A Search and Retrieval Based Approach to Music Score Metadata Analysis

Jamie Gabriel

FACULTY OF ARTS AND SOCIAL SCIENCES
UNIVERSITY OF TECHNOLOGY SYDNEY

A thesis submitted in fulfilment of the requirements for
the degree of Doctor of Philosophy

April 2018

CERTIFICATE OF ORIGINAL AUTHORSHIP

I certify that the work in this thesis has not previously been submitted for a degree nor has it been submitted as part of requirements for a degree except as part of the collaborative doctoral degree and/or fully acknowledged within the text.

I also certify that the thesis has been written by me. Any help that I have received in my research work and the preparation of the thesis itself has been acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

This research is supported by the Australian Government Research Training Program.

Signature: Production Note:
Signature removed prior to publication.

Date: 1/10/2018

ACKNOWLEGEMENT

Undertaking a dissertation that spans such different disciplines has been a hugely challenging endeavour, but I have had the great fortune of meeting some amazing people along the way, who have been so generous with their time and expertise. Thanks especially to Arun Neelakandan and Tony Demitriou for spending hours talking software and web application architecture. Also, thanks to Professor Dominic Verity for his deep insights on mathematics and computer science and helping me understand how to think about this topic in new ways, and to Professor Kelsie Dadd for providing me with some amazing opportunities over the last decade.

On the music side of things, I would like to thank David Smith for sharing his profound musical expertise: our discussions of harmony, melody and voice-leading have been pivotal in my understanding of how jazz and modern orchestral music can come together, and have also deeply influenced the requirements and design of much of my music software. I would also especially like thank Carl Orr for amazing support and endless creativity and giving me new ways to understand music.

I want to acknowledge and thank my supervisor for this dissertation, Professor Mark Evans, who has been absolutely fantastic. He has tirelessly read my unfinished drafts and always patiently put me back on track which I am grateful for. I would also like to thank Dr Liz Giuffre and Professor Ola Stockfelt for reading draft chapters and providing such indispensable feedback and suggestions.

Above all, I want to thank my beautiful wife Paula for all her love and support during this very long journey. I am not sure how you have put up with me during this, but you know that without you I would never have finished. My son Luke and daughter Stefanie have also been amazing and helped me keep some perspective during this whole process, about the things that are truly important. Undertaking a PhD part time has of course taken far longer than I thought it might, but without Paula, Luke and Stefanie (the Ste, Lu and Pa in Stelupa) I would never have gotten to the end.

ABSTRACT

Music metadata is the body of information that music generates, or leaves behind. It is the notes written on an orchestral score by a composer hoping to ensure his or her longevity; a jazz lead sheet or pop music chart that gives musicians basic instructions of what can be played; the informational encoding of bytes on storage devices (such as CDs or MP4 files), that can be used to capture music recordings; the catalogues of information about collections of recordings held by music streaming services.

This thesis will chart the use of this metadata in creating models of music theory and analysis, and its use in creating prescriptive rules around music practice and creation. It will examine new approaches being taken in music score metadata search and retrieval to understand how these might be leveraged in order to allow a rethinking of music score metadata use. Such approaches can reposition music theory and analysis frameworks as sites of dynamic search and retrieval, which can be highly adaptable to an underlying corpus of music scores.

The dissertation features an extended case study demonstrating how such an approach can be applied to ten Keith Jarrett jazz solos that have been transformed into a single large dataset. It will show how this can provide deep insights and new knowledge into Jarrett's improvisational style, and uncover structures that are not possible to find using more traditional models of music analysis.

Reimagining the music score as metadata challenges both how music theory can be understood, and how it can be presented. In responding to this, the dissertation will show how music theory can be viewed as a crowd sourced phenomenon, related to an underlying corpus and other users. To this end it will present a software application, Stelupa, a nuanced search engine to explore music score

metadata, that leverages off many of the features found in other modern music metadata applications such as Spotify and iTunes.

TABLE OF CONTENTS

Certificate of original authorship		i
Acknowledgement		ii
Abstract		iv
Table of contents		vi
List of tables		vii
List of figures		ix
	Introduction	1
Chapter 1	The use of music score metadata in traditional music theory and analysis	11
Chapter 2	Music as a problem of information	51
Chapter 3	Jazz improvisation and the style of Keith Jarrett	93
Chapter 4	Tools and technologies used for the case study	123
Chapter 5	Jazz improvisation analysis case study: Ten Keith Jarrett jazz solos	138
Chapter 6	Conclusion and future work	250
Appendix 1	Transcriptions of Keith Jarrett solos	271
Appendix 2	Notes for software related to this dissertation: Music Metadata Builder, Jupyter Analysis Notebooks and Stelupa	325
Bibliography		326

LIST OF TABLES

4.1	Technologies used in dissertation	123
4.2	Steps for preparing data for case study	135
4.3	Sample record of prepared data	136
5.1	General characteristics of the dataset	144
5.2	Sample record taken from the dataset	145
5.3	Characteristics of pitches (as midi numbers) used in dataset	147
5.4	Counts of different types of durations used in the dataset	151
5.5	Average and median notes per measure and standard deviation, grouped by title	164
5.6	Average amount of notes played in a measure, grouped by chord type and title	165
5.7	Three sample records of phrases found in the dataset	167
5.8	Most commonly occurring phrases described by midi number sequence, ignoring rhythm	168
5.9	Count of phrases in each solo, and percentage of phrase in each measure	169
5.10	General characteristics of phrase length in all solos	171
5.11	Short phrase lengths in the dataset	172
5.12	Phrases over 80 notes in length and commencing measure	177
5.13	Count of different length microphrases that can be constructed from the dataset	201
5.14	Top two-note microphrases with note names and no rhythm	208
5.15	Top five three-note microphrases with note names and no rhythm	208
5.16	Top five four-note microphrases with note names and no rhythm	209

5.17 Top five five-note microphrases with note names and no rhythm 209
5.18 Top five six-note microphrases with note names and no rhythm 209
5.19 Top five seven-note microphrases with note names and no rhythm 210
5.20 Midi number counts with and without durations 221
5.21 Count of microphrases with the midi sequence "77, 75, 74, 72" 223
5.22 Most commonly occurring four-note microphrases ignoring rhythm and transposed to start on middle C 224
5.23 Names of harmonic degrees with an example on the root note C 231
5.24 The use of the flat-seventh on beat 2.5 on a dominant chord 237
5.25 Examples of major seventh being used on a dominant chord 238
5.26 Preparation of the major seventh on a dominant chord 238
5.27 Examples of the sharp ninth being used on a major seventh chord 242
5.28 Examples of the fifth being used on a diminished seventh chord 245
5.29 Cross tabulation of harmonic degrees and position in the measure in which they are used on the dominant seventh chord 246
5.30 Cross tabulation of harmonic degrees and position in the measure in which they are used on the diminished seventh chord 247
5.31 Cross tabulation of harmonic degrees and position in the measure in which they are used on the minor seventh chord 248

LIST OF FIGURES

1.1	Example of data from iTunes Database Search API	2
2.1	Example of two notes encoded in MusicXML	65
2.2	Two element n-gram	75
2.3	Use of midi and audio files in Jazzomat	84
2.4	Discography, chordal progressions, and biography information in Jazzomat	85
2.5	Aggregated statistics of Jazzomat	86
4.1	Transcribe software screenshot	125
4.2	Jupyter notebook screenshot	129
4.3	Example of Music21 and Lilpond rendered score	130
4.4	JSON output from Music Metadata Builder	131
4.5	JSON output from iTunes database	132
4.6	JSON output from Music Metadata Builder (annotated)	133
4.7	Music Metadata Builder Score Visualisation	134
4.8	Excerpt from Stella By Starlight transcription	136
5.1	Original phrase (Days Of Wine And Roses)	139
5.2	Phrase ignoring rhythm (Days Of Wine And Roses)	139
5.3	Phrase transcribed to start on middle C (Days Of Wine And Roses)	139
5.4	Phrase transcribed to start on middle C ignoring rhythm	140
5.5	Phrase and microphrase	141
5.6	Pitch classes used in all solos	148
5.7	Notes used across all solos	149
5.8	Midi numbers used across all solos	150

5.9 Count of different chord roots in all solos 152
5.10 Count of different chord types in all solos 153
5.11 Number of notes played over time measured in seconds (All The Things You Are, Groovin High) 155
5.12 Number of notes played over time measured in seconds (Autumn Leaves) 156
5.13 Number of notes played over time in seconds (If I Were A Bell, In Love In Vain) 157
5.14 Number of notes played over time measured in beats (All The Things You Are, Groovin High) 158
5.15 Number of notes played over time measured in beats (Autumn Leaves) 158
5.16 Number of notes played over time measured in beats (Stella By Starlight, If I Were A Bell) 159
5.17 Number of notes played over time measured in beats (Someday My Prince I Will Come 160
5.18 Count of notes played in each measure (All The Things You Are) 162
5.19 Count of notes played in each measure (If I Were A Bell) 163
5.20 Count of notes played in each measure (Groovin High) 164
5.21 Different phrase lengths in all solos 171
5.22 Melodic phrase excerpt (In Love In Vain) 173
5.23 Different phrase lengths across all solos (All The Things You Are) 174
5.24 Different phrase lengths across all solos (Groovin High) 174

5.25 Different phrase lengths across all solos (Stella By Starlight) 175
5.26 Different phrase lengths across all solos (Someday My Prince Will Come) 175
5.27 Number of notes in phrase vs. commencing measure 176
5.28 Phrase starting locations within measures across all solos 178
5.29 Phrase starting locations within measures 179
5.30 Phrase ending locations within measures across all solos 180
5.31 Phrase ending locations within measures 181
5.32 Melodic phrase excerpt (Days Of Wine And Roses) 182
5.33 Melodic phrase excerpt (Groovin High) 183
5.34 Melodic phrase excerpt (Autumn Leaves) 184
5.35 Melodic phrase excerpt (My Funny Valentine) 184
5.36 Percentage of unique musical frequencies used in phrase in solos 186
5.37 Count of notes in phrase were all pitches are unique 187
5.38 Melodic phrase excerpt (Autumn Leaves) 188
5.39 Melodic phrase excerpt (My Funny Valentine) 188
5.40 Percentage of unique musical frequencies in phrases greater than ten notes 190
5.41 Pitch classes used in melodic phrases in all solos 191
5.42 Pitch classes in in melodic phrases in phrases with more than 20 notes 192
5.43 Pitch classes used in melodic phrases in phrases with more than 40 notes 193

5.44 Pitch classes used in melodic phrases in phrases with more than 60 notes 194
5.45 Percentage of unique musical durations used in phrase 195
5.46 Melodic phrase excerpt (Days Of Wine And Roses) 196
5.47 Comparison of leaps and steps in phrases greater than 40 notes in length 197
5.48 Comparison of positive and negative movements in phrases greater than 40 notes in length 198
5.49 Range measured in semitones 199
5.50 Most commonly occurring eight-note microphrases 201
5.51 Melodic phrase excerpt (Days Of Wine And Roses) 202
5.52 Melodic phrase excerpt (Days Of Wine And Roses) 202
5.53 Melodic phrase excerpt (In Love In Vain) 203
5.54 Melodic phrase excerpt (In Love In Vain) 203
5.55 Melodic phrase excerpt (In Love In Vain) 203
5.56 Melodic phrase excerpt (Groovin High) 204
5.57 Melodic phrase excerpt (Groovin High) 204
5.58 Melodic phrase excerpt (Groovin High) 204
5.59 Most commonly occurring two-note micro-phrases 205
5.60 Most commonly occurring two-note microphrases ignoring rhythm 206
5.61 Most commonly occurring two-note microphrases ignoring rhythm and transposed to start on middle C 207

5.62 Most commonly occurring four-note microphrases ignoring rhythm and transposed to start on middle C 211
5.63 Most commonly occurring five-note microphrases ignoring rhythm and transposed to start on middle C 211
5.64 Most commonly occurring six-note microphrases ignoring rhythm and transposed to start on middle C 212
5.65 Most commonly occurring seven-note microphrases ignoring rhythm and transposed to start on middle C 213
5.66 Most commonly occurring eight-note microphrases ignoring rhythm and transposed to start on middle C 214
5.67 Melodic phrase excerpt (Stella By Starlight) 215
5.68 Melodic phrase excerpt (Stella By Starlight) 216
5.69 Melodic phrase excerpt (Days Of Wine And Roses) 216
5.70 Melodic phrase excerpt (Days Of Wine And Roses) 216
5.71 Melodic phrase excerpt (Days Of Wine And Roses) 217
5.72 Melodic phrase excerpt (Autumn Leaves) 217
5.73 Melodic phrase excerpt (If I Were A Bell) 217
5.74 Melodic phrase excerpt (In Love In Vain) 218
5.75 Melodic phrase excerpt (My Funny Valentine) 218
5.76 Melodic phrase excerpt (Stella By Starlight) 218
5.77 Decision tree for the probability of choosing a note given a D5 has just been played 219
5.78 All possible outcomes following the note D5 220

5.79 All possible outcomes following the note C4 (middle C) 221
5.80 Melodic phrase excerpt (Someday My Prince Will Come) 222
5.81 Melodic phrase excerpt (My Funny Valentine) 222
5.82 Melodic phrase excerpt (Days Of Wine And Roses) 224
5.83 Melodic phrase excerpt (My Funny Valentine) 225
5.84 Melodic phrase excerpt (If I Were A Bell) 225
5.85 Melodic phrase excerpt (Days Of Wine And Roses) 225
5.86 Melodic phrase excerpt (Days Of Wine And Roses) 226
5.87 Melodic phrase excerpt (Days Of Wine And Roses) 226
5.88 Melodic phrase excerpt (Days Of Wine And Roses) 226
5.89 Melodic phrase excerpt (Days Of Wine And Roses) 227
5.90 Melodic phrase excerpt (My Funny Valentine) 227
5.91 Melodic phrase excerpt (All The Things You Are) 227
5.92 Melodic phrase excerpt (In Love In Vain) 228
5.93 Melodic phrase excerpt (In Love In Vain) 228
5.94 Melodic phrase excerpt (If I Were A Bell) 228
5.95 All possible outcomes following the note sequence G4 - Bb4 229
5.96 All possible outcomes following the note sequence C4 - Eb4 230
5.97 Different chord types used across all solos 232
5.98 Different harmonic degrees used on dominant chords across all solos 233

5.99 Resolution of the flat-seventh in the dominant chord 235
5.100 Resolution down one semitone of the flat-seventh in the dominant chord 236
5.101 Different harmonic degrees used on major seventh chords across all solos 240
5.102 Resolution of the sharp ninth in the major seventh chord 241
5.103 Different harmonic degrees used across all solos 243
5.104 Resolution of the fifth in the diminished seventh chord 244
6.1 Spotify discovery visualisation 253
6.2 Stelupa landing page 256
6.3 Search panes of application 257
6.4 World filtering metadata 258
6.5 Range filtering metadata 259
6.6 More nuanced searching 260
6.7 Phrase sequence searching using a sunburst partition 261
6.8 Piano-roll visualisation to render results 262
6.9 Pinning a result in the style of Pinterest 263
6.10 Building collections 263
6.11 Annotation a pinned excerpt 264
6.12 The built in Javascript synth 265
6.13 Stelupa Data API 266
6.14 Searching for the data in the Stelupa Data API 267

Introduction

The different kinds of data that can be derived from music are far reaching and ubiquitous. In this dissertation, I will refer to this data, information that can be drawn from music, as *music metadata*. Music metadata is best thought of as the body of information that music generates, or leaves behind: the notes written on an orchestral score by a composer hoping to ensure his or her longevity; a jazz lead sheet or pop music chart that gives musicians basic instructions of what can be played, which assumes domain specific knowledge; the informational encoding of bytes on storage devices (such as CDs or MP4 files), that can be used to capture music recordings; or the catalogues of information about collections of recordings held by music streaming services.

Music metadata surrounds us. In an increasingly networked world, it is this metadata that can inform and dictate our interactions with music itself. Examples of this include the kind of metadata found in services such as a Google Play, Spotify or iTunes, each of which are vast databases holding information about music. Figure 1 demonstrates an example of this kind metadata, the details of a single song (out of an estimated 40 million songs) held on iTunes (https://affiliate.itunes.apple.com/resources/documentation/itunes-store-web-service-search-api/#searchexamples/(2018)).

Figure 1.1. Example of data from a song in the iTunes Database Search API


```
{"wrapperType":"track",
 "kind":"song",
 "artistId":909253,
 "collectionId":120954021,
 "trackId":120954025,
 "artistName":"Jack Johnson",
 "collectionName":"Sing-a-Longs and Lullabies for the Film Curious George",
 "trackName":"Upside Down",
 "collectionCensoredName":"Sing-a-Longs and Lullabies for the Film Curious George",
 "trackCensoredName":"Upside Down",
 "artistViewUrl":"https://itunes.apple.com/WebObjects/MZStore.woa/wa/viewArtist?id=909253",
 "collectionViewUrl":"https://itunes.apple.com/WebObjects/MZStore.woa/wa/viewAlbum?i=120954025&id
=120954021&s=143441",
 "trackViewUrl":"https://itunes.apple.com/WebObjects/MZStore.woa/wa/viewAlbum?i=120954025&id=1209
54021&s=143441",
 "previewUrl":"http://a1099.itunes.apple.com/r10/Music/f9/54/43/mzi.gqvqlvcq.aac.p.m4p",
 "artworkUrl60":"http://a1.itunes.apple.com/r10/Music/3b/6a/33/mzi.qzdqwsel.60x60-50.jpg",
 "artworkUrl100":"http://a1.itunes.apple.com/r10/Music/3b/6a/33/mzi.qzdqwsel.100x100-75.jpg",
 "collectionPrice":10.99,
 "trackPrice":0.99,
 "collectionExplicitness":"notExplicit",
 "trackExplicitness":"notExplicit",
 "discCount":1,
 "discNumber":1,
 "trackCount":14,
 "trackNumber":1,
 "trackTimeMillis":210743,
 "country":"USA",
 "currency":"USD",
 "primaryGenreName":"Rock"}
```


Within this context, music metadata is focused primarily towards listeners. It can provide them with a body of information to facilitate search and retrieval tasks, and allow listeners to iterate through a vast body of information about music in order to easily find the things they wish to hear. Data analytics and machine learning techniques can be applied to this type of information to explore listening patterns and make inferences about personal tastes.

But there are other types of music metadata too. One with a far longer history, is the music score. Music score metadata is comprised of the dots on a page whose purpose is to provide instructions about how music should be performed (which will be the working definition this thesis adopts going forward). At its most basic level, music score metadata provides time-series information about the playing of sound, specifying the pitch of notes, the time at which these notes should be played, and how long they should be played.

Some metadata from music scores can be highly elaborate, containing far more information than others. In a Mahler orchestral score, for example, there are highly prescriptive instructions provided for the players to faithfully execute the composer's intentions in as accurate manner as possible. In other music scores, things are more minimalistic. The figured bass notations found in the Baroque period used numbers to indicate the characteristics of harmony, yet left the specific choices of chord voicings to the performer's discretion. In jazz, rock and folk settings, the music score is often little more than a signpost. Much of the the rhythmic nuance and harmonic complexity is left aside, and there is an implicit assumption that the domain specific knowledge of performers will ensure that music is interpreted in a way that is appropriate to the genre.

The example of iTunes metadata is certainly different from the music score metadata. The former is orientated predominantly towards the facilitating of curated listening, to expedite the search and retrieval of audio files. The latter makes implicit assumptions of additional domain specific knowledge, such as the ability to play an instrument, to understand how the dots on the page relate to the pitches that are playable on an instrument, and to grasp the expectations of the stylistic idioms particular to the music under consideration.

In this dissertation, my intent is not to provide an overarching theory of music or analytical model. I am not setting out to transform music score metadata into a specific set of rules that might explain how music can or should be constructed within any or all genres and periods. Examining the long history of music analysis reveals countless examples that seek to do precisely this. Across multiple disciplines, it is certainly possible (and I will survey many of these in chapter one) to find so called laws of music, prescriptive codifications of best practices, or repositories of melodic structures and ideal chord progressions, that together might typify what should and should not happen in music.

These attempts, however, are problematic: while they can partially capture the characteristics of how music practices in a given time and place tend to operate, they are often confounded by the exceptions. Though patterns can certainly (and easily) be found in music, its creation is often a process of transgression: one generation's dissonance and is another's consonance. The accepted norms of how music should be structured can change radically over time. As such, I will employ an alternate approach and present a framework for analysis designed to be robust enough to allow for a music theory that can be customised to a given corpus of music in a given genre and flexible enough to be changed over time.

From the very outset of this dissertation, I want to emphasis that music score metadata (and music metadata more generally) is qualitatively different from music itself. Across the literature, the music score and music often become conflated into the same entity. However the dots on the page of a music score capture almost none of the nuance of music itself. They are nothing more than a log of information, an attempt to catalog sonic events that take place over time. This dissertation will show that, despite this limitation, analysing music score metadata and creating a framework for its interactive exploration can still provide deep insights into our understanding of music.

The idea of approaching the problem of music analysis within a metadata search and retrieval framework is also a reflection of my own experiences of studying music over a long period. When I first started to reason about how music worked, over thirty years ago now, what struck me most was its apparent logic. It seemed to be full of patterns, structures and symmetries that made intuitive sense to me as listener. At that time, (and, perhaps overly idealistically), music seemed to me to be a kind of mathematics, whose foundational elements were more sophisticated than numbers, in that their meaning could transform depending on time and context.

In setting out to study how it was that music worked, I followed a fairly conventional and well-trodden path: instrumental study, score analysis, harmony, counterpoint, voice-leading, and orchestration. I examined the notes on music scores, trying to come to an understanding of why they appeared, why a composer might choose one note instead of another. I wanted to understand how dissonance was created and resolved, how different approaches to modulation took place in different periods, how voice-leading could function as a mechanism to allow movement between unrelated tonalities. And as a typical student of music, I approached this, as if, underneath it all, there was some kind of formula that might explain things.

Later I become more interested in jazz: the complexity of its melodies and harmonies seemed to have far reaching implications for what I understood from score based analysis. But in the course of studying jazz (and especially when transcribing jazz solos) my question was the same: I wanted to understand why an improvisor played one note rather than another, as if there was an underlying reason that could be found.

Reflecting on this experience of wanting to know how music works has led to two primary outcomes, both of which have been unexpected. First, (at least in the context of composing or improvising music) the burning question of how music works has become largely unimportant. A consequence of spending so much time listening to and studying music has led to the development of strong intuitions around knowing what notes are the most appropriate to play and when to play them. This is particularly telling for jazz improvisation: when I first started to try and play jazz I was frustrated that it did not seem to sound like the jazz I was hearing: it sounded contrived and unconvincing. But the process of transcribing and learning to play so many jazz solos, and memorising so many jazz standards eventually allowed me to converse in the structures appropriate to this style of music. My experience as a jazz musician (which is similar to other musicians I have spoken to), is that, eventually you have no idea which note or chord you are playing. It just sounds appropriate, and you have developed a deep enough intuition to know what should happen next, whether playing in an ensemble or solo context. So in answer to the question as why a composer might choose one note rather than another, or why a jazz musician chose one chord voicing as opposed to another, is simply because it is what the situation calls for, a logical outcome of that musician's taste and expertise emanating from domain specific knowledge accumulated over time.

The second outcome was the realisation that my study of music, and the process of becoming a musician, was something that could, above all, be characterised as a problem of the search and retrieval of metadata. When I set out to examine music scores, or transcribe solos, or to find particular locations in recordings that highlighted composers utilising different techniques, the biggest problem I faced was the difficulty of finding things. Coming across, for example, a particular brass section chord voicing in a Mahler symphony that seemed atypical turns out to be profoundly difficult problem to explore in terms of information retrieval. It requires manually searching through very large scores for similar things that might be in different keys. Such an exploration could lead to further, difficult questions: I may want to explore if the chord voicing only appears in certain tempos, or examine if it is indicative of Mahler's early career or late career. To allow this kind of examination, what is needed is the ability to easily search across a wide corpus of scores (and related information to those scores) in order to ascertain if this was something idiomatic of a particular style or even a geographical location, or original to a particular composer.

Of course, this experience is both highly personal and limited to a very small subset of Western music and post-bebop jazz study. But regardless of how one approaches the study music, the problem of finding information to enable this study is a profound one.

In this dissertation I will address these problems by reframing the metadata of the music score in order to demonstrate how locating it within a search and retrieval framework can inform new insights into music analysis. The information found on the music score will be reimagined as a site of scaleable metadata that can be easily interrogated, and one that is optimised for exploratory investigation regardless of the genre under consideration. Specifically, the research question I will address is as follows:

> Can adopting a search and retrieval based approach to music score metadata change the way music theory and analysis is practiced?

Dissertation Structure

This dissertation contains six chapters. Chapter one will examine the history of music analysis, framing it as a problem of metadata, and exploring how various disciplines have sought to extract information from music to gain a deeper understanding of it. I will explore how music is understood through the lens of those fields which have explored it in different ways, such as music analysis and theory, musicology, semiotics, psychology, mathematics and statistics. The approaches taken by these fields can vary greatly, yet they have all sought to convert the sonic nature of music into some kind of body of information or metadata that is amenable to different styles of interrogation. This chapter will begin by examining approaches taken to understand musical thought and practice in Ancient Greece, and culminate in the numerous schools of thought that arose throughout the twentieth century. It will demonstrate that what characterises music, above all else, is a lack of consensus around the way it is examined.

More recently, much exploration of music metadata has been taking place in the field of Music Information Retrieval (MIR). MIR is a relatively recent discipline, having its first academic conference in Plymouth, Massachusetts in 2000, and has sought to fuse together ideas from music theory, computer science, psychology, neuroscience, library science, electrical engineering, machine learning, information theory, and digital signal analysis. Though there has been relatively limited work in MIR regarding the use of music score metadata, its approaches can be utilised to understand how to frame music score metadata as a problem of information. Chapter two will explore how this field positions music as a problem of information retrieval, locating its origins in the twentieth century relationship between music theory and information theory. The field has a particular interest in music metadata, but rather than being focused on music score metadata, it has often explored different music types of metadata, such as the kind of data that is used to inform products such as Pandora, Spotify, and Shazam, which heavily utilise search and retrieval methods. The purpose of this chapter will be to position the music score as a scalable metadata, and demonstrate how existing MIR approaches to data might be leveraged off to achieve this.

One particularly disruptive idea whose origins can be located in MIR is that knowledge about music, rather than being curated by an expert, can be aggregated through crowd sourced data. Spotify, for example, utilises recommender systems and machine learning approaches that allow the views of the many to be aggregated into individual recommendations. I am particularly interested in applying this idea to music analysis, and this dissertation will explore how music analysis might be mediated by crowd sourced focused technologies, allowing music theory to be customised for specific users.

The case study to be undertaken in this dissertation will explore jazz improvisation practice and, to this end, chapter three will examine issues relating to the analysis of jazz improvisation. This chapter will explore some of the profound challenges that have arisen in the analysis of jazz improvisation, which fuses highly complicated melodic and harmonic structures with a lack of availability of music score metadata. Even defining what jazz improvisation is can be notoriously difficult, and any definition seems heavily dependent on its proponents at different times, highlighting how diffuse the process of analysis can become. I will also provide some specific background on the metadata to be used in the case study, taken from transcriptions of ten improvised solos of by jazz pianist Keith Jarrett.

Chapter four will provide a methodology for the case study, and will outline the different software applications I have created to be used to undertake the search and retrieval of music score metadata. The chapter will also provide some background on the process by which the jazz transcriptions were prepared for the case study, and provide a summary of the various tools and technologies that were be used to facilitate the creation of a search and retrieval method framework.

Chapter five will undertake a case study to explore music score metadata in jazz improvisation. Keith Jarrett has been chosen as the subject of the case study as he poses a profound problem for music analysis: there is virtually no repetition in his playing (in that exact melodic phrases almost never appear twice). Of the ten solos that I will explore in the case study (comprised of over 15,000 notes) no melodic phrase appears more than once. Applying more traditional models of analysis (such as exploring what scales Jarrett might employ, or what " jazz licks" he employs) does not make sense due to the lack of repetitive structures found in his playing. The chapter will demonstrate how a search and retrieval approach can allow deep insights into the nature of the improvised solos. Additionally, Jarrett's playing has had almost no analysis carried out on it (examples include Strange 2003, Terefenko 2009, and Page 2009) and this chapter will also be used to provide a new insights into his improvisational style. Finally, the chapter will seek to demonstrate how any theory of music must be tied to a particular corpus and is dependent on this corpus for its evidence base.

Chapter six will examine possible future work around a search and retrieval approach to music metadata. It will present a proof of concept open source web application, Stelupa, a music score search engine that can be used for scaleable music metadata exploration. It will show how filters can be applied in a multi-modal networked environment to locate specific musical structures, and demonstrate how this exploration might be linked to audio representations and multiple data visualisations and track the behaviour of users. It will provide a framework from which a crowd-sourced theory of music could be derived and capture it evolving over time.

Chapter 1
The use of music score metadata in traditional music analysis

The history of music theory and music analysis can be characterised, above all, by disagreement. Yet underneath the lack of consensus is a powerful consistency, the implicit belief, borne out by practice, that it is possible for information or metadata to be drawn from music, and used to make inferences about its meaning.

In this chapter I will provide a historical summary of music theory and music analysis. I will begin by focusing on the texts of antiquity and trace this lineage through to works found in the twentieth century. The investigation will be limited to those historical writings about music that utilise metadata from music: the treatises, frameworks, commentaries, and pattern analysis of musical works. These are the works that overwhelmingly draw metadata from music in the form of information extraction from music scores.

One of the challenges in exploring the history of music theory and music analysis is locating where these fields begin and end (particularly before the mid nineteenth century). As such, this chapter will cover works found in the fields of aesthetics, philosophy, the natural sciences, music psychology, music semiotics, and musicology. The disciplines of mathematics and statistics also have strong connections to the search for models of music design and analysis, however this discussion will be deferred to the next chapter because of their relationship to the field of music information retrieval. At the same time, I will exclude those works that explore the nature of sound exclusively, without reference to specific musical works or practices.

The earliest Western record of music theory can be found in Greek antiquity (West, 1992, p. 1). Pythagoras (570 - 495 BCE) wrote about how frequency (or pitch) was used in music practice by conducting explorations into the nature of both consonance and dissonance. He explored how the frequency of a sound could be altered depending on the size of a vibrating physical phenomena (such as the plucking a string of different lengths). He also discovered that changing the length of a string using simple ratios (i.e. 2:1 or 3:2) would produce frequencies that could be regarded as consonant with each other, based on the subjective view of consonance and dissonance at the time. From making these observations, Pythagoras is credited with the discovery of the first diatonic scale (a set of notes between an octave whose relationship could be characterised by simple numerical ratios). The idea of this set of notes that each had a relationship to each other would go on to to have a profound influence in the creation of Western music (Joost-Gaugier, 2009, p. 13).

Whereas Pythagoras' observations were focused on the nature of sound, it was Aristoxenus of Tarentum (375-360 BCE), who would develop the first substantial surviving work of music theory. He extracted and explored information about musical practices of the time and this work is one of the first examples of the use of music metadata as defined by this dissertation. Aristoxenus was the son of a musician and follower of Aristotle. Though the fragmented nature of his surviving writings make it difficult to piece together a clear picture of his overall theory (Gibson, 2005, p.11), his aim was to rationalise the musical thinking and practices of the time. Aristoxenus explored pitch, rhythm and melody as separate musical attributes that could each be altered to create variation in a music performance (Gibson, 2005, p. 44). He also catalogued a summary of techniques that musicians of the time were utilising, though he stopped short of putting forward an overarching theory of music.

Later, in the third century AD, Aristides Quintilianus wrote a more comprehensive text, consisting of three volumes, entitled On Music. Unlike the writings on music before it, much of this text has survived and the work is regarded as the first treatise on music theory (West, 1992, p. 11). The first volume of On Music explored the place of music in relation to other disciplines being explored at the time, (such as mathematics and philosophy) as well as technical aspects of music and the way in which it was practiced; the second volume examined the relationship between ethics and the human soul; and the third volume explored music and its wider relationship to the cosmos.

