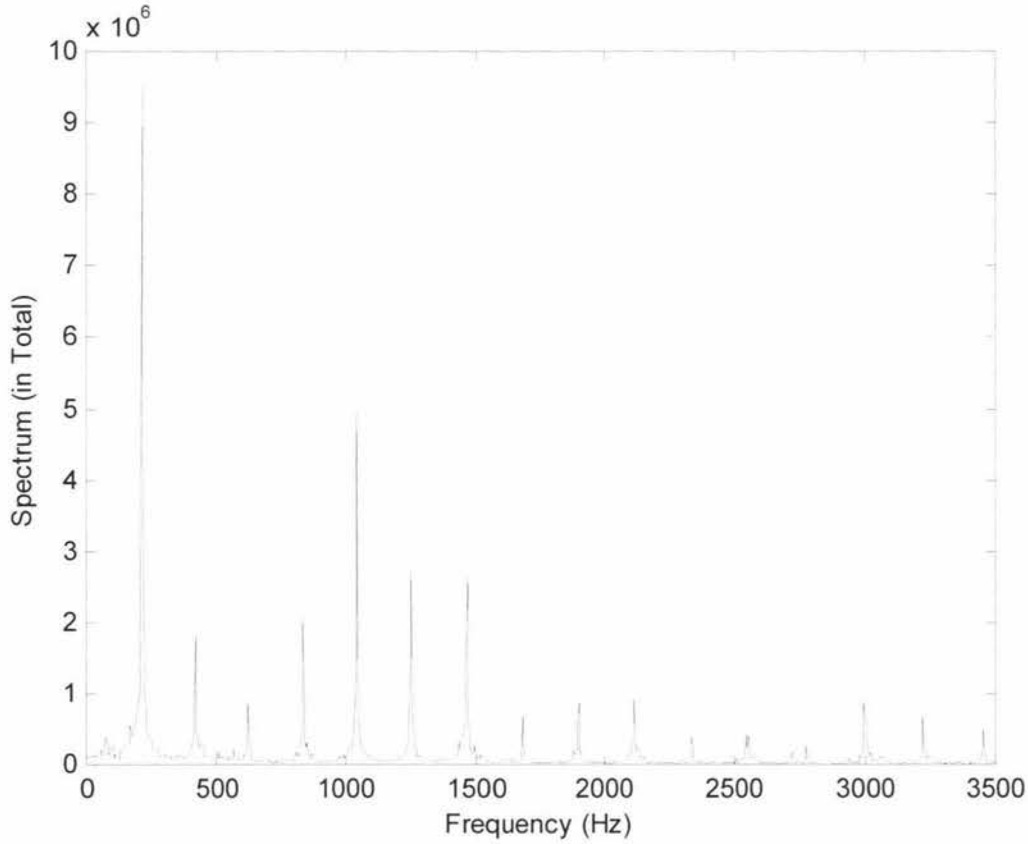

Figure B.17 The results of Fourier Transform when using different starting points on a bright acoustic piano when playing the note G#3 (207.65 Hz) after totalling all spectra.

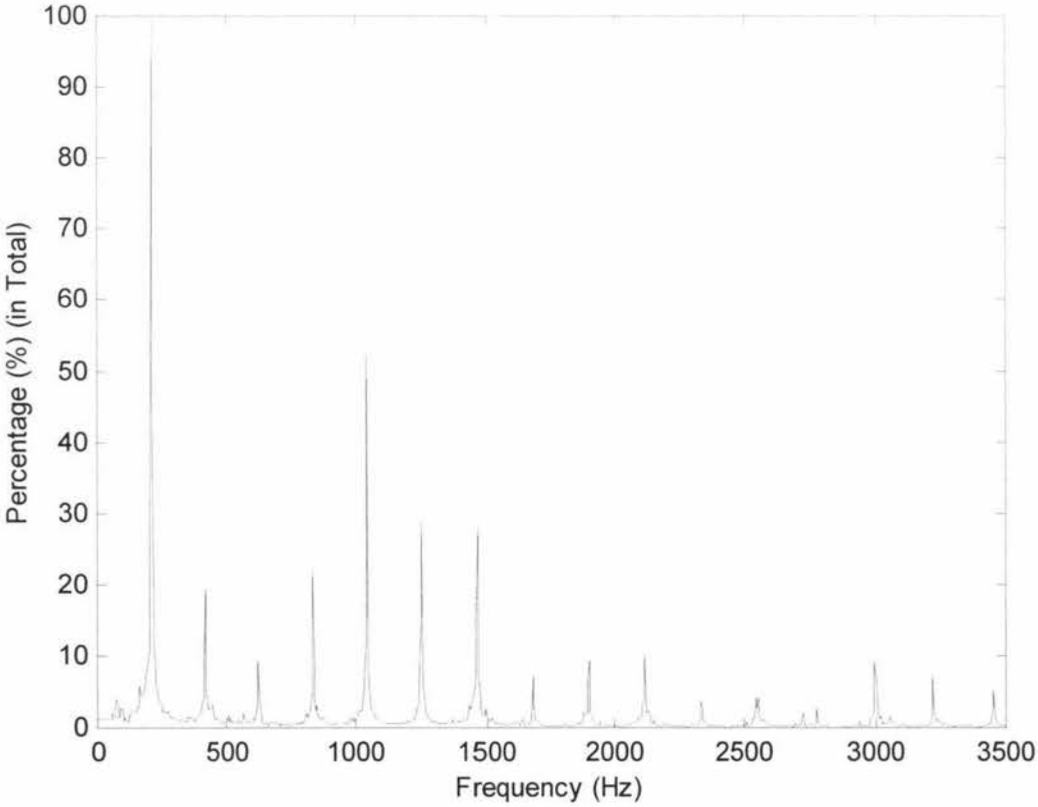

Figure B.18 The results of Fourier Transform when using different starting points on a bright acoustic piano when playing the note G#3 (207.65 Hz) in percentages after totalling all spectra.

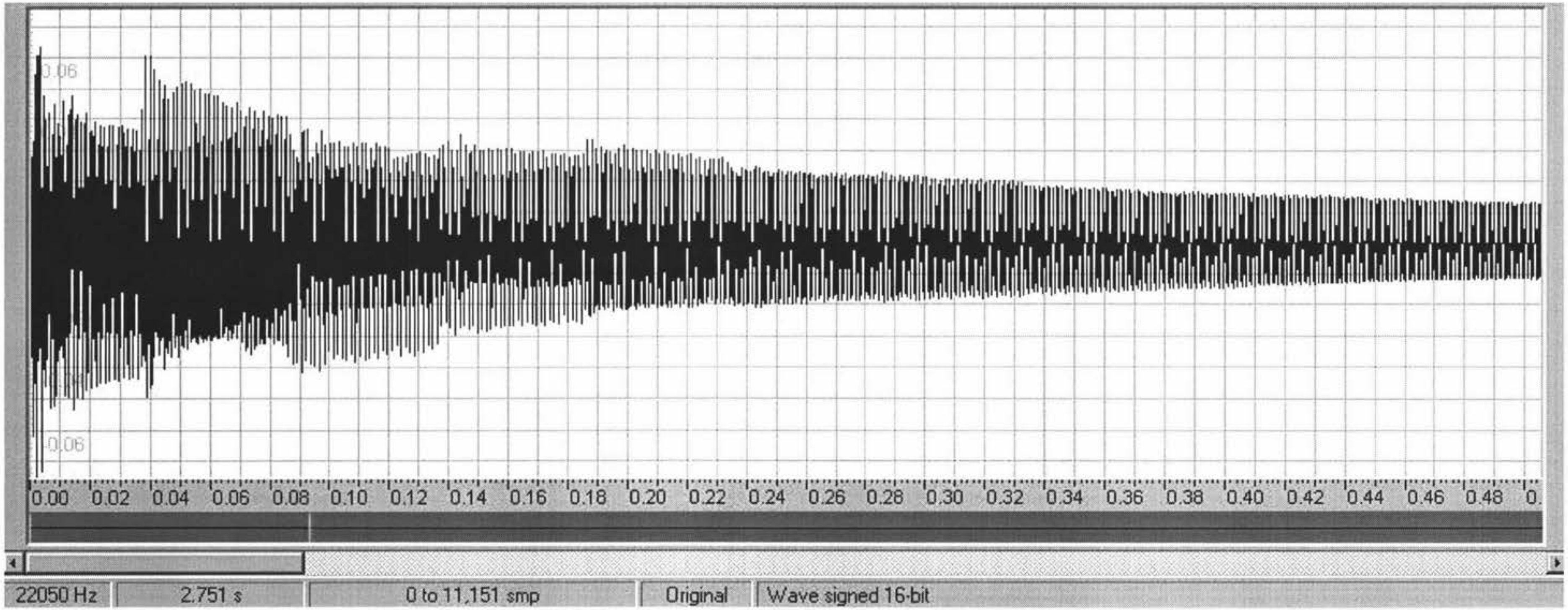

Figure B.19 The sample of an electric acoustic piano when playing the note E5 (659.26 Hz).