The works of Aristides Quintilianus' treatise sought to present a thorough account of the music practices of the time, and connect this to a deeper spiritual meaning. It set out to provide an "overarching vision of the divine order of things", which could elaborate, "the divine source of musical structures in their three major instantiations: in the audible music of human practice, in the soul, and in the natural universe at large" (quoted in Barker, 1984, p. 392). Aristides Quintilianus also noted the complicated relationship human beings have to music, as a phenomenon that elegantly manifests itself both in the physical world, and within consciousness.

It was not until much later however, that these ancient texts started to grow in influence. It was during the second half of the fourteenth century that interest in them grew markedly as part of the humanistic revival. Music theory texts of the ancient Greek world became the subject of interest across western Europe (West, 1992, p.5). There emerged (particularly in Italy) a "mania for music theory" (Giger & Mathiesen, 2003, p. 8). There was a growing fascination with uncovering theoretical truths that might explain the relationship between music practice and the apparent patterns that could be seen in the information, or metadata, that could be derived from music.

At this time, the disciplines of music theory and music analysis were still loosely defined, however their exploration had begun to take place within a wider epistemological framework that sat uneasily between rationalism and empiricism (Christensen, 2002, p. 21). The logic was that if musical works were to be analysed and understood through a rationalist lens, it followed that they could be viewed as being comprised of building blocks which could be formed into more complex structures. The creation of music could then be understood within a modular, theoretical framework. Music could produce information that could be analysed and recreated based on an analytical model. The alternate, empirical view, was that music could not be understood without understanding the complexities of the human experience.

The music theory treatises of this time had also started to address more practical concerns, such as the problems composers and instrumentalists faced when plying their craft. Marchetto of Padua, (fl. 1309-18) produced two influential music treatises, Lucidarium in arte musice plane, and the Pomerium in arte musice mensurate which addressed a range of practical issues such as notating rhythmic values, interval measurement, and the ideal tuning of chromatic intervals. Marchetto appealed to Aristotelian systematics, but presented this as a very different application, exploring music or ‘modulated sound’ and also isolated timbres as a way to discuss the ‘genus’ of notes discoverable in the overtone series. Another theorist of this period was Franchino Gaffurio (1451-1522), also a well known composer at the time. Gaffurio produced three major works of music theory and analysis, Theorica musice (1492), Practica musice (1496), and De harmonia musicorum instrumentorum opus (1518) which explored topics such as tempo, rhythmic notation, vocal polyphony, and counterpoint, by extracting information from music scores.

While these early writings did not yet seek to present a comprehensive model of analysis or provide full blown theories of music, they nevertheless took a metadata driven approach. They extracted information from music scores and merged this with other available contextual information to understand music's meaning.

One of the challenges faced by many music theorists at this time was simply keeping up with the the rapid pace at which music practice and music pedagogy had started to evolve. Already during Gaffurio's lifetime, the printing press has become a viable vehicle from which to produce musical manuscripts. The paper based score provided a powerful way to compress the information of music, store it, and allow its distribution. Music scores were becoming far more available than ever before (Christensen, 2002, p. 33) and could now be explored to examine evolving music practices. Accompanying this was a marked growth in music education and increasing access to musical instruments. What was possible in music (both from a composer and instrumentalist standpoint) was being reinvented at a rapid pace.

The pace of change held steady through the the seventeenth century too, and by this time far more diversity could be seen in texts of music theory and music analysis. There was still an emphasis on instructional works, such as Thomas Campion's A New Way of Making Fowre Parts in Counterpoint by a Most Familiar, and Infallible Rule produced in 1618. This highly prescriptive work drew numerous examples from music scores as a vehicle by which to provide a rigid sets of rules to ensure music was created with appropriate care and skill (according to Campion at least). Campion's work set out to show that, if one followed some simple rules, well formed bass lines and harmonic progressions to emulate the popular songs of the day would follow. Campion's work aimed to present its readers with a formula, passed to the reader by extracting score metadata, that could then be relied upon for creating music of quality.

In the same year, Rene Descartes published Musicae Compendium. Adopting a radically different approach to that of Campion, Descartes presented a rationalist treatment of how intervals might be measured, demonstrating the geometric relationships that could be found in musical works and music practice. His inquiry drew some similar findings to the writings of Pythagoras, though Descartes's attempt can be located as part of a much wider project to integrate geometry and algebra and use mathematical tools to explain worldly phenomena, in this case music.

Descartes was not seeking to explain the musical works of the time, or provide an insight into music practices. He was instead aiming to demonstrate that, regardless of how musical works and performance might evolve, they could still be grounded in certain universal norms that were susceptible to mathematical investigation, and even conducive to an overarching model. DeMarco notes that music, for Descartes, was, "as it were, frozen mathematics, a kind of congealed intelligibility" (quoted in Sweetman, 1999, p. 22).

Positioning the complexities of music as a future conquest for mathematicians was not unique to Descartes. Leibniz (1646-1716) shared the belief, claiming that beauty of music could be found "only in the agreement of numbers and in counting, which we do not perceive but which the soul nevertheless continues to carry out" (quoted in Sweetman, 1999, p. 18). This idea has a powerful lineage that can be seen in many later texts, for example the Mathematical Basis For the Arts, (1948) by Joseph Schillinger, who claimed that, at some yet to be determined point in the future, there would be a “logical end of [to the study of] music… as physiology becomes a branch of electrical engineering in the study of brain functioning, and aesthetics becomes a branch of mathematics” (quoted Sweetman, 1999, p. 39). The idea that the creation of music might be susceptible to mathematical models is a powerful one, and will be explored more in the next chapter.

The occult philosopher, Robert Fludd (1574-1637), also wrote about music theory in the early seventeenth century, as part of his wider writings on cosmology. Fludd provided yet another variation on the meaning of music compared to that of Descartes and Campion. His intended audience was not practicing musicians however, and he rejected the tenets of Cartesian rationalism. On the nature of harmony, Fludd claimed:

> As one string moves to another tuned to the same or a consonant note, so the jewels which are replete with the nature of the sun, may be moved by the sound of the voice of man, if he knows the true sound of Apollo. (cited in Amman, 1967, p.33)

Exploring similar themes in 1650, Athanasius Kircher presented the Musurgia Universalis, a work in which dissonance and consonance were presented as being in deep connection to the functioning of the harmonic balance found in the universe at large. The text included richly detailed images of the notation of birdsong, a summary of existing instruments in use, and extensive references to Greek mythology. Again, information was taken from music scores to make inferences about their meaning.

Though such texts may seem anachronistic with the benefit of a more contemporary gaze, the theories they presented were both widely disseminated and deeply influential. Bach and Beethoven, for example, both regarded Kirchner's work as providing a deep insight into the meaning of music (Christensen, 2002, p. 21). These types of treatises (which included many examples of music scores) also demonstrate the somewhat ambiguous nature of music theory at the time, in which music and the music score had become conflated, and whose study moved between both "sensible and suprasensible" domains (Christensen, 2002, p. 133). The discipline of music theory in the seventeenth century could be variously located in rationalism, empiricism, and mysticism.

The practical problems of how instruments should be tuned, and to what exact frequency, was also of growing interest during this time, and increasingly permeated music texts. In 1636, Marin Mersenne wrote Harmonie Universelle, in which he utilised a Pythagorean conception of music to demonstrate the ideal tunings of instruments. Mersenne derived a formula to generate equally tempered semitones, and his work came to be particularly influential on the instrumentalists of the time, especially in France (Shirali, 2013, p. 228). Mersenne's work was also indicative of the changing approaches being adopted in music theory (Shirali, 2013, p. 230): in mid life he had moved away from the speculative approaches used by Fludd and Kirchner to embrace a mathematical methodology, driven by practical necessities of music performance. Whatever the meaning of music might be, it seemed more closely related to mathematics and rationalism that empiricism and mysticism.

The music theorists of the seventeenth century who were responding to the practical problems faced by composers and performers, were also beginning to face another challenge: trying to account for the ever increasing availability of music scores. The circulation of music scores had by this time become prolific, making this early form of music metadata increasingly available. The problematic duality between music score and music itself would come much later, and at this time the music score offered a highly convenient way both of storing and analysing music, and, for the theorists, to derive laws to inform meaning.

Another complication faced by music theorists was the growing complexity of both musical works and instrumental techniques. Music theorists who were writing pedagogically oriented treatises were required to deliver increasingly complex explanations that could account for both the changes in music practice, and the new techniques used by composers. Harmony and voice leading in particular, had become more complicated. Christensen notes of the period that, "more and more energy seemed to be devoted to systematising and regulating the parameters of a rapidly changing musical practice and poetics" (Christensen, 2002, p. 22).

For theorists, the manual problem of search and retrieval had begun: theoretical works began to take the form of exhaustive catalogues of minutiae, and in depth treatises appeared that could equip musicians with long lists of what they should and should not do in an increasing list of musical situations. The examination of turning systems and the nature of sound had by this time moved away from the discipline of the music theory toward the natural sciences, becoming more concerned with the "pedantic" study of intervals and tuning systems (Christensen, 2002, p. 40) and music analysis had become increasingly focused on score analysis.

By the close of the seventeenth century, the increased availability of music scores as a vehicle of convenience upon which analysis could take place, along with the multitude of new instrumental techniques, saddled the discipline of music theory with an unexpectedly modern problem: an overload of information. Music theorists seeking to encode music practices had to contend with the very practical problem of iterating through an increasing amount of data, much of which was disruptive to existing beliefs regarding the nature of musical works and performance.

Despite the difficulties faced by the music theorists of the seventeenth century, music practice and composition was enjoying a period of rapid growth, and this was the era that would go on to prove so influential on modern Western music (Atcherson, 2001, p. 4). By this time, the Baroque style of music had been deeply embedded, only starting to decline in the early to mid-eighteenth century. Composers such as Bach, Handel, Rameau, Scarlatti and Telemann were producing works of growing complexity that showcased new techniques of modulation, voice leading, and leveraging off an increased consensus in the tuning and construction of instruments (Wang, 2011, p. 23). The models of counterpoint seen in the medieval period had given way to a new conception of harmony and new explanations were sought by music theorists, composers and instrumentalists.

One of the most enduring musical treatises written around this time was The Study of Counterpoint (1725) by Joseph Fux. In the opening pages of this treatise, Fux laments the the declining quality of the music compositions of the time. In setting out to remedy such a state of affairs, Fux promises to equip readers with a series of rules that can be used to ensure that music can be correctly constructed. Regarding the state of music treatises in Vienna, Fux claimed that, although there was an “abundance of works on the theory of music” most of these “have said very little, and this little is not easily understood”. Fux’ agenda aimed to present a right way to do things, and an excerpt taken from the second chapter of the text is typical of the style of presentation to be found in the work.

> The second species results when two half notes are set against a whole note. The first of them comes on the downbeat and must always be consonant; the second comes on the upbeat, and may be dissonant if it moves from the preceding note and to the following note stepwise. However if it moves by a skip it must constant. If [in] this species a dissonance may not occur, except by diminution, i.e., by filling out the space between the two notes that are a third distance from each other. (Fux, 1725 (ed. 1965), p. 23)

In presenting the reader with a highly prescriptive set of instructions, Fux recast the process of music composition as something that was either correct or in need of correction. In providing a rigorous set of rules, Fux' intention was not to present a scientific work however. He was instead leveraging his own, extensive knowledge of the craft of composition, which had been endorsed by many of his contemporaries, to address practical shortcomings in the way compositions of his time were being constructed. His music treatise is a forerunner of both the tone and approach that characterises so many of the later works of music theory. The writing is not grounded in science or logic, yet has the tone of scientific rationalism. The subject matter is presented as highly technical, as if a theory is being presented, and the author is positioned as the technical expert who can decide on the artistic merit of a musical work.

Of course, others did not always concur with Fux' expertise. Reflecting on the approach used by his father, C.P.E Bach (Clarke, 1997, p. 57) claimed that the early species of Fuxian counterpoint were not at all useful and it was far more valuable to provide students and amateur musicians with tasks that were of more practical value in the pursuit of music skills and knowledge. The approach championed by Bach had students commence by learning four-part thorough bass, then chorales, and then move through a series of exercises to add one of the four parts.

Another important music theorist at this time was Jean-Philippe Rameau (1683-1764). Rameau is still regarded as one of the most historically important music theorists (Girdlestone, 1969. p. 23) and his theory of fundamental bass can still be found in many modern composition pedagogy programs (Girdlestone, 1969, p. 18).

The most striking difference between Rameau's music theory and that of his predecessors was his treatment of dissonance. Before Rameau, the fundamental chord (for example, the C major chord in the key of C major) was regarded as the most important building block of music composition. Rameau presented a radically alternate view, elevating the status of the dominant seventh chord as the most important harmonic structure that can be used to explain music (for example a G dominant seventh chord in the key of C major).

Rameau's claim was radical for its time, and led to a conclusion that consonance is a product and outcome of dissonance, rather than dissonance being a product of consonance. Christensen claims Rameau's conception of dissonance is the most important feature of his entire theory (Christensen, 2002, p. 144). Though seemingly subtle, it is a view of dissonance that prevails in so many later music theory texts, which often demonstrated how the dominant chord could be used as a means of modulating to different tonal centres.

Though Rameau was regarded as a rigidly deductive thinker, his approach to music theory was tempered, like Fux, by his own taste as composer. His Nouveau

Système presented a structured view of how harmony should be used, but he also noted that the final choice should not be driven by rationalism alone: “At least this is what the ear decides, and no further proof is necessary” (Christensen, 2004, p. 96). Christensen comments that this approach undermined Rameau’s wider project:

> What is striking is not just the peremptory and final appeal to the ear, but the fact that if the principle of interchangeability is to be taken seriously then much of the apparatus of generation becomes redundant. (Christensen, 2002, p. 222)

Rameau’s view of music theory is one in which the artist dominates nature and any theory must be subordinate to the needs of an artwork which can be understood by the expert composer. Though the study of the construction of musical works may reveal patterns and techniques which can be reused to construct new works, the final choice of notes in a musical work is above all, to be found in the domain of artistic taste.

Rameau’s view lays bare an enduring problem in music theory: rather than being the result of a scientific application of general principles, the construction of musical works is driven by taste in a particular time period. The intuitive expertise of figures such as Rameau and Fux (backed up by their reputation as experts in the field) allowed them to present a mechanical system of rules that others might use, who had little of the expert’s knowledge. It is a pragmatic approach to theory, and foreshadows the model that is so prevalent in music instructional texts of the modern era. Amateur musicians (and even in Rameau's time, there was a growing market of amateur musicians) were presented with a formula for music creation that could be trusted as it had been devised by an expert.

A similar approach to that of Rameau can be seen in the work of Leonhard Euler who published An attempt at a new theory of music, exposed in all clearness according to the most well-founded principles of harmony in 1739. Euler sought to provide a mathematical basis for music. His agenda was ambitious, aiming not only to explain the music of his own time, but also the music of the future. Like Rameau, Euler problematised dissonance, but went about this in a different way. He rejected the idea that consonance and dissonance were discrete states, re-imagining them as highly stratified structures.

Euler faced a similar problem to Fux and Rameau however, when it came to how to account for human taste. In responding to this problem he adopted a similar position to Rameau:

> The musician must act like the architect who, worrying little about the bad judgements which the ignorant multitudes pass on the buildings, builds according to unquestionable laws based on nature, and is satisfied with the approval of the people who are enlightened in this matter. (quoted in LaRue, 1966, p. 33)

This notion of the composer (or an elite group of experts) as the ultimate judge of the quality of an artwork, and the consequently subordinate place of music theory, is the powerful and enduring legacy that begins to emerge from the time of Rameau and Fux. Over a century later, Schoenberg would take up this same theme, yet far more aggressively in his Theory of Harmony (1910) defending the role of the artist, and demand music theory speak to directly to works of art rather than its own end:

> And the theorist, who is not usually an artist, or is a bad one (which means the same), therefore understandably takes pains to fortify his unnatural position. He knows that the pupil learns most of all through the example shown him by the masters in their masterworks. And if it were possible to watch composing in the same way that one can watch painting, if composers could have ateliers as did painters, then it would be clear how superfluous the music theorist is and how he is as harmful as the art academy. (Schoenberg, 1910 (ed. 2010), p. 17)

The practical application of music theory by composers and musicians in the eighteenth century also coincided with the more complicated landscape of music pedagogy, which increasingly needed to cater for musicians with a range of different skill levels. Writing for a more varied audience of aspiring musicians, Johann Nikolaus Forkel (1749-1818) produced a range of pedagogically orientated music theory texts suitable for amateur musicians. Topics covered included tones, scales, keys, modes, melodic patterns, rhythmic patterns, existing musical styles, and form. Forkel applied a broad brush in his writings, covering speculative music theories that had emerged in the seventeenth and eighteenth centuries, as well as the physical nature of sound. His treatises also introduced another new component into the discipline of music theory, the idea of critical analysis.

Music theory was at this time being repositioned as a discipline that could provide the means for musicians to further their skills in composition and performance. Following Forkel, "no longer was music theory a preliminary or metaphysical foundation to practice. One the contrary, it was practical pedagogy that was now a subset of theory" (Christensen, 2016, p. 217). The search for a theory of music had been pragmatically transformed into a discipline that increasingly formulated and catalogued practical solutions to the problems faced by musicians.

The new pedagogy that informed music theory at this time had also to contend with the profound shift in musical style that was taking place in the eighteenth century. Composers had moved away from the contrapuntally dense lines and instrumental style of Baroque music and looked towards new instrumental groupings and techniques. An emphasis on music with a singular melodic phrase accompanied by harmony had also emerged in the Classical Period (1730 - 1820).

One of the first treatises that explored this new style was Heinrich Christoph Koch's (1749–1816) three volume work, Versuch einer Anleitung zur Komposition (1782, 1787 & 1793). Though the first volume provided a more traditional treatment of harmony and counterpoint, the second volume was devoted entirely to melody. While Koch did not locate himself as an expert in the manner of Fux, he noted that ultimately, the creation of melody is dependent on genius: "only taste, the ultimate eighteenth century arbiter, can be the final judge of what is beautiful" (Baker, 1977, p. 185). Koch also differentiated the notion of what he termed the "inner nature" and "outer nature" of musicians that could account for the intermingling of genius and the skills (embodied in music theory) that might assist it. The "inner nature" of music cannot be taught, but can be given rise to through the study of the "outer nature" (Baker, 1977, p. 190).

By the end of the eighteenth century, much of the information in these treatises had started to be institutionalised. There had been a sharp increase in music conservatories and music schools throughout Europe by this time, and music theory texts were increasingly setting the standards for the way musicians should be trained. Music theory had become a canon of knowledge, and the study of sound, acoustics had now been subsumed into the the natural sciences.

Despite the evolution of the discipline of music theory with its focus on the practical problems of music composition and performance, the search for a theory of music was still in play. By now, however, it had aligned itself to a much larger question that sought to understand the very meaning meaning of art itself in relation to the human condition. This enquiry was markedly different from the mythology infused writings of earlier theorists such as Kirchner and Fludd, and there was still a belief that the complexity of music might be grounded in some kind of scientific or philosophical basis.

As part of the exploration of art and its relationship to the human condition, the highly emotive connection that human beings appeared to have in relation to music also came under scrutiny. By the latter half of eighteenth century, music had come to be regarded as "the most publicly emotional of [all the] arts" as well as the "most infectious" (Cowart, 1989, p. 88). Observing the French music of his time, Rousseau marvelled at the, "lively and brilliant accompaniments that the better performances harrow and enrapture the soul and carry away the spectator" (Cowart, 1989, p. 89).

The emotional state of both the composer during a work's creation, and that of the performer during its performance, became objects of inquiry that could potentially shed light on the nature of art. Johann Georg Sulzer published his General Theory of the Beautiful Arts, in 1774 which explored these themes, and proved deeply influential for Koch's pedagogical orientated works. Sulzer rejected the notion that meaning in music might be deduced in a scientific manner, and criticised the idea that music could lend itself to the deduction of empirical axioms that might be susceptible to systemisation (Bent, 1998, p. 168).

Sulzer went on to pose a much more open ended question regarding the effect of music on the human condition: "Whence comes this extraordinary intensity of the soul and how can it affect such happy results?" (quoted in Cowart, 1989, p. 87).

The human ability to translate such intense emotional content when creating or performing art works also came under consideration. Marpurg marvelled at this ability in performers, claiming:

> The musician must play a thousand different roles as dictated by the composer, and for this reason, he must possess the greatest sensitivity and happiest powers of divination to execute every piece. (quoted in Cowart, 1989, p. 180)

Daniel Webb echoed such sentiments, noting in his 1769 treatise, Observations on the Correspondence between Poetry and Music, that, "the gifted composer has the ability to transport and delight audiences into a sublime state" (Christensen, 2002, p. 67).

There was also an increased interest into the philosophical foundations of music, and that of art more generally. Although Descartes had written about music in his Compendium Musicae in 1618, he had at the time rejected the connection between musical phenomena and its emotional impact on the brain and had not taken up art as a philosophical problem. It was not until the close of the eighteenth century that a Western philosopher took up this enquiry, locating music in a wider framework of aesthetics. Immanuel Kant's Critique of Judgement (1724-1804) explored the place of music in within a wider framework of Aesthetics, a term Kant used to denote the "critical analysis of perception" (Schueller, 1955, p. 220). In this work, Schueller notes:

> Kant, then, stresses the uniqueness of the art-work and the inner rule which genius employs. He stresses also the exemplary nature of the standard or rule which genius works by. Though this rule is not scientific, it seems to come from nature itself, and the master-composer does not even know how it has occurred to him nor can he invent similar ideas if he wishes; and he cannot give precepts to others so that they can create works of genius also. He can only exemplify possibilities through works appearing to have inevitability. (Schueller, 1955, p. 221)

Problematically, Kant too did not provide a theory of music or art more generally. He instead located art as something that appears to emanate from the interaction between the genius and the phenomena that the genius encounters in the world. Further, not even the genius can understand, in a rational sense, the meaning of art, or catalog the conditions in which it may be recreated.

By the nineteenth century, great stylistic changes could be seen in the music composition. The discipline of music theory had by now been embedded into educational institutions and acted as a legitimised mechanism through which deep insights into both music composition and music performance might be gained. The large orchestral form had also emerged (typified in the works of composers such as Berlioz, Schumann, Mahler, and Brahms, who enjoyed an increased access to a growing palette of instruments from which they could pick and choose orchestral textures, along with a freedom to explore harmony and dissonance in new ways (Christensen, 2002, p. 222).

There was still a flood of music theory treatises during this time, and many had a strong pedagogical emphasis. One of these was by Simon Sechter (1788–1867) who had taken up a professorial position at the Music Conservatory of Vienna in the mid 1850s. His written works were later published by Carl Muller under the title, The Correct Order of Fundamental Harmonies: A Treatise on Fundamental Bass, and their Inversions and Substitutes. Sechter's theories and teaching methods had a deep influence on later music theorists, and he expanded on the theories of Rameau. Sechter's work, quoted below, is typical of how technical the exploration of music composition had become, and it took the form of a rigid set of rules that sought to cover almost any situation a composer might encounter:

> The chromatic alteration of the chords of the seventh , and of the seventh and ninth, of A minor, into chords of the seventh, and of the seventh and ninth, of relative scales, may be easily made, if the directions given for the chromatic alteration of the triads are adhered to. It should not be forgotten, however, that no raised degree can ever become a seventh or a ninth. (Sechter, 2013 edition, p. 11)

Sechter, a teacher of Bruckner and Marxson (also a teacher of Brahms), stressed the importance of studying strict counterpoint, and doing exercises rather than compositions (Christensen et al, 1992, p. 17). He claimed that anything in a music composition could be explained by appealing to the diatonic nature of scales and their capacity for modulation and voice-leading, rather than chromaticism.

Around the same time, in 1845, Alfred Day (1810-1849) published his Treatise on Harmony. Day was regarded as the "first truly original voice of English music theory" (Herissone, 2000, p. 33), and his music theory put forward a view that all chord voicings comprised of stacked thirds (such a 9th, 11th and 13th chord voicings) can be derived from seventh chords, and their behaviour can be traced to the properties of the harmonic series. Day located harmony in two discrete categories, diatonic and chromatic, and his treatise explored the capacity for modulation in both of these categories. Day's treatise was regarded as both dense and difficult (and originally garnered negative criticism) (Christensen, 2002, p. 333), but it displayed a view of English thinking about harmony at the time, that would be influential to later English theorists and composers (Herissone, 2000, p. 40).

One of the more disruptive treatises that appeared in the mid-nineteenth century was <u>On the Sensations of Tone as a Physiological Basis for the Theory of Music</u> by Hermann von Helmholtz (1821-1894) in 1863. This work recast the problem of music theory as an exploration on the the effect of sound on the human ear, which might be explained by the laws of physiological acoustics. Helmholtz believed the way in which a physical sound (be it any noise including something as simple as a sine wave) was heard by the human ear (which could be verified by experiment) could prove to be a compelling basis for theory of music. In the preface to his work, Helmholtz problematised existing approaches to music theory as lacking a basis in the natural sciences, and claimed his treatise would rectify this:

> All attempt will be made to connect the boundaries of two sciences [music theory and natural science], which, although drawn towards each other by many natural affinities, have hitherto remained practically distinct; I mean the boundaries of physical and physiological acoustics on the one side, and of musical science and aesthetics on the other. (Helmholtz 1863 (ed. 1954) , p. 2)

Helmholtz also questioned the increasingly narrow concerns of music theory, that had become too pedagogically orientated, and could not provide a sound basis for music:

> The horizons of physics, philosophy, and art have of late been too widely separated, and, as a consequence, the language, the methods, and the aims of any one of these studies present a certain amount of difficulty for the student of any other of them; and possibly this is the principal cause why the problem here undertaken has not been long ago more thoroughly considered and advanced towards its solution. (Helmholtz 1863 (ed. 1954) , p. 1)

Helmholtz' treatise did not locate notions of dissonance and consonance as entities that might be encoded on music score. He instead saw these as verifiable physical states (Steege, 2012, p. 285). Dissonance, rather than being located in the domain of a composer or expert, was instead "the coincidence and proximity of the overtones and difference tones that arise when simultaneously sounded notes excite real nonlinear physical resonators, including the human ear" (Helmholtz & Ellis, 1954, p. 28). This positioning of dissonance and consonance as physical entities also allowed the possibility for a theory of dissonance that could be altered depending on the timbre of an instrument.

Helmholtz' work was also instrumental in providing a scientific basis for the validity of equal temperament (i.e. the hypothesis that an octave that could be divided into 12 equal pitch steps). He observed that creating small amounts of detuning in certain intervals within an octave could allow musical works to be created in multiple keys, without undermining the sonorous properties of the intervals. In the last section of his treatise, Helmholtz turned to more practical questions of music theory, exploring the place of music scales and tones within this framework.

Although Helmholtz provided a scientific basis for the nature of overtones and a possible relationship they had to dissonance, his work had a limited impact on the music theory of the time. Both score analysis and music composition had become far more technical undertakings, and explanations of dissonance had increasingly come to be located in the domain of the pedagogically orientated music theorist and the music score itself. Hartmann, in 1887, noted with a sense of disappointment that the positivist approach taken by Helmholtz had not been embraced or led to further discoveries: “on the contrary, no progress of any kind has been made” (Steege, 2012, p. 288). Dissonance and consonance had become self evident realities by this time whose scientific basis was far less importance than the views held by music experts and practitioners. The complicated questions of how music might work were no longer rooted in the scientific basis of sound, but instead focused on increasingly complex patterns that could be found in music scores.

Hartmann’s comments were exacerbated by the fact that, as the nineteenth century drew to a close, the pedagogically informed music theory which had been created by experts in the field, had evolved to look and feel like a rationalist scientific endeavour in its own right. It increasingly used the language of scientific positivism (Christensen, 2002, p. 355), and any evidence for or against a hypothesis was now only to be found in patterns present in music scores. By the beginning of the twentieth century, the search for a model of music analysis or design that had a scientific basis from which one might derive musical works had largely been abandoned. This effort had been absorbed into other disciplines.

The growth and institutionalisation of music theory had also led to the creation of other disciplines as music theory became both increasingly professionalised and compartmentalised. In 1884, Friedrich Chrysander, Philipp Spitta, and Guido Adler (the latter is often referred to as the founder of musicology) founded the first journal of musicology which cast a wide gaze across the materials and context of both music composition and music performance.

Adler had written his own music theory treatise in 1883, History of Harmony. In it he had stressed the importance of taking a scientific approach (Mugglestone, 1981, p. 5), though the scope of musicology was to be instead focused on the context and social practices that surrounded the creation of musical works and music performance (Mugglestone, 1981, p. 9). Adler regarded "the palaeo-logical dating of a work of art" (Mugglestone, 1981, p. 5) as a critical step in musicological investigation, along with having ready access to a musical score in order to undertake analysis:

> If a work of art is under consideration, it must first of all be defined palaeo-logically. If it is not written in our notation it must be transcribed. Already in this process significant criteria for the determination of the time of origin of the work may be gained. Then the structural nature of the work of art is examined. We begin with the rhythmic features: has a time signature been affixed, and if so, which; which temporal relationships are to be found in the parts; how are these grouped and what are the characteristics of their periodic recurrence? (Mugglestone, 1981, p. 15)

It becomes increasingly difficult to track the search for a model of music analysis into the twentieth century. The meaning of musical works and music performance had come to be examined across multiple schools of thought and multiple disciplines that each had different foundational questions and specialised languages. The human relationship to sound and music is taken up heavily in psychology and, later, semiotics. The effect on the human body of performance and music improvisation is explored through performance studies and nature of gesture. The social practices that give rise to musical works are examined in fields such as musicology, ethnography and sociology. The deeper meaning of art and artistic expression becomes a complicated question of philosophy. Cultural studies would explore music creations as cultural artefacts and examine their potential to create social and political structures of meaning.

The idea of relying so strongly on the music score to understand art becomes problematised at this time, over shadowed by more complicated explorations of the complex relationship between human beings and music. The patterns found on musical scores however, increasingly became the subject of mathematical studies, and later the field of computer science explored the possibility of generative music algorithms.

Pedagogically focused music theory was still very much in abundance however. And as an established and institutionalised discipline, it had also become susceptible to criticism. One very vocal critic of existing approaches taken in music theory was Arnold Schoenberg (1874-1951). On Schoenberg, Christensen notes:

> Arnold Schoenberg would castigate the pretensions and conservatism of academic music theorists; indeed, the whole preface to the third edition of Schoenberg’s own Harmonielehre (1921) opens with a blistering assault on the hidebound discipline of “Musiktheorie” and its stultified pedantry.

(Christensen, 2002, p. 10)

Arnold Schoenberg was both a deeply influential composer and music theorist, who wrote his first major treatise, Theory of Harmony in 1910. The content and tone of the work is similar to so many of theory texts that had appeared before it, utilising the music score as a means from which to equip aspiring musicians with new ways of exploring voice-leading and harmony. Schoenberg explicitly problematised the study of music theory as a scientific endeavour, but was also pragmatic, acknowledging that there is "hardly any other way" to seek an understanding of music, other than observing what happens in music scores, and deriving laws from this these observations (Schoenberg, 1910, (ed. 1978), p. 11). Schoenberg criticised much of the existing music theory, however, noting that it erroneously "professes to have found the eternal laws" (Schoenberg, 1910, (ed. 1978), p. 11). In this treatise he notes that music theory:

> Observes a number of phenomena, classifies them according to some common characteristics, and then derives laws from them. That is of course correct procedure, because unfortunately there is hardly any other way. But now begins the error. For it is falsely concluded that these laws, since apparently correct with regard to the phenomena previously observed, must then surely hold for all future phenomena as well. And, what is most disastrous of all, it is then the belief that a yardstick has been found by which to measure artistic worth, even that of future works. (Schoenberg 1910, (ed.1978), p. 11)

In both this work, and his later writings, Schoenberg presented music theory as a means to an end, a vehicle that can guide aspiring composers in the acquisition of skills needed to become composers. For Schoenberg, any theory or set of laws that might underpin music should always be subordinate to the study of masterworks: “the pupil learns most of all through examples in masterworks” (Schoenberg ed. 1978, p. 13). He rejected any aspect of music theory that was not practical or whose application could not be evidenced in the masterworks. These masterworks were the foundational corpus upon which quality should be measured. Schoenberg was not speaking generally: in his writings, references are made to the masterworks as comprising the collected compositions of Beethoven, Bach and Mozart (Schoenberg, ed. 1975, p. 78).

Although Schoenberg is often portrayed as one of the most progressive composers of the twentieth century, his use of language and overall approach to music theory is still quite traditional. He wrote prescriptively and at length about what should and should not happen in musical works, using a style similar to earlier theorists such as Sechter, Fux and Rameau. In his writing there is an expectation that the rules he presents are to be followed. Consider a typical example: “consonances, such as simple triads, if faulty parallels are avoided, can be connected unrestricted, dissonances require special treatment” (Schoenberg 1978, p. 21). For Schoenberg, the rules he presented were made to be broken, but only in the pursuit of art by the true artist.

Despite the view that Schoenberg’s thinking and approach to composition evolved to become “atonal”, a label he rejected (Dahlhaus, 1987, p. 5), Schoenberg viewed dissonance as a consequence of pushing harmony and voice-leading to its limits, rather than abandoning it (Dahlhaus, 1987, p. 9). The tendency of notes within a diatonic scale to imply tonality was challenged by Schoenberg’s conception of an “emancipation of dissonance”. He envisaged musical works in which tonality might come to be “concealed by the vagueness of the contention that emancipated and unresolved dissonance is immediately comprehensible” (Dahlhaus, 1987, p. 10).

Another influential music theorist of the twentieth century, Heinrich Schenker (1868-1935) had published a treatise on harmony in 1906. Schenker presented a different approach to that by Schoenberg, and highlighted the use of passing notes (a notion rejected by Schoenberg) to create musical variations in underlying musical forms (Christensen et al, 1992, p. 77). Schenker believed it was possible to look beneath the surface of musical structures to uncover different layers within a composition. This iterative process of exploring the various layers would eventually lead downward to a foundational layer of the musical work, which Schenker referred to as the "Ursatz". The Ursatz was the basic elaboration of a tonic chord. The purpose of Schenker's investigation was not intended to be reductive but instead to provide a framework through which the growing complexity modern music might be navigated (Christensen et al, p. 87). It allowed very different works to be examined as alternative developments of an common underlying Ursatz structure, and thus be seen through a similar lens.

Like Schoenberg, Schenker viewed the pursuit of music theory as a science as problematic. In the The Masterwork in Music, he writes:

> I am keenly aware, that my theory, extracted as it is from the very products of artistic genius, is and must remain itself art, and so can never become 'science'. While in no sense a scheme for breeding up geniuses, it does address itself to practicing musicians, and only the most gifted of those at that.
> (Schenker, ed. 1994, p. 2)

Schenker also complained that existing notions of music theory were incorrect, and the discipline suffered from "centuries old errors" (Schenker, ed. 1994, p. 5).

This was where any consensus between Schoenberg and Schenker ended however. Their theories were at odds both with the existing tenets of music theory, and with each other. Of their differences, Dudeque notes:

> Thus, while Schoenberg demands that the consequence for the harmonic progression of even the most fleeting dissonance must be taken account of, Schenker postulates the exact opposite: that the dissonant nature of even the harshest vertical combinations must be disregarded in order to penetrate the superficial layer and arrive at the horizontal progression upon which musical coherence depends. (Dudeque, 2005, p. 11)

The disagreements between Schoenberg and Schenker, which, in part, can be attributed to an intentional misunderstanding of each other's work (Dahlhaus, 1987, p. 33) are typical of the lack of consensus that comes to characterise music theory in the twentieth century. It is a lack of consensus, however, that does not take issue with the foundations of the discipline, or even problematise the music score as a site where music theory investigation should take place. The disagreement between Schoenberg and Schenker is a powerful example of the problem that faces modern music theory, which so often descends into polemic debates that have no end, and where truth is located in personal points of view.

Both Schenker and Schoenberg have become important influences on the evolution of music theory and the way music composition appeared in the university curriculum. By the 1950s, Schenker's influence had grown markedly, particularly in North America, where it heavily influenced undergraduate theory instruction (Christensen et al, 1992, p. 66). While not setting out to provide a theory of music as such, Schenker nevertheless provided a methodology from which to explore complex musical works.

While Schenker responded to this complexity by providing a methodology that could categorise the complexities found on the score, the composer and theorist Paul Hindemith (1895-1963) adopted an alternative approach. In seeking a simpler way from which to understand the creation of musical works, Hindemith sought a theory that might explain how musical works differed depending on their genre and period. Commenting on Hindemith's Craft of Musical Composition, in 1940, Virgil Thomson noted:

> I call it the most comprehensive procedure I have yet encountered because it is based on acoustical facts rather than on stylistic conventions. At least, it proposed an analytic method that can be applied to the tonal structure of all the written music of Europe from medieval to modern times. (quoted in Luttman, 2009, p. 11)

Rather than basing his enquiry on the works of particular composers, or utilising his own expertise, Hindemith claimed that a gradual increase in dissonance can be seen in the overtone series itself and musical works could be explained by appealing to its structure. Instead of musical works being characterised by the presence of tonality or lack of tonality, or diatonicism and chromaticism, the structure of the overtone series showed how dissonance could be increased and decreased. This notion could be applied to any genre of musical works, and even used to explain musical works that utilised alternate tunings. Forte noted that “at a time in which the world was becoming more and more chaotic and threatening, [Hindemith] represented for many musicians a way out of the seeming chaos of twentieth century music practice” (Forte, 1998, p. 3).

Hindemith provided a link between the approach to dissonance and consonance by Helmholtz, and its location within an explanation of complex musical works. Rather that seeking to explain the works or techniques that could be utilised to create musical works, his theory allowed for the location of consonance and dissonance in any type of music. Like Schoenberg and Schenker, the influence of the writings of Hindemith has been lasting, particularly in the latter half of the twentieth century throughout American universities.

The search for models or theories of music analysis becomes a more fractured affair in the twentieth century, because its exploration increasingly takes place across different disciplines. The remainder of this chapter will provide a brief survey of the fields of musicology, music psychology and music semiotics, which draw metadata from music, but often not from a music score.

The discipline of musicology has a far wider agenda than that of music theory, seeking to understand the “inherent duality” between the “both separate and related constructs” of musical works and music performances, and the environment in which they exist (Beard & Gloag, 2005, p. 21). While music theory predominantly explored the technical problems located in the patterns found on music scores, musicology utilised a far wider lens, exploring the social practices that informed the production of musical works and music performances. It is a discipline concerned with both “the musical and the extra musical” (Ruwet & Everist, 1987, p. 11) at the same time.

The musical and extra musical aspects of musicology include: the study of the motivations behind the composition of musical works; the social milieu in which musical works and music performances reside; a musical work’s significance to the society in which it is created; a musical work's critical reception and reception by a wider audience and; the social demographic profile of this audience. Whereas the music theory of previous centuries had enjoyed the patriarchal convenience of the select few deciding on the merits of a musical work, musicology, to an extent, broke through these barriers. Western music was no longer to be regarded as the narrow lineage of concert music encoded on music scores, but any kind of music, produced by any part of society.

In exploring everything about the human condition and its connection to music, musicology quickly came to question the way music had previously been studied and understood, which lead to the problematising of pedagogical music theory. Musicologist Philip Tagg has claimed that score based analysis is not a valid way by which to examine music at all, but actually something qualitatively different altogether. It is instead, he argues, merely an analysis of a system of storage, an examination of ordered dots on the page (Tagg, 1982, p. 1). For Tagg, utilising a score based approach to examine music ignores the musical expressions that emanate from human existence. He claims that it is the musicians themselves who are guilty of this approach, often displaying an “exclusive guild mentality” expressed by the refusal in relating “items of musical expression” to extra-musical phenomena (Tagg, 1982, p. 1). This state of affairs, he notes, is compounded by a “time honoured adherence to notation as the only viable form of storing music, and a culture-centric fixation on only the parameters of music which are susceptible to notation” (Tagg & Brackett 1998, p. 13). Given such limitations, “music notation cannot be the analyst's main source of material” (Tagg, 1982, p. 28).

Tagg calls for a complete rethinking of the study of music to include more music genres and different tools and methodologies, that can allow for the inclusion of other, non-traditional music (Tagg, 1982, p. 70). Musicology should instead explore “how the musical statement of implicit attitudes prevalent in society at large affects those listening to such culturally eclectic and heterogeneously distributed types of music [such] as title tunes and middle-of-the-road pop" (Tagg, 1982, p. 70).

Musicology becomes problematic primarily because of its scope. There has never been clear agreement in the field regarding the way tools that examine music might be used, or even how they might be constructed. It is a discipline that cuts across ethnography, history, and sociology, and variously utilises the different methodologies specific to these fields. From the 1980s its scope is further enlarged again with the rise of "new" musicology which sought to explore how music exists in areas such as gender studies, postcolonial theory and cultural studies.

Despite this scope, musicology has not been successful in putting forward a model of analysis (and to be fair, this it not its intention). However, its agenda demands that, whatever a model of analysis might look like, it must be far more inclusive than anything put forward in the discipline of music theory, and respond to the problematic reliance on the music score.

Whereas the discipline of music theory allowed experts to put forward a view on how it was that musical works come into existence, musicology problematises our subjective relationship to music and its place in our culture. In asking these far wider questions, the study of music moves away from finding a model or formula, to an exploration of the way music exists in the world. On musicology, Kerman notes that, "though considerably larger and better organized other fields of music analysis in terms of the "rigors of its approach", it has nevertheless "produced signally little of intellectual interest" (Kerman 1985, p. 14). Charles Rosen is far more aggressive in his criticism of musicology, claiming that much of its output has no meaning at all, and certainly no significance.

The field of music psychology explores the way in which the human brain processes sound, as well as its role in both creating and listening to musical works. The field has evolved to have strong links into neuroscience, but its concerns can be dated as far back Aristoxenus, who was not only seeking to understand the mathematical ratios of music intervals, but the effect that listening to these had on the brain (Levitin, 1994, p. 3). Gjerdingen describes music psychology as "a subfield of psychology that addresses questions of how the mind responds to, imagines, controls the performance of, and evaluates music" (Gjerdingen, 2008, p. 55). He further notes that, going back at least to the seventeenth century, examples in the field of music theory can be found that have a strong relationship with music psychology, in their effort to understand the effect of a musical work on its listeners.

Early work in music psychology included the examination of the ways in which tones were heard and processed by the human brain. The growing availability of instruments in the eighteenth century made it feasible for them to be explored in a laboratory setting (a practice termed "brass instrument psychology"), which allowed controlled experiments of interval and tonality recognition. As an example, Carl Lorenz recorded 110,000 observations regarding the nature of tones around 1885, which led to fierce debates around the way in which the brain processes tone and its ability to apprehend specificity (Gjerdingen, 1988, p. 936).

Music psychology also has powerful ties into the idea of creating a theory of music. Understanding the way in which the human brain might differentiate tones and tonality shed light on how such a process might be assisted by a theoretical approach. Early studies that explored this included The Measurement of Musical Talent (1915) and The Psychology of Musical Talent (1919) by Carl Seashore (1866-1949). Seashore believed that there would be no end to the "scientific procedure in the interpretation, evaluation and education of the musical mind" (Gjerdingen, 1988, p. 938), and that a complete theory of talent, aesthetics and criticism might be found through this approach, whose tenets could be utilised by musicians (Gjerdingen, 1988, p. 938).

Another, more recent work along these lines, was Fred Lerdahl and Ray Jackendoff's, Generative Theory of Tonal Music (1983). In it they claimed to create a "comprehensive theory of music [that] would account for the totality of the listener's musical intuitions" (Lerdahl & Jackendoff 1983, p. 8). In the preface to the text, Leonard Bernstein highlighted the importance of such an enterprise which he believed could be in the form of a "formal description of the musical intuitions of a listener who is experienced in the musical idiom" (Lerdahl & Jackendoff 1983, p. 3). The work attempted to formalise and categorise musical intuitions about harmony and rhythm, similar to the construction of a generative grammar in linguistics.

On the field of Music Semiotics, Monelle notes:

> Rigorously scientific, [music] semiotics offers a new and radical theory for the basis for analysis and criticism. (Monelle 1992, p. 24)

The above statement, taken from the Raymond Monelle text, Linguistics and Semiotics in Music, indicated the philosophical departure that took place in the 1970s, away from the more traditional and descriptive models of music analysis.

Again moving away from the music score as a site of analysis, music semiotics explored foundational questions regarding both the creation and understanding musical works. It explored how information could be encoded between the musical work and the listener. It purported to locate this enquiry in a scientific framework which codified music information.

The idea that a musical work might be a producer of information was a powerful forerunner to the enquiries seen in the field of music information retrieval. Music semiotics also directly challenged the author-as-expert model seen in more traditional forms of analysis. It rejected the idea of an authoritative view of music held by an expert. The meaning of a musical work was "not to be found in the emotions of the composer or performer, or in the reactions of the listener, because these emotions are not real emotions" (Monelle 1992, p. 30). Meaning emanated from the fabric of the music itself, and the musical work acted as an artefact onto which attributes could be codified and shared to those interacting with it.

Typical methodologies used in music semiotics located an observer who would take action that would lead to encoding musical works as signs. The observer could then examine how these signs interacted with each other. Worthen explains:

> To make a chart of what I hear, I proceed in the following manner. If what I hear is new, I assign it a letter. When I hear something that is different, I give it a new letter, placed to the right of the previous one. If it is something I have heard before, I identify it with the same letter as before, placing the letter below its former entry. Measure numbers are in subscript, and a variation of a previous element or sign is in superscript. (Worthen 1992, p. 2)

In Music and Discourse (regarded as a critical early text of music semiotics) Jean Jacques Nattiez claimed that the musical work is not merely a “text”, or simply a music score. It should not be regarded simply as a tangible object composed of underlying structures. Rather, the musical work is also constituted by the procedures that engendered its creation, and it is possible to codify these as an observer. Nattiez complains that ‘in conventional analysis, the musical work may be reduced completely to its imminent properties” (Nattiez, 1990, p. 33). Music semiotics moves away from this structuralist position, allowing the observer to codify the poetic, immanent and aesthetic variables found in a musical work. This information can then be made the object of scientific analysis.

Because of the disagreements in the field, it is difficult to ascertain both the success of music semiotics and the validity of its methodologies. Monelle claimed that there was not a “single book you could send people to” and although there was a “proliferation of theoretical models, there was little consensus amongst practitioners” (Monelle, 1992, p. 33). Criticising the current state of the field Tagg claimed:

> Unfortunately, a great deal of linguistic formalism has crept into music semiotics…[which has led to the] extra generic question of relationships between musical signifier and signified and between the musical object under analysis and society being regarded as suspect, a problem of needing more information. (Tagg 1991, p. 6)

Seeking to quantify the totality of information that emanates from a musical work in the presence of an observer, even if these interactions are reduced into signs, music semiotics became faced with the observer’s seemingly infinite capacity to experience information. Having an “increased reluctance to locate musical wholeness, its identity, purely in terms of cultural norms [inevitably] must lead to more and more comprehensive description" (Dunsby 1983, p. 29). Criticising one of the key figures in the field, Nicolas Ruwet claimed that Nattiez "failed to realise [his] theory had no basis in experiment; it is intuitive" (Monelle 1992, p. 31). Monelle also noted that "the progress of musical semiotics has been retarded by a desire for irrefutability" (Monelle 1992, p. 31).

The difficulties of music semiotics also emanate from the limits of scientific enquiry itself. Piaget notes:

> If one tries deal with structures within an artificially circumscribed domain, and any given science is just that, one soon hits on the problem of being unable locate multiple entities one is studying, since structure is so defined that it cannot coincide with any system of observable relations. (Piaget, 1971, (ed. 2016), p. 17).

Despite its difficulties, the field of music semiotics speaks directly to the uneasy dichotomy between the intuitive and scientific aspirations of those seeking to understand music. It seeks to be inclusive with regard to the complexity of music, but rigorous in its analysis and data collection. Music semiotics is critical in setting the academic stage for a radically different way of thinking, and positioning the musical work as an agent of information production.

Reflecting on the vast body of work that had come to inform the investigation of music towards the close of the twentieth century, Nicolas Cook makes the troubling comment that there is a still a "good deal of muddled thinking on this topic" (Cook 1987 p. 271). Despite the plethora of approaches that have been taken in a variety of different disciplines, Cook notes that, in the end, most examination of music had little variation in terms of the questions it posed:

> Whether is is possible to chop up a piece of music into a series of more-or-less independent sections. They ask how the components of the music relate to each other, and which relationships are more important than other. They ask how these components derive their effect from the context they are in. (Cook 1987, p. 39)

Cook also reflected on the difficulty of adopting a strictly scientific approach, which could undermine the utility of an analytical model for those seeking to create musical works:

> Personally I dislike the tendency for analysis to turn into a quasi-scientific discipline in its own right, essentially independent of the practical concerns of musical performance, composition or education. Indeed I do not believe that analysis stands up to a close examination when viewed in this way: it simply doesn't have a sufficiently sound theoretical basis. (Cook 1987, p. 3)

All of this suggests that, in creating a theory, or an analytical framework, from which to understand music, we find ourselves faced with a subject that "notoriously resists its own history, constantly shifting over time" (Dahlhaus, 1987, p. 2). Gjerdingen claimed that, "Whenever I attend a meeting of music theorists, I am struck by the conviction with which old beliefs are invoked as eternal verities" (Gjerdingen 2008, p. 163) goes on to say that:

> Although music theory may endorse experiments, and grants the presumption that [these] experiments are skilfully performed and accurately reported, the interpretation of experimental results takes place in a no man's land between disciplines, with very different histories, mores, central subject matters, and professional goals. (Gjerdingen, 2008, p. 165)

Examining the history of music theory and music analysis shows that, while there certainly may be "something fascinating about the very idea of analysing music" (Cook, 1987, p.1), there is also a complete lack of consensus around how it might take place. It shows that our relationship to music is volatile. It is opinionated, changeable and deeply individual. Music takes place at the forefront of our emotional lives and this clouds our judgement. Nietzsche famously remarked that "without music, life would be a mistake" (quoted in Ball, 2010, p. 8). Schopenhauer claimed that music is "completely and profoundly understood [in our] innermost being as an entirely universal language" (Schopenhauer, 1818 (ed. 2010), p. 33). Oliver Sacks claims, "music, uniquely among the arts, is both completely abstract and profoundly emotional" (Sacks, 2007, p. 13). Such sentiments confound consensus.

Even though it may be impossible to reach agreement on what music is and how it can be understood, an alternative approach can be taken. It is possible to treat the information that is derived from music as completely decoupled from music itself, and explore it on its own terms. This approach, seen in Music Information Retrieval, will be taken up in the next chapter.

Chapter 2
Music as a problem of information

The focus of this chapter will be on the field of Music Information Retrieval (MIR), and its potential to provide an alternate framework for the analysis of musical works and music practices by extracting metadata. Rather than placing the musician at the centre of music analysis, or examining the socio-cultural context of musical works, MIR has instead focused on the study of information that music generates when human beings interact with it.

Adopting an information oriented approach has allowed MIR to elegantly sidestep some of the more thorny issues of music analysis. MIR does not purport any particular underlying meaning of music, or seek to contextualise music in a fixed way, being more closely aligned to disciples such as mathematics, which seeks meaning through the conclusions drawn from manipulation of patterns, rather than a derived meaning.

The MIR focus is on the patterns that can be found in any music related data. This data can be drawn from a range of sources, such as music scores, audio files, user preference data in music streaming services, or curated playlists. The data can be any and all of these things. Research in MIR often relies on the fact that when human beings create and interact with music, they will leave traces of information behind. It is these traces of information that can be examined and explored.

This chapter will begin by surveying some of the early work that preceded MIR, and highlight the field's reliance on an increased availability of networked computational technologies that have enabled the study of large data sets to become more feasible. I will then examine the way in which data is positioned in the field of MIR in terms of finding effective ways to search and retrieve it, to ensure it is of high quality, and to develop techniques for music data generation (such as optical music recognition and automated music transcription).

I will also provide a survey of the tools and methodologies that have been employed for pattern analysis in the field, and highlight their links to more traditional music theory approaches (such as Schenkerian analysis). MIR differs markedly from music theory however, in that it views the music score (or what it terms as a symbolic representation of music) as just one of many possible metadatas that can be derived from music, and it does not privilege the music score above any other type of information.

The origins of the idea that music might be related to information can be traced back to the early twentieth century. In 1928, Ralph V.L. Hartley published the paper, Transmission of Information, in which he set out to understand the properties of information. Hartley's paper presented three core ideas: firstly, that any system of communication (and an example might be a music listener receiving audio data from a music performer) can exist independently of the human sender and human receiver; secondly, that information could be understood as a commodity that can be represented by some sequence of physical signals and; thirdly, that the meaning of information was not important, it was only the structure of information (being the speed of the signal transmission and the relationships between repeating and non-repeating signals) (Hartley, 1928, p. 45).

A short time later these ideas had begun to find their way into music. A pivotal moment that preceded this was in 1951 when Claude Shannon published A Mathematical Theory of Communication, which was heavily influenced by Hartley's theories. This paper (which consolidated Shannon's place as the founder of the field of information theory,) put forward the notion of "entropy", a mathematical measure of the amount of uncertainty in the information between a sender and receiver (Shannon, 1951, p. 12). Although Shannon's work focused on problems in electrical engineering (such as data compression), both his ideas and methodologies soon came to permeate many other fields, including the study of music.

In 1957, music psychologist Leonard Meyer published Meaning in Music and Information Theory. In this work he proposed there existed a relationship between music and information, claiming that deep similarities existed between the problems of understanding music, and solutions offered in the field of information theory. Meyer claimed:

> In that analysis of musical experience many concepts were developed and suggestions made for which I subsequently found striking parallels, indeed equivalents in information theory. Among these were the importance of uncertainty in musical communication, the probabilistic nature of musical style, and the operation in musical experience of what I have since learned. (Meyer, 1957, p. 417)

Hiller also claimed that the field of information theory could be used to provide insight both into the structural details of musical works, and as a means of developing a deeper understanding of how human beings communicated music-related signals to one another (Hiller, 1966, p. 96). Properties that can be found in music, such as variation, repetition, and novelty, were perfectly suited to investigation in an information theory framework. It became possible to characterise the vast majority of musical works that are created by human beings (regardless of their location of origin or era), as being "neither totally organised, nor totally disorganised, but [falling] somewhere between these extremes" (Hiller, 1966, p.121). The process of measuring entropy in music related information (a process which often utilised music score data) also revealed that musical works tend to exhibit an "average information level" (Hiller, 1966, p. 123) during their overall duration, and increases and decreases in the level of information can be related to structural elements of the musical work. Speaking about how such measurements might be made, Meyer noted:

> Information is measured by the randomness of the choices possible in a given situation. If a situation is highly organised and the possible consequents in the pattern process have a high degree of probability, then information (or entropy) is low. If, however, the situation is characterised by a high degree of shuffled-ness so that the consequences are more or less equally probable, information (or entropy) is said to be high. (Meyer, 1957, p. 19)

The early studies involving music and information theory can be categorised into two areas. The first utilised mathematical techniques and statistical methods in order to obtain quantitative results, often positioning the music score as an "objective specimen that could be used to derive a rigorous set of musical processes" (Hiller, 1966, p. 133). The second type were far more speculative in nature, and predominantly located in the field of music psychology (Hiller, 1966, p. 133). These examinations sought to understand how information theory might further the understanding of psychological responses to music listening (Hiller 1966, p. 138), and were concerned with the different ways in which human beings used music (for example, in the role of listener, composer, performer, and theorist).

Examples of early investigations included Information Theory and Melody (Pinkerton 1956) which computed the monogram distribution of diatonic scale degrees in a corpus of 39 monophonic nursery rhymes, and derived a redundancy estimate of 9% (being related to the repetition that existed in the overall corpus). In 1958, in Style as Information, Youngblood calculated the difference between different musical styles by comparing twenty songs from the Romantic period (composed by Schubert, Mendelssohn and Schumann), with a selection of Gregorian chants (Youngblood, 1958, pp. 24-35). Kraehenbuehl and Coons published Information as a Measure of the Experience of Music Information a year later, which had a stronger emphasis on music psychology (Kraehenbuehl & Coons, 1959). Of the connection between information theory and music they note:

> Information theory has been applied most successfully to small finite sets of events where all possible events in any particular set could be designated and a reliable probability established for the frequency with which each event would occur in samples of sufficient length. In music both the twelve-tone chromatic and seven-tone diatonic scales are such sets of events. (Kraehenbuehl & Coons, 1959, p. 518)

In 1966, Hiller and Bean published Information theory analyses of four sonata expositions, exploring the differing levels of entropy in a selection of sonatas of Mozart, Beethoven, Berg and Hindemith. Entropy was here framed as a level of uncertainty that can be derived when mathematically predicting notes that would occur in the sonatas. This work confirmed the intuitive belief of its authors, that musical works which spanned the classical and modern era were becoming increasingly complex, and this complexity could be defined and measured mathematically. Using techniques from information theory, the authors were able to chart this increase of entropy between composers in subsequent eras.

These early articles had access to a very limited amount of data from musical works, such as text files holding pitch related information and basic rhythmic divisions. However, for the first time, it became possible to speak about structure and complexity in music within a measurable and objective computational framework, that could also be located in human communication. Information theory provided a common measure with which to view musical works and the relationships between musical works from any time period. Rather than being internally descriptive or seeking an underlying understanding of what music was, the meaning of music could be now be viewed as a product of the information it generated and related to the patterns that could be found in this information.

Such studies also show an early strategic response to a problem that was increasingly facing music analysis: the difficulty of working with larger amounts of information. Some early music archiving projects also began at this time, such as Barlow and Morgenstern's Dictionary of Musical Themes (Barlow and Morgenstern 1948) as well as a number of later projects that sought to store music information on magnetic tape (see Hudson 1970).

These articles demonstrated that the analysis of music could only take place in regard to the information that music could generate. There was little to be gained in seeking an understanding beyond this, which risked being biased and subjective. This early approach also spoke to the possibility of locating a theory of beauty or art within a wider scientific framework, without losing its meaning. On the application of scientific principles to art, Arthur Eddington claimed in his 1927 Gifford lectures that “there are the strongest grounds for placing entropy alongside beauty and melody”.

The rise of MIR has also been fuelled by the increased access to computational power and digital storage. When reflecting on the current state of affairs in 1974, Patrick claimed that "computer-aided study is meagre in its scope" for music analysis (Patrick, 1974, p. 322). Since that time however, both the availability of technology and the increasingly intuitive ways by which it can be accessed, have proved critical in setting a foundation for the emergence of MIR.

Early work in computer music related research can be traced to the 1960s. It had a mathematical focus, and utilised computational power in order to speed up pattern analysis. Examples of early works in the field included Forte's theoretical framework for segmentation (1966), a method that employed rigorous logic and pattern recognition procedures in order to model the human ability to read music scores. In 1969, John Rothgeb published his dissertation on automated realisation of un-figured basses, using the SNOBOL symbolic computing language. Nancy Rubinstein created a program in the FORTRAN programming language that could detect patterns found in the music of the German region of Franconia in 1969. Raymond Erickson published Rhythmic Problems and Melodic Structure in Organum Purum: A Computer-assisted Study in 1970 to explore patterns in plainchant melody. An interest in the relationship between artificial intelligence and music also emerged, and can be seen in Denis Baggi's 1974 dissertation entitled, Realisation of the Un-figured Bass by Digital Computer. Baggi has gone on to write write widely in the field, exploring neural networks and AI applications in music. In 1979, Polansky also put forward the proposal of a computer model for the perception of hierarchical memory in music (which emerges again in the field of MIR), based on theories developed by the experimental electronic composer, James Tenney.

These early attempts to fuse techniques found in music, technology, engineering and mathematics were, like those related to information theory, basic compared to the computational analysis that has come to be undertaken today. They were attempts that faced the difficulty of not only preparing the data that might be examined, but lacked the computational power to explore it in depth. Yet such attempts laid the groundwork for not only how music might be explored, but also the mediums by which it is created and transferred. These attempts indicate that, at some point in the future at least, technology might enable the automated creation of musical works, that would be indistinguishable from those created by a human, both in their structure and perceived emotional content.

An early champion of a project to bring together composers, musical aesthetics, and technology for the purpose of artistic creation, was David Cope. In the 1980s, Cope became interested in building a computer program which could encode a composer's musical style, and might be utilised to generate musical works. Cope claimed:

> My initial idea involved creating a computer program which would have a sense of my overall musical style and the ability to track the ideas of a current work such that at any given point I could request a next note, next measure, next ten measures, and so on. My hope was that this new music would not just be interesting but relevant to my style and to my current work. Having very little information about my style, however, I began creating computer programs which composed complete works in the styles of various classical composers, about which I felt I knew something more concrete. (Cope, 1991, p. 11)

The idea that technologically driven processes can be embedded into human consciousness, to emulate and interact with the the creative process, is a profound challenge to the way human beings interact with music. It also challenges the process of creating music and questions the notion of originality. Cope has claimed that, "The genius of great composers, I believe, lies not in inventing previously unimagined music but in their ability to effectively reorder and refine what already exists" (quoted in Doornbusch, 2010, p. 73).

By the beginning of the twenty first century, technology had become ubiquitous in music. It was not only a critical tool for researching the patterns and meanings that might be found in music related information, but also the preeminent medium through which music was created and transferred.

The academic field of MIR emerged in the late twentieth century, starting as an informal research group, and the group held its first formal symposium in October 2000, in Plymouth, Massachusetts, USA. Research in the field is explicitly concerned with exploring the data that can be derived from music. It crosses over a number of disciplines, and MIR conference papers can be located in areas such as digital signal processing, musicology, machine learning, robotics, recommender systems, and music psychology. There is a pronounced technical emphasis in MIR, and a heavy utilisation of mathematical methods that are used to explore music data, along with a number of engineering and commercial applications (such as Shazam, Spotify and Pandora). While some work has been carried out in relation to generative and automated composition of musical works, there is a stronger emphasis on the automation of other manual processes such as automatic transcription of audio (i.e. the conversion between audio and MIDI data).

There are strong links between MIR and many of the problems seen in music theory. Efforts in MIR that seek to understand melodic similarity across a corpus of works can also be located as a critical theme in the work of Schoenberg, in ethnomusicology (Nettl, 1983) and in music analysis more generally, (Quinn, 2000,). The availability of big data storage and use of data iteration techniques, along with the rise of personal computing, has made it feasible to undertake this work across a growing corpus of musical works.

As an emerging field, MIR also has its share of challenges. Some of these are practical. In the early 2000s especially, researchers were still struggling with the limitations of technology and problems of bandwidth, storage and processing power. There were few established and widely available techniques in the early years of MIR that could be used for big data processing, yet at the same time the volume of data had become unwieldy. There was also a wider philosophical issue in play too, regarding the best way to locate the scope of enquiry in the field, and how to position the user of MIR research. In 2003 it was observed that, "MIR is beginning to emphasise certain areas of research without having identified user communities and evaluated whether the techniques developed will meet the needs of those communities" (Futrelle & Downie, 2003 p. 124). In a 2001 keynote, Jeff Raskin took up this theme, saying the field had a distinct bias toward computer science and audio engineering. (Futrelle & Downie, 2003, p. 124).

At the very heart of the field of MIR however, is the problem of music data, and the way data can be effectively searched and retrieved. Examining the papers that have been written in field since 2002, it is possible to identify four broad categories of data under investigation.