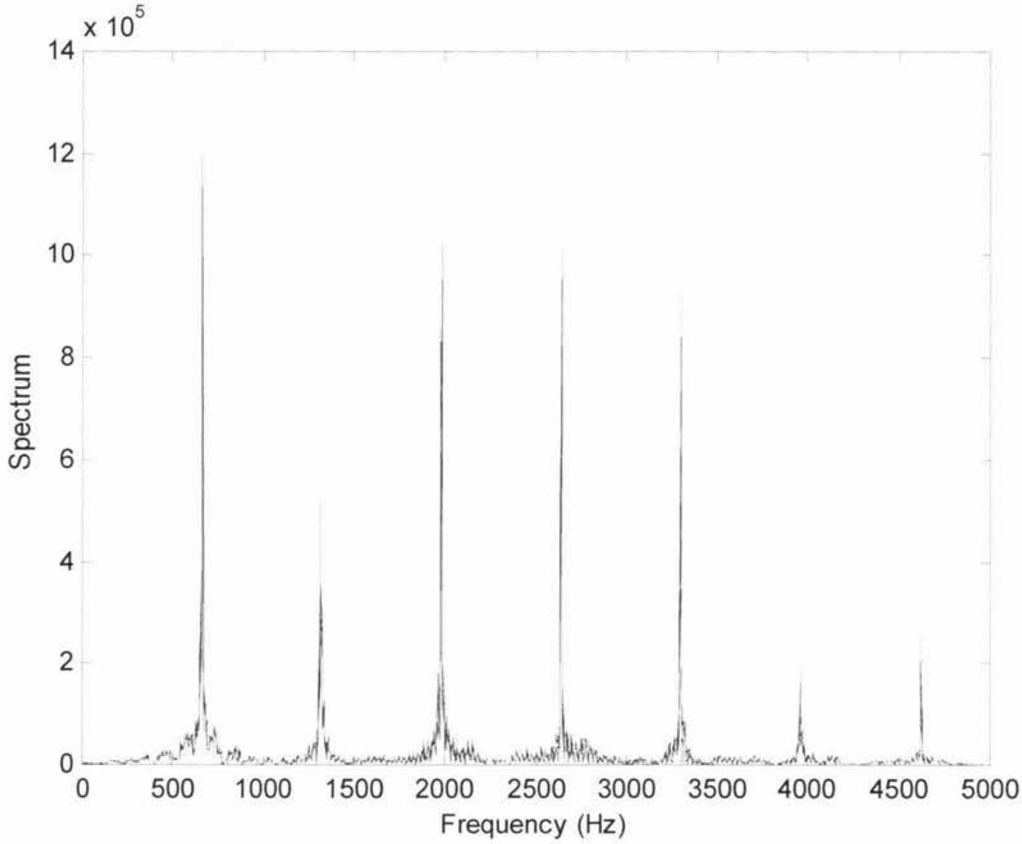

Figure B.20 The results of Fourier Transform in 2D when using different sample sizes on an electric acoustic piano when playing the note E5 (659.26 Hz).

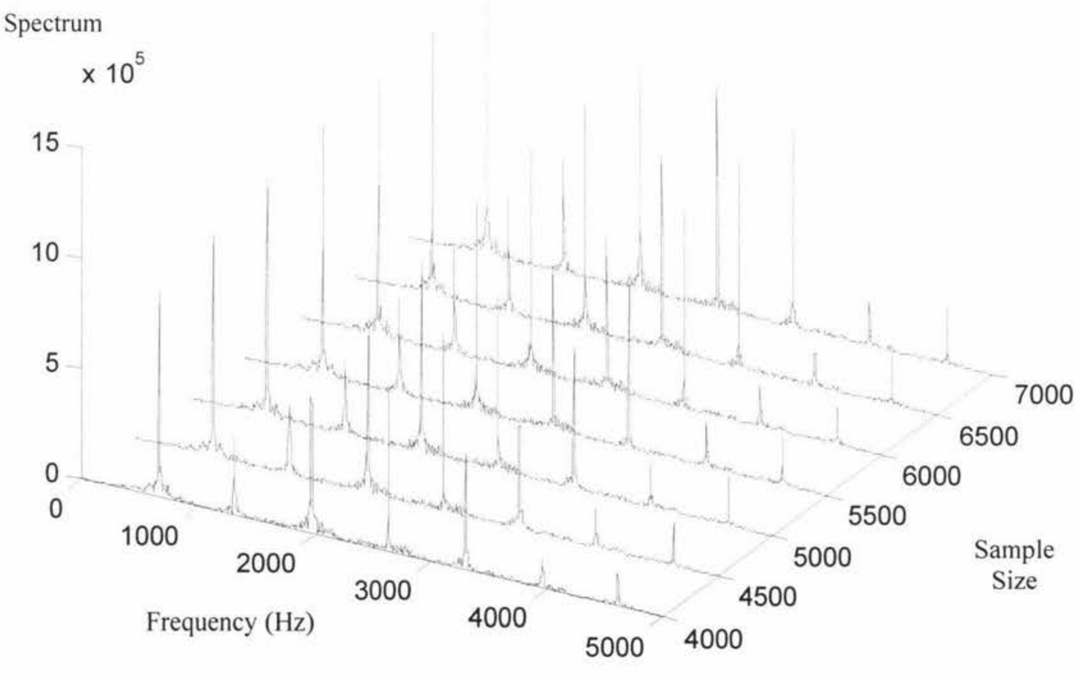

Figure B.21 The results of Fourier Transform in 3D when using different sample sizes on an electric acoustic piano when playing the note E5 (659.26 Hz).

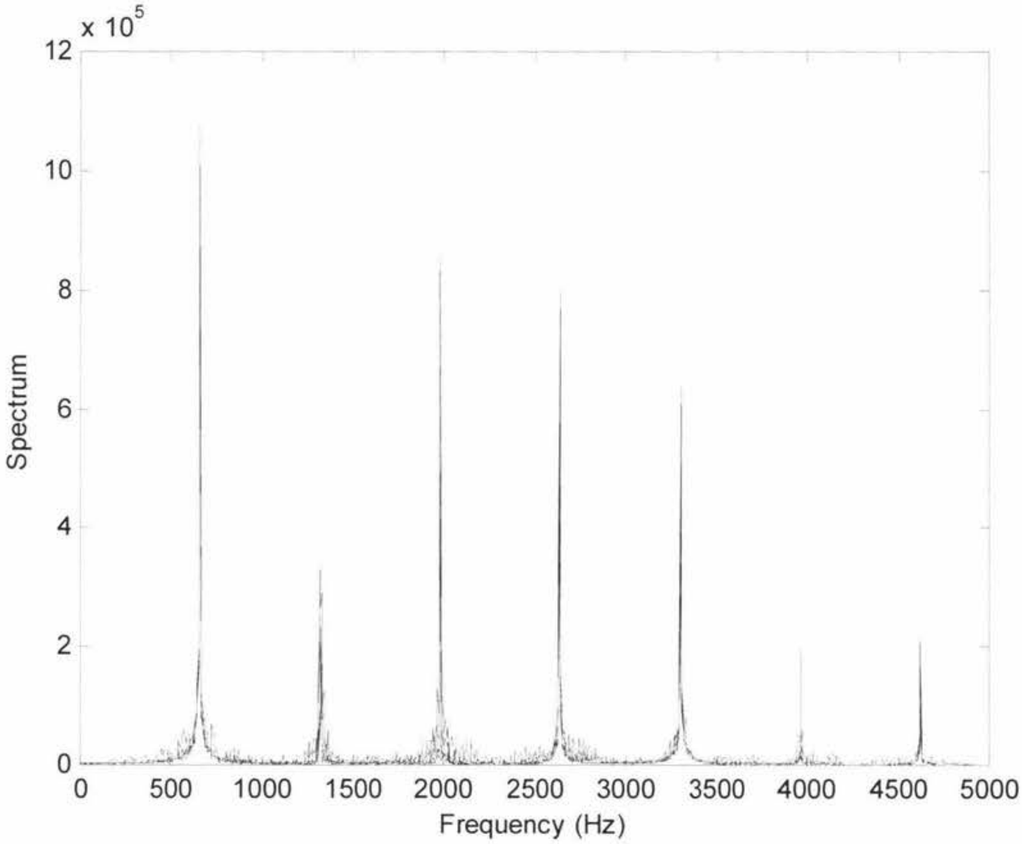

Figure B.22 The results of Fourier Transform in 2D when using different starting points on an electric acoustic piano when playing the note E5 (659.26 Hz).

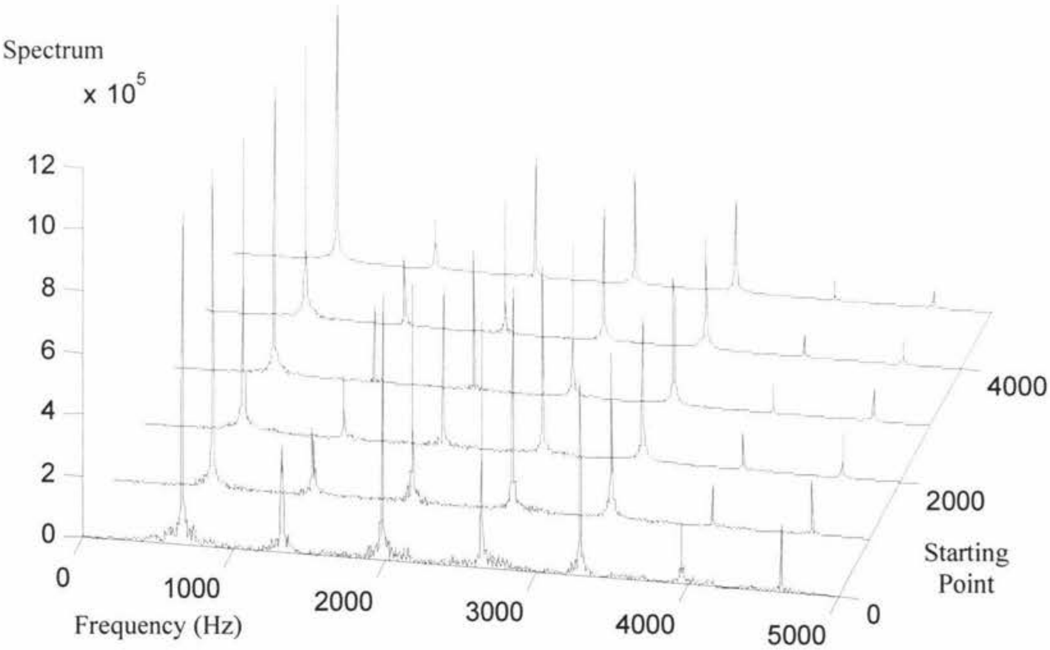

Figure B.23 The results of Fourier Transform in 3D when using different starting points on an electric acoustic piano when playing the note E5 (659.26 Hz).