The first of these is data relating to the symbolic representation of music (how MIR refers to the music scores). An early example of this is the New Zealand Digital Library project, MELDEX. This project is web based, and was designed to allow users to perform both text and sung queries. The MELDEX repository includes over one thousand melodies from popular songs that have been converted into duration, location and frequency data from the music scores, using optical music recognition techniques. The collection also contains 10,000 additional folksongs and over 100,000 MIDI files. Another, more well known example, is the IMSLP/Petrucci repository of public domain scores (though much of this in PDF format and difficult to extract into useable data). These kind of repositories have allowed MIR to undertake longitudinal pattern analysis across music scores from different styles and time periods.

A second type of data is the music metadata associated with audio music. A popular example of this is the MusicBrainz database, an online repository of information that includes such attributes as genre, artist name, release date, compact disc ID number, track length and album name. MusicBrainz currently has over 16 million indexed tracks and has developed retrieval methods to search for tracks that include acoustic fingerprinting, where a sample of of the audio can be used as a track identifier.

A third type of data used heavily in MIR is user preference data. User preference data can be generated whenever a user interacts with a tangible representation of music. Sandvold notes that this data can be generated when transactions occur such as buying a new song or album to add to an existing music collection, participating in a music related discussion forum on the internet, choosing and sharing music playlists through an online community, or stopping and starting playback of music in networked software (Sandvold et al, 2006, p. 1). It is possible to track and record data regarding an individual user interaction with music, or in a group, in order to examine trends across listener communities. Sandvold also notes that the behaviour exhibited in relation to music can create communities, bring together individuals with similar taste, and it is even possible to explore the patterns that arise when these communities interact (Sandvold et al, 2006, p. 1).

The last type of data is the analog and digital representation of audio information itself. Recent examples of this type of data include the stored data repositories held in music streaming services such as Spotify, Pandora, and Apple Music. These types of data sets are held in a number of music data formats, including Compact Disc, MP3, WAV, and AAC. These are formats which can encode audio information in similar ways, but their main point of difference is related to the size of the file in which the information can be held. MP3 and AAC file formats utilise strategies to remove the frequencies outside the standard human hearing range, in order to reduce the amount of information needing to be stored, making the file smaller). Audio files are utilised in MIR for a range of tasks related to audio signal processing, and research problems include automatic music transcription and musical instrument separation. To give an indication of the amount of data that is is held as audio data in various repositories, in 2013 the music streaming service Spotify released data showing the twenty million songs being currently held on the its servers, four million of which had never been played at all.

Increasingly in the research of MIR, all of these different data types can be found together. One of the benefits of the MIR approach is that qualitatively different types of information (such as music scores and audio files) can be explored in similar ways, leading to more multimodal and scaleable approaches to analysis. An example of this type of work can be seen in Peeling, Cemgil, and Godsill's *A* Probabilistic Framework For Matching Music representations (2007), which created a “probabilistic framework for matching different music representations (score, MIDI, audio) by incorporating models of how one musical representation might be rendered from another” (Peeling, Cemgil, and Godsill, 2007, p. 1). In the article, the authors also highlight how different types of information can be used to form an understanding of music:

> Musical information is roughly represented in one of three ways: a score, which is a symbolic representation, a MIDI file, which represents discrete musical events with more precise timing information, and sampled audio, which is the most faithful representation of the sound produced. (Peeling, Cemgil & Godsill, 2007, p. 1)

They go on to note that a possible application for their research could be the automatic annotation of audio databases, where the score data is known, that would allow automatic syncing between audio files and music score information. This is a powerful idea that demonstrates how music analysis might become more multimodal, and one that I will revisit later in the dissertation.

It is not only the type of data, but the structure of data which is of critical concern in the field of MIR. As noted in the previous chapter, Philip Tagg criticised the practice of using a music score as an object for music analysis as it has limited value beyond being a system of storage. MIR does not take issue with Tagg’s viewpoint of the music score, but instead problematises how the music score might be converted into a dataset that is more conducive for analysis.

Some of the more popular data specifications used in the field to encode music score information include Music Information Digital Interface (MIDI) and MusicXML. The MIDI specification has been in use since 1982 and encodes basic note on/off information to allow for the encoding of limited additional metadata. It has proved critical as an early data source for music, and is a common technology utilised for music playback in digital devices due to its small storage footprint (Wiil, 2005, p. 1). Lemstrom and Laine have noted however, that using MIDI for data analysis can be problematic, especially in more complicated retrieval tasks (Wiil, 2005, p. 1). Much of the information that would be found on a typical music score (such as slurs, mordents, arpeggiations etc.) cannot be explicitly encoded in the MIDI data specification.

MusicXML was partly a response to many of the problems faced by MIDI in terms of the limitations in rendering the visual complexity of music scores. First appearing in 2003, MusicXML was designed to be a comprehensive data representation of a music score that can be easily ported between different software applications. MusicXML is a subset of Extensible Markup Language (or XML) which is a markup language that defines a set of rules for encoding documents in a format that is both human-readable and machine-readable. Ganseman et al note that “the ability to use the countless mature software tools that are available for XML parsing and processing, is the main reason to prefer XML-based formats over others” (quoted in Ganseman, Scheunders, & D'haes, 2009, p. 1).

In its current specification, MusicXML can encode over 600 different types of elements that can be found on a music score. This includes not only pitch and rhythmic information, but attributes such as lyrics, expressions, dynamics, attributes, instrument fingerings, transpositions, etc. An example of two whole notes (in this case a C note and D note), encoded in MusicXML can be seen below in Future 2.1.

Figure 2.1. Example of two notes encoded in MusicXML

```
<score-partwise version="3.0">
 <part-list>
  <score-part id="P1">
   <part-name>Music</part-name>
  </score-part>
 </part-list>
 <part id="P1">
  <measure number="1">
   <attributes>
    <divisions>1</divisions>
    <clef>
     <sign>G</sign>
     <line>2</line>
    </clef>
   </attributes>
   <note>
    <pitch>
     <step>C</step>
     <octave>4</octave>
    </pitch>
    <duration>4</duration>
    <type>whole</type>
   </note>
   <note>
    <pitch>
     <step>D</step>
     <octave>4</octave>
    </pitch>
    <duration>4</duration>
    <type>whole</type>
   </note>
  </measure>
 </part>
</score-partwise>
```

Because MusicXML was principally designed to encode visual components of music scores, the resulting datasets can contain highly prescriptive information about how a music score should look (and can even include the relative x and y coordinates of visual components of the page).

Although MusicXML was not specifically designed for use in data analytics, it is increasingly being used to explore patterns found on music scores (and the case studies in the following chapters will use information taken originally from MusicXML files). Speaking about the types of analyses that might be carried out, Good notes:

> Say we want to investigate whether Bach's pieces really have 90% of notes in one of two durations—e.g., quarters and eighths, or eighths and sixteenths. We can do this by plotting a distribution of note durations on a bar chart, displayed together with a simple spreadsheet. (Good, 2000, p. 2)

Good goes on to characterise the problem of music score analysis as a 'Tower of Babel' problem (Good, 2000, p. 2), and positions MusicXML as an ideal way of tackling it, claiming: "developing converters between existing formats and a single MusicXML language could greatly simplify the tasks of music information retrieval" (Good, 2000, p. 2).

MusicXML does have some drawbacks however. One of these is that it only stores the note order and note length, rather than the absolute position in the score at which the note occurs (Ganseman, Scheunders, & D'haes, 2009, p. 664). This can be particularly problematic as, often in music score data analysis, there is a need for "absolute timestamp[ing] in order to know at any given time where we are in the score" (Ganseman, Scheunders, & D'haes, 2009, p. 664). This lack of absolute positioning can be seen above in Figure 2.1: the position of the C note is not explicitly provided, but implied as it occurs before the D note.

Another problem with MusicXML is the file sizes it tends to generate. Ganseman et al note that "common uncompressed [MusicXML] files contain easily up to 250KB of text for a single A4 size page of piano solo music" (Ganseman, Scheunders, & D'haes, 2009, p. 664).

Both of these issues can make it problematic to undertake data analysis and information retrieval tasks. For the purpose of this dissertation, I have created my own MusicXML converter (called Music MetaData Builder), which can explicitly encode timestamp information for all duration and location information on the music score, and substantially reduces the file size, and is suited for rendering in SVG format (using data visualisation libraries such as D3.js) found in many web applications. The converted data is also far less nested than MusicXML, making it more convenient for analysis tasks.

Although they are the most popular specifications, MusicXML and MIDI are not the only data specifications that are used to encode music data from music scores. Furthermore, the popularity of these formats is to an extent driven by their use in commercial software applications such as Logic Pro, Finale and Sibelius.

An alternative specification is the Music Encoding Initiative (MEI), created by Perry Roland, which was purpose designed for content based searching, analysis and visual presentation, and uses a hybrid specification including MIDI and MusicXML. MEI differs from MusicXML, in that it "seeks to encode information and its intellectual content in a structured and systematic way". It privileges the semantics above the representation found in MusicXML, and offers exciting possibilities for the structures needed in data analytics[1].

Another specification, GUIDO, also focuses on searching music data and seeks to address the "multidimensional, often complex structure of [music] data" aiming to capture general musical concepts as well as other information traditionally found on the music score (Hoos, Renz & Gorg, 2001, p. 1).

The other critical data related task in the field of MIR is data generation and, more specifically, the problem of creating tools to ensure high quality data generation. Fujinaga and Riley note that "the quality of the data itself is a critical part of the retrieval system, as content-based retrieval cannot work on inferior content" (Fujinaga & Riley, 2002, p. 1).

Another way that data is generated in MIR is by using optical music recognition (OMR) techniques. OMR techniques are related to the more general problem of optical character recognition, which seeks to convert images of typed or handwritten text into digital formats. In MIR, this usually means processing a music score (usually in PDF format) in order to extract the critical visual components that can be encoded into a machine-readable format such as MusicXML or MIDI. The ability to analyse large bodies of symbolic music information is dependent on having the tools that can convert images of symbolic data into formats suited for data analysis. There are currently large repositories of music scores that are held online, which could potentially be made available as datasets if the technology existed to facilitate their conversion (for example, the International Music Score Library Project (IMSLP) currently holds 93,000 music scores by over 12,000 composers).

---

[1] Although the focus of this dissertation is very much an transforming MusicXML, the future work does have more of a focus on MEI. Thought it is not as widely used as MusicXML, its decoupling of semantics and presentation make it more amenable to analytics and machine learning tasks.

Fujinaga and Riley note that “large scale digitisation projects” in MIR will allow the creation of “larger collections, [and] linkage between data types, and different modalities (Fujinaga & Riley, 2002, p. 1). Yet it remains a difficult problem in the field because, as Fujinaga and Riley claim, “musical scores are difficult to properly digitally capture and deliver for several reasons. They contain small details such as staff lines, dots, and bars that are essential to the meaning of the notation” (Fujinaga & Riley, 2002, p. 1).

The other, practically infinite, source of data generation in the field is the automatic transcription of audio files (i.e. the automatic conversion of audio data to MIDI data). Developing reliable automatic transcription tools is regarded as something of a holy grail in the field of MIR, because the datasets related to the the symbolic representation of music (found in forms such as MIDI and MusicXML) are far more amenable to data analysis techniques and indexing than is audio data. There has been extensive work in MIR with regard to automated transcription over the last 15 years, and much of this has focused on different audio data extraction tasks, such as methods for extracting rhythm, frequency or timbre (Raphael, 2001, p. 3). Much of the work in the space “can be roughly sorted into two categories: parameterised, such as statistical model based methods and non-parameterised, such as non-negative matrix factorisation based methods” (Gao, Dellandrea & Chen, 2013, p. 1). This includes the use of statistics, probability and stochastic methods for analysing audio files, often with a view to understanding what elements of sound files are most likely to consist of (i.e. by identifying a musical pitch made up of a fundamental and overtones, in various timbral and rhythmic settings). Other investigations in this space involve sound wave analysis, pitch correlation, and the position of the sound and acoustic modelling (Bello, Guiliano & Sandler, 2000).

Overall, the challenges in managing data in MIR are related to wider concerns around the way that information should ideally be indexed and archived. New approaches to these problems have been put forward, such as Lee's multi-feature index structures which have significantly sped up searching through a multimodal corpus (Lee & Chen, 2000).

Moving away from the storage, structure and generation of data in MIR, the next critical issue to address is how any kind of meaning in music might be derived from all of this data. The field predominantly utilises statistical and pattern analysis techniques to do this, and in the following section I will provide a survey of different approaches that have been used to analyse various types of music data. I will start by surveying the techniques used to analyse audio data, before turning to examples of analysis that utilise symbolic representations of music (such as Midi, MusicXML, and n-gram/text analysis), and will also examine the increasing number of automated music analysis projects that are appearing in the field.

The examination of audio data in MIR can be difficult to disentangle from the more general problem of the automated transcription techniques discussed above. Furthermore, using audio analysis to understand musical works can be a far more complicated process than examining the data taken from music scores. This is because the music score has a relatively limited number of non-ambiguous descriptors (encoding information such as frequency, duration and location, and various other metadata), whereas audio files can reveal far more information. Audio information contains the frequency of each note, but will also capture information pertaining to the overtones of all instruments that are present. It also encodes precision in rhythm (for example capturing timing information, where notes might be played just after or just before the beat).

Audio analysis tasks in MIR often utilise algorithms derived from other fields, such as digital signal processing, statistics and speech recognition. An example of this is Automatic Segmentation for Music Classification using Competitive Hidden Markov Models (Batlle & Cano, 2000), which utilises hidden Markov models, to track how notes move from one to another, aim to logically segment data so that labels can be applied.

Responding to the challenges of audio analysis, Pachet and Zils claim that “the exploding field of music information retrieval has recently created extra pressure [on] the community of audio signal processing, for extracting automatically high level music descriptors” (Pachet & Zils, 2001, p. 1). Unlike the music score, there are no agreed upon conventions that can be used to ascertain the relative importance of different aspects of data. As such, Pachet and Zils claim, “interestingness [in audio analysis] rather lies, extrinsically, in the confrontation [and] compromise between several music similarities or descriptors” (Pachet & Zils, 2001, p. 4).

A strong theme found in audio analysis research in MIR is the notion of similarity. Exploring similarity within audio files often starts with an examination of the relative differences found in various parts of audio data, such as interrogating audio spectrograms generated from different audio excerpts. This can uncover examples of audio data that are more similar to each other in some way, and categorisation can take place based on these similarities. Cliff and Freeburn (2000) claimed that this notion of similarity is “an intuitive criterion for indexing and classification of digital audio files in music information retrieval systems” (Cliff & Freeburn, 2000, p. 1).

In addition to the notion of similarity, audio analysis is concerned with uncovering structure. The exploration of structure however, is quite different to the way structure is explored in music theory, whose investigations were often informed predominantly by cultural and aesthetic assumptions about musical works. The structural investigation of audio files is instead concerned with applying mathematical techniques to find long term patterns. An example of this is Foote's Retrieving Orchestral Music by Long-Term Structure, which defines structure as the longitudinal presence of loud and soft passages within an audio file (Foote, 2000). This analysis attaches an "energy profile" (created from categorising the loud and soft passages) and ranks each audio document by a measure derived from the energy profile score, which can then be used to ascertain similar structural parts within audio files.

Another example of this is Jiang and Muller's Automated Methods for Analyzing Music Recordings in Sonata Form (Jiang & Muller, 2013). They problematise structure in music by claiming, "because of different structure principles, the hierarchical nature of structure, and the presence of musical variations, general structure analysis is a difficult and sometimes a rather ill-defined problem" (Jiang & Muller, 2013, p. 1). In addressing this, Jiang utilises audio analysis techniques to locate clusters of frequencies that can be used to infer modulation between different tonal centres, and can be seen as indicative of changing sections occurring within musical works.

Audio analysis has also allowed the examination of many aspects of music which were not feasible in traditional music theory or musicology, such as the mathematical comparison of similar timbral combinations, or the examination of specific techniques used by individual performers. An example of this can be seen in Bendor and Sadler's Time Domain Extraction of Vibrato from Monophonic Instruments, which sought to understand how vibrato worked in "slight oscillations in the pitch and/or volume of the musical tone" (Bendor & Sandler, 2000, p. 1). This kind of work can have applications into both real-time teaching tools and also be applied to some of the pre-processing steps that are required for tasks in automated music transcription.

Some audio analysis examples aim to limit the investigation to certain aspects of musical works, such as rhythmic patterns or tonal centres. This can be seen in Dixon et al's Towards Characterisation of Music via Rhythmic Patterns (Dixon, Gouyon & Widmer, 2003) which examined only the rhythmic patterns that might be extracted from audio data. Dixon completed an analysis of 698 musical works (in the genre of ballroom dance), locating temporal patterns as features, which could then be used to categorise other audio examples (Dixon, Gouyon & Widmer, 2003). Bello has also placed scope around audio analysis by limiting investigation to chord progressions that can be extracted from audio. This work used chroma features to isolate and categorise sounds into scale systems and Hidden Markov Models to probabilistically derive string representations of progressions. Success was then measured by the ability to locate similar audio passages across different audio files (Bello, 2007).

The derivation of tonality is also an important problem in audio analysis. Here, audio analysis research has focused on finding clusters of certain fundamental frequencies. Izmirli employed a "similarity metric between predetermined reference features and the analysed features from the audio" (Izmirli, 2009, p. 1) in order to derive tonality. This idea of key or tonality estimation can "inform many other tasks including music analysis, segmentation…song detection, modulation tracking, local key finding and chord recognition" (Izmirli, 2009, p. 3). Another example of this can be seen in Exploring African Tone Scales (Cornelis, Leman & Moelants, 2009), which explored the possibility that scale identification might be used to index large databases of music collections for ethnographic research.

A final example in the area of audio analysis is Flexer's A Closer Look on Artist Filters for Musical Genre Classification (2007). Music genre is not usually discussed in works of music theory (and is seen more in musicology), and this example demonstrates that, due its focus on data, the dualities between music theory and musicology can be revisited. Flexer proposes the "automatic classification of audio signals into user defined labels describing pieces of music" (Flexer, 2007, p. 1).

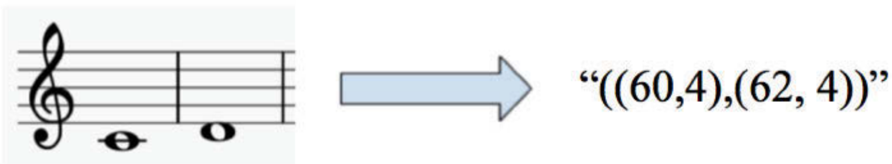
Undertaking pattern analysis using information from music scores (the symbolic representation of music) has the closest corollary to the examinations seen in music theory and analysis examples from the previous chapter. These types of investigations have usually involved converting information that can be encoded on a music score into text, and examining the patterns that might be found. Much of the symbolic data that is used tends not to be derived from MusicXML, but uses the MIDI specification, which is then encoded into a time-series representation, such as a sequential list of musical event data, holding attributes such as frequency, location, duration, volume and instrument name.

Investigations of the patterns found in strings of text has its origins in information retrieval techniques more generally, and has been used increasingly since the 1950s. This approach has the ultimate aim of "using computers to automatically search collections of unstructured online text" (Pickens, 2000, p. 1) to uncover meaning. It is possible that the "musical score can be viewed as a string" (Crochemore et al, 2000, p. 2), and there is an increasingly available body of MIDI and MusicXML data being made available online to inform this type of investigation (Rizzo et al, 2006, p. 1). Many of the methods of working with string representations of music are quite similar:

> The process for turning a music query into a text query is similar to that of turning a music document into a text document. The query "wrapper", the syntactic sugar, differs for each target system, but the basic method is the same. (Pickens, 2000, p. 5)

One commonly used string representation of music data is the n-gram. An n-gram is a structure that can be used to encode various aspects of musical information in a string-of-text format, and is often derived from MIDI or MusicXML. Figure 2.1 above demonstrated a MusicXML representation of two sequential notes, C and D, both of which were whole note durations. An alternate way to encode this information, using an n-gram, could be effected by using the string seen in Figure 2.2 below.

Figure 2.2. Two element n-gram



The above string consists of a list containing two elements, each enclosed in parentheses and separated by a comma. The first element contains information about a frequency (the number 60 is the MIDI number denoting middle C) and the number 4 denotes the duration in quarter notes. This is followed by a second element which has a frequency value of 62 (being the MIDI number denoting a D note above middle C) and the number 4 denoting duration in quarter notes. This type of encoding allows the creation of data sets that can hold specific information regarding musical works, in a way that is human and machine readable and creates a smaller data footprint than MIDI or MusicXML data.

An early example of n-gram pattern analysis in MIR was the SEMEX project, which first appeared in 2000 (Lemstrom & Perttu, 2000). It used n-grams to attempt to resolve some of the complications that can exist in musical works, such as music passages that appear in different keys. The authors set out a criteria that was aimed to establish similarity between music passages in different tonal centres. The SEMEX project addressed this problem by using a bit-parallel algorithm that focused on the numerical differences between each pitch, rather than on the individual pitches themselves, and also sought to isolate melodic phrases in polyphonic settings.

There have other examples in this area. Crochemore et al (2000) applied a similar method to the SEMEX data, in order to extract motifs and find melodic gaps (the time spans that elapse between melodic motifs). They represented the data in strings, encoding frequencies as MIDI numbers and intervals as the number of semitones between subsequent frequencies. This system also allowed the user to set the parameters of what constituted a gap (being a numerical value indicating an elapsed time) and returned melodic subsections that could represent motifs within a musical work (Crochemore et al, 2000).

The structure of n-grams has become increasingly complicated in the field. As early as 2001, Doraisamy and Ruger had sought to "encode rhythmic as well as interval information, using the ratios of onset time differences between two adjacent pairs of pitch events", in order to uncover structural elements in polyphonic music (Doraisamy & Ruger, 2001, p. 1). Debate also exists around how many elements should be included in an n-gram for use in pattern analysis. Pickens limits the length of n-grams to three elements, which can be used to explore smaller melodic figures. In these types of investigations, it is important that both the structure of n-grams, and the information they encode, are carefully managed in order to maximise the possibility of good results. In the case study chapter of this dissertation I will demonstrate how the ideal length of n-gram can be derived from the corpus under investigation.

The n-gram related studies highlight the difficulty of uncovering the nuanced structures that can be found in musical works. Because the human brain is highly adept at fuzzy pattern matching, it will easily uncover patterns in a music score, such as melodic motives that might be in different keys, or related rhythmic variations. Approaching these problems through n-gram pattern matching highlights the complexity of trying to automate such a process, but has the advantage of being able to utilise far larger datasets than the human brain can cope with. As an example of this, Caplin has used symbolic data from the music of Haydn to propose a formal set of features that might be used to encode symbolic data, and differentiate between melody and harmony in polyphonic music, that could lead to a useable algorithm for automated music analysis (Caplin, 2000).

Other symbolic pattern analysis work can be directly related to the field of music theory. Kirlin and Jensen have proposed that Schenkerian analysis exhibits "statistical regularities that can be represented, discovered, and reproduced" (Kirlin & Jensen, 2011, p. 1) and that it may be possible to create an algorithmically based methodology that could be used to apply Schenkerian analysis to an arbitrary corpus of music.

A final example of the use of n-grams to explore basic music information (such as location, duration and frequency) is Vladimir Viro's Peachnote project (Viro, 2011, p. 4). Viro describes the ambitious project in the following manner:

> Our system takes the scores in PDF format, runs optical music recognition (OMR) software over them, indexes the data and makes them accessible for querying and data mining. The search engine is built upon Hadoop and HBase and runs on a cluster. Our system has already recognized more than 250 million notes from about 650 thousand sheets. (Viro, 2011, p. 1)

Viro built in an n-gram search capability built into the system, (Ngram Viewer), which lets users "select the time range and get the list of scores composed during this time which contain the given note sequence [from the user]" (Viro, 2011, p. 2).

Increasingly, hybrid approaches are being taken in musical analysis within MIR, in which investigation can be both wide ranging and large scale. Analysis can include exploration of such things as "scores, lyrics, photography and artwork, and other associated metadata" (Weigl & Guastavino, 2011, p. 1). This signifies a substantial shift from the kinds of investigations encountered in the last chapter, which took the form of a curated examination of a small group musical works or the study of the practices found in particular time periods.

One of the advantages of undertaking an increasingly hybrid form of analysis that includes symbolic music analysis along with other metadata forms, as opposed to more traditional music score analysis in the field of music theory, is that the attributes under consideration (things such as frequency, duration, location, dynamics etc) can easily be scaled. Encoding symbolic data as a time-series data set can be extended to other types of data also. It becomes possible to incorporate other information, (attributes such as the year of composition, the cycle of works in which a musical work belongs, geolocation data, personal information about the composer) and undertake more nuanced queries. It is even possible, for example, to collect the reviews of a work's public performance and undertake sentiment analysis that could indicate a musical work's popularity, which can then be tied back to the compositional devices used by composers, and influence their theoretical importance. Taking a more longitudinal approach to data means that it is possible to evidence test many of assumptions that exist in music. This could include the tracking of instrumental combinations, and harmonic progressions longitudinally through time and place. The software application that accompanies this dissertation is designed to meet these types of requirements.

Taking this approach allows the domains traditionally inhabited by music theory and musicology to be blurred. Both musical works, and the environment in which they exist, can be regarded as non-ambiguous sites of different yet compatible metadatas, all of which can be analysed with similar techniques, and this can lead to results about how a large corpus of music tends to behave.

Work can be seen in the field already that is heading this goal. In Calculating Similarity of Folk Song Variants with Melody-Based Features (Bohak & Marolt, 2009), the authors claim that it is "possible to classify folk song melodies into correct variant types based on statistical features of their melodies alone" (Bohak & Marolt, 2009, p. 1), and use various melodic and rhythmic attributes (as well as a notion of entropy) to cluster the various examples together. They arrive at the powerful conclusion that, just by examining melody in a longitudinal fashion without reference to its context, it is still possible to categorise different types of folk music that originate from different places, thus providing an evidence based historiographical dimension to music understanding.

A second example is Kiernan's Score-based style recognition using artificial neural networks. This study applied machine learning techniques to differentiate the geolocations of compositions of musical works (Kiernan, 2000). The work demonstrated that the compositions of Frederick II, Quantz, and Bach could be traced to different geolocations confirming "that statistical data is sufficient in the identification of individual musical characteristics" (Kiernan, 2000). This work is an important crossover into the field of musicology, and found a stronger similarity between the works of Frederick II and Quantz than to those of C.P.E Bach, "thus supporting historical speculation concerning musical allegiances" (Kiernan, 2000, p. 1).

This type of hybrid and longitudinal music analysis also has the potential to be large and increasingly automated. Examples of this include Design and Creation of a Large-Scale Database of Structural Annotations (Smith, Burgoyne & Fujinaga, 2011), a project which aims to "produce structural analyses for a very large amount of music, over 300,000 recordings" (Smith, Burgoyne & Fujinaga, 2011, p. 1). The work is aimed at partitioning large amounts of data into different sections. Rather than examining structure at the note level (such as the individual durations, frequencies and locations of note events) this research explores music at a more abstract level, identifying similar sections that might occur within different musical works.

The use of large scale analysis can also be seen in Antila and Cumming's article, The Viz Framework: Analyzing Counterpoint in Large Datasets. The authors created the framework specifically to undertake big data queries of symbolic music data, claiming:

> Until recently, musicologists' ability to accurately describe polyphonic textures was severely limited: any one person can learn only a limited amount of music in a lifetime, and the

> computer-based tools for describing or analysing polyphonic music in detail are insufficiently precise for many repertoires. (Antila & Cumming, 2014, p. 1)

The authors also problematised personal expertise being used as way to undertake music analysis, because of its tendency to limit investigation to "intuitive impressions and personal knowledge of repertoire" (Antila & Cumming, 2014, p. 2). Additionally, the authors note that assumptions made in traditional score analysis are seldom tested, and when they are, these assumptions can often be seen as incorrect. On their investigation of musical works from the renaissance period they note:

> Certain patterns that musicologists consider to be common across all Renaissance music are in fact not equally common in our three test sets. For example, motion by parallel thirds and tenths appears to be more common in certain style periods than others, and in a way that does not yet make sense. (Antila & Cumming, 2014, p. 5)

The above example demonstrates that the ability to verify assumptions of how music behaves is a powerful strength in MIR. However, it is important to temper this strength too: abandoning individual expertise is problematic in MIR, in that it can render the purpose of an investigation ambiguous. The challenge in the field will be to create verification frameworks that can work in tandem with individual expert understanding. This is also related to an issue of how users are constructed and function in MIR, to be discussed later.

This increasingly hybrid research makes it possible to come full circle, to merge both audio data analysis and symbolic data analysis. An example of such an attempt can be seen in the article Sparse Music Decomposition onto a MIDI Dictionary driven by Statistical Musical knowledge that aims to "sparsely decompose the music signal onto a MIDI dictionary made of musical notes" (Gao, Dellandrea, & Chen, 2013, p. 1). The authors claim that:

> Large amounts of digitalised music available drive the need for the development of automatic music analysis, for example automatic genre classification, mood detection and similarity measurement. (Gao, Dellandrea, & Chen, 2013, p. 1)

The authors also position the discrete information that can be encoded onto music scores (such as what is encoded in MIDI or MusicXML) as being ideal in providing "the most comprehensive information, since music is indeed sound poetry comprised of notes played by instruments" (Gao, Dellandrea, & Chen, 2013, p. 1). Thus, it is not only large volumes of data, and different types of data which are important in MIR, but also their quality and suitability to data analysis tasks.

The myriad of different approaches in MIR analysis has inevitably had an impact on how music analysis should look. Data visualisations in MIR have become increasingly complicated, which can be seen in both commercial and research settings. They explore music information that contains both large and small structures, as well as numerous integrated metadatas.

The way that music should look to the human eye has a long and varied history, and there are many examples of composers and music theorists who have sought to use alternative visualisations to encode musical information. This is also an important issue in MIR that has been explored. In Visualising Music: Tonal Progressions and Distributions, Mardirossian and Chew claim:

> Music visualisation literature can be broadly grouped into two categories: visualisation of individual pieces of music (our focus), and of collections of pieces. It can be said that the first form of music visualisation created for individual pieces was music notation itself. An experienced musician can often look at the score of a piece and "see" what the music sounds like. (Mardirossian & Chew, 2007, p. 1)

The authors go on to problematise the difficulty of working with traditional music notation visualisations, calling for alternates that are both more intuitive, and which can better capture the hierarchical information that tends to be generated from music. They note that "it can take years of training to learn to decipher the subtleties of the encoded information" (Mardirossian & Chew, 2007, p. 2) and a principle barrier of entry to existing music visualisations is the music score itself. They address this with an attempt to "create a more intuitive visualisation that can reveals important features of the music that may not be readily audible to the inexperienced ear" (Mardirossian & Chew, 2007, p. 2), by "using visualisations that include dimensionality, colour, and animation" (Mardirossian & Chew, 2007, p. 2).

There are a number of existing, large scale projects and applications, that bring together many of these approaches. They are an important showcase of the potential of music theory and analysis to be multimodal, to utilise numerous different types of data, to work with hierarchical information, and to use a range of different visualisation techniques.

The first of these projects is the commercial application, Chordify. Chordify is is an online web application that provides an "automatic chord extraction service where users can create their own personalised chord sequences" (Bas de Haas et al, 2012, p. 1). It provides users access to a large repository where "different chord label sequences of popular songs [can be] obtained" (Bas de Haas et al, 2012, p. 1). Chordify does not provide a theory about how chord progressions should ideally be structured. Instead, this expertise is crowd sourced (through the act of users accessing chord progressions, and uploading their own chord progressions). Users can also share what they are exploring and which progressions they are learning and easily share this to various social media platforms. The site is multimodal and allows users to hear and see progressions (in a format similar to a piano roll) and play the audio of the original recording. This suggests a rethinking of how harmony works in music. Its rules are being inferred in real time by the activities that take place on the website by users.

A second example is the Jazzomat project. This again, is a multimodal music analysis project that commenced in 2011, which aims, according to its website, to "investigate the creative processes underlying jazz solo improvisations with the help of statistical and computational methods" as a means of exploring "the cognitive and cultural foundations of jazz solo improvisation". Researchers collected various metadata on 299 jazz solos including transcriptions, midi files (seen in Figure 2.3), discographic information, chord changes and biographical information (Figure 2.4.). Additional basic statistical information was also included about the time-series information in the transcription (examining location, duration and pitch (Figure 2.5)

Figure 2.3 Use of midi and audio files in Jazzomat

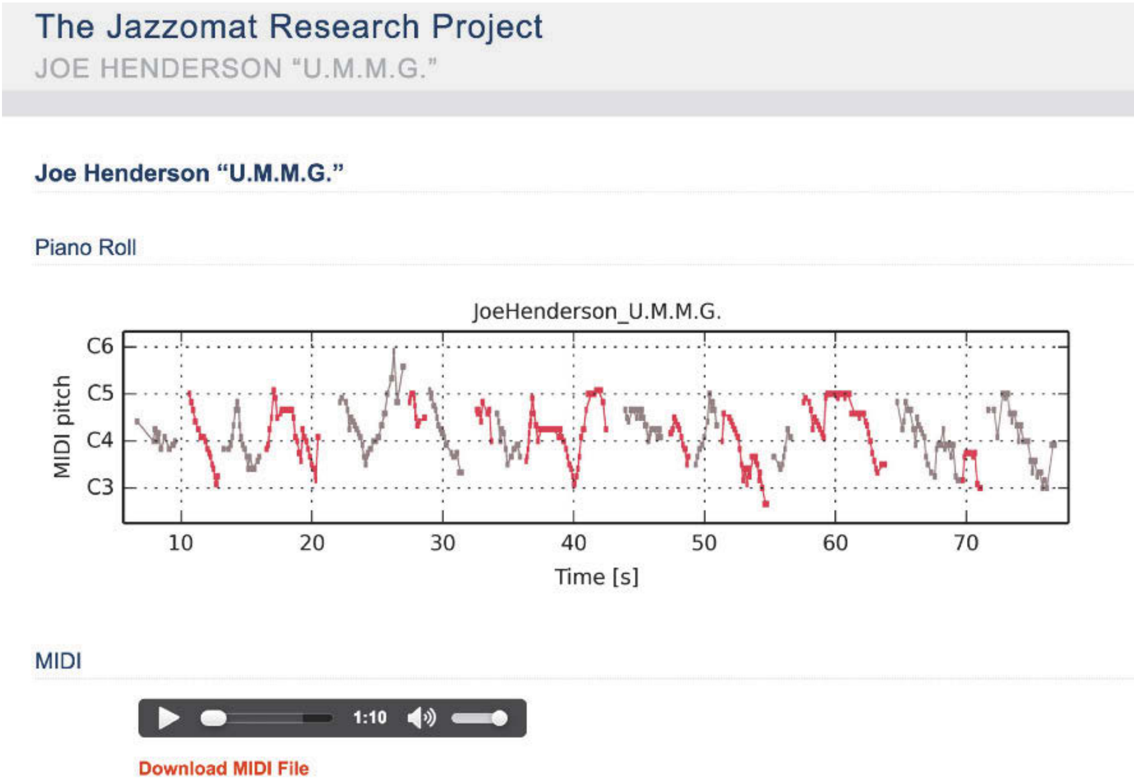


Figure 2.4. Discography, chordal progressions, and biography information in Jazzomat

Discographic Information

Instrument	ts
Total Duration	70.3 s
Position in Track	00:45-01:56
Album Title	Lush Life: The Music of Billy Strayhorn
Label	Verve (511 799-2)
Recording Date	03.-08.09.1991
Line Up	Joe Henderson (ts), Gregory Hutchinson (dr), Wynton Marsalis (tp), Christian McBride (b), Stephen Scott (p)
Style	POSTBOP
Key	Db-maj
Signature	4/4
Groove/Feel	SWING
Tonality Class	FUNCTIONAL
MusicBrainz ID	e301192d-6f25-4775-874b-a9d1d43f49b9
More Info	Wikipedia

Chord changes of solo:

```
A1: ||Fm7b5   |Bb7   |Eb-7   |Ab7   |Dbo7 Db6 |Dbj7   |Db-7   |Gb7   ||
A2: ||Fm7b5   |Bb7   |Eb-7   |Ab7   |Dbo7 Db6 |Dbj7   |Ab-7 Db7 |Ab-7 Db7 ||
B1: ||Gm7b5   |C79b   |Fj7   |Fj7   |Abm7b5   |Db79b   |Gb-   |Ab7   ||
C1: ||Fm7b5   |Bb79b   |Eb-7   |Ab7   |Dbo7   |Dbo7   |Db6   |Db6   |Dbo7   |Dbo7   |Db6   |Db6   ||
```

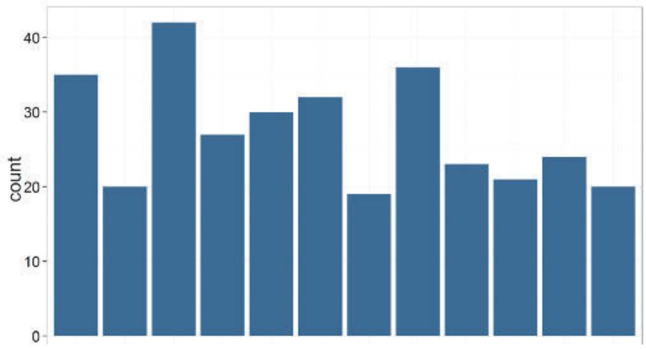
Figure 2.5. Aggregated statistics in Jazzomat

Statistics

Features ¶

Number of Notes	329	
Number of Bars	64 bars	
Number of Choruses	2	
Mean Tempo	219.9 (UP)	
Event Density	4.68 Notes/s	
Metrical Event Density	5.12 Notes/bar	
Median Swing Ratio	1.55:1	Ratio of longer to shorter eighth of beats with binary subdivision.
Metrical Centroid	2+	Mean concentration of events in the bar, rounded and normalized to 4/4.
Syncopicity	24.6 %	Share of syncopated notes.
Ambitus	44-83 (39)	Min/max and range of MIDI pitches.
Extrema ratio	41.0 %	Share of notes with direction reversal (i.e, minima and maxima of pitch contour).

Chordal Pitch Class Histogram



The aggregations in Jazzomat are currently limited. Yet this project, like Chordify, signals a potentially powerful move in music theory and analysis. It positions the music score as just one of a number of different sets of metadata which can be added together and interrogated. Chronological information, geolocation information, and biographic information, can all be data mined in the same way as the music score. The manner in which the data is collected is also scaleable.

A final example is the popular music recommendation service, Spotify. This is another web application that allows large numbers of users to implicitly encode their opinions about what they view as good and bad in music, and compile and access their own curated playlists. They do this simply by choosing to listen to certain pieces of music rather than others. The psychological mechanics of what might underpin these preferences are not the focus here. Spotify can generate data about user behaviours in regard to music and this data can be mined to find meaning in music. Zhang et al note:

> We found that in Spotify, not only session arrivals, but also session length and playback arrivals exhibit daily patterns. For individual users, we first studied the behavior of switching between desktop and mobile devices for using Spotify. Second, we found that Spotify users have their favorite times of day to access the service. Third, we observed clear correlations between the session length and downtime of successive user sessions on single devices. (Zhang et al 2013, p. 17)

The collected data of these online streaming services has the potential to be unlimited. As of June 2016 Spotify had 100 million registered users, who were actively listening on a daily basis. The data can be used to ascertain not only the things that particular individuals regard as preferable and not preferable. It can also be used to view the trends across an entire population of listeners. This approach makes it possible to utilise this data in order to make recommendations to users of the application. Spotify also provides a weekly playlist to all of its users, which Matthew Ogle (of the Spotify discovery playlist) claims:

> There's two parts to it. First, we look at all the music you've been playing on Spotify but we give more emphasis to the stuff you've been jamming on recently. Something that you played yesterday is probably more interesting to you than something you played six months ago. But the real core of it is looking at the relationships between songs based on what other users are playlisting around

> the songs that you've been listening to and essentially finding the missing ones – the ones you haven't heard yet, or maybe haven't heard much. (Ogle 2016, para 3)

The Spotify model (which is also seen in services as Pandora and Apple Music) of allowing an aggregated user to determine what is good and bad in music, again challenges the author as expert model seen in more traditional forms of music analysis. Instead of positioning an individual who will provide a judgement on what is good or bad music, this judgement is generated from an aggregated outcome of behaviours exhibited across the population of users.

Though the idea of drawing information from the interactions human beings have with music in order to understand its meaning is an attractive one, it can also be problematic. The specificity in the kinds of studies seen in the previous chapter, such as theoretical works of Hindemith, Schoenberg, Rimsky-Korsakov and Rameau have given way in MIR to analyses that can be far more wide ranging, and whose scope is scaleable, yet whose audience is somewhat ambiguous. Services such as Spotify, Chordify and Jazzomat cater for very different audiences, none of whom are defined, and who will be seeking out music related information for different ends. This can leave the MIR in the position of revealing a great deal about about music, but it also runs the risk of revealing it to no one in particular. As individuals, the questions we pose toward music are deeply personal, and for end users in MIR systems, it is not clear how these will be answered. Guastavino and Weigl have claimed that the field has a "system-centric" focus (which they see as having been motivated, to some extent by textual information retrieval which have influenced the field dating back from 1950s) which problematises the role of the end user in the field (Weigl & Guastavino, 2011, p. 1).

Part of this problem relates to the complexity that characterises the human relationship with music, and the large space in which MIR operates. Much of the work undertaken in music theory held the assumption that music was the product of a creative artist, and the perfection of its construction was mediated by this truth. However this is not at all the case. Music is not something that has a fixed relationship to us or means any particular thing. Our relationship to music changes over time, and will reveal profoundly different things in different contexts. Weigl and Guastavino capture this eloquently when they claim "an ethnomusicologist's analytical requirements are likely served by queries of a different nature to those used by a party host compiling a playlist" (Weigl & Guastavino, 2011, p. 1). Huron also notes:

> Music is used for an extraordinary variety of purposes: the restaurateur seeks music that targets a certain clientele; the aerobics instructor seeks a certain tempo; the film director seeks music conveying a certain mood; an advertiser seeks a tune that is highly memorable; the physiotherapist seeks music that will motivate a patient; the truck driver seeks music that will keep him/her alert. (Huron, 2000, p. 1)

It can be difficult even to begin teasing out the surface of this relationship. For example, A Cross-cultural investigation of the perception of emotion in music: psychophysical and cultural cues (Balkwill & Thompson, 1999), has sought to explore the role that cultural background plays in music perception. The authors interviewed people from different cultural backgrounds, who listened to excerpts of Hindustani ragas, specifically chosen as the works were from a relatively unfamiliar tonal system. They asked participants to identify emotions they believed would be associated with the music. Findings showed that while the emotions of joy, sadness, and anger, were “identifiable by the listeners and the emotional judgments were significantly related to psychophysical characteristics of the pieces”, pain was not (Balkwill & Thompson, 1999, p. 64). The authors followed up with a second paper that explored the differences between American, Korean and Chinese responses to musical works. They discovered that American and Chinese listeners perceived music in noticeably different ways, and Korean listeners seem to share traits of both Chinese and American listeners. They also noted that gender was a key differentiator between American and Korean groups, whereas age differentiated Korean and Chinese groups. This suggests that our relationship to music is extremely complicated, and it is these complications that somehow need to be taken into account.

The challenge this leaves for MIR is how to conceive of an end user who can interact with the analytical models put forward in the field. Verco and Chai (2000) posed the following questions and answers with regard to users in MIR:

> How to model the user? User-programmed, machine learning and knowledge-engineered methods can be used. 2) What information is needed to describe a user for [MIR] purposes? It may include both the user’s indirect information (e.g. age, sex, citizenship, education, music experience, etc.) and direct information (e.g. user’s interests, definition of qualitative features, appreciation habit, etc.). (Verco & Chai, 2000, p. 2)

In their 2011 article, User Studies in the Music Information Retrieval Literature, Weigl and Guastavino argued that there needs to be more work carried out in determining the user requirements in the field, making the reflection that, “articles reflecting on the state of MIR have repeatedly called for a greater focus on the potential users of MIR systems" (Weigel & Guastavino, 2011, p. 335). Downie has also noted that this "multi-experiential" challenge in MIR, relates to the "subjective musical experiences varying not only between, but also within, individuals" (cited in Weigel & Guastavino, 2011, p. 335).

In 2000, at the time of MIR's infancy, Bonardi provided a prescriptive account of what he believed the field might contribute to the kinds of music analysis seen in more traditional models. He noted:

> The musicologist is facing a computer screen, while handling scores and books. This terminal allows him, among many other possibilities, to listen to music, to access musical databases and hypermedia analyses. The musicologist is handling several devices on several media at the same time. First of all, the listener needs a framework that takes him/her into account. The purpose is to set the conditions of possibility of listening by restricting the heuristics of "forms". It is therefore necessary to set a listening framework for the musicologist, to assist him in discovering the "intentions" of music. The main feature of this listening environment is thus its capacity to enable its user to vary the music representation. (Bonardi, 200)

At this time, Bonardi called for systems to be constructed that would allow real-time interaction and feedback. They must "enable rapid changes of the representation of abstract objects" (Bonardi, 2000). Such systems should propose to the "listener/musicologist to build [his or her] own adequate structures to look for forms using specific languages to encode the patterns, either global or local. (Bonardi, 2000).

It is an ambition that poses a daunting challenge to music theory analysis in MIR: in order for any model of music theory or framework of analysis to be viable, it needs to be both attuned to the requirements of its users, related to a specific corpus of musical works, and be responsive to changes in both. The model in framework should be able to change depending on who is using it.

Locating and constructing a user in MIR who can be positioned to explore music on many levels is a critical problem. Weigl and Guastavino claim that:

> If the "Grand Challenge" of the field is to provide a fully integrated system providing all manners of MIR access, a firm focus on user requirements is important. using it and the musical works it refers to (Weigl & Guastavino, 2011, p. 337).

Though building these structures may seem a daunting task, it can become possible. To do this, musical works need to be understood as producers of potentially wide-ranging metadata and the user interaction must be integrated into this information. In this way, models of music theory and frameworks of music analysis can become customised to individuals, and mediated through groups of individuals.

Chapter 3

Jazz Improvisation and the style of Keith Jarrett

The chapter will begin by examining some of the practical problems that are encountered when seeking to undertake analysis of jazz improvisation, and the lack of information that this is often characterised by. It will survey various approaches taken in jazz analysis and relate them to more traditional models of music theory and analysis. It will frame some of the difficulties of jazz analysis as foundational problems related to the often opaque definitions and shared understandings of jazz improvisation. Finally, it will locate the improvisational style of Keith Jarrett (whose improvisations will be examined in Chapter 5) within this context and summarise both his personal views on improvisation, and the various analytical approaches that have been taken to explore his music.

Although the application of music analysis within other genres has certainly been more prolific than that of jazz, since the mid-1980s there has been an, "enormous grown in jazz theory scholarship" (Larson, 2009, p. 2). Some of the approaches used can find strong parallels to the approaches taken in jazz, and many models focusing of jazz analysis can be viewed within a context whose lineage can be traced back to the writings of Aristoxenus. Yet at the same time, jazz improvisation is something altogether different. Martin couches the challenge by saying, "groups of related and overlapping theoretical models delimit sub styles within broader musical genres", (Martin, 1995, p. 16) suggesting a connection between the type of music analysis and its genre, which will have an an impact on the model used, and this seems particularly applicable to jazz. According to Martin, the goal of musical analysis in jazz:

> Largely concerns itself with discovering (and sometimes inventing) sets of rules that model various kinds of musical

> structure. These models attempt to show how a piece "works" or how music in some given style is written or performed (Martin, 1996, p. 1)

The concerns and challenges that inform the analysis of jazz improvisation have shown that it is fundamentally different from other models so far encountered. Unlike many of the music analysis models encountered in chapter one that leveraged off highly structured information (predominantly being complex music scores), the majority of jazz music is not notated. It is instead found in recordings, and has no associated music score. As such, it is often not at all practical to use a vehicle such as the music score to interrogate what happens in jazz. Unlike much western music in which the music score precedes the performance or recording, and aims to provide as detailed instructions as possible for performers to recreate it, the jazz score functions only as an optional extra, optimised to the wide ranging interpretations of different jazz sub-genres. As such, the use of the score in jazz is a highly simplified affair, capturing only partial information, and usually from only some of the instruments that are present. A complete transcription of all the instruments within an jazz ensemble is also extremely rare. Smith notes the resulting analytical challenge as follows:

> Since music lacks specific meaning and grammatical categories of the sort found in language, the [jazz] musical analyst is deprived of the tools with which linguistic formulas are discovered. Unless comparable tools are devised for isolating recurrent melodic ideas, the formulaic analysis of melody is condemned to census-taking, to tallying up the literal repetitions of randomly encountered pitch sequences. (Smith, 1983, p. 11)

Despite the problems with regard to the ways jazz improvisation might be encoded on the score, there still exists an extensive body of literature and materials that claims to understand how jazz improvisation works, which utilises music score information. As well as academic writings, much of this is found in the form of instructional texts aimed at aspiring jazz musicians. These types of resources summarise skills and techniques that can be transferred in a digestible fashion, and are often backed up by recordings of the concepts under discussion. Against such a backdrop the music score is not so much of an authoritative text, but rather an incidental convenience that can facilitate training. Scores are often found in the form of lead-sheets that players will interpret in a way that they deem suitable. Thus, in jazz, more often than not, there is simply "no score to examine" (Dean, 1992, p. 28).

All of this raises practical difficulties when undertaking any kind of analysis of jazz improvisation: there is no score to examine, and the techniques to automatically transcribe jazz audio recordings to not yet exist. In order to even begin a process of analysis, the theorist must first decide how the aural information is to be dealt with, and if it can be converted in some way to make it more amenable to analytical tasks. This is most often achieved by the painstakingly manual task of transcribing the notes of a recording. Reflecting on the process, Hodson notes that though, typically, "an analyst will need to create a transcription to aid the discussion of a recorded performance", a process which presents significant barriers to accessing a corpus for analytical purposes. (cited in Dean, 1992, p. 2).

As an illustration of just how difficult it can be to obtain pre-prepared transcriptions of jazz improvisation in some kind of score based format, out of the ten solos to be explored in case study chapter of this dissertation, none of them have been published elsewhere (there are no professionally published jazz transcriptions of Keith Jarrett jazz improvisations over jazz standards). Of the ten solos (comprising around 16,000 notes) Only three of the transcriptions could be found via the internet, and these differed markedly from my own transcriptions. Additionally, although these were taken from jazz trio performances, there is no information pertaining to the double bass and drums, and the piano transcription is the right hand only, making it impossible to view these transcriptions as a traditional score which might be used to recreate the exact performance in any meaningful way.

Dean claims that there is something “fundamentally different in the transcribed solo” (Dean, 1992, p. 7), and Hodson, echoing the sentiment, claims that “with regard to the issue of whether a transcribed improvisation is comparable to a composed score and can be analysed as such, a number of authors express differing viewpoints” (Hodson, 2007, p. 2). Hodson also casts doubt on the possibility that existing and accepted analytical models might be applicable to jazz improvisation (and points to what he views as the problematic Schenkerian analysis that has been undertaken on solos by Bill Evans, Oscar Peterson and Thelonius Monk in Larsen’s Schenkerian Analysis of Modern Jazz).

The foremost problem of accepting that a jazz transcription could have an equivalent validity to a more traditional music score in terms of the aural information it can hold, is that it simply lacks so much of the nuance of the recorded performance. Music notation of rhythm, being “simply a symbolic representation based on mathematical ratios” (Busse, , 1999, p. 444) cannot hope to capture the subtle rhythmic structures that are so idiomatic of jazz[2]. Although in previous chapters I have raised the issue of the music score’s status as a metadata, this problem becomes particularly vexed when it comes to jazz as the

[2] To highlight how little of the nuance the jazz practice transcription actually captures, consider the track at https://soundcloud.com/jamie3103/all-the-things-you-are . This uses the transcription of All The Things You Are which will be featured in the analysis chapter, but the notes have been assigned to modern synthesised instruments, tempo slowed for the purposes of ear training.

same metadata can be drawn from different music styles in jazz. Different performers will approach the same jazz standard in extremely different ways, which are often highly dependent on both the other musicians present, and sub-genre of jazz in which they play (Busse cites examples of performance evaluation from Boyle, 1992, Cooksey, 1982, Fiske 1983, and George, 1980).

Much jazz theory and analysis however, does make extensive, yet pragmatic, use of score based transcriptions. Examples in this space also includes analysis that leverages off more traditional approaches such as Schenkerian and Neo-Reimannian music theory. In his 1998 article, A Schenkerian Analyses of Modern Jazz, Larson applied Schenkerian techniques to transcriptions of Oscar Peterson, Bill Evans and Thelonious Monk and, when juxtaposing differences between the musicians, claimed:

> [Peterson's and Evans's] solutions elevate the relationship-between-the-parts of Monk's theme to the level of a premise: the linking motive's hidden repetitions become a premise of Peterson's performance, and the closing motive's delay of dissonance resolution becomes a premise of Evans' performances (Larson, 1998, p. 210)

The idea that a Schenkerian approach is implicit in the improvisation process is something that others find difficult. Heyer takes issue with Larson's approach, noting it to be somewhat problematic that "improvising musicians really intend to create the complex structures shown in Schenkerian analyses" (Heyer, 2012, p. 4). For Heyer, Larson's argument that Bill Evans "has in mind an improvisational approach based in Schenkerian principles, which Evans applies consciously, and in real time, to his improvising" is simply not viable (Heyer, 2012, p. 4). While Martin praises this work as a rich and expansive treatment (and also a "tour de force" of transcription) he finds it problematic that Larsen could apply Schenkerian principles so rigidly to the analysis.

Another example, strongly rooted in an existing music analysis framework, can be seen in Briginshaw's work A Neo-Riemannian Approach to Jazz Analysis. According to Briginshaw, the Neo-Riemannian theory has particular relevance in jazz analysis as it "originated as a response to the analytical issues surrounding Romantic music that was both chromatic and triadic while not "functionally coherent" (Briginshaw, 2012, p.57). The complexity of jazz harmony, in that it is characterised by upper chordal extensions and intricate voice-leading, along with melodic phrases that utilise all twelve pitch classes, was well suited as an extension to the Neo-Riemannian "Tonnetz" a geographical rendering of pitch-space that aided in the explanation of rapid modulatory passages (Briginshaw, 2012, p. 59).

Other applications of this type of analytic approach have included Strunk's Notes on Harmony in Wayne Shorter (Strunk, 2005) which claimed that the Neo-Riemannian representation of transformations among tetrachords was ideal when examining jazz music, as it offered a conceptual basis from which to accommodate dominant sevenths and half-diminished sevenths in the context of a larger harmonic design. A final example can be seen concerning Pat Martino's style in The Nature of the Guitar: An Intersection of Jazz Theory and Neo-Riemannian Theory, (Capuzzo, 2006). This paper explored the teaching materials used by jazz guitarist Pat Martino and placed them in a framework of Neo-Riemannian theory, positing that it was highly correlated to the way Martino explains the complexity of his music when teaching, for the purpose of helping students access novel methods of instrumental practice (Capuzzo, 2006).

Like much of traditional theory music theory and analysis, jazz analysis is often problematised by deeper questions around its meaning, author intent and opinion, and is criticised on the basis of apparent writer bias. A telling example of this can be seen in Gunther Schuller's analysis of the Sonny Rollins solo on the jazz standard *Blue Seven*. Here, Schuller posits that that the entire solo "organically grows" out of a two-note motive stated at the solo beginning (Schuller 1958, p. 8). He explains that although, amongst some improvising musicians,"there appears a tendency to bring thematic, motivic, and structural unity into an improvisation", the average improvisation is "mostly a stringing together of unrelated ideas". But this "lack of structural coherence is not altogether deplorable" according to Schuller (Schuller, 1958, p. 9). Schuller cites Rollins as the exception to the rule, whose improvisational abilities are "symptomatic of the growing concern by an increasing number of jazz musicians for a certain degree of intellectuality" (Schuller, 1958, p. 10). For Schuller, Rollins' approach signals a move toward the thematic unity that improvisation so sorely needs.

Some took issue with the article, arguing that it misrepresented Rollins' intentions, and read too much into structures that were not really there. Walser located the work as being more concerned with inculcating jazz improvisation into the language of musicology than uncovering any implicit structural meaning, and claimed:

> Though it is clear that Schuller, along with everyone else, hears much more than that in this recording, his precise labelling of musical details and persuasive legitimation of jazz according to longstanding musicological criteria caused many critics to hail this article as a singular critical triumph (Walser, 1993, p. 344)

Walser also questioned the weight of Schuller's conclusions. While the Rollins improvisation made sense, in that the jazz improvisation appeared coherent to those who have the relevant, domain specific knowledge, the depths which Schuller claimed are simply not there:

> All it really tells us about Rollins, however, is that his improvisations are coherent; it says nothing about why we might value that coherence, why we find it meaningful, or how this solo differs from any of a million other coherent pieces of music. (Walser, 1993, p. 350).

In a similar vein, Smith problematised Frank Tirro's analysis of Charlie Parker which explored the saxophonist's "syntactic coherence and hierarchical structure". Smith claimed that, "not once does Tirro demonstrate the syntactic function of the reworking of previous material or how it contributes to the structural coherence of the music" (Smith, 1983 p. 55).

Overall, these articles speak to the problem of relating a performer's intent to the data under consideration. Walser notes that:

> One of Davis's biographers asserted that the "My Funny Valentine" solo demonstrates "no readily apparent logic," while another waxed enthusiastic about its "dramatic inner logic." Each critic found it a powerfully moving performance, but both lacked an analytical vocabulary that could do justice to their perceptions (Walser, 1993, p. 49)

As well as the lack of the availability of scores that make jazz analysis difficult, getting accurate data from critical aspects of the music is also highly problematic. While it is often feasible to approximate pitch when manually transcribing jazz, finding the exact point at which a note is played in regard to the underlying beat can be extremely difficult. Yet the placement of notes in regard to the beat is one of the most important aspects in describing jazz improvisation. On this issue, Smith notes that in jazz, “it is the rhythms, not the pitches, that create the resistances, and the pulse or beat, not harmony, that provides the points of resolution (Smith, 1983, p. 94). However there is very little jazz analysis related literature that explores this.

Mazzola and Cherlin take the problem further, suggesting that jazz, opposed to other genres of music, actively emancipates the problem of time in music. Of the changes that have taken place in the way musicians conceptualise time, they note that:

> [Time] made the move from facticity to the level of making: time became a thing to be constructed from scratch. No more tyrannic clocks, no more eternal lines, no lines at all. We make time, we are the new hands, and the clock, and the gestures, which mould time. Not surprisingly, such expressive making also changed the time’s stature: physics’ anorexic timeline was transmuted into a voluminous body of time as shaped by the powerful hands of working musicians (Mazzola & Cherlin, 2008, p. 52)

Attempting to locate consensus and even a limited evidence base from which undertake analysis of jazz improvisation can be profoundly challenging. Even if more transcriptions are made available, the music score as a structure is not equipped to hold critical information needed to explore jazz.

These difficulties can also be linked a more foundational problem: it is not readily agreed how jazz improvisation should be defined and understood. Even if one accepts that a music score might be pragmatically accepted as a medium through which to meaningfully access a corpus of jazz improvisation for the purpose of analysis, a definition of what jazz improvisation actually is proves elusive. Jazz improvisation is variously discussed in the literature and its related resources as a practice, process or a product. Its meaning is ambiguous.

In the Grove online music dictionary, improvisation is defined as follows:

> The creation of a musical work, or the final form of a musical work, as it is being performed. It may involve the work's immediate composition by its performers, or the elaboration or adjustment of an existing framework, or anything in between. To some extent every performance involves elements of improvisation, although its degree varies according to period and place, and to some extent every improvisation rests on a series of conventions or implicit rules (http://oxfordindex.oup.com/view/10.1093/gmo/9781561592630.article.13738/ (2018)).

By locating improvisation within the context of "conventions" and "implicit rules", the definition shares a similar language found in more traditional music theory and analysis. But it is a definition that is one of many however, and it is this ambiguity that makes it hard to pin down any lasting agreement on what jazz improvisation really is. When reflecting on the differing definitions of jazz improvisation, Smith points to the problematic dichotomy that underpins it: for some, it is understood as a creative process, and by others the result of a creative process. The upshot of the dichotomy is that it is "not always clear, therefore, if one means by "improvisation" the way the music is created, or the music that is created" (Smith, 1983, p. 88).

Furthermore, it is also often not clear if jazz improvisation refers to something solitary (which examines the activities of only one musician either playing alone or in an ensemble), or if it should be regarded as a collaborative affair. On this, Hodson claims:

> Most technical writings on jazz focus on improvised lines and their underlying harmonic progressions. These writings often overlook the basic fact that when one listens to jazz, one almost never hears a single improvised line, but rather a texture, a musical fabric woven by several musicians in real time. (Hodson, 2007, p. 1)

In the end, the difficulty at arriving at a definition of jazz improvisation becomes predominantly one of scope. For jazz improvisation, the “terminology is lacking for a comprehensive description of the relationship between improvisation and recreative processes of music-making” (Smith, 1983, p. 44). There is, “no word to express the performance of music transmitted person-to-person and retained through memory” (Smith 1983, p. 44).

One attempt to reconcile these definitional problems is by locating jazz improvisation as a multi-layered cognitive process. Citing a 1974 study by Pike, de Bruin denotes jazz improvisation as:

> Idea generation from the projection of 'tonal imagery' as the fundamental process in improvisation, whereby improvisers express themselves from a perceptual field of creative consciousness (de Bruin, 2015, p. 91).

In Pike’s approach, sonic phenomena is understood as “memory based tonal images”, from which the brain has the capacity to create an “inner continuum [integrated with] external musical events, to create a perceptual insight or intuitive cognition from which ideas are generated”. Jazz improvisation in this sense is a kind of sonic coupling of the self and other. From the individual's standpoint at least, the improvisatory process is “perceptual and consists of a layer of tonal impressions, a consciousness-flux of percepts and feelings" (cited in de Bruin, 2015, p. 91).

It is possible to trace these cognitive ideas back further. Charles Keil's article, Motion and Feeling through Music, which first appeared in 1966, was concerned with the problem of finding a viable way to speak about performance, and attempted to locate the performer within a nexus of musical processes which could be reliably codified. Keil drew upon Leonard Meyer's influential text, Emotion and Meaning in Music, and sought a definition of jazz improvisation which was underpinned by psychological principles from which would emanate meaning and expression. At the same time, he sought to extend this idea further. For Keil, Meyer's "syntactically-focused notion of embodied meaning" was too imprecise and though the results it yielded might have value for "through-composed, harmonically oriented styles of our own Western tradition", they did not generalise well to other non-Western styles (Keil, 1966, p. 340). Instead, Keil proposed an alternative set of musical characteristics that contributed to what he called an "engendered feeling" (Keil, 1966, p. 341) which sought to understand jazz improvisation holistically in which content, form and expression were all taken into account.

A more recent example that attempts to examine the cognitive processes that can work together to reconcile a definition of jazz improvisation is David Sudnow's longitudinal self reflective study (2001), which examined the personal learning process of becoming a jazz pianist. As he acquired jazz improvisation skills, he documented these and the thought processes underpinning them. The documented observations and reflections allowed an understanding of how cognitive abilities could be developed to a level of being able to generate music improvisations. He located critical phases of improvisational development, such as "beginnings", centred around the acquiring of an appropriate vocabulary of sounds which could be heard in jazz and developing the accompanying motor skills; "going for sounds", which sought to document the struggle towards "reasonably acceptable places" in jazz improvisation proficiency (Sudnow, 2001, p. 3). Sudnow's work presents jazz improvisation as a process of becoming: an evolving self directed learning which differs from the process of playing music in real time, that has the capacity to create products in the form of music recordings.

These difficulties of finding a working definition of jazz improvisation are only exacerbated when exploring the differing viewpoints of its practitioners, analysts, and audience. And despite the increase in jazz analysis that has taken place within academic circles, the bulk of it is practically orientated, and found in the commercial sphere. It exists in the form of instructional texts, videos, play along recordings and interactive online software. The analysis of jazz improvisation has to an extent become a multi-modal endeavour accessible by those seeking to learn how to do it.