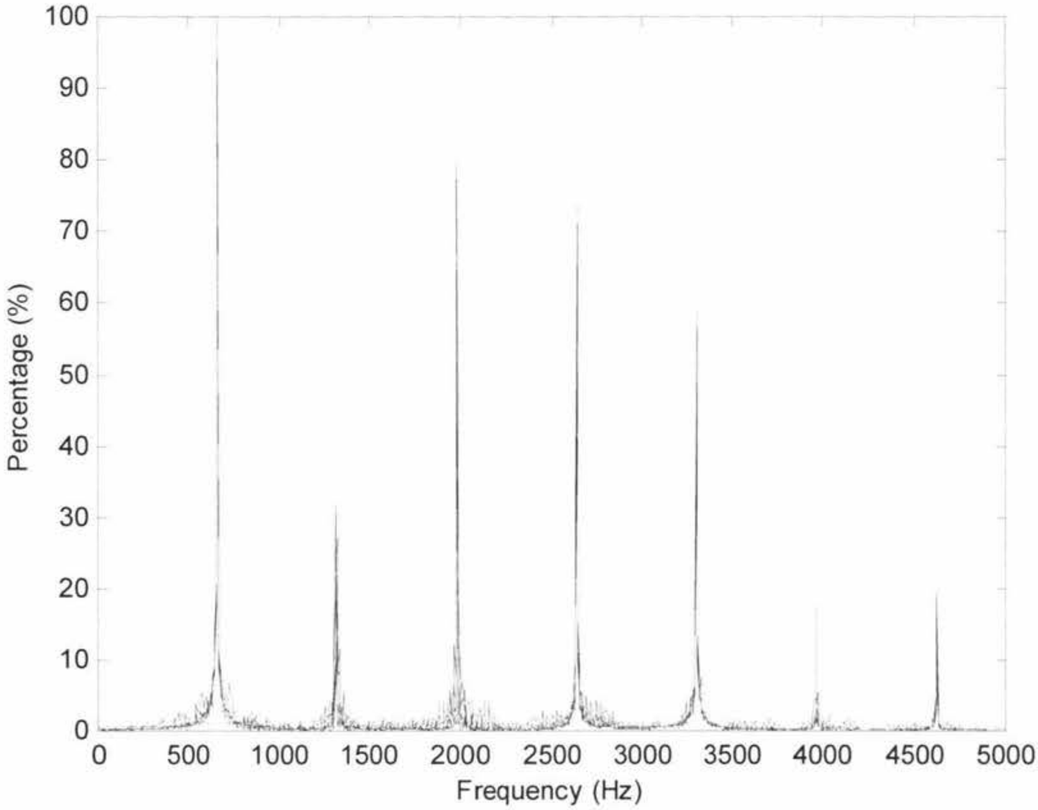

Figure B.24 The results of Fourier Transform in 2D when using different starting points on an electric acoustic piano when playing the note E5 (659.26 Hz).

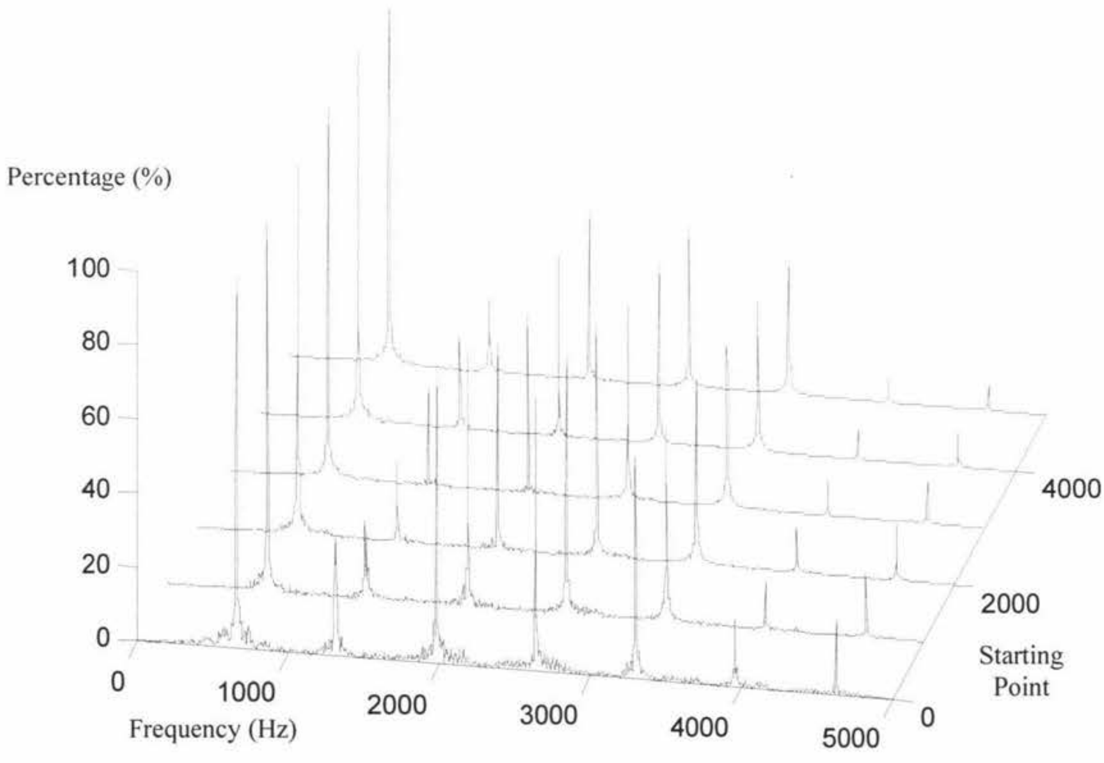

Figure B.25 The results of Fourier Transform in 3D when using different starting points on an electric acoustic piano when playing the note E5 (659.26 Hz).

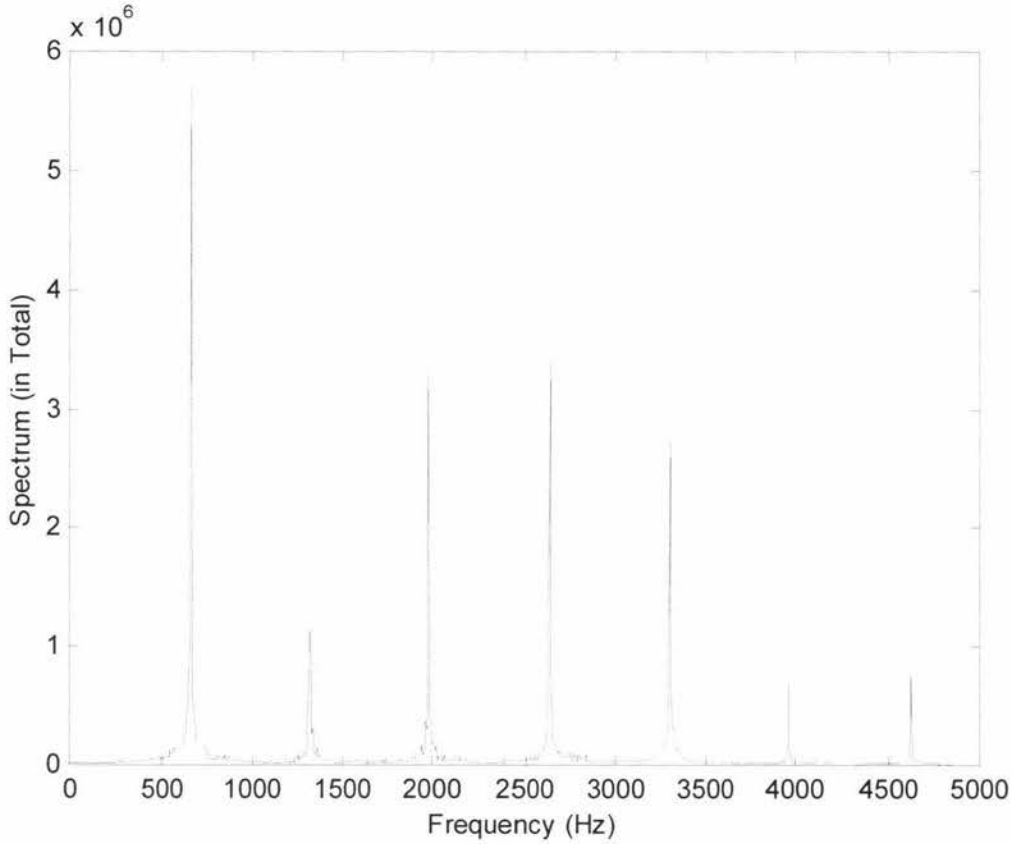

Figure B.26 The results of Fourier Transform when using different starting points on an electric acoustic piano when playing the note E5 (659.26 Hz) after totalling all spectra.

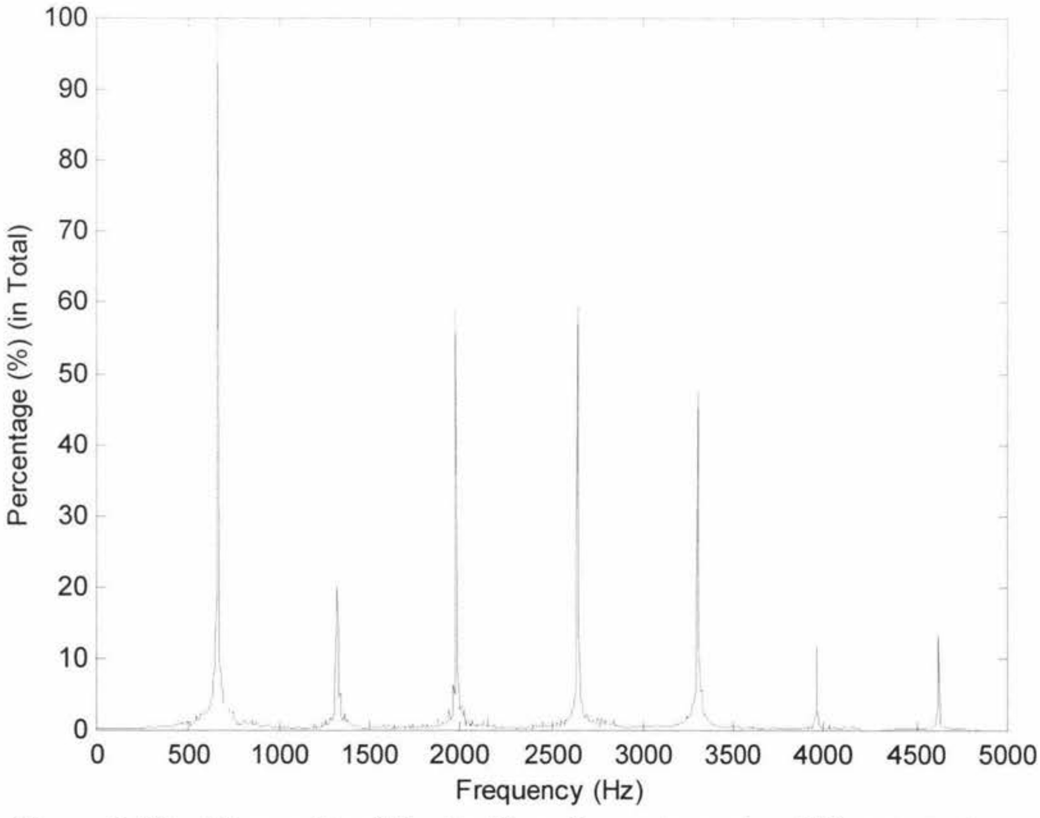

Figure B.27 The results of Fourier Transform when using different starting points on an electric acoustic piano when playing note E5 (659.26 Hz) in percentages after totalling all spectra.

In Table B.1, the top section of "No. Of Cases" shows the number of operations performed. The first column of table shows the number used by MIDI to represent each different instrument (Table 1.3 on page 7). Then column 2 shows the sum of all the results within each instrument and columns 3 to 7 show the results in the ranges we investigated (from C2 to B6). The second part (columns 8 to 13) shows the same results in percentages of the number of operations. And from the bottom line of the whole table we can see the averages of all instruments. This averages show about 50%~60% of the notes can be found by locating the highest peak of the Fourier Transform.

	Results in each octave						Percentage respect to each octave					
	All	2	3	4	5	6						
No. Of Cases	**88**	**12**	**12**	**12**	**12**	**12**	**All**	**2**	**3**	**4**	**5**	**6**
MIDI	62	3	10	12	12	12	70%	25%	83%	100%	100%	100%
Instrument	67	7	8	12	7	12	76%	58%	67%	100%	58%	100%
Number	29	0	0	0	7	9	33%	0%	0%	0%	58%	75%
4	74	7	11	12	12	12	84%	58%	92%	100%	100%	100%
5	78	12	12	12	12	10	89%	100%	100%	100%	100%	83%
6	80	12	12	12	11	12	91%	100%	100%	100%	92%	100%
7	52	0	11	12	9	9	59%	0%	92%	100%	75%	75%
8	0	0	0	0	0	0	0%	0%	0%	0%	0%	0%
9	81	12	12	12	12	12	92%	100%	100%	100%	100%	100%
10	81	12	12	12	12	12	92%	100%	100%	100%	100%	100%
11	46	4	5	10	7	5	52%	33%	42%	83%	58%	42%
12	81	12	12	12	12	12	92%	100%	100%	100%	100%	100%
13	81	12	12	12	12	12	92%	100%	100%	100%	100%	100%
14	77	10	12	10	12	12	88%	83%	100%	83%	100%	100%
15	0	0	0	0	0	0	0%	0%	0%	0%	0%	0%
16	32	0	1	6	2	10	36%	0%	8%	50%	17%	83%
17	37	5	8	12	5	2	42%	42%	67%	100%	42%	17%
18	2	0	0	0	0	1	2%	0%	0%	0%	0%	8%
19	11	0	0	0	2	3	13%	0%	0%	0%	17%	25%
20	24	0	2	5	4	7	27%	0%	17%	42%	33%	58%
21	38	5	6	3	3	7	43%	42%	50%	25%	25%	58%
22	64	10	12	12	8	6	73%	83%	100%	100%	67%	50%
23	17	3	4	3	4	0	19%	25%	33%	25%	33%	0%
24	25	0	2	1	3	5	28%	0%	17%	8%	25%	42%
25	28	6	12	10	0	0	32%	50%	100%	83%	0%	0%
26	43	6	7	7	7	7	49%	50%	58%	58%	58%	58%
27	79	10	12	12	12	12	90%	83%	100%	100%	100%	100%
28	28	0	0	0	9	8	32%	0%	0%	0%	75%	67%
29	60	6	9	12	9	10	68%	50%	75%	100%	75%	83%
30	36	0	0	5	12	7	41%	0%	0%	42%	100%	58%
31	8	0	0	2	4	2	9%	0%	0%	17%	33%	17%
32	0	0	0	0	0	0	0%	0%	0%	0%	0%	0%

Table B.1 The comparison between the highest peak from Fourier Transform and the expected results when using different instrument on MIDI (Continued on next page).

Table B.1 (cont.)

	Results in each octave						Percentage respect to each octave					
	All	2	3	4	5	6						
No. Of Cases	88	12	12	12	12	12	All	2	3	4	5	6
33	79	12	12	12	12	11	90%	100%	100%	100%	100%	92%
34	73	12	12	12	10	10	83%	100%	100%	100%	83%	83%
35	11	0	0	0	1	4	13%	0%	0%	0%	8%	33%
36	77	12	12	10	10	12	88%	100%	100%	83%	83%	100%
37	17	0	0	4	5	1	19%	0%	0%	33%	42%	8%
38	35	0	3	7	9	6	40%	0%	25%	58%	75%	50%
39	72	11	12	10	10	9	82%	92%	100%	83%	83%	75%
40	81	12	12	12	12	12	92%	100%	100%	100%	100%	100%
41	53	9	11	9	3	6	60%	75%	92%	75%	25%	50%
42	45	12	12	11	2	0	51%	100%	100%	92%	17%	0%
43	55	3	10	9	9	10	63%	25%	83%	75%	75%	83%
44	57	7	10	9	9	4	65%	58%	83%	75%	75%	33%
45	68	12	12	11	7	9	77%	100%	100%	92%	58%	75%
46	69	7	6	11	12	12	78%	58%	50%	92%	100%	100%
47	42	0	7	12	10	5	48%	0%	58%	100%	83%	42%
48	76	12	12	12	8	12	86%	100%	100%	100%	67%	100%
49	54	0	5	12	12	12	61%	0%	42%	100%	100%	100%
50	43	0	2	6	10	12	49%	0%	17%	50%	83%	100%
51	57	7	4	7	12	12	65%	58%	33%	58%	100%	100%
52	54	0	6	12	12	11	61%	0%	50%	100%	100%	92%
53	33	0	0	1	9	11	38%	0%	0%	8%	75%	92%
54	81	12	12	12	12	12	92%	100%	100%	100%	100%	100%
55	78	9	12	12	12	12	89%	75%	100%	100%	100%	100%
56	0	0	0	0	0	0	0%	0%	0%	0%	0%	0%
57	26	0	0	0	3	10	30%	0%	0%	0%	25%	83%
58	35	0	0	6	9	9	40%	0%	0%	50%	75%	75%
59	29	0	2	9	7	5	33%	0%	17%	75%	58%	42%
60	33	7	12	6	0	0	38%	58%	100%	50%	0%	0%
61	36	0	0	0	11	12	41%	0%	0%	0%	92%	100%
62	17	0	0	0	5	4	19%	0%	0%	0%	42%	33%
63	69	10	12	12	12	10	78%	83%	100%	100%	100%	83%
64	61	4	11	12	11	11	69%	33%	92%	100%	92%	92%
65	72	9	12	12	9	11	82%	75%	100%	100%	75%	92%
66	45	9	6	2	4	6	51%	75%	50%	17%	33%	50%
67	33	0	1	2	6	12	38%	0%	8%	17%	50%	100%
68	51	4	8	10	8	10	58%	33%	67%	83%	67%	83%
69	2	1	0	0	0	0	2%	8%	0%	0%	0%	0%
70	20	0	6	4	5	1	23%	0%	50%	33%	42%	8%
71	0	0	0	0	0	0	0%	0%	0%	0%	0%	0%
72	50	12	12	10	6	1	57%	100%	100%	83%	50%	8%

(Continued on next page)

Table B.1 (cont.)