Many of the instructional orientated approaches to jazz improvisation position an external teacher or author as critical to skills acquisition. While the approaches share similarities to the work of Keil, Pike and Sudnow above, the process of skills acquisition is here mediated through an implied student-teacher relationship. An early example of this type of approach is Pressing's Improvisation:Methods and Models (1988). This work, which found parallels in the developmental approaches suggested by Kratus (de Bruin 2015, pp. 91-93) located an analytic framework in the context of a shared learning experience. It sought to show how the psychology of learning might be integrated into the acquisition of the ability to improvise, and drew parallels to methods used in the teaching of music in Baroque and classical times. The work aimed to utilise a "spectrum of pedagogies that merged facets of physiology, neuropsychology, motor programming and skill development, with a discourse on intuition and creativity" (de Bruin, 2015, p. 91). Pressing's work presented five stages aimed at transforming an aspiring novice into a fully fledged jazz improvisor, that seem reminiscent of the Fuxian approach to species counterpoint. It is an approach that views jazz improvisation as collaborative process, variously locating the collaboration between musician and ensemble, and teacher and student. Of the role of the teacher in the process, Hickey claims that "teacher directed learning and freer forms of improvisation that represent a student oriented enculturation can be depicted within a continuum of learning opportunity" (Hickey, 2009, p. 292). Jazz improvisation here again becomes a kind of process of becoming, in which the learner achieves expertise through relationships of trust operates between musicians and experts.

This field of jazz expert practitioners and those that aspire to expertise is a wide one and often plays out in commercial applications, in the form of instruction texts and related resources. Examples of this include Jerry Coker's Improvising Jazz, a work which sets out to explain the "real" theoretical principles of jazz, (listing them as intuition, intellect, emotion, and a sense of pitch), which can be honed into habits following correct practice methods. For Coker, the overarching aim to is to develop the "student's ability to translate the music he hears in his head into sounds on his instrument" (Coker, 1964, p. 3). Jazz theorist and educator David Baker also places an emphasis on aural skills, outlining a similar model to help the student "translate the sounds he hears on recordings directly to his instrument, dispensing as soon as possible with the step of writing them down" (Baker, 2005, p. 63). There is an industry of these types of texts and they are beyond the scope of this dissertation.

Overall then, the challenges posed by jazz improvisation for both the more traditional approaches to music analysis seen in chapter one, and to the approaches seen in MIR, are profound. The current transmission of any understanding of jazz improvisation is predominantly mediated through an author/practitioner as expert paradigm, similar to what was seen in chapter one. But at the same time, enhancing this paradigm by utilising data is extremely difficult. In jazz, music score data capturing what is happening in jazz improvisation is scarce and minimal at best. Audio data from jazz improvisation, though it may be ubiquitous, is simply not well suited for the exploration of questions of music analysis explored in this chapter (again problematising the issue of who the user is in MIR, within this context at least). The intent of the analysis chapter of this dissertation is to show that, even when having such scarce music score metadata available, by using an information retrieval approach there are profound insights into jazz improvisation to be discovered.

The analysis chapter of this dissertation will examine ten Keith Jarrett improvised solos and, as such, the following section will provide a brief biography of Jarrett, and canvas his views on improvisation. Notably, though Jarrett is outspoken about the nature of jazz improvisation and music more generally, his views do not serve to clarify the issues raised above. If anything, the opposite is true: Jarrett is openly critical of tendencies to intellectualise music or even the use language of as a viable way of describing it.

Keith Jarrett was born on May 8, 1945 in Allentown Pennsylvania (Carr, 1992)[3]. At an early age, his musical abilities were were noticed by his parents (particularly his mother), and by the age of three Jarrett had started taking classical piano lessons. By the age of seven, Jarrett had begun giving recitals, some of which included his original compositions. He became interested in jazz as a teenager, and has cited some early pivotal experiences of listening to Dave Brubeck and Bill Evans. He also expressed an interest in composing and at eighteen was given an offer to study composition with Nadia Boulanger in Paris

[3] The only biographical account published about Keith Jarrett is Keith Jarrett, The Man and His Music by Ian Carr (1992). The biographical details have been drawn from this text.

which he chose not to take up, and instead attended Berkeley College of music in Boston.

Jarrett attended Berkeley for a year, and largely disagreed with both the teaching approach and curriculum which he found overly rigid. In 1964 he moved to New York, and had his first professional breakthrough when drummer Art Blakey heard him play at a Village Vanguard, and offered him a spot in the Art Blakey Jazz Messengers. This engagement lasted only four months, during which time Jarrett played on the record Buttercorn Lady.

Jack DeJohnette, who would later become Jarrett's long time collaborator in his jazz trio recommended Jarrett to saxophonist Charles Lloyd's quartet, a position which Jarrett held until 1970. The group played modal music tunes, avant-garde jazz and had some cross over into rock influences, which for Jarrett was a dramatic departure from the more mainstream jazz sound of Art Blakey's group.

After leaving the Charles Lloyd quartet, Jarrett played and recorded with Miles Davis during the height of Davis' fusion period. Around this time, Jarrett also started performing improvised solo concerts for which he has become well known. During the mid to late seventies, he also became band leader of two groups, the European Quartet, and the American Quartet, recording a number of recordings with both groups.

In 1983, Jarrett started playing in a jazz piano trio format, often referred to as the "Standards" trio with drummer Jack DeJohnette and bassist Gary Peacock. The group predominantly plays songs from the "standard" jazz repertoire, being the popular American songs from movies and musicals of the twenties, thirties and forties, as well as some of the compositions of bebop players from the late forties and fifties. The group has also released three free jazz recordings. Jarrett announced the trio had finished performing together in 2017, and after a long hiatus from performing solo piano concerts has returned to this format. Together the trio released 22 recordings.

There are only handful of existing analysis' of Jarrett's work. Examples include Strange's Keith Jarrett's Up-tempo Jazz Trio Playing: Transcription and Analysis of Performances of "Just in Time", a doctoral thesis by Dariusz Terefenko, Keith Jarrett's Transformation of Standard Tunes, in 2004, and, in 2009, Page's Master's thesis Motivic Strategies in Improvisations by Keith Jarrett and Brad Mehldau.

Terefenko's work is heavily influenced by Schenker, and locates the notion of a phrase model at the centre of his analysis. The phrase model is a fundamental structure that can capture the "the tonal motion of a phrase… in terms of its underlying melodic, contrapuntal, and harmonic structure" (Terefenko 2004, p. 28). This analysis aims to demonstrate two essential features of Jarrett's approach to jazz improvisation. The first is Jarrett's ability to make large-scale harmonic and melodic connections with the original version of the standard, and the second is his sophisticated sense of formal organisation which allows Jarrett to apply a notion of form in the solo piano improvisations (Terefenko, 2004, p. 312). Terefenko provides both a highly detailed theoretical Schenkerian framework and a dense descriptive context to explore Jarrett's playing. A typical example (here related to Jarrett's performance on the jazz standard It Never Entered My Mind) can be seen below:

> In mm. 1-24, Jarrett mostly relies on the original melody. In the last A section, Jarrett takes liberties while rendering the melody. Not only does he vary the melodic content rhythmically (as he did in mm. 1-24), but he also transforms its basic framework.

> The original repeated notes in m. 25 are embellished by upper neighbours (Terefenko, 2004, p. 229).

For me, Terefenko's approach is problematic. It presents a rigorous theoretical work, but moves uncomfortably between the statistical and descriptive, in order to show that, above all, that Jarrett's music is coherent and highly structured. Though it locates Jarrett's work in a strong theoretical framework, the work also highlights the problems of using the language of traditional music theory in capturing rapidly changing harmonic phenomena on a score. An example of this density can be seen in the following commentary on Stella By Starlight:

> The structure of the dominant 7th features an impressive array of formations derived exclusively from the DNC: the Mixolydian (mm. 10, 14, 24, and 30); the Mixolydian b13 (m. 17 and m. 26); the Altered b9 (mm. 2, 6, 13, 16, 18, and 28); and the Altered #9 (m. 27)…the Lydian (m. 4 and m. 19), the melodic minor (mm. 8, 11, and 29), and the Locrian #2 (mm. 10, 15, and 25). Jarrett's noteworthy alterations of the quality of the minor 7(b5) occur in mm. 25-32. Here, Jarrett transforms its quality into Em7, D7alt, and Ebm(ma7), (m. 25, 27, and 29, respectively). The last harmonic change, Ebm(ma7), adheres to the original version. (Terefenko, 2004, p. 259)

This is certainly not incorrect on its own terms, but highlights one of the critical challenges that I am seeking to address: the use of language, labels, and categorisation that informs music analysis is not well suited to large amounts of music score data with rapid movement through different tonalities.

A later work by Page (2009), juxtaposes Jarrett's style with that of Brad Mehldau. Its focus is on comparative motivic analysis, taking its cue from "European art music…[which was] especially prevalent in various early to late-mid 20th-century analytical circles, to examine how motive informs form". Page sets out to demonstrate the "unity" of works to be analysed, and explores the "organic growth" of motives found in melodies (Page, 2009, p. 2).

Page also draws heavily on Schenker, when discussing the myriad ways in which a melodic motive might repeat itself at different structural levels of a composition. He utilises a notion of "motivic parallelism" (Page, 2009, p. 14), an umbrella term for a variety of phenomena discussed by Schenker, and later explored by Burkhart. Using this approach, a given pitch is deemed more or less "structural" based on its harmonic and contrapuntal importance relative to an underlying harmony or harmonic progression (Page, 2009, p. 19). Page develops the idea of a "motivic chain association" that can capture "any kind of audible motivic relatedness between elements of a melodic line" (Page, 2009, p. 14).

One of the difficulties facing Page can be seen when he attempts to apply a Schenkerian perspective to highly intricate melodic lines which often use all pitch classes of the octave. This makes it difficult to ascertain which pitches in a given melodic passage might be considered as structural. Page notes that the harmonic degrees in chordal structures that are regarded as stable in jazz harmony, such as sevenths, ninths, elevenths, and thirteenths, are often not resolved to related adjacent consonances, such as thirds, fifths, sixths, and octaves (a number of Schenkerian analysts of bebop acknowledged this problem also, such as Strunk 1996, Larson 1998, and Martin 1996).

Page mitigates the issues by changing the focus to a comparative study, showing that in Jarrett's jazz improvisations, there is more likelihood of "dovetailing from the end of immediately preceding phrases than references to earlier phrase beginnings" (Page, 2009, p. 9), which is in contrast to Mehldau's approach. There is a "constant forward developmental motion on display in Jarrett's solo in comparison to the Mehldau's" (Page, 2009, p. 38). Page explains the comparison by claiming:

> When interpreted with an eye to process, motivic chain association analyses of the two solos studied lead to clear evidence of Jarrett's relative propensity, compared to Mehldau, for tightly woven motivic work characterised by forward-moving transformation of small motivic fragments. (Page, 2009, p. 48)

Other articles that explore Jarrett's work are not focused on explicit score analysis or extracting metadata from his music. However they explore other aspects of his approach, tending to locate his music within wider sub-genres related to jazz. These include Moreno's, Body 'n Soul: Voice and Movement in Keith Jarrett's Pianism in 1999, Blume's Blurred Affinities: Tracing the Influence of North Indian Classical Music in Keith in 2003, Elsdon's 2008 article, Style and the Improvised in Keith Jarrett's Solo Concerts in 2008.

Moreno's study examines the role of the body and gesture and examines Jarrett's movements and singing when in a solo piano setting. Moreno claims:

> I believe that by this procedure he reveals the presence of a conscious thought process. He makes explicit the fact that imagining sound and structuring it around the chord progressions and melodies of the songs he improvises on entails embodying it in mind, soul, and body (here, body signifies the voice). The sound of his voice unleashes what in the critics' minds should be a

> metaphysical presence, which is to say, an invisible or repressed Other (Moreno, 1999, p. 79)

For Moreno, the role of the body and the way it moves are critical to understanding Jarrett's improvisations, and he claims that:

> Jarrett's body appears to take flight and his voice seems to sing, it is because he believes in the priority of the improviser as a person whose imagination rolls and tumbles...whose body is not only instrument, expression, and locus of self, but self itself (Moreno, 1999, p. 89).

While it may be counter productive to link Moreno's article to more specific questions of analysis that utilise metadata, it highlights the difficulty faced by jazz: even extracting large amounts of metadata from transcriptions and audio files, there are other important considerations to Jarrett's playing.

Blume's article explores notions of place and genre in Jarrett's playing, again focusing on Jarrett's solo performances. He relates to the solo concerts "long form improvisations" that gradually build elaborate rhythmic structures and motivic structures (Blume, 2003, p. 118). Blume finds parallels between Jarrett's music and North Indian classical music, noting in particular the rubato section of the Koln Concert, 'Part I', which features "tambura-like drones and frequent mohra- like cadential figures (Blume, 2003, p. 132).

In interviews, Jarrett himself has also discussed the problem of geographical place in music (often when reflecting on the differences between European and American music forms) and I will take this up later this in the chapter. Blume claims that Jarrett's ability to work across different genres, "adds to a shimmering ambiguity that makes Jarrett's products attractive to audiences not readily identified with jazz" (Blume, 2003, p. 119).

An article by Elsdon's briefly touches on questions of analysis, but more generally locates Jarrett's work in the framework of different sub-genres of music through which Jarrett can effectively traverse. Elsdon alludes to some questions that are amenable to analysis, highlighting Jarrett's use of "ballad passages" which can act to avoid establishing a definitive tonal centre, that are "always breaking off to move in a new direction as soon as any cadential inference might be drawn" (Elsdon, 2008, 58). He also explores "long vamp-driven sequences" that often appear in Jarrett's playing, noting that, in contrast to passages that more through different tonalities rapidly, they are typified by the removal of conventional harmonic or rhythmic progressions typically found in jazz standards, and often Jarrett juxtaposes these different approaches to great effect (Elsdon, 2008, p. 61)

For Elsdon, even locating Jarrett in the genre of jazz is problematic, and he positions Jarrett as signalling a departure from more traditional modalities of jazz which focuses on the intersection of geographies and socio-demographic space: Jarrett accesses a genre that "no longer presents a single, unified vision of a bucolic America" (Elsdon, 2008, p. 62). Eldson claims that:

> Quite the contrary, in fact, they express and explore a broad range of styles and attitudes. What unifies this body of music—and this is the point I want to emphasise in this paper—is the shared idealisation of non-urban spaces and lifestyles (Elsdon, 2008, p. 62)

Finally, a more recent analytical work has appeared on Jarrett in Blake's Improvising Optimal Experience: Flow Theory in the Keith Jarrett Trio, in 2016. This work locates Jarrett's playing in the the trio in the context of Mihály Csíkszentmihályi's Flow theory which can be be characterised as follows:

> The concept of flow describes a set of conditions that allow a person to engage in optimal experience in the course of an activity. These conditions require that the activity be goal-oriented and rule-bound, that the challenge presented by the activity is balanced with the participant's ability and… the presence of intentionality on the part of the person performing the activity. (Blake, 2016, p. 8)

Again, this work is a departure from both music theory and analysis approaches, or focusing on extracting metadata. But it reinforces the complexity of the information that is generated by jazz improvisation and the problematic nature of capturing this in the vehicle of a music score in order to interrogate it.

Jarrett himself has strong and often expressed opinions on jazz improvisation, through he almost never speaks of music theory or even specific things that he practices. Further, he takes the view that language itself is not equipped with the means to articulate the meaning of jazz improvisation (see https://www.youtube.com/watch?v=fDbOKHOuy9M/2018). Generally, Jarrett positions jazz improvisation as a holistic process that constantly challenges his creativity, noting:

> For me, if I don't play something that doesn't challenge my concept of what I liked before that second, something's wrong. So what you do is you create a "cell," let's call it. And that cell is

> your voice. And then you want that cell to replicate in whatever direction it wants to per microsecond. And that's when you expand it, and it becomes not a personality anymore, it becomes a biofeedback mechanism.
> https://ethaniverson.com/interviews/interview-with-keith-jarrett/ (2008)

This type of sentiment is typical of the way many jazz musicians tend to speak about the process of jazz improvisation. Miles Davis has claimed, "when you start playing just try and finish what someone else has left" (https://www.theguardian.com/music/2012/nov/06/miles-davis-interview-rocks-backpages/ (2018)). In the closing moments of Eric Dolphy's Last Date, he can be heard saying, "music is in the air and when you play it is gone" (Dolphy, 1964). Finally, pianist Bill Evans notes, "the art of improvisation, and the art of music, for that matter, lies in mastering the ability to take an idea and treat it as such—to respond to it musically, according to the context in such a way as to say what you want to say, which for me is to try to get to a slightly deeper feeling" (https://www.allaboutjazz.com/breakfast-with-bill-evans-bill-evans-by-bob-kenselaar.php?page=1/ 2018)). For jazz improvisers at least then, it seems the definition of jazz improvisation is deeply personal.

Jarrett also views jazz and jazz improvisation as being fundamentally different from other music genres. This is because jazz improvisation takes place in real time. It is a response to the conditions of a precise moment and the stimuli of this moment. This idea is often presented in a conception of self with a disposition towards the world. On the substantive difference between classical music and jazz he claims:

> If a player gets used to not disappearing into the music completely and starts thinking about the kind of details you have to think about in classical performance, that's not what you should be doing when you play the blues (cited in Rosenthal, 1996, para 12)

Jarrett has also spoken at length regarding non-improvised music. A difficulty that emerges when analysing this is that he talks interchangeably about classical composition, the process of performing classical compositions, and roles of musicians in classical performance. All of this is juxtaposed against the fundamentally different real-time creation of music that informs jazz. Speaking of the difference he notes:

> Because I think [jazz] may be the only art form at this point in time that asks the player…not the conductor, not any detached entities from the actual playing…that asks the player to find out who he is and then decide if it's good enough to speak from that self. (cited in Panken, 2018 para 23)

Jarrett does not intend this as a criticism of classical music as opposed to jazz improvisation, and has performed wide ranging classical repertoire and has released recordings of Bach, Mozart and Shostakovich. However the implication is that the two approaches are simply qualitatively different. He claims that one must "become a musicologist when you become a classical player [which can] undermine one's ability to improvise effectively in jazz" (Rosenthal, 1996, para. 3)

Jarrett views improvisation very much as a process undertaken in real-time, and a response to the surrounding world. He believes that creativity is not about the self creating from nothing, but the self becoming nothing and allowing music to flow through, which he sees as a profoundly spiritual phenomenon. In a 1984 Downbeat interview with Art Lange he claims:

> Really, I've been feeling in the last few years, even while improvising, I am playing other people's music, or other music. It isn't mine. (Lange, 1984, para. 5)

In a later 1996 interview, he revisits the theme, specifically in the context of the of the jazz trio, in a Ted Rosenthal interview:

> Gary (Peacock) said to me once, "every time the trio plays, it's like we are taking in more history each time we play". It isn't like people will say I'm using so and so's licks, but if you let something enter, then there's a bunch more possibilities. So a line would end up being longer. But if you tighten up a little, it will shorten up. If you let more air in, then the pulse gets freer. Then you play five notes in a two beat area and have it sound fine, you know? (Rosenthal, 1996, para. 17)

Adopting Jarrett's view suggests that the only way to understand meaning of his jazz improvisations is to locate them in the context of a much larger corpus of works and account for changes in improvisation practice over time, which poses a profound challenges for any kind of analysis.

Interestingly, this is similar to the view of metadata adopted in MIR, in which there is not a fixed theoretical foundation, just changing meaning based on the body of music and the listener. Related to this idea is that, for Jarrett, jazz improvisation has nothing to do with originality or creativity, but is about playing into a history. He claims it is not about “using so and so's licks, but if you let something enter, then there's a bunch more possibilities. So a line would end up being longer. But if you tighten up a little, it will shorten up” (Rosenthal, 1996, para 17).

In many interviews, Jarrett’s views about the self and process of jazz improvisation becomes difficult to track. The process of jazz improvisation is presented as a higher state of being, that is not currently amenable to analysis at all. In a 2015 NPR interview Jarrett claimed:

> I'm trying to think of the right way to put this: It's potential limitlessness that I'm feeling at that moment. If you think about it, it's often in a space between phrases, [when I'm thinking,] How did I get to this point where I feel so full? (Martin, 2015, para 12)

When viewing these types of comments (and there are many of them) it seems that for Jarrett, jazz improvisation is an extremely complicated ontology located somewhere between the bounds being and nothingness. As such, this leaves a practical problem for this dissertation in terms of how to understand his music.

There are however, some practical and pragmatic concerns that Jarrett does allude to in his interviews. Specifically, he talks about the effect of geographical space, some practical issues of playing an instrument such as a piano, and some (albeit limited) discussion and what he practices at the piano.

In terms of the relationship between geographical space and music, Jarrett believes that jazz improvisation is fundamentally different in different locations. This relates to improvisation practice drawing on different music traditions, and he cites a fundamental difference between European and American improvisation practices, saying:

> It is hard to be a European jazz player, I think it is hard to be an American composer. It's not hard to be an American jazz player. But we didn't invent composing, and it's a tough country to draw a large-scale anything of, because everyone is so [much] themselves. In jazz, you are not expecting anybody to do anything they can't do, and you aren't expected to be able to analyze a symphony (Rosenthal, 1996, para. 36)

From the point of view of analysis, this is something that can be explored in an evidence based way, and I will discuss this further in the following chapters.

Jarrett also often speaks about the practical difficulties and limitations of the piano, and how this affects the process of improvisation. He notes that the piano is a "relatively boring" instrument (Rosenthal, 1996, para 21), a "really structured thing, basically a percussion instrument". He notes that, "even when a piano is in perfect operating condition it does not have much personality" (Panken, 2008 para 1). Jarrett goes as far to say that ideally the piano should not be a part of the improvisatory process:

> There's a fluidity in an instrument that uses air. I've always wanted to get as close as possible to subtracting the mechanism of the piano from the whole affair. (Panken, 2008, para 15)

Jarrett also claims that saxophone players have influenced him far more than piano players, and notes a key difference in his piano trio setting as characterised by the move away from “thick textures in the rhythm section”, and approach he describes as more “Brubeckian” in nature.

Despite the limitations of the instrument however, Jarrett views it as a far better option than the electric alternatives. While on tour with Miles Davis during 1970-71, Jarrett played electric instruments the first and only time in his career. Reflecting on the experience, Jarrett has claimed:

> Keyboard players got enamoured of electric instruments, and never could go back. These are artistic decisions, and you can’t make them lightly. It’s like a painter throwing away their paint, saying, ‘Well, I want to get these,’ but they’re all monotone, and then, ‘Well, no, I want my old paints back.’ Sorry. They went out in the garbage. (Rosenthal, 1996, para. 39)

Using Jarrett’s collected interviews to discuss music theory is impossible. He cites the importance of Bebop which is “somehow centre stage to what modern jazz has done even since” (Iverson, 2009, para 5) in terms of harmony and melody, and claims:

> Voice-leading is melody-writing in the centre of the harmony. If you can do it, you’re lucky enough to get to a moment where you can actually find more than one thing happening and trace those things at the same time to a logical next place… or illogical place (Iverson, 2009, para 6)

In an interview with Panken, Jarrett does make a passing reference to tonality. He claims that there is no such thing as atonality, and music can only be regarded as “multi-tonal” (Iverson, 2009) and believes his music pushes the boundaries of moving between tonalities, a playing style that that took him a long time to develop (Panken, 2008, para 11).

Finding a way to understand and contextualise jazz improvisation then, is highly problematic. There is often not enough information to work with, and no clear consensus about how to contextualise the meaning of the findings. In the analysis chapter of this dissertation, I will present a metadata driven approach to address this problem. I will show that, even without knowing what jazz improvisation might or might not mean it is possible to access the minimal traces of information it leaves behind to gain deep insight.

Chapter 4
Tools and Technologies used for the Case Study

In the previous chapters I have explored the challenges that arise when analysing metadata taken from the music score, and different approaches to analysis that have been adopted in different periods and disciplines. In the analysis chapter chapter, I will address these challenges by using extracted metadata from music scores to facilitate music analysis.This chapter will provide a summary of the tools and technologies that have been used both in the analysis chapter, and the associated software application of the dissertation.

The software packages, computer programming technologies, and data specifications used in this dissertation are listed in Table 4.1 below.

Table 4.1. Technologies used in this dissertation

Technology	Description
Transcribe	Transcribe is a desktop software program used for the manual transcription of recorded music
MuseScore	MuseScore is a desktop software program used for high quality music engraving and printing
JavaScript	JavaScript is a programming language, used heavily in web applications.
Node	Node is a programming framework that utilises JavaScript, suited for building networked orientated software applications and web application.
JSON	Javascript Object Notation (JSON) is a key-value orientated data specification commonly used in web applications.
D3.js	D3 is a data visualisation library, written in JavaScript for custom data data visualisation and in-browser SVG manipulation.
Python	Python is multi-purpose programming language derived from the C programming language, used heavily in data science.
Jupyter Notebook	Jupyter Notebook is a scientific computing environment used heavily in exploratory data science.

Pandas	A Python based library for data cleaning, transformation, and analysis.
Muisc21 and LilyPond	A Python based library used for music analysis applications and to render music scores from code
React, Django and PostgreSQL	A technology stack used to build web applications

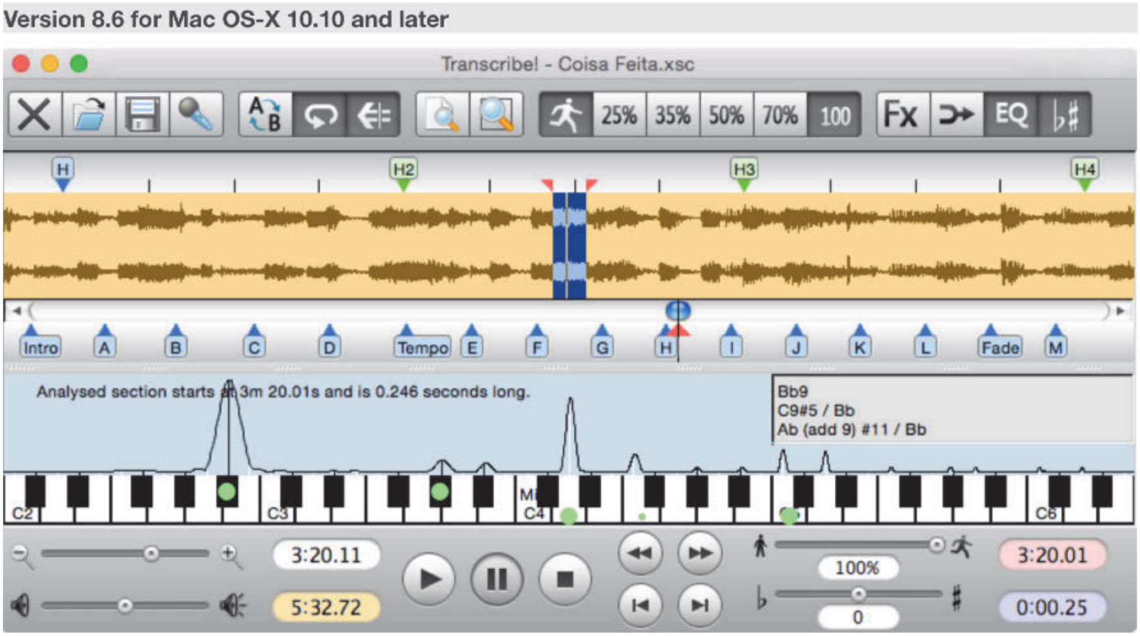
Transcribe

The practice of transcription is often regarded as a critical part of the learning process in jazz improvisation. It involves listening to a given recording, and ascertaining which notes are being played, along with where they have been rhythmically placed. To facilitate the process, I have used a desktop software program called Transcribe. Transcribe software is not intended to automate the transcription process, rather it provides functionality to play sound files at different speeds without altering the pitch, and also allows users to set markers at different positions in the audio, to facilitate repeated listening to the same passage multiple times. On its website, developers of Transcribe note its purpose as follows:

> The Transcribe application is an assistant for people who want to work out a piece of music from a recording, in order to write it out, or play it themselves, or both. It doesn't do the transcribing for you, but it is essentially a specialised player program which is optimised for the purpose of transcription. It has many transcription-specific features not found on conventional music players. (Seventh String Software, 2017)

Figure 4.1 below displays a screenshot of the Transcribe user interface, showing some of its features. A visualisation of an audio wave-form can be seen in the toolbar, which can be used to set markers and loops in order to allow repeated playing of small sections of audio. The toolbar itself allows the changing of speed of the recording, with or without affecting pitch. Different sections of the music (denoted by the blue markers) can also be set, to easily allow movement between different section of the audio file.

Figure 4.1. Transcribe software screenshot



MuseScore

MuseScore is a desktop software program whose purpose is to provide "high quality print renditions of the music scores" (https://musescore.org/en/(2018)). It is freely available open source software, and provides similar functionality to commercial music engraving software programs such as Sibelius and Finale.

One of the features of MuseScore (which is similar to the other commercially available software programs) is the range of different output formats for the music score data it holds. These include PDF format (which is ideal for printing, and transferring files between different operating systems) as well PNG format (which renders individual pages of a music score as high quality images, and the music scores found in Appendix 1 are all exported PNG files from MuseScore). There are also alternative output formats more suited to machine related data, such as MIDI and MusicXML.

Javascript

JavaScript is a programming language. It is used predominantly as a way of managing events and interactivity in web applications. Increasingly, JavaScript is being used in a number of non-browser/web environments, (such as Node.js and Apache CouchDB) in order to manage programmatic event handling in networked applications (Mozilla Developer Network, 2017, para 1). Despite its name, JavaScript is not a scripting language for the carrying out of small programmatic tasks. It is a fully-fledged programming language which is a "prototype-based, multi-paradigm, dynamic language, supporting object-oriented, imperative, and declarative (e.g. functional programming) styles" (Mozilla Developer Network, 2017, para 2). Like many other programming languages, JavaScript is extendable, and there are many additional libraries available that can be utilised in order to extend the language's core functionality.

Node

Node is a platform used to create network orientated applications. Its original release was in 2009, and it has since become a popular framework upon which to build complex web applications that require event handling such as data transfer, authentication, user payments and chat functionality (https://nodejs.org/en/ (2018)). Examples of Node being used in web applications include software developed by PayPal, Netflix, Uber, LinkedIn and Walmart Node utilises an "event-driven, non-blocking I/O model", which aims to be lightweight and well suited to highly complicated web applications (https://nodejs.org/en/(2018)).

For this dissertation, I have used Node as a framework on which to build the software module that converts MusicXML data into JSON data and allows JSON data to be easily integrated with other music metadata. This software could have been built in any number of languages, however my choice of Node was influenced by the requirement to easily be able to integrate this software into a companion web application (whose front end is built in React.js) that allows users to upload their own MusicXML.

D3

D3 is a JavaScript library whose purpose is "manipulating documents based on data" (https://d3js.org/(2018)). The D3 library provides a range of functions and methods that work with existing browser technologies (such as HTML, SVG and CSS) which together can be used to create highly interactive data visualisations for users. I have used D3 in this dissertation to provide data visualisations for the software that converts music score data, and it has also been heavily used to build the music data visualisations that will be discussed in Chapter 6.

Python

Python is a popular programming language based on the C programming language. It is particularly well suited to scientific computing, data analysis and data-modelling. Like most programming languages, Python has a basic instructions set, allowing users to accomplish a wide variety of computation tasks. However its functionality can also be extended by using additional Python software libraries. It has been used to carry out all the analysis tasks in the upcoming case study chapter.