	Results in each octave						Percentage respect to each octave					
	All	2	3	4	5	6						
No. Of Cases	88	12	12	12	12	12	All	2	3	4	5	6
73	80	12	12	12	12	11	91%	100%	100%	100%	100%	92%
74	80	12	12	12	12	12	91%	100%	100%	100%	100%	100%
75	81	12	12	12	12	12	92%	100%	100%	100%	100%	100%
76	78	12	12	12	9	12	89%	100%	100%	100%	75%	100%
77	80	11	12	12	12	12	91%	92%	100%	100%	100%	100%
78	33	0	0	0	8	12	38%	0%	0%	0%	67%	100%
79	77	8	12	12	12	12	88%	67%	100%	100%	100%	100%
80	81	12	12	12	12	12	92%	100%	100%	100%	100%	100%
81	63	3	8	12	12	12	72%	25%	67%	100%	100%	100%
82	80	12	12	12	12	12	91%	100%	100%	100%	100%	100%
83	79	12	11	11	12	12	90%	100%	92%	92%	100%	100%
84	81	12	12	12	12	12	92%	100%	100%	100%	100%	100%
85	41	0	0	5	11	12	47%	0%	0%	42%	92%	100%
86	37	0	0	4	11	11	42%	0%	0%	33%	92%	92%
87	54	9	11	8	7	5	61%	75%	92%	67%	58%	42%
88	71	12	12	12	9	7	81%	100%	100%	100%	75%	58%
89	77	8	12	12	12	12	88%	67%	100%	100%	100%	100%
90	81	12	12	12	12	12	92%	100%	100%	100%	100%	100%
91	57	1	2	12	12	12	65%	8%	17%	100%	100%	100%
92	78	9	12	12	12	12	89%	75%	100%	100%	100%	100%
93	81	12	12	12	12	12	92%	100%	100%	100%	100%	100%
94	49	4	7	8	6	9	56%	33%	58%	67%	50%	75%
95	67	2	12	12	12	12	76%	17%	100%	100%	100%	100%
96	66	9	11	12	12	11	75%	75%	92%	100%	100%	92%
97	81	12	12	12	12	12	92%	100%	100%	100%	100%	100%
98	80	12	12	12	12	11	91%	100%	100%	100%	100%	92%
99	26	5	1	3	2	4	30%	42%	8%	25%	17%	33%
100	81	12	12	12	12	12	92%	100%	100%	100%	100%	100%
101	42	0	0	9	10	11	48%	0%	0%	75%	83%	92%
102	81	12	12	12	12	12	92%	100%	100%	100%	100%	100%
103	40	5	0	4	6	9	45%	42%	0%	33%	50%	75%
104	81	12	12	12	12	12	92%	100%	100%	100%	100%	100%
105	12	1	0	0	2	5	14%	8%	0%	0%	17%	42%
106	1	0	0	0	0	1	1%	0%	0%	0%	0%	8%
107	0	0	0	0	0	0	0%	0%	0%	0%	0%	0%
108	0	0	0	0	0	0	0%	0%	0%	0%	0%	0%
109	81	12	12	12	12	12	92%	100%	100%	100%	100%	100%
110	2	0	0	0	0	0	2%	0%	0%	0%	0%	0%
111	51	9	7	6	7	5	58%	75%	58%	50%	58%	42%
112	10	0	0	0	0	2	11%	0%	0%	0%	0%	17%

(Continued on next page)

Table B.1 (cont.)

	Results in each octave						Percentage respect to each octave					
	All	**2**	**3**	**4**	**5**	**6**						
No. Of Cases	**88**	**12**	**12**	**12**	**12**	**12**	**All**	**2**	**3**	**4**	**5**	**6**
113	56	10	7	8	5	9	64%	83%	58%	67%	42%	75%
114	0	0	0	0	0	0	0%	0%	0%	0%	0%	0%
115	11	2	4	0	0	0	13%	17%	33%	0%	0%	0%
116	3	1	0	0	0	0	3%	8%	0%	0%	0%	0%
117	0	0	0	0	0	0	0%	0%	0%	0%	0%	0%
118	0	0	0	0	0	0	0%	0%	0%	0%	0%	0%
119	0	0	0	0	0	0	0%	0%	0%	0%	0%	0%
120	0	0	0	0	0	0	0%	0%	0%	0%	0%	0%
121	0	0	0	0	0	0	0%	0%	0%	0%	0%	0%
122	81	12	12	12	12	12	92%	100%	100%	100%	100%	100%
123	0	0	0	0	0	0	0%	0%	0%	0%	0%	0%
124	0	0	0	0	0	0	0%	0%	0%	0%	0%	0%
125	0	0	0	0	0	0	0%	0%	0%	0%	0%	0%
126	0	0	0	0	0	0	0%	0%	0%	0%	0%	0%
127	0	0	0	0	0	0	0%	0%	0%	0%	0%	0%
128	0	0	0	0	0	0	0%	0%	0%	0%	0%	0%
						Average	51%	44%	52%	58%	59%	61%

Range Started		Range Ended		Number of Cases	One Correct		Two Correct		Three Correct	
Acoustic Grand Piano (MIDI 1)										
36	C2	95	B6	8120	7808	96.16%	5306	65.34%	1484	18.28%
48	C3	95	B6	4048	4048	100%	3402	84.04%	1266	31.27%
60	C4	95	B6	1632	1632	100%	1552	95.1%	728	44.61%
72	C5	95	B6	440	440	100%	440	100%	326	74.09%
84	C6	95	B6	40	40	100%	40	100%	40	100%
Acoustic Guitar (steel) (MIDI 26)										
36	C2	95	B6	8120	7432	91.53%	3772	46.45%	540	6.65%
48	C3	95	B6	4048	3684	91.01%	1860	45.95%	266	6.57%
60	C4	95	B6	1632	1470	90.07%	754	46.2%	120	7.35%
72	C5	95	B6	440	408	92.73%	236	53.64%	36	8.18%
84	C6	95	B6	40	36	90%	12	30%	0	0%
Acoustic Bass (MIDI 33)										
36	C2	95	B6	8120	8120	100%	8086	99.58%	6666	82.09%
48	C3	95	B6	4048	4048	100%	4030	99.56%	3238	79.99%
60	C4	95	B6	1632	1632	100%	1624	99.51%	1246	76.35%
72	C5	95	B6	440	440	100%	440	100%	324	73.64%
84	C6	95	B6	40	40	100%	40	100%	20	50%
Violin (MIDI 41)										
36	C2	95	B6	8120	8052	99.16%	6316	77.78%	898	11.06%
48	C3	95	B6	4048	3980	98.32%	2930	72.38%	346	8.55%
60	C4	95	B6	1632	1564	95.83%	924	56.62%	66	4.04%
72	C5	95	B6	440	372	84.55%	140	31.82%	4	0.91%
84	C6	95	B6	40	38	95%	20	50%	2	5%
Recorder (MIDI 75)										
36	C2	95	B6	8120	8120	100%	8120	100%	8120	100%
48	C3	95	B6	4048	4048	100%	4048	100%	4048	100%
60	C4	95	B6	1632	1632	100%	1632	100%	1632	100%
72	C5	95	B6	440	440	100%	440	100%	440	100%
84	C6	95	B6	40	40	100%	40	100%	40	100%

Table B.2 The results of playing three notes together.

Instrument 1 (Acoustic Grand Piano)									
Note Value in MIDI	Starting Point						Max.	Min.	Average
	Periods			15 Periods plus number of samples					
	0	30	45	500	1000	2000			
36	144%	73%	60%	97%	94%	89%	144%	60%	93%
37	153%	73%	48%	96%	94%	87%	153%	48%	92%
38	165%	64%	48%	95%	91%	82%	165%	48%	91%
39	170%	65%	49%	96%	92%	81%	170%	49%	92%
40	143%	72%	48%	96%	91%	84%	143%	48%	89%
41	118%	75%	52%	97%	94%	87%	118%	52%	87%
42	105%	97%	90%	101%	101%	100%	105%	90%	99%
43	125%	96%	90%	100%	102%	99%	125%	90%	102%
44	139%	90%	84%	101%	99%	94%	139%	84%	101%
45	135%	97%	84%	100%	101%	100%	135%	84%	103%
46	146%	92%	86%	100%	94%	94%	146%	86%	102%
47	130%	82%	71%	90%	87%	85%	130%	71%	91%
48	125%	89%	77%	98%	96%	92%	125%	77%	96%
49	138%	97%	89%	100%	100%	98%	138%	89%	104%
50	93%	88%	79%	100%	97%	91%	100%	79%	91%
51	98%	88%	80%	100%	97%	90%	100%	80%	92%
52	90%	86%	70%	99%	96%	86%	99%	70%	88%
53	109%	92%	84%	98%	96%	92%	109%	84%	95%
54	103%	96%	92%	99%	97%	95%	103%	92%	97%
55	101%	90%	81%	99%	95%	88%	101%	81%	92%
56	100%	90%	81%	97%	94%	87%	100%	81%	92%
57	100%	89%	81%	97%	93%	85%	100%	81%	91%
58	99%	94%	86%	98%	95%	90%	99%	86%	94%
59	95%	98%	91%	100%	99%	94%	100%	91%	96%
60	101%	93%	85%	98%	95%	88%	101%	85%	93%
61	100%	93%	88%	98%	94%	89%	100%	88%	94%
62	106%	91%	81%	97%	92%	83%	106%	81%	92%
63	110%	89%	82%	94%	89%	82%	110%	82%	91%
64	101%	94%	85%	98%	94%	84%	101%	84%	93%
65	101%	96%	89%	99%	95%	89%	101%	89%	95%
66	102%	86%	73%	92%	85%	71%	102%	71%	85%
67	102%	87%	71%	94%	84%	67%	102%	67%	84%
68	102%	83%	66%	90%	77%	61%	102%	61%	80%
69	107%	83%	67%	89%	78%	61%	107%	61%	81%
70	105%	91%	83%	94%	88%	76%	105%	76%	89%
71	104%	84%	68%	88%	75%	58%	104%	58%	80%

Table B.3 The relationship on an acoustic grand piano using different starting point and compared to the original curve (15 Periods of wavelength) in percentage (continued on next page).

Table B.3 (cont.)

Instrument 1 (Acoustic Grand Piano)									
Note Value in MIDI	Starting Point						Max.	Min.	Average
	Periods			15 Periods plus number of samples					
	0	30	45	500	1000	2000			
72	104%	92%	80%	93%	85%	65%	104%	65%	87%
73	110%	87%	71%	89%	75%	57%	110%	57%	82%
74	114%	73%	54%	75%	58%	36%	114%	36%	68%
75	104%	72%	50%	73%	52%	37%	104%	37%	65%
76	96%	89%	74%	89%	74%	47%	96%	47%	78%
77	103%	71%	49%	69%	48%	40%	103%	40%	63%
78	99%	83%	67%	81%	61%	35%	99%	35%	71%
79	102%	73%	54%	68%	48%	24%	102%	24%	61%
80	130%	61%	51%	54%	53%	47%	130%	47%	66%
81	118%	76%	58%	70%	54%	35%	118%	35%	68%
82	121%	81%	62%	73%	47%	22%	121%	22%	68%
83	94%	96%	93%	95%	85%	69%	96%	69%	89%
84	103%	82%	71%	75%	63%	58%	103%	58%	76%
85	101%	90%	83%	84%	80%	72%	101%	72%	85%
86	112%	75%	61%	62%	40%	27%	112%	27%	63%
87	122%	86%	72%	72%	64%	40%	122%	40%	76%
88	124%	70%	68%	69%	77%	48%	124%	48%	76%
89	115%	72%	71%	70%	49%	36%	115%	36%	69%
90	119%	68%	60%	54%	60%	39%	119%	39%	67%
91	110%	74%	58%	59%	50%	34%	110%	34%	64%
92	116%	84%	72%	66%	37%	40%	116%	37%	69%
93	106%	100%	100%	100%	87%	61%	106%	61%	92%
94	110%	82%	71%	64%	49%	38%	110%	38%	69%
95	107%	93%	91%	91%	80%	68%	107%	68%	88%
Max.	170%	100%	100%	101%	102%	100%	170%		
Min.	90%	61%	48%	54%	37%	22%		22%	
Average	113%	85%	73%	88%	80%	69%			85%

Instrument 26 (Acoustic Guitar (steel))									
Note Value in MIDI	Starting Point						Max.	Min.	Average
	Periods			15 Periods plus number of samples					
	0	30	45	500	1000	2000			
36	577%	128%	91%	101%	101%	101%	577%	91%	183%
37	451%	140%	107%	99%	99%	112%	451%	99%	168%
38	567%	186%	153%	107%	125%	150%	567%	107%	215%
39	517%	172%	145%	114%	131%	151%	517%	114%	205%
40	247%	93%	68%	100%	101%	103%	247%	68%	119%
41	421%	118%	105%	105%	108%	113%	421%	105%	162%
42	88%	103%	105%	102%	102%	103%	105%	88%	101%
43	91%	103%	104%	101%	101%	102%	104%	91%	100%
44	80%	101%	100%	100%	104%	102%	104%	80%	98%
45	94%	104%	106%	103%	102%	103%	106%	94%	102%
46	100%	103%	106%	101%	102%	103%	106%	100%	102%
47	104%	88%	75%	98%	97%	91%	104%	75%	92%
48	113%	88%	77%	96%	94%	90%	113%	77%	93%
49	109%	88%	77%	97%	95%	89%	109%	77%	93%
50	108%	87%	77%	97%	94%	88%	108%	77%	92%
51	109%	86%	76%	96%	93%	87%	109%	76%	91%
52	94%	88%	78%	97%	94%	88%	97%	78%	90%
53	110%	86%	80%	96%	92%	86%	110%	80%	92%
54	113%	98%	83%	98%	98%	97%	113%	83%	98%
55	99%	88%	79%	95%	91%	86%	99%	79%	89%
56	112%	89%	80%	96%	92%	86%	112%	80%	93%
57	101%	83%	69%	94%	89%	77%	101%	69%	86%
58	106%	97%	96%	96%	97%	96%	106%	96%	98%
59	96%	91%	82%	97%	93%	87%	97%	82%	91%
60	95%	91%	82%	96%	93%	87%	96%	82%	91%
61	104%	86%	78%	95%	89%	80%	104%	78%	89%
62	127%	86%	75%	92%	87%	77%	127%	75%	91%
63	143%	84%	77%	90%	83%	77%	143%	77%	92%
64	94%	99%	95%	100%	99%	95%	100%	94%	97%
65	127%	87%	100%	91%	87%	101%	127%	87%	99%
66	97%	100%	97%	100%	99%	96%	100%	96%	98%
67	104%	93%	85%	96%	91%	83%	104%	83%	92%
68	105%	93%	86%	95%	91%	83%	105%	83%	92%
69	108%	89%	82%	93%	86%	79%	108%	79%	90%
70	104%	96%	93%	97%	95%	89%	104%	89%	96%
71	107%	90%	84%	93%	87%	80%	107%	80%	90%

Table B.4 The relationship on an acoustic guitar (steel) using different starting point and compared to the original curve (15 Periods of wavelength) in percentage (continued on next page).

Table B.4 (cont.)

Instrument 26 (Acoustic Guitar (steel))									
Note Value in MIDI	Starting Point						Max.	Min.	Average
	Periods			15 Periods plus number of samples					
	0	30	45	500	1000	2000			
72	102%	95%	90%	96%	91%	83%	102%	83%	93%
73	109%	91%	83%	93%	85%	74%	109%	74%	89%
74	99%	99%	96%	100%	97%	87%	100%	87%	97%
75	96%	99%	95%	99%	95%	82%	99%	82%	94%
76	97%	102%	101%	102%	101%	97%	102%	97%	100%
77	106%	94%	89%	94%	88%	78%	106%	78%	91%
78	101%	97%	93%	97%	91%	84%	101%	84%	94%
79	105%	95%	91%	94%	89%	78%	105%	78%	92%
80	102%	97%	91%	95%	87%	74%	102%	74%	91%
81	100%	99%	95%	98%	92%	86%	100%	86%	95%
82	98%	102%	103%	103%	102%	97%	103%	97%	101%
83	102%	98%	96%	97%	95%	90%	102%	90%	96%
84	103%	97%	93%	96%	85%	71%	103%	71%	91%
85	117%	94%	90%	91%	82%	62%	117%	62%	89%
86	100%	100%	100%	100%	97%	90%	100%	90%	98%
87	105%	95%	90%	90%	90%	82%	105%	82%	92%
88	104%	96%	93%	93%	82%	64%	104%	64%	88%
89	102%	98%	96%	96%	91%	81%	102%	81%	94%
90	103%	97%	94%	94%	83%	64%	103%	64%	89%
91	106%	94%	93%	94%	93%	91%	106%	91%	95%
92	101%	99%	98%	97%	90%	75%	101%	75%	93%
93	101%	99%	97%	95%	91%	85%	101%	85%	95%
94	104%	96%	94%	91%	77%	64%	104%	64%	88%
95	108%	91%	84%	77%	66%	49%	108%	49%	79%
Max.	577%	186%	153%	114%	131%	151%	577%		
Min.	80%	83%	68%	77%	66%	49%		49%	
Average	140%	99%	92%	97%	94%	89%			102%

Instrument 33 (Acoustic Bass)									
Note Value in MIDI	Starting Point						Max.	Min.	Average
	Periods			15 Periods plus number of samples					
	0	30	45	500	1000	2000			
36	131%	61%	41%	97%	90%	79%	131%	41%	83%
37	153%	65%	51%	94%	87%	80%	153%	51%	88%
38	149%	68%	58%	96%	91%	80%	149%	58%	90%
39	149%	69%	63%	94%	85%	73%	149%	63%	89%
40	170%	65%	64%	93%	87%	77%	170%	64%	93%
41	146%	70%	64%	93%	87%	79%	146%	64%	90%
42	131%	65%	59%	94%	87%	76%	131%	59%	85%
43	133%	61%	53%	94%	87%	73%	133%	53%	83%
44	141%	77%	68%	95%	91%	82%	141%	68%	92%
45	147%	70%	66%	94%	86%	74%	147%	66%	89%
46	154%	80%	71%	89%	76%	79%	154%	71%	92%
47	156%	53%	61%	90%	77%	59%	156%	53%	83%
48	152%	80%	77%	93%	87%	82%	152%	77%	95%
49	146%	90%	74%	97%	99%	94%	146%	74%	100%
50	139%	63%	60%	88%	77%	64%	139%	60%	82%
51	156%	66%	62%	88%	76%	67%	156%	62%	86%
52	137%	70%	61%	83%	71%	70%	137%	61%	82%
53	134%	87%	84%	94%	89%	88%	134%	84%	96%
54	116%	89%	78%	97%	93%	88%	116%	78%	94%
55	120%	93%	88%	98%	97%	92%	120%	88%	98%
56	138%	77%	76%	90%	82%	77%	138%	76%	90%
57	113%	84%	78%	93%	88%	82%	113%	78%	90%
58	120%	81%	80%	93%	85%	80%	120%	80%	90%
59	121%	76%	68%	90%	81%	71%	121%	68%	85%
60	121%	104%	102%	104%	104%	103%	121%	102%	106%
61	114%	85%	79%	93%	87%	80%	114%	79%	89%
62	115%	91%	87%	96%	93%	88%	115%	87%	95%
63	112%	90%	91%	94%	90%	91%	112%	90%	95%
64	110%	88%	84%	92%	88%	84%	110%	84%	91%
65	137%	109%	104%	108%	108%	104%	137%	104%	112%
66	127%	97%	99%	93%	97%	96%	127%	93%	101%
67	123%	84%	83%	88%	84%	84%	123%	83%	91%
68	101%	94%	95%	97%	93%	98%	101%	93%	96%
69	115%	87%	85%	90%	87%	83%	115%	83%	91%
70	112%	87%	80%	90%	83%	75%	112%	75%	88%
71	127%	96%	95%	95%	96%	90%	127%	90%	100%

Table B.5 The relationship on an acoustic bass using different starting point and compared to the original curve (15 Periods of wavelength) in percentage (continued on next page).