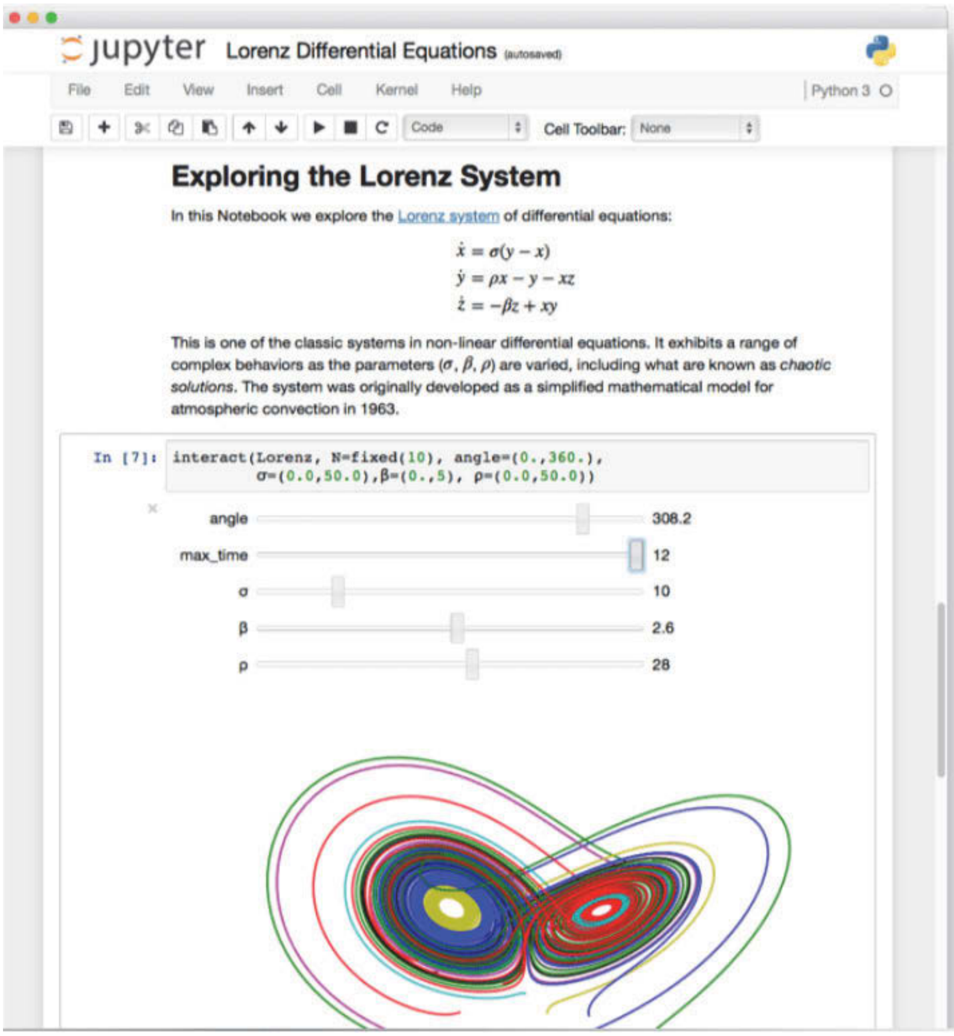
Jupyter Notebook

Jupyter Notebook is an interactive environment in which Python code can be executed (and it also supports a number of other languages commonly used for scientific computing) and is used heavily for statistics and data related tasks. According to the Jupyter Notebook website, (http://jupyter.org/(2018)):

> The Jupyter Notebook is an open-source web application that allows you to create and share documents that contain live code, equations, visualisations and explanatory text. Uses include: data cleaning and transformation, numerical simulation, statistical modelling, machine learning and much more.
> (http://jupyter.org/(2018))

A screenshot of a Jupyter Notebook is listed in Figure 4.2 below from jupyter.org. It highlights the technology’s ability to allow developers to quickly create markdown text, mathematical notation, interactivity and visualisations.

Figure 4.2. Jupyter notebook screenshot



Pandas

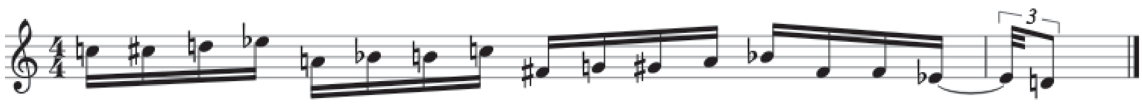
Pandas is a software library that can be used in conjunction with the Python programming language and can be used within a Jupyter notebook. Its purpose is to extend the Python language to include a comprehensive set of data preparation and statistical analysis tools. It is used heavily in various scientific analysis and financial analysis applications. Many of the Pandas library features are designed to mimic those found in the 'R' programming language, which is also used widely by statisticians.

The Pandas library allows information to be held in ‘data-frames’. A data-frame can best be conceptualised as a list of rows, where each row contains information about one object in the data set. The data-frame can then be heavily manipulated to accomplish a wide variety of statistical tasks.

Music21 and LilyPond

Music21 is a Python library that can be used to accomplish a wide variety of music related tasks (and includes its own converter from MusicXML to a Python data structure). However its use in this dissertation is limited to the rendering of music scores within the Jupyter Notebook. To accomplish this, Music21 can be used used in conjunction with an open source score visualisation library, LilyPond. Together these two software modules allow for rendering of music score excerpts to be produced programmatically based on code. An example of a music score excerpt rendered from Python code can be seen in Figure 4.3.

Figure 4.3. Example of Music21 and Lilipond rendered score



Django, PostgreSQL, and React

There are many different technologies currently available for building large scale web applications, and for the purposes of exploring further work to come out of this dissertation I have used Django, PostgreSQL, and React. Django is a “high-level Python Web framework that encourages rapid development and clean, pragmatic design” (https://www.djangoproject.com/(2017)), and handles tasks such as setting up different pages of websites, user authentication and database interaction. PostgreSQL is an Structured Query Language (SQL) database, which is well suited to for storing and query large amounts music metadata in a web application environment. React is a javascript library that is a front end web framework (specifically for designing the user experience) created by Facebook for the purpose of building rich interactive user experiences that are computationally efficient.

Music Metadata Builder: Software to extract metadata from a music score

To create the software needed to extract the music data from scores, the Node framework was used. The software works by iterating through all parts of a music score and extracting all score related attributes (such as time, duration and pitch information, score notations, dynamic markings etc.) and then converts the information into a flattened list of notes, linking all attributes to an underlying note or rest structure. For the Keith Jarrett solos that will be explored in the case study, the following informational attributes were extracted from the score and the recording below. Figure 4.4 displays the first record, a rest from the score.

Figure 4.4. JSON output from Music Metadata Builder


```
{
  "Midi number": -1,
  "Harmony note flag": false,
  "Current measure": 1,
  "Duration": 192,
  "Duration due to tied notes": 192,
  "Location": 0,
  "Location in measure": 0,
  "Measure location": 0,
  "Time signature numerator": 3,
  "Time signature denominator": 4,
  "Quarter beats per minute": 60,
  "Time stamp": "1685-03-20T13:00:00.000Z",
  "Instrument": "P1-I1",
  "Voice": 1
},
```


The software also allows additional metadata to inputted by a user which can be combined with the information taken from the music score (this could include additional attributes such as title, recording location, track number listing). The additional metadata can either be manually provided by the user, or sourced through a standard data API. For example, it is possible to provide the software with a query from the iTunes database (which can return the kind of information found in Figure 4.5, here being an example of information about a Jack Johnson track) so it can be integrated with the score metadata extracted by the software.

Figure 4.5. JSON output from iTunes database


```
{"wrapperType":"track",
 "kind":"song",
 "artistId":909253,
 "collectionId":120954021,
 "trackId":120954025,
 "artistName":"Jack Johnson",
 "collectionName":"Sing-a-Longs and Lullabies for the Film Curious George",
 "trackName":"Upside Down",
 "collectionCensoredName":"Sing-a-Longs and Lullabies for the Film Curious George",
 "trackCensoredName":"Upside Down",
 "artistViewUrl":"https://itunes.apple.com/WebObjects/MZStore.woa/wa/viewArtist?id=909253",
 "collectionViewUrl":"https://itunes.apple.com/WebObjects/MZStore.woa/wa/viewAlbum?i=120954025&id
=120954021&s=143441",
 "trackViewUrl":"https://itunes.apple.com/WebObjects/MZStore.woa/wa/viewAlbum?i=120954025&id=1209
54021&s=143441",
 "previewUrl":"http://a1099.itunes.apple.com/r10/Music/f9/54/43/mzi.gqvqlvcq.aac.p.m4p",
 "artworkUrl60":"http://a1.itunes.apple.com/r10/Music/3b/6a/33/mzi.qzdqwsel.60x60-50.jpg",
 "artworkUrl100":"http://a1.itunes.apple.com/r10/Music/3b/6a/33/mzi.qzdqwsel.100x100-75.jpg",
 "collectionPrice":10.99,
 "trackPrice":0.99,
 "collectionExplicitness":"notExplicit",
 "trackExplicitness":"notExplicit",
 "discCount":1,
 "discNumber":1,
 "trackCount":14,
 "trackNumber":1,
 "trackTimeMillis":210743,
 "country":"USA",
 "currency":"USD",
 "primaryGenreName":"Rock"}
```

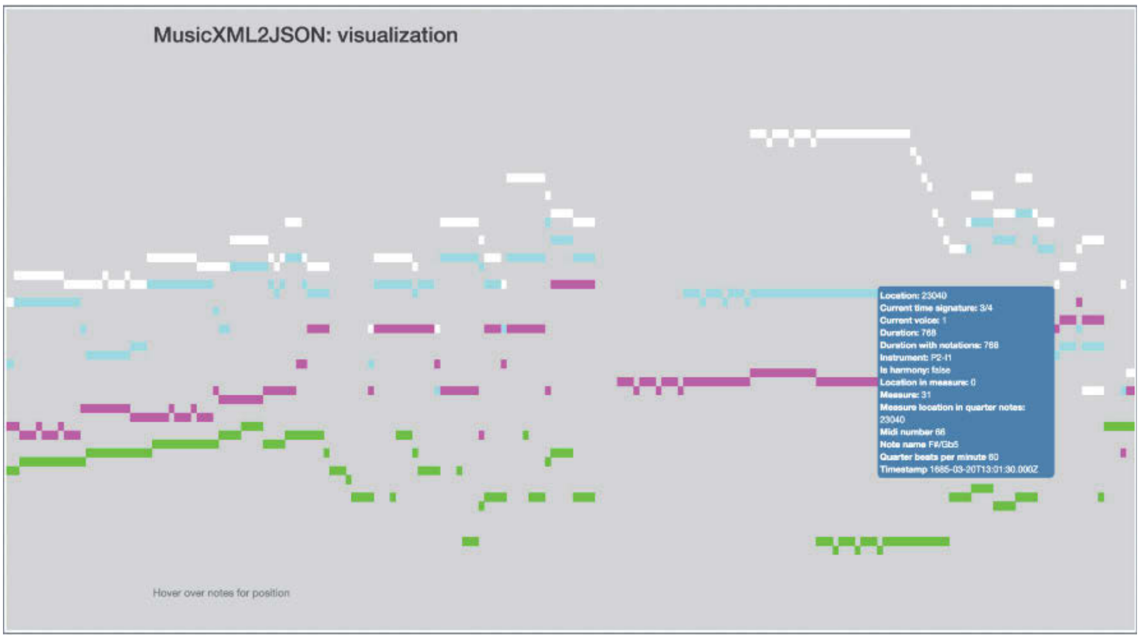

For the case study, additional data specific to the jazz standards under analysis was manually integrated with the basic metadata from Figure 4.5 (which included additional data about the jazz standards under consideration, the place of recording etc.), and an example of a resulting record can be seen in Figure 4.6 below.

Figure 4.6.   JSON output from Music Metadata Builder (annotated)


```
{
  "Midi number": -1,
  "Current chord root": "0",
  "Current chord root as int": 0,
  "Current chord type": 0,
  "Harmony note flag": false,
  "Current measure": 1,
  "Duration": 720,
  "Duration due to tied notes": 720,
  "Location": 0,
  "Location in measure": 0,
  "Measure location": 0,
  "Time signature numerator": 3,
  "Time signature denominator": 8,
  "Quarter beats per minute": 60,
  "Time stamp": "1685-03-20T13:00:00.000Z",
  "Instrument": "Part 0",
  "Voice": 1,
  "Genre": "Acoustic Jazz",
  "Composer name": "Victor Young",
  "Composer nationality": "US",
  "Lyricist name": "Ned Washington",
  "Lyricist nationality": "US",
  "Performer name": "Keith Jarrett",
  "Performer nationality": "US",
  "Title": "Stella By Starlight",
  "Composer collection": "The Uninvited",
  "Performer collection": "Standards Live",
  "Date composed": 1944,
  "Date recorded": 1983,
  "Recording location": {
    "lat": [
      48.8566,
      "N"
    ],
    "lon": [
      2.3522,
      "E"
    ]
  }
},
```


The software also has inbuilt data visualisation capability, built using D3, which can render the data into a piano roll style visualisation. Figure 4.7 shows an example this visualisation, and here I have used the software to extract information from a Beethoven String Quartet movement.

Figure 4.7. Music Metadata Builder Score Visualisation



This software has been designed to function as a stand alone application (and can be deployed as what is known as a Node module, or to be used in web application environments so users can upload music scores, have this information extracted into a format well suited for a wide range of analysis. Details of all software used in the discretion can be found in Appendix 2.

For the purposes of the analysis chapter, the steps listed in Table 4.2 were taken, and a summary of each of these is provided below the table.

Table 4.2. Steps for preparing data for the case study

Step	Description
1	Ten Keith Jarrett improvised solos were transcribed by hand, with the aid of Transcribe software.
2	The handwritten scores were inputed into into MuseScore software.
3	The scores were exported from MuseScore in a standard format of MusicXML.
4	The scores were converted to a flattened data structure using the MusicXML2JSON software and combined with additional metadata related to the jazz standards.
5	The data was imported into Jupyter Notebook, into a Pandas Data Frame for the purpose of exploration and analysis.

Step 1

For jazz musicians, the transcription process tends to be viewed as a convenience from which to capture basic information about a solo that can then be then used to learn how to play it. As such, there is often a degree of accepted approximation during the transcription process. Following these conventions, the decision was made to simplify all chords to having no more than an extension of a seventh, and use the chords typically found in a standard real book. Additionally, eighth-notes with a swing feel were transcribed as straight eighth notes. Transcribing rhythm in Jarrett's playing can be particularly challenging, as he will often play passages during which he will shift the part of the beat he is playing notes on. These were notated to the approximate closest standard rhythmic subdivision. Additionally (and very occasionally), Jarrett plays two notes at once in the course a melodic line (and this happens less than ten times across over 15000 notes of melody). In these cases, I have taken the melody note to be the one that I feel best represents Jarrett's melodic intention, based on my experience of transcribing many of his solos.

Steps 2 and 3

The ten handwritten scores were then inputted into the music engraving software, MuseScore. An excerpt of the opening bars of Jarrett's solo on Stella By

Starlight, can be seen in Figure 4.8. After the scores was entered into MuseScore, it was exported in the MusicXML format.

Figure 4.8. Excerpt from Stella By Starlight transcription

Step 4

The MusicXML files of the ten solos were then converted into a JSON dataset using the Music Metadata Builder software application, producing ten JSON files holding extensive information about each note and rest of the solo in a flattened list structure.

Step 5

The JSON data structure was then directly imported into a Jupyter Notebook using the Python Pandas library. The first record of the resulting data-frame is below.

Table 4.3. Sample record of prepared data

Composer collection	Very Warm For May
Composer name	Jerome Kern

Composer nationality	US
Current chord root	G
Current chord root as int	7
Current chord type	min7b5
Current location in seconds	0.242915
Current measure	1
Date composed	1939
Date recorded	1983
Duration	1.0
Duration as string	Quarter note
Duration due to tied notes	480
Duration of one second	1976
Genre	Acoustic Jazz
Instrument	Part 0
Location	0
Location in measure	0
Lyricist name	Oscar Hammerstein II
Lyricist nationality	US
Measure location	0
Midi number	-1
Midi number as string	Rest
Midi number as string without octave	Rest
Performer collection	Standards, Vol. 1
Performer name	Keith Jarrett
Performer nationality	US
Quarter beats per minute	247
Recording location	{'lat': [40.7831, 'N'], 'lon': [73.9712, 'W']}
Time signature denominator	4
Time signature numerator	4
Time stamp	1685-03-20T13:00:00.000Z
Title	All The Things You Are

Chapter 5

Jazz Improvisation Analysis Case Study: Ten jazz solos of Keith Jarrett

This chapter will demonstrate how music score metadata can be used to explore the improvisational style of Keith Jarrett. Ten of Jarrett's jazz trio solos (transcriptions of which can be found in Appendix 1) have been converted into a single dataset. The intent of this approach will be to demonstrate how using search and retrieval methods can afford a new understanding of Jarrett's improvisations than utilising more traditional approaches. It can also provide an evidence based understanding of the underlying structures that characterise Keith Jarrett's approach to improvisation. The case study will demonstrate how the music score can be re-imagined as a site characterised by an ease of data transformation and pattern exploration, and can inform multiple data visualisations, only one of which is the traditional music score.

In regard to the choice of Keith Jarrett, this has been made as there is virtually no repetition in Jarrett's playing. Almost every melodic phrase in these improvisations is unique (and every melodic phrase greater than three notes is unique). Additionally, though Jarrett is regarded as being an improvisor in the lineage of modern jazz, his playing is very different to other comparable musicians. As such, typical frameworks used to explore jazz improvisation cannot be used to effectively explain his improvisational approach.

All of the methods used in this chapter are independent of the data under consideration. They can be be applied to any dataset that has the same structure. The case study is also intended to be scaleable: its methods can be used to explore any number of solos by any number of musicians. In doing this, I will also utilise a number of different data visualisations and the use of the traditional music score should be regarded as simply one possible data visualisation among many. Its use will be employed when it is the most appropriate view of the data under consideration. The appearance of the score as a visualisation is here used as a convenience, particularly when exploring how Jarrett uses phrases and microphrases.

All music score visualisation used in this chapter are been rendered directly from the dataset via Python code, which allows for many of the issues facing traditional score analysis to be alleviated (such as the need to transpose passages for analysis). Figures 5.1 through 5.4 provide a number of different visualisations of a single phrase taken directly from the data set (from the Days Of Wine And Roses solo), that can be used to explore different aspects of melodic structure.

Figure 5.1. Original phrase (Days Of Wine And Roses)

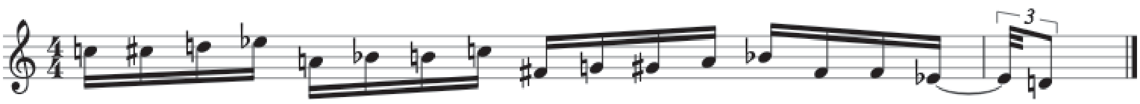


Figure 5.2. Phrase ignoring rhythm (Days Of Wine And Roses)

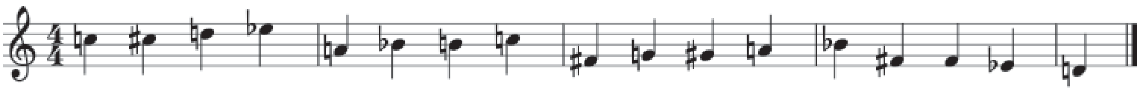


Figure 5.3. Phrase transcribed to start on middle C (Days Of Wine And Roses)

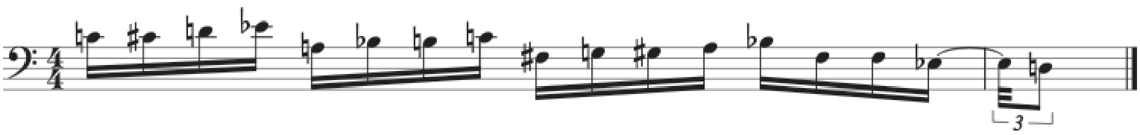
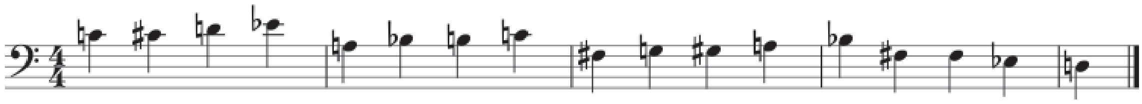


Figure 5.4. Phrase transcribed to start on middle C ignoring rhythm (Days Of Wine And Roses)



In this chapter I will utilise some basic nomenclature from more traditional approaches taken to music score analysis, employing terms such as tonality, key centre, scale, phrase, and chordal structure. However my choice of employing this language has been made for the sake of convenience of labelling and convention, rather than referencing any kind of external framework. In all cases this nomenclature relates directly to the data under consideration, and nothing external.

With regard to notions of scales and tonality, it is problematic to employ these kinds of structures to explain the melodic lines of the dataset under consideration. However the dataset also contains the underlying chords of the jazz standards and, when examining these, it quickly becomes clear that some pitch classes have far more prominence than others. In the Days Of Wine And Roses solo for example, the pitch classes F, G, A Bb, C D and E are far more prominent than other notes in comprising the underlying chords. Additionally, at the beginning and end of the this jazz standard, the pitch classes F, A and C appear. Similar patterns can be found in the other solos, and the movement between chord voicings is also very similar across the dataset. It is the only the presence of this information that is used to inform any notions of tonality or scale.

The case study will also utilise the idea of a melodic phrase, however its definition will be more flexible than existing definitions. The Grove Music Dictionary defines the melodic phrase as follows:

A term adopted from linguistic syntax and used for short musical units of various lengths; a phrase is generally regarded as longer than a motif but shorter than a period. It carries a melodic connotation, insofar as the term 'phrasing' is usually applied to the subdivision of a melodic line. As a formal unit, however, it must be considered in its polyphonic entirety, like 'period', 'sentence' and even 'theme' (http://www.oxfordmusiconline.com/grovemusic/(2017))

For the purposes of this case study, I will be defining a phrase simply as a group of subsequent notes that have no rest between each note. I will also explore the idea of what I denote as a "microphrase" in Jarrett's playing, here defined as a part of a melodic phrases. An example of a microphrase (and its relationship to a phrase) can be seen in Figure 5.5, where a four note microphrase is located within the blue box.

Figure 5.5 Phrase and microphrase

Though this definition is somewhat problematic as it is certainly overly simplified, (in that it misses so much nuance of what a melodic phrase is intended to be), the amount of data under consideration can allow for this broader definition to exist without negatively impacting on the findings.

Before commencing the analysis, there are three final caveats to raise. The first is to reiterate that this analysis is based on an analysis of music score metadata. The previous chapters have discussed the tendency to conflate the music score and music itself. This analysis will say nothing about a jazz musician's intention, or overall philosophy of playing, or become entangled in questions that explore the human relationship to music more generally. The metadata used in this chapter tells us almost nothing about this: it is simply a log of time-series information. This analysis will still show however that, even with such limited access to information, we can obtain such deep insights in the nature of jazz improvisation.

Secondly, the case study will reframe the question from a definition of jazz improvisation which was so problematic, to one of many possible definitions drawn from a corpus. Insights gained in this case study can have far reaching implications for the other work of Keith Jarrett, and are comparable to other jazz improvisors. However the insights can only be evidenced from this particular dataset. I can find out, for example, about the very specific ways in which a major seventh note can appear in a melodic phrase when there is an underlying dominant chord (i.e. a B note being played on a C dominant seventh chord) and can use this to help inform a very detailed picture of Jarrett's improvisational approach, however this can only be regarded as true in the context of the dataset at hand.

Finally, the case study is intended to be exploratory. Having access to the data in this form makes it feasible to transform the data in any way one might imagine, and allows a high level of freedom to explore. But the practical implication is that the case study becomes very long. Every behaviour of every note can be easily explored. This challenges the feasibility of this type of analysis being presented in a linear fashion, and the next chapter will discuss how this problem can be revisited by the use of alternate user interfaces.

The analysis and visualisations for this chapter have been carried out with the Python programming language, and all the code that produces different views of the data and different visualisations is available in several Jupyter notebooks. The details of of these can be seen in Appendix 2 with details of how they can be viewed and downloaded.

The case study will begin by describing some general characteristics before exploring different ways in which musical time can be described and how this can be influenced by tempo. I will then transform the base dataset into one that isolates all the separate melodic phrases, and examine the characteristics of these.

Following this, I will examine one of the major challenges of analysing Jarrett's work: the apparent lack of repetition in his playing of melodic phrases. What quickly becomes apparent when examining all of the melodic phrases that comprise the ten solos is that almost all of them appear only once. Often in jazz improvisation analysis, it is common to speak of typical patterns or "licks" that might characterise a player's style, however this is not possible with Jarrett. However this case study will also show that, although phrases may be unique in Jarrett's playing, it is possible to isolate small microphrases which can be viewed as building blocks of larger phrases, and within these can be found a high level of structure, repetition, and even predictability.

Finally, the case study will explore harmony and voice leading. For this dataset, the way in which notes in melodic phrases are prepared and resolved appears critical, and relates directly to Jarrett's ability to transition between microphrases.

The dataset under consideration has been taken from ten improvised jazz solos (comprised of 16,174 records in the dataset). Basic details about the solos can be seen in Table 5.1.

Table 5.1 General characteristics of the dataset

Title	Performer collection	Date recorded	Composer collection	Date composed	Quarter beats per minute	Tonality	Number of records
All The Things You Are	Standards, Vol. 1	1983	Very Warm For May	1939	247	Ab major	2027
Autumn Leaves	Still Live	1986	Les Portes De La Nuit	1945	251	G minor	1826
Autumn Leaves	Tokyo 96	1996	Les Portes De La Nuit	1945	224	G minor	1243
Days Of Wine And Roses	Keith Jarrett At The Blue Note, The Complete R...	1994	Days Of Wine And Roses	1962	160	F major	1424
Groovin High	Whisper Not	1999	Shaw Nuff	1945	289	Eb major	1811
If I Were A Bell	Up For It	2002	Guys And Dolls	1950	167	Ab major	1982
In Love In Vain	Standards, Vol. 2	1983	Centennial Summer	1946	147	Bb major	1280
My Funny Valentine	Still Live	1986	Babes In Arms	1937	122	C minor	1254
Someday My Prince Will Come	Up For It	2002	Snow white and the seven dwarfs	1937	148	Bb major	1815
Stella By Starlight	Standards Live	1983	The Uninvited	1944	151	Bb major	1512

The improvised solos listed above are all from well known jazz standards, and are all taken from jazz piano trio performances with Gary Peacock playing double bass and Jack DeJohnette playing drums. The solos were recorded over a nineteen year period, between 1983 and 2002. Two of the improvised solos taken from the dataset are from the same jazz standard, Autumn Leaves, and these versions were recorded ten years apart.

Jazz standards typically utilise keys with flats in the key signature (such as F major, Bb major, Eb major etc.) and this is the case here. Five of the ten solos are in a key signature with two flats (being the improvised solos in Bb major and G minor) and other keys include Ab, Eb and C minor.

In the methodology, I demonstrated how metadata can be taken from a music score and integrated with other, related information. This method was used to create the dataset and Table 5.2 displays the first record of the dataset. Each record within the dataset represents a note or rest in a music score, and contains a large number of attributes that describe the note and its relationship to the dataset. As well as basic information around pitch and duration taken from the music score itself, all records have also been encoded with additional attributes such as title, year recorded, year composed, and even the latitude and longitude of the location in which the note was recorded.

Table 5.2. Sample record taken from the dataset

Composer collection	Very Warm For May
Composer name	Jerome Kern
Composer nationality	US
Current chord root	G
Current chord root as int	7
Current chord type	min7b5
Current location in seconds	0.242915
Current measure	1
Date composed	1939
Date recorded	1983
Duration	1.0
Duration as string	Quarter note
Duration due to tied notes	480
Duration of one second	1976

Genre	Acoustic Jazz
Instrument	Part 0
Location	0
Location in measure	0
Lyricist name	Oscar Hammerstein II
Lyricist nationality	US
Measure location	0
Midi number	-1
Midi number as string	rest
Midi number as string without octave	rest
Performer collection	Standards, Vol. 1
Performer name	Keith Jarrett
Performer nationality	US
Quarter beats per minute	247
Recording location	{'lat': [40.7831, 'N'], 'lon': [73.9712, 'W']}
Time signature denominator	4
Time signature numerator	4
Time stamp	1685-03-20T13:00:00.000Z
Title	All The Things You Are

From the above example, it can be seen that pitch and duration are represented in the dataset in a different ways. Pitch is represented both as a string (for example "rest", "C4", or "C#/Db5") or a midi number (60, or 61). Duration is encoded both as a string (such as "quarter note") and a number (where 480 is the equivalent of a quarter note).

When filtering the records that indicate a rest (denoted in the set as having a midi number of -1) this leaves 14,537 improvised notes from the ten improvised solos. Some basic characteristics and summary statistics of all of these notes can be seen in Table 5.3 below.

Table 5.3. Characteristics of pitches (as midi numbers) used in dataset

Total count of notes	14537.000000
Average midi number	71.076357
Standard deviation	7.045010
Minimum midi number	46.000000
First quartile	66.000000
Second quartile	71.000000
Third quartile	75.000000
Maximum midi number	103.000000

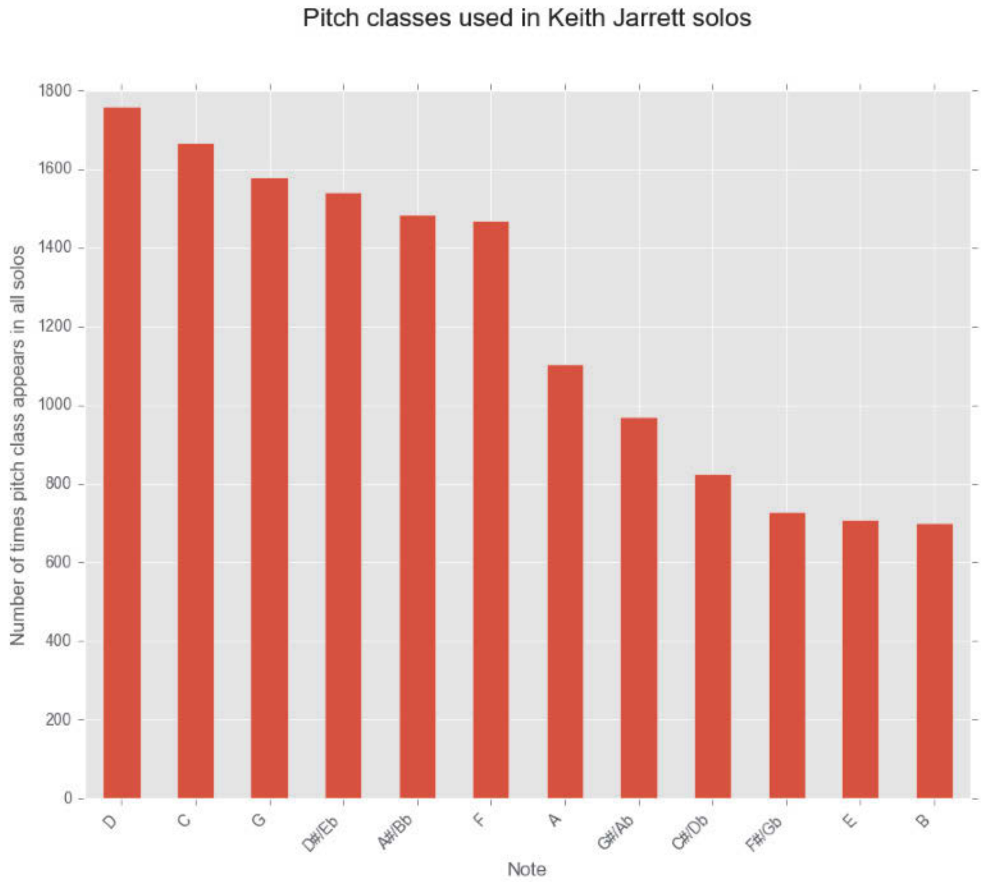
The above table shows that the lowest midi number used in the ten improvised solos is 46 (being the note G#2/Ab2) and the highest midi number is103 (being the note G7). Further, 50% of all notes played fall between the midi note 66 (being the note G#4/Ab4) and the midi number 75 (being the note D#5/Eb5).