Table B.5 (cont.)

Instrument 33 (Acoustic Bass)									
Note Value in MIDI	Starting Point						Max.	Min.	Average
	Periods			15 Periods plus number of samples					
	0	30	45	500	1000	2000			
72	110%	90%	88%	92%	88%	86%	110%	86%	92%
73	116%	89%	90%	90%	89%	89%	116%	89%	94%
74	109%	95%	97%	95%	97%	99%	109%	95%	99%
75	117%	85%	78%	85%	79%	75%	117%	75%	87%
76	106%	95%	93%	95%	93%	88%	106%	88%	95%
77	111%	94%	98%	94%	99%	97%	111%	94%	99%
78	103%	102%	96%	101%	95%	96%	103%	95%	99%
79	101%	98%	101%	99%	102%	100%	102%	98%	100%
80	115%	90%	98%	90%	102%	105%	115%	90%	100%
81	112%	91%	91%	90%	94%	87%	112%	87%	94%
82	112%	86%	80%	81%	78%	79%	112%	78%	86%
83	100%	101%	101%	102%	99%	98%	102%	98%	100%
84	105%	100%	102%	101%	101%	100%	105%	100%	102%
85	98%	101%	103%	102%	107%	114%	114%	98%	104%
86	97%	101%	101%	101%	92%	87%	101%	87%	96%
87	103%	97%	93%	94%	95%	95%	103%	93%	96%
88	105%	99%	103%	103%	112%	107%	112%	99%	105%
89	110%	90%	87%	88%	93%	95%	110%	87%	94%
90	106%	95%	94%	94%	101%	99%	106%	94%	98%
91	104%	98%	100%	100%	96%	91%	104%	91%	98%
92	97%	102%	104%	105%	97%	93%	105%	93%	100%
93	101%	100%	101%	102%	101%	95%	102%	95%	100%
94	103%	97%	96%	93%	94%	92%	103%	92%	96%
95	100%	101%	104%	107%	110%	116%	116%	100%	106%
Max.	170%	109%	104%	108%	112%	116%	170%		
Min.	97%	53%	41%	81%	71%	59%		41%	
Average	122%	86%	83%	95%	91%	87%			94%

Instrument 41 (Violin)									
Note Value in MIDI	Starting Point						Max.	Min.	Average
	Periods			15 Periods plus number of samples					
	0	30	45	500	1000	2000			
36	52%	106%	68%	96%	97%	100%	106%	52%	87%
37	56%	92%	76%	100%	100%	101%	101%	56%	88%
38	55%	87%	85%	101%	99%	99%	101%	55%	88%
39	50%	83%	63%	102%	103%	100%	103%	50%	83%
40	66%	96%	111%	100%	98%	95%	111%	66%	94%
41	51%	95%	79%	101%	102%	102%	102%	51%	88%
42	68%	122%	107%	102%	105%	114%	122%	68%	103%
43	74%	116%	107%	102%	102%	91%	116%	74%	99%
44	50%	72%	67%	99%	100%	89%	100%	50%	79%
45	76%	118%	95%	102%	111%	125%	125%	76%	105%
46	61%	88%	66%	94%	82%	85%	94%	61%	79%
47	81%	121%	100%	88%	94%	118%	121%	81%	100%
48	67%	80%	64%	99%	94%	76%	99%	64%	80%
49	73%	109%	75%	90%	99%	105%	109%	73%	92%
50	82%	129%	84%	87%	106%	126%	129%	82%	102%
51	92%	143%	91%	110%	131%	143%	143%	91%	118%
52	86%	114%	77%	91%	93%	114%	114%	77%	96%
53	85%	73%	79%	90%	80%	72%	90%	72%	80%
54	75%	96%	86%	104%	102%	94%	104%	75%	93%
55	100%	95%	90%	88%	99%	86%	100%	86%	93%
56	84%	98%	77%	106%	105%	91%	106%	77%	93%
57	79%	100%	95%	102%	101%	98%	102%	79%	96%
58	89%	97%	95%	100%	99%	96%	100%	89%	96%
59	83%	102%	93%	105%	104%	96%	105%	83%	97%
60	103%	106%	125%	106%	105%	120%	125%	103%	111%
61	86%	102%	94%	102%	102%	97%	102%	86%	97%
62	89%	90%	92%	98%	92%	92%	98%	89%	92%
63	88%	130%	139%	111%	128%	140%	140%	88%	123%
64	92%	105%	102%	104%	105%	102%	105%	92%	102%
65	93%	97%	92%	99%	96%	91%	99%	91%	95%
66	90%	85%	67%	94%	83%	64%	94%	64%	80%
67	86%	95%	107%	100%	94%	107%	107%	86%	98%
68	87%	95%	90%	99%	93%	93%	99%	87%	93%
69	85%	102%	90%	102%	99%	106%	106%	85%	97%
70	91%	110%	122%	107%	117%	120%	122%	91%	111%
71	89%	104%	104%	104%	104%	105%	105%	89%	102%

Table B.6 The relationship on a violin using different starting point and compared to the original curve (15 Periods of wavelength) in percentage (continued on next page).

Table B.6 (cont.)

Instrument 41 (Violin)									
Note Value in MIDI	Starting Point						Max.	Min.	Average
	Periods			15 Periods plus number of samples					
	0	30	45	500	1000	2000			
72	99%	100%	104%	100%	103%	95%	104%	95%	100%
73	82%	102%	104%	102%	103%	112%	112%	82%	101%
74	93%	103%	107%	103%	106%	116%	116%	93%	105%
75	85%	108%	118%	107%	116%	117%	118%	85%	109%
76	99%	104%	106%	104%	106%	105%	106%	99%	104%
77	87%	106%	119%	107%	119%	101%	119%	87%	106%
78	103%	118%	134%	121%	136%	152%	152%	103%	127%
79	84%	111%	117%	113%	118%	110%	118%	84%	109%
80	101%	95%	114%	98%	133%	146%	146%	95%	115%
81	85%	119%	131%	125%	129%	115%	131%	85%	117%
82	103%	107%	110%	108%	111%	85%	111%	85%	104%
83	100%	100%	101%	103%	92%	109%	109%	92%	101%
84	94%	105%	108%	102%	127%	127%	127%	94%	110%
85	95%	100%	98%	96%	103%	123%	123%	95%	102%
86	105%	88%	94%	93%	92%	105%	105%	88%	96%
87	95%	103%	100%	100%	104%	83%	104%	83%	97%
88	95%	98%	92%	92%	112%	144%	144%	92%	106%
89	98%	99%	101%	102%	103%	119%	119%	98%	104%
90	99%	107%	110%	112%	123%	156%	156%	99%	118%
91	98%	100%	106%	111%	135%	145%	145%	98%	116%
92	101%	102%	99%	97%	101%	125%	125%	97%	104%
93	99%	100%	107%	111%	139%	176%	176%	99%	122%
94	94%	103%	110%	119%	134%	155%	155%	94%	119%
95	100%	107%	113%	109%	110%	110%	113%	100%	108%
Max.	105%	143%	139%	125%	139%	176%	176%		
Min.	50%	72%	63%	87%	80%	64%		50%	
Average	85%	102%	98%	102%	106%	110%			100%

Instrument 75 (Recorder)									
Note Value in MIDI	Starting Point						Max.	Min.	Average
	Periods			15 Periods plus number of samples					
	0	30	45	500	1000	2000			
36	26%	106%	44%	102%	104%	106%	106%	26%	81%
37	33%	99%	57%	101%	102%	104%	104%	33%	83%
38	37%	102%	77%	101%	103%	104%	104%	37%	87%
39	36%	97%	86%	102%	102%	102%	102%	36%	87%
40	44%	95%	89%	101%	101%	100%	101%	44%	88%
41	41%	103%	100%	102%	103%	104%	104%	41%	92%
42	40%	107%	106%	103%	104%	107%	107%	40%	95%
43	48%	110%	117%	102%	104%	106%	117%	48%	98%
44	52%	101%	95%	102%	102%	104%	104%	52%	93%
45	52%	103%	101%	102%	103%	103%	103%	52%	94%
46	74%	87%	81%	96%	93%	88%	96%	74%	86%
47	60%	97%	97%	100%	100%	97%	100%	60%	92%
48	64%	95%	90%	100%	100%	97%	100%	64%	91%
49	77%	104%	98%	100%	101%	104%	104%	77%	97%
50	82%	114%	121%	106%	109%	114%	121%	82%	108%
51	76%	109%	117%	102%	103%	107%	117%	76%	102%
52	85%	94%	96%	99%	98%	94%	99%	85%	94%
53	74%	93%	89%	100%	99%	93%	100%	74%	91%
54	69%	108%	107%	103%	106%	108%	108%	69%	100%
55	82%	93%	86%	98%	96%	91%	98%	82%	91%
56	75%	107%	110%	102%	105%	108%	110%	75%	101%
57	71%	108%	106%	104%	107%	108%	108%	71%	101%
58	76%	105%	101%	102%	104%	103%	105%	76%	99%
59	76%	109%	110%	105%	108%	110%	110%	76%	103%
60	97%	105%	117%	97%	102%	113%	117%	97%	105%
61	80%	105%	101%	103%	105%	103%	105%	80%	99%
62	87%	94%	85%	98%	95%	87%	98%	85%	91%
63	87%	93%	82%	98%	93%	83%	98%	82%	89%
64	83%	100%	93%	101%	100%	93%	101%	83%	95%
65	84%	113%	121%	108%	114%	121%	121%	84%	110%
66	86%	99%	93%	100%	99%	91%	100%	86%	95%
67	84%	104%	102%	103%	104%	100%	104%	84%	99%
68	92%	116%	127%	110%	120%	127%	127%	92%	115%
69	90%	96%	85%	98%	93%	76%	98%	76%	90%
70	90%	97%	88%	98%	93%	82%	98%	82%	91%
71	87%	104%	100%	104%	103%	96%	104%	87%	99%

Table B.7 The relationship on a recorder using different starting point and compared to the original curve (15 Periods of wavelength) in percentage (continued on next page).

Table B.7 (cont.)

Instrument 75 (Recorder)									
Note Value in MIDI	Starting Point						Max.	Min.	Average
	Periods			15 Periods plus number of samples					
	0	30	45	500	1000	2000			
72	93%	95%	84%	96%	88%	71%	96%	71%	88%
73	98%	89%	75%	91%	80%	79%	98%	75%	85%
74	86%	105%	105%	105%	105%	103%	105%	86%	102%
75	85%	107%	102%	107%	103%	92%	107%	85%	99%
76	107%	98%	113%	98%	113%	132%	132%	98%	110%
77	87%	115%	120%	116%	120%	119%	120%	87%	113%
78	97%	93%	82%	91%	78%	84%	97%	78%	87%
79	93%	93%	78%	90%	79%	84%	93%	78%	86%
80	88%	109%	113%	110%	116%	129%	129%	88%	111%
81	90%	114%	122%	117%	127%	134%	134%	90%	117%
82	86%	116%	128%	122%	139%	163%	163%	86%	126%
83	97%	97%	90%	93%	85%	65%	97%	65%	88%
84	90%	110%	114%	113%	118%	124%	124%	90%	111%
85	108%	88%	87%	88%	78%	81%	108%	78%	88%
86	105%	92%	104%	102%	116%	100%	116%	92%	103%
87	94%	110%	116%	115%	115%	104%	116%	94%	109%
88	96%	103%	103%	103%	104%	100%	104%	96%	101%
89	98%	104%	107%	106%	108%	118%	118%	98%	107%
90	98%	103%	103%	101%	88%	80%	103%	80%	95%
91	103%	94%	83%	78%	80%	80%	103%	78%	86%
92	94%	103%	106%	109%	121%	140%	140%	94%	112%
93	100%	93%	90%	91%	76%	68%	100%	68%	86%
94	101%	101%	101%	98%	90%	59%	101%	59%	92%
95	97%	102%	100%	97%	90%	103%	103%	90%	98%
Max.	108%	116%	128%	122%	139%	163%	163%		
Min.	26%	87%	44%	78%	76%	59%		26%	
Average	80%	102%	98%	102%	102%	101%			97%

Instrument Value	Known note from the same instrument					Known note from 5 different instrument				
	Before Subtraction	Using filtered FT results		Using complete FT results		Before Subtraction	Using filtered FFT results		Using complete FT results	
		After the first subtraction	After the second subtractions	After the first subtraction	After the second subtractions		After the first subtraction	After the second subtractions	After the first subtraction	After the second subtractions
Average Frequency										
1	3999	3589	2033	3805	2895	3454	2783	1559	3042	2027
26	3795	3593	3096	3692	3619	3256	2913	2113	3191	2781
33	3922	3554	3027	3791	3537	2598	2142	1828	2450	2064
41	4050	4010	3569	4045	3947	2678	3090	2554	3270	2938
75	2562	3221	3321	3342	3465	1918	2662	2912	2859	3247
Average	3665.6	3593.4	3009.2	3735	3492.6	2780.8	2718	2193.2	2962.4	2611.4
No. of Cases	4060	4060	4060	4060	4060	4060	4060	4060	4060	4060
Percentage	90.29%	88.51%	74.12%	92.00%	86.02%	68.49%	66.95%	54.02%	72.97%	64.32%
Lower Frequency										
1	4027	3846	2486	3829	3052	3668	3104	1719	3005	1840
26	4039	3968	3269	4002	3853	3808	3093	1660	3126	1969
33	4040	4020	3987	4045	4007	3491	2799	2547	2809	2557
41	4057	4035	3848	4035	3974	3748	3434	3028	3450	3175
75	3358	3347	3279	3378	3523	2727	2873	2804	2938	3138
Average	3904.2	3843.2	3373.8	3857.8	3681.8	3488.4	3060.6	2351.6	3065.6	2535.8
No. of Cases	4060	4060	4060	4060	4060	4060	4060	4060	4060	4060
Percentage	96.16%	94.66%	83.10%	95.02%	90.68%	85.92%	75.38%	57.92%	75.51%	62.46%

Table B.8 Summary of the number of correct results when three notes are played together.

	Number of Errors						Percentage						
Inst. No.	1	26	33	41	75	Average	1	26	33	41	75	Average	Rank
Average frequency starting point (using the known note from the same instrument)													
Oct. 2	0	84	8	0	203	59	0%	12%	1%	0%	29%	8%	4
Oct. 3	1	0	8	0	245	51	0%	0%	1%	0%	35%	7%	3
Oct. 4	5	4	26	0	45	16	1%	1%	4%	0%	6%	2%	1
Oct. 5	35	15	47	0	59	31	5%	2%	7%	0%	8%	4%	2
Oct. 6	107	51	69	11	97	67	15%	7%	10%	2%	14%	9%	5
Lower frequency starting point (using the known note from the same instrument)													
Oct. 2	0	1	3	0	66	14	0%	0%	0%	0%	9%	2%	2
Oct. 3	0	0	2	0	117	24	0%	0%	0%	0%	17%	3%	5
Oct. 4	3	3	5	3	68	16	0%	0%	1%	0%	10%	2%	4
Oct. 5	2	1	2	0	73	16	0%	0%	0%	0%	10%	2%	3
Oct. 6	11	0	2	4	3	4	2%	0%	0%	1%	0%	1%	1
Average frequency starting point (using the known note from 5 different instruments)													
Oct. 2	40	172	65	49	210	107	6%	24%	9%	7%	30%	15%	4
Oct. 3	38	27	114	44	275	100	5%	4%	16%	6%	39%	14%	2
Oct. 4	42	29	146	95	104	83	6%	4%	21%	13%	15%	12%	1
Oct. 5	119	44	175	81	106	105	17%	6%	25%	11%	15%	15%	3
Oct. 6	247	69	186	133	169	161	35%	10%	26%	19%	24%	23%	5
Lower frequency starting point (using the known note from 5 different instruments)													
Oct. 2	23	8	28	5	90	31	3%	1%	4%	1%	13%	4%	1
Oct. 3	38	15	33	41	140	53	5%	2%	5%	6%	20%	8%	2
Oct. 4	47	26	101	84	122	76	7%	4%	14%	12%	17%	11%	3
Oct. 5	98	120	135	22	113	98	14%	17%	19%	3%	16%	14%	4
Oct. 6	125	150	138	49	69	106	18%	21%	19%	7%	10%	15%	5

Table B.9 Number of Errors occurs using the absolute value method with complete FT results when two notes are played together.

Appendix C

Expected Solution (Frequency in Hz and Note Name)		Results with the Known Notes from						
		Real Instrument		MIDI 1		5 MIDI Instruments		
						Instrument Program Used		
246.94	B3	246.94	B3	123.47	B2	33	246.94	B3
293.66	D4	293.66	D4	293.66	D4	1	293.66	D4
293.66	D4	293.66	D4	293.66	D4	1	293.66	D4
293.66	D4	293.66	D4	293.66	D4	1	293.66	D4
261.63	C4	261.63	C4	523.25	C5	75	261.63	C4
329.63	E4	329.63	E4	329.63	E4	1	329.63	E4
329.63	E4	329.63	E4	329.63	E4	1	329.63	E4
329.63	E4	329.63	E4	329.63	E4	1	329.63	E4
369.99	F#4	369.99	F#4	185	F#3	75	369.99	F#4
369.99	F#4	369.99	F#4	185	F#3	75	369.99	F#4
329.63	E4	329.63	E4	329.63	E4	1	329.63	E4
369.99	F#4	369.99	F#4	185	F#3	75	369.99	F#4
392	G4	392	G4	392	G4	1	392	G4
440	A4	440	A4	440	A4	75	440	A4
493.88	B4	493.88	B4	493.88	B4	1	493.88	B4
246.94	B3	246.94	B3	123.47	B2	33	246.94	B3
293.66	D4	293.66	D4	293.66	D4	1	293.66	D4
293.66	D4	293.66	D4	293.66	D4	1	293.66	D4
293.66	D4	293.66	D4	293.66	D4	1	293.66	D4
261.63	C4	261.63	C4	523.25	C5	1	523.25	C5
329.63	E4	329.63	E4	329.63	E4	1	329.63	E4
329.63	E4	329.63	E4	329.63	E4	1	329.63	E4
329.63	E4	329.63	E4	329.63	E4	1	329.63	E4
369.99	F#4	369.99	F#4	185	F#3	75	369.99	F#4
369.99	F#4	369.99	F#4	185	F#3	75	369.99	F#4
329.63	E4	329.63	E4	329.63	E4	1	329.63	E4
369.99	F#4	369.99	F#4	185	F#3	75	369.99	F#4
392	G4	392	G4	392	G4	1	392	G4
293.66	D4	293.66	D4	293.66	D4	41	293.66	D4
369.99	F#4	392	G4	392	G4	1	392	G4
Number of Correct Answer (in 30 cases)								
		29	96.67%	19	63.33%		28	93.33%

Table C.1 The results using different types of known notes when playing a song on a real grand piano.