This suggests that, although Jarrett has access to all 88 notes of the keyboard for the purpose of improvising melodic lines with his right hand, he utilises a far smaller range. Half of all the notes he plays takes place within the range of only one octave and a fifth. 75% of notes played fall in a two octave range. Table 3 also notes a standard deviation of 7.04. This means when Jarrett plays a note in a solo it is, on average, only 7 semitones to either side of the average pitch used in the dataset which is 71 (being the note B4).

Although these considerations are at a very high level, these metrics still suggest that, whatever meaning can be found in jazz improvisation, it is problematic to characterise it as something exhibits constant change, in which a musician has the freedom to play any particular note. Instead (at this level at least), jazz improvisation appears to be characterised by very strict limitations.

It is possible to ignore the octave of any given note in order to focus on the pitch classes being used. Figure 5.6 displays the count of pitch classes across the dataset, and shows that they do not appear in a uniform distribution. Some pitch classes, such as C, and D, are more than twice as likely to appear than other pitch classes, such as B, E and F#.

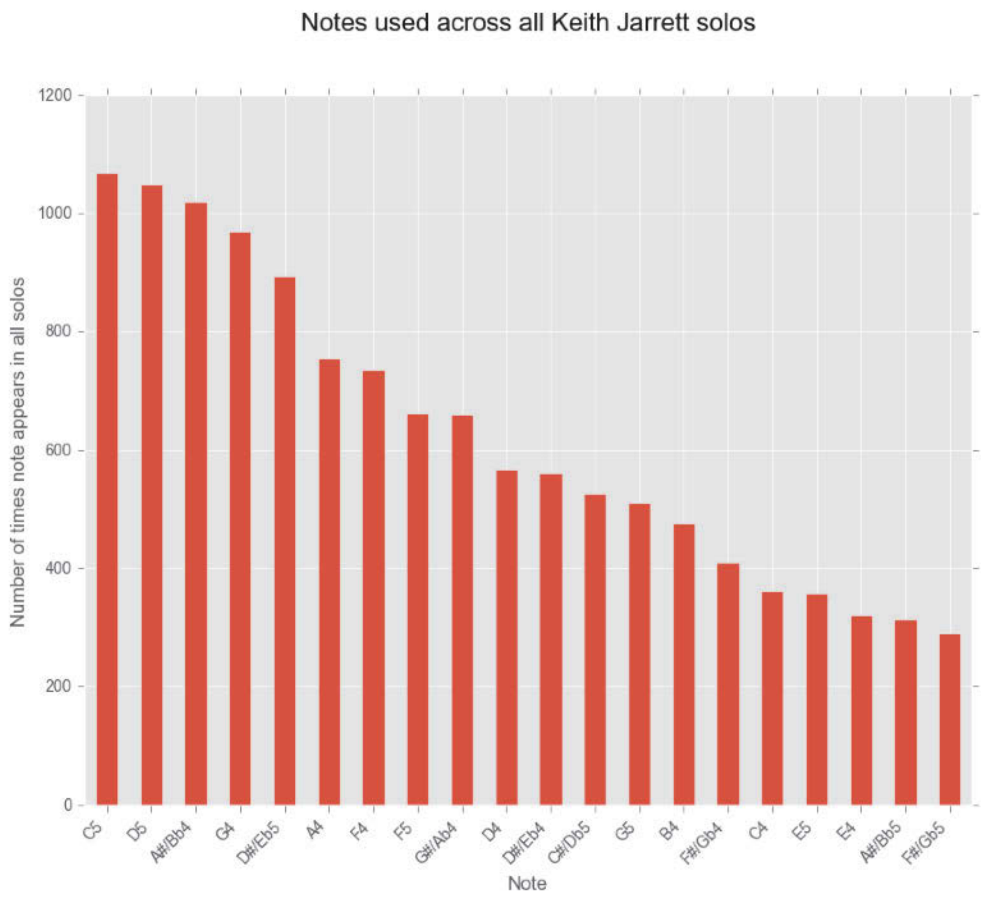
Figure 5.6. Pitch classes used in all solos



This suggests that there must be some kind of correlation between the notes of the solos and and notes in the underlying chords. Five of the ten solos are in either Bb major or G minor and the notes that predominantly inform the chord progressions of both these keys are Bb, C, D, Eb, F, G and A. Thus, the chords must somehow be influencing the played notes. However this does not necessarily imply that Jarrett plays chord tones of underlying chord tones in solos, but rather that, on the whole, Jarrett tends to favour notes found in the underlying tonality.

Figure 5.7 provides a more nuanced view of pitch, this time including octave information. Only the top 20 highest occurring pitches are displayed. Again, the most commonly occurring notes are those that are more closely related to the tonality of the solos.

Figure 5.7. Notes used across all solos



Though there is a correlation between the notes of the underlying chords and notes used in the solos, it can still be seen in the data that Jarrett will use all twelve pitch classes (often in the course of a single phrase), and utilises them regardless of chord root and chord type. This suggests that a correlation between underlying chords and notes in the solo does not provide the whole story. There is far more complexity in play here, in understanding how the notes in the solos interact with the notes of underlying chords.

Figure 5.8 provides a visualisation of the different pitches used across all the solos. This is in the form of a histogram, and shows the frequency of use of different notes (here seen as midi numbers) across the entire dataset. The note choice is normally distributed (meaning it has a bell shape) indicating that Jarrett not only plays in a limited pitch range but also balances the playing of higher pitches with lower pitches.

Figure 5.8. Midi numbers used across all solos

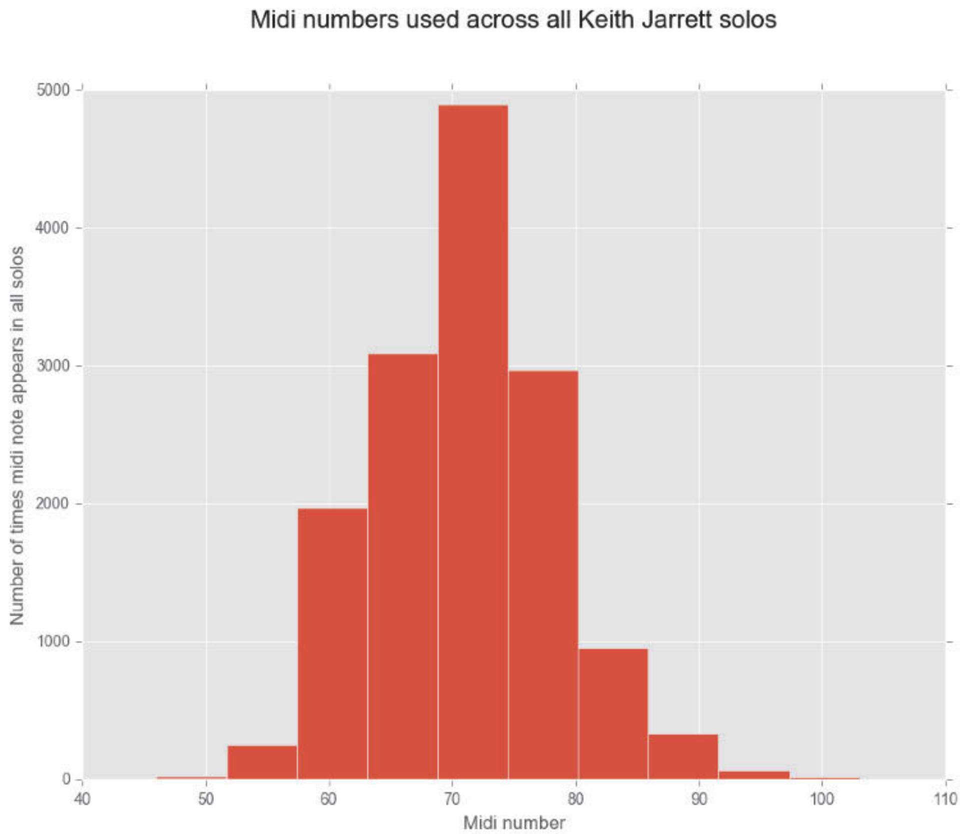


Table 5.4 provides the counts of the different rhythmic units that Jarrett utilises while improvising. It reflects what is typical of much Western music, that duration in music tends to divided into symmetrical units whose precise timing is influenced by a given tempo (this is denoted in the dataset in quarter beats per

minute, and can be seen in Table 5.1). Almost 50% the notes played by Jarrett have eighth note durations. Taken together, more than two thirds of the durations are eighth notes and sixteenth notes. A complete breakdown of the note duration types can be seen in Table 5.4 below.

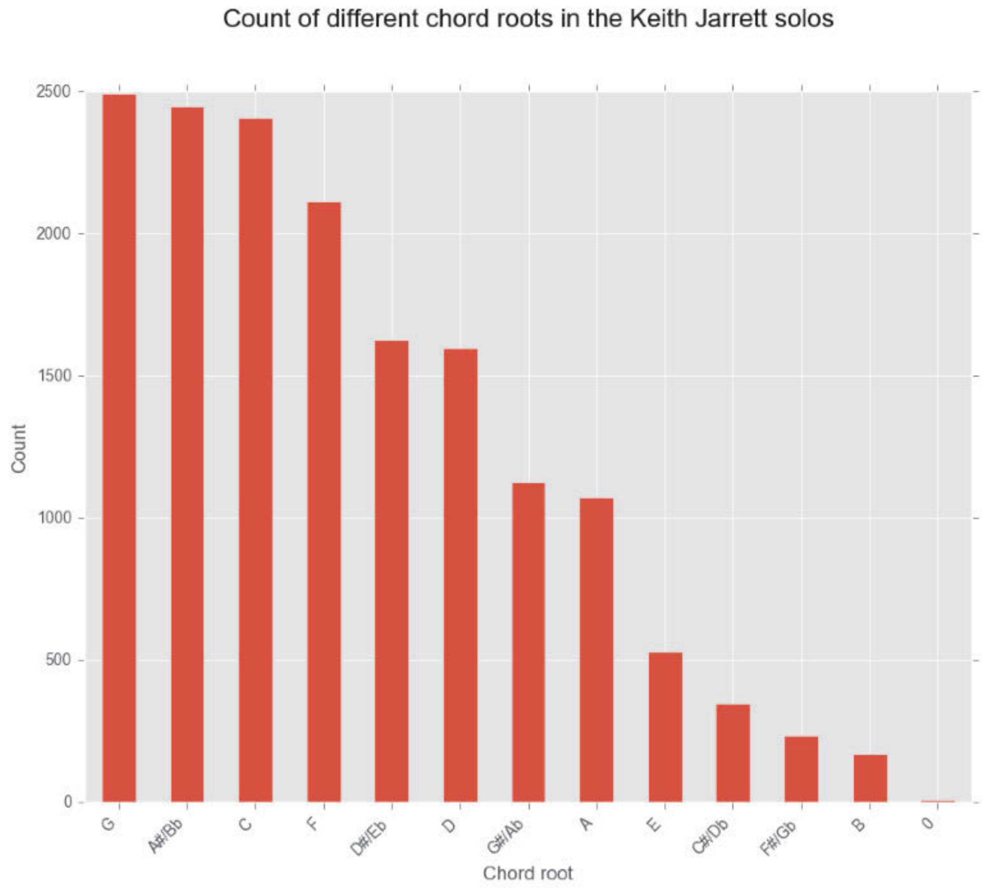
Table 5.4. Counts of different types of durations used in the dataset

Duration type	Number of times duration type appears in all solos
Eighth note	7430
Sixteenth note	3841
Twelfth note	1448
Quarter note	784
Thirty-sixth note	217
Twentieth note	217
Dotted quarter note	191
Sixth note	86
Half note	67
Thirty second note	56

Similar to what was found regarding the range of notes in the solos, it seems that, for Keith Jarrett at least, improvised melodies in jazz are not characterised by high levels of freedom and endless inventiveness in rhythm. On the contrary, the above suggests that there are severe limits being placed on what is possible. Not only is the range of playing extremely limited, but so is the rhythmic choice.

Every record in this dataset, be it a note or rest, is accompanied by a underlying chord type and chord root. This makes it possible to examine some general characteristics of the chords in the dataset and the distribution of the chord roots. Figure 5.9 shows the different chord roots found in the dataset.

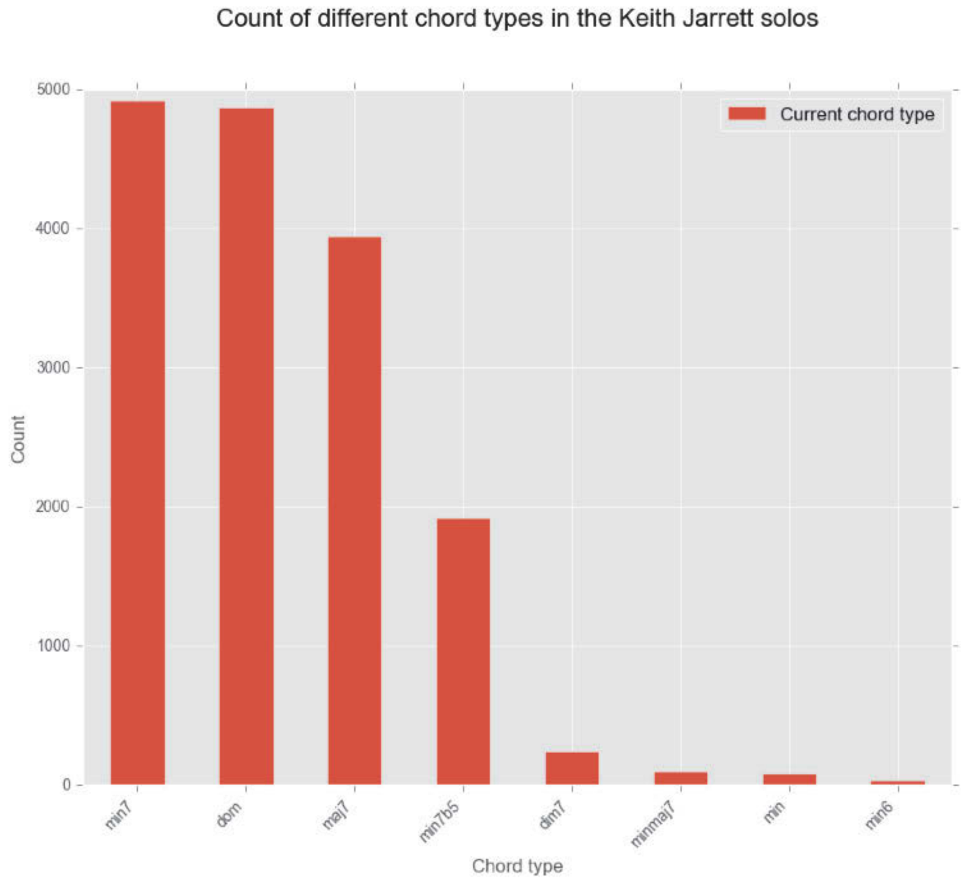
Figure 5.9. Count of different chord roots in all solos



The chord roots also reveal a strong relationship to the most commonly used key centres. The chords roots of I, IV and V in both G minor and Bb major are the six most common chord types, which also indicates that the chord progressions have some kind of underlying structure. Also, chord roots such as G and A#/Bb are more than ten times more likely to occur than B and F#.

Figure 5.10 displays the distribution of the chord types being used across the dataset.

Figure 5.10. Count of different chord types in all solos



In this dataset, chord types such as minor, dominant and major chords are far more prevalent than other chords types. Typical jazz standards are often characterised by the appearance of II-V-I progressions in sequence (for example the sequence of D minor-seventh, G dominant-seventh, and C major-seventh). Other chord types, such as the diminished seventh, are more rare and this is reflected in the dataset, indicating that this dataset represents a typical sample of jazz standards. Note that the appearance of a small number of minor, minor-major-seventh and minor-sixth chords. This is due to the opening chords of <u>My Funny Valentine</u>, a progression limited to only one the solos represented in the dataset. Having such a large sample of notes that occur on a limited number of chord types will inform the exploration of voice-leading and harmony later in this case study.

It might be argued that what is being seen so far is, self evident and intuitive. Putting forward the claim that notes in the solo tend to reflect the notes in underlying chords is not a radical one. However, the purpose here is to show this in an evidence based way. Tracking metrics such such as range and note duration and progressions are critical ways to establish comparative benchmarks to other solos and jazz improvisors. For the purpose of this case study, they can empirically establish that the lack of repetition in Jarrett's playing the apparent endless inventiveness is taking place in a highly structured and predictable environment.

Much score analysis is predominantly concerned with notions of musical time (in which an overarching tempo indicates allowable subdivisions of duration) rather than actual time. However this dataset allows the posing of a more foundational question, which asks if there is a relationship between the elapsing of time and playing of notes. It is possible to explore this in the dataset and even derive a notes-per-second rate of playing that Jarrett tends to adopt regardless of tempo and duration.

Figure 5.11 below plots the number of notes that are played on the x-axis, and the number of seconds that have elapsed during the course of the solo on the y-axis. The dataset is here filtered to show only those records found in All The Things You Are and Groovin High.

Figure 5.11. Number of notes played over time measured in seconds

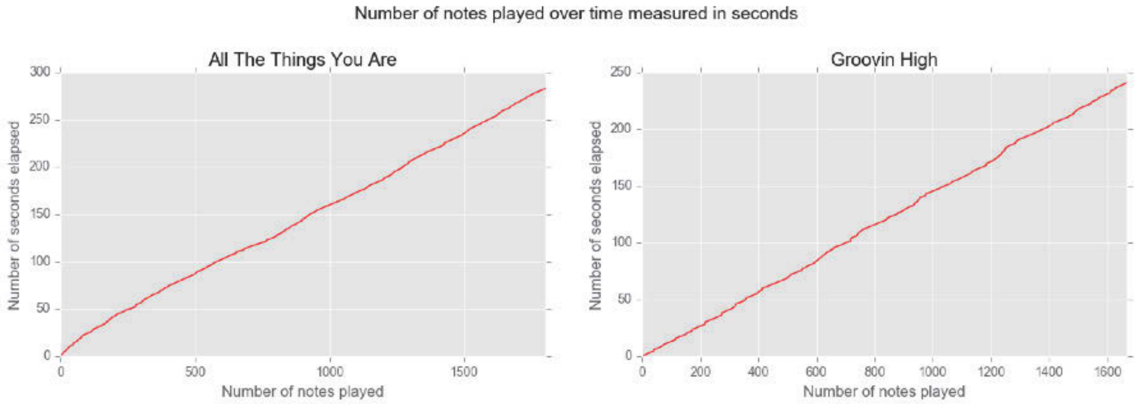
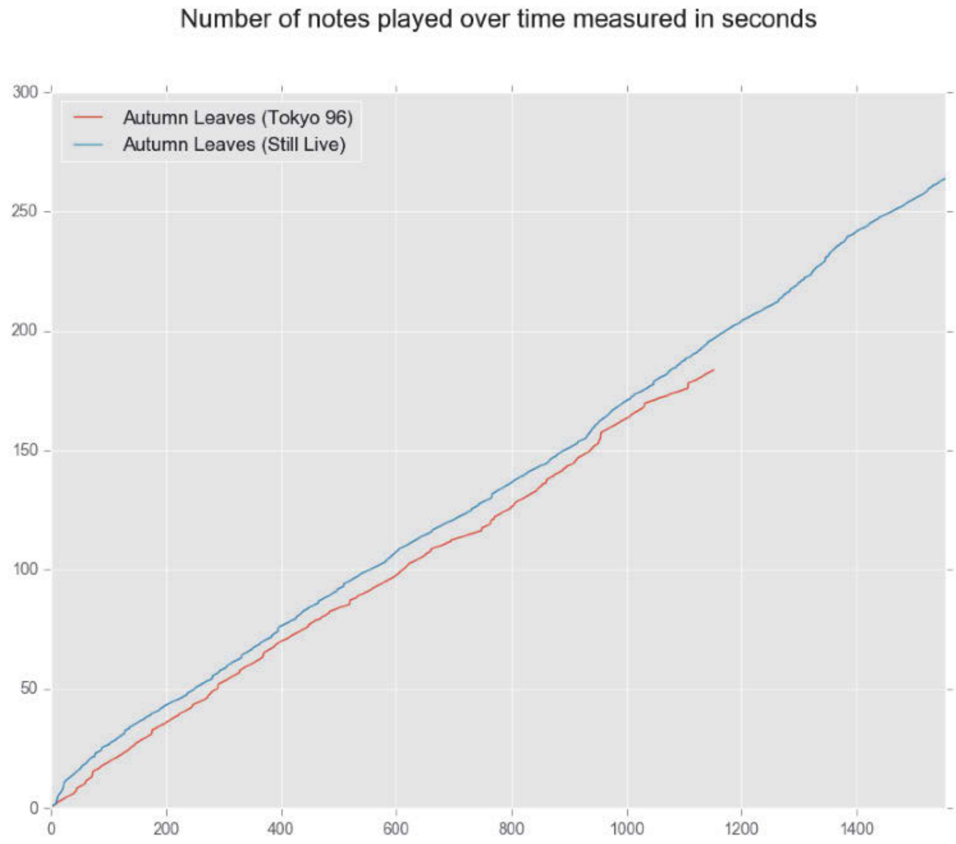


Figure 5.11 shows that, for these two solos, Jarrett tends to improvise at a constant rate of time. He does not play more notes in some parts of the improvisation and less in others: as the time elapses the rate of playing stays constant. Yet the playing rate is different for each solo. In the All The Things You Are solo, Jarrett plays approximately 1000 notes in 160 seconds. In Groovin High, he plays a 1000 notes in just under 150 seconds.

Figure 5.12 displays a side by side comparison of both the Autumn Leaves improvisations (and note that the *Still Live* version is a longer solo, so this line extends over more seconds). It again reveals Jarrett to be using a fairly constant rate of playing in terms of notes per second. However the rate in both is not the same: the *Tokyo 96* version of Autumn Leaves appears to have a slower playing rate than than the *Still Live* version.

Figure 5.12. Number of notes played over time measured in seconds



The rate of playing in Jarrett's soloing can be explained when examining the tempo (here measured in quarter note beats per minute or bpm). When examining the Autumn Leaves improvisations, the version from *Still Live* has a tempo of 251 bpm and the *Tokyo 96* version has a tempo of 224 bpm. Jarrett's heavy use of eighth and sixteenth note durations leads to the playing of more notes at faster tempos. This can also explain the different playing rates found in All The Things You Are and Groovin High.

This might imply that improvisations with slower tempos should have slower rates of playing. But, as is shown in Figure 5.13 which plots the same data If I Were A Bell and In Love In Vain, things are more complex.

Figure 5.13. Number of notes played over time measured in seconds

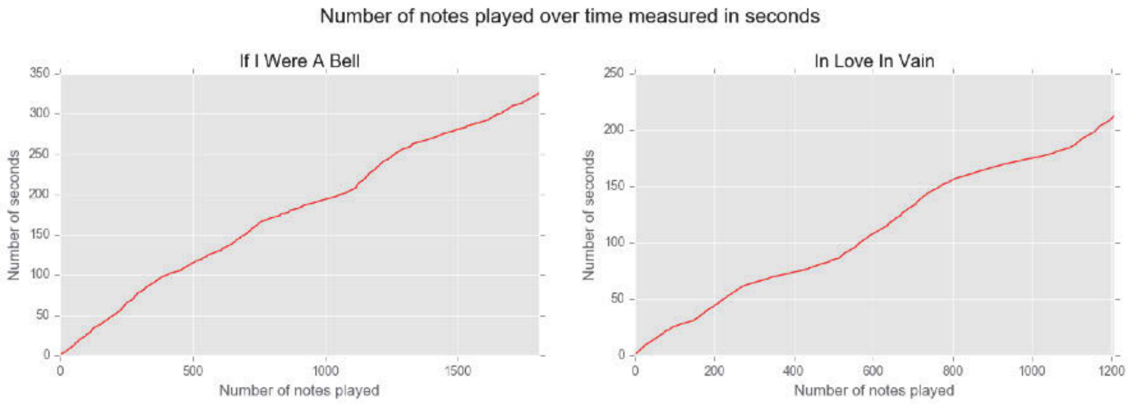
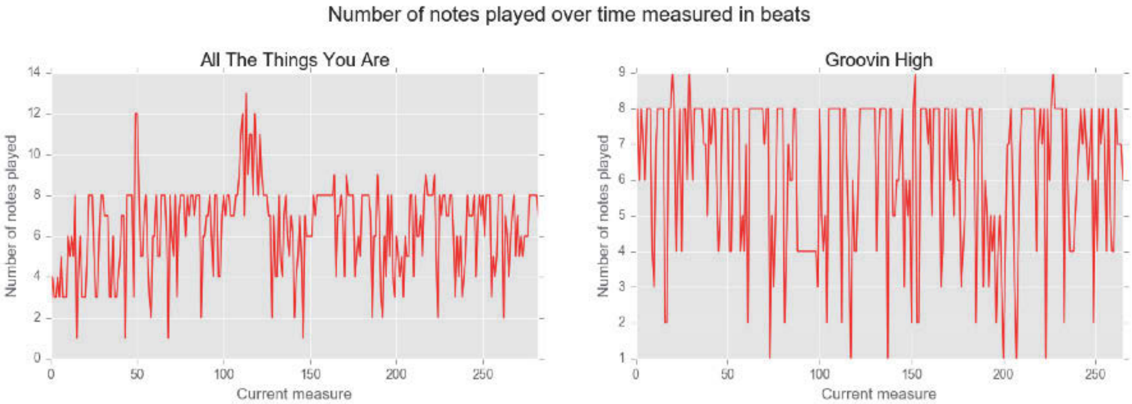


Figure 5.13 shows that, at slower tempos, the rate the playing is not nearly as constant. The plotted line becomes less linear, and instead undulates, indicating that at certain times in the solo Jarrett plays more densely, and at other times plays less notes. Examining the data, this is because, at slower tempos, Jarrett will start to move more freely between eighth note and sixteenth notes. The solos If I were a Bell and In Love In Vain have a bpm of 167 and 147 respectively, indicating that the slower the tempo the higher the rhythmic variation in his playing, expressed as curvature in the figure above. Slower tempos allow Jarrett to play more notes as time elapses, and he does this by accessing smaller rhythmic subdivisions.

It is more typical in music analysis to examine time in terms of how measures and beats can be subdivided, and how they relate to time signatures. Figure 5.14 below provides a line plot of the count of notes played on the y-axis, but time is now correlated to measures, to examine how many notes are played in each measure during a solo and whether this changes over time. The data has again been filtered to consider the examples of All The Things You Are and Groovin High.

Figure 5.14. Number of notes played over time measured in beats



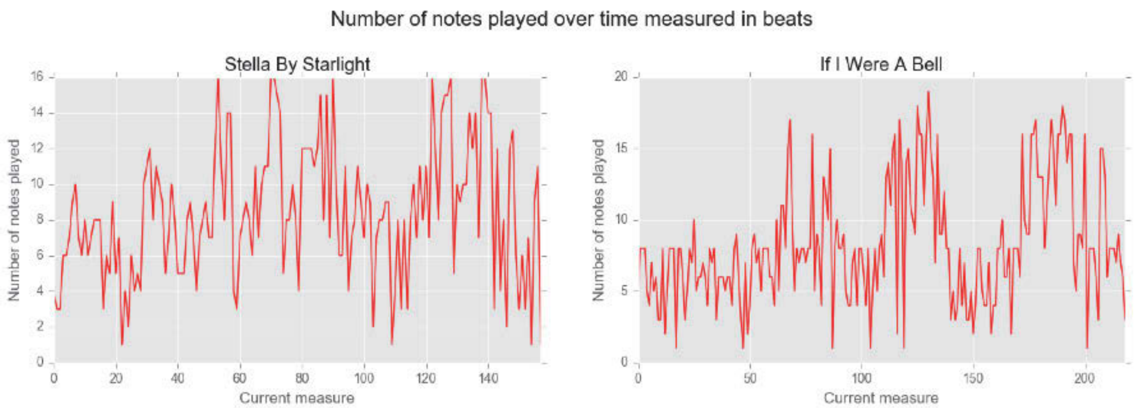
This reveals a strong tendency from Jarrett to play measures which have eight notes each within them. Though there are certainly some occasions when there are less than eight notes in the bar, and even fewer occasions when there are more than eight notes in the bar, overall it is kept at eight. This visualisation also shows an overall balance in the solo. It is not the case that there are always eight notes at certain times, or always more at other times, rather Jarrett varies this throughout the course of the solo.

Figure 5.15 shows the number of notes per measure plotted for both versions of <u>Autumn Leaves</u>. Again, things are are similar. Both show a tendency toward eighth notes mixed with variation throughout.

Figure 5.15. Number of notes played over time measured in beats

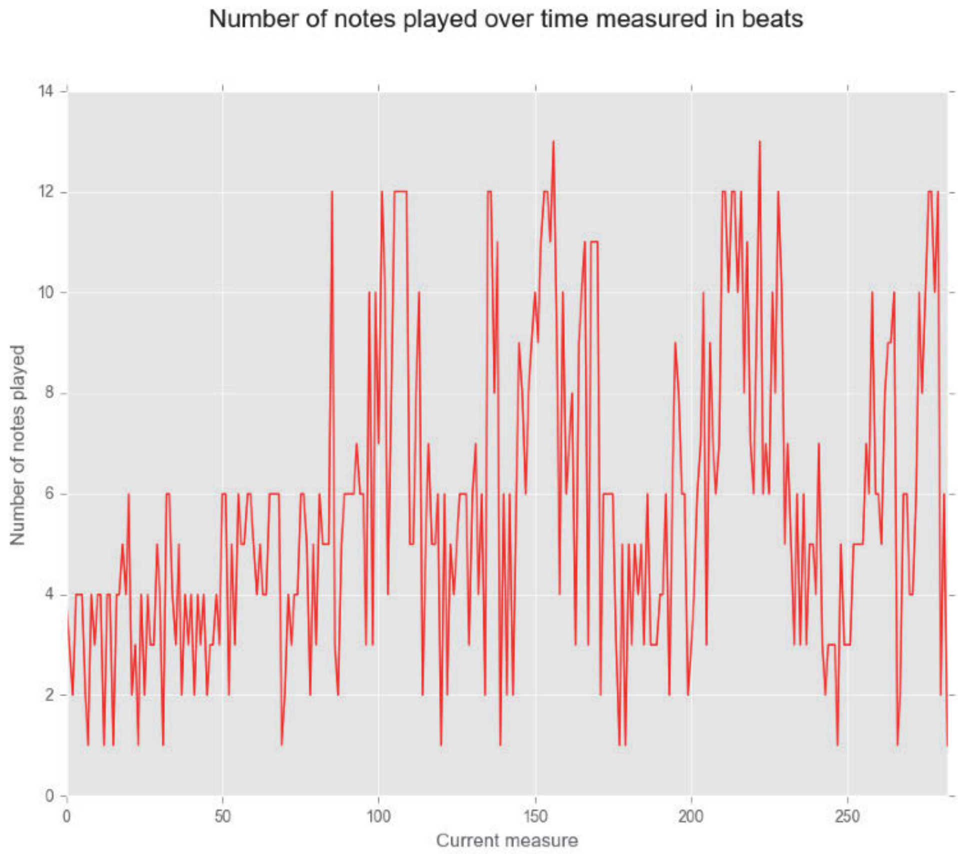
The solos on <u>If I Were A Bell</u> and <u>Stella By Starlight</u>, have slower tempos and have been seen to have more variation in rhythmic choice, and support the argument that slower tempos disrupt the regularity of notes played over time. Examining the notes played per bar, there is no longer a clear trend to be seen in regard to the use of eight notes in a bar. Measures containing eight, twelve and sixteen notes can seen interchangeably. Again, there appears to be a sense of balance within the solos, and the number of notes constantly moves: measures with many notes are balanced by sparser measures.

Figure 5.16. Number of notes played over time measured in beats



A final visualisation is given in Figure 5.17 which shows the number of notes being played correlated to any given measure for <u>Someday My Prince Will Come</u>.

Figure 5.17. Number of notes played over time measured in beats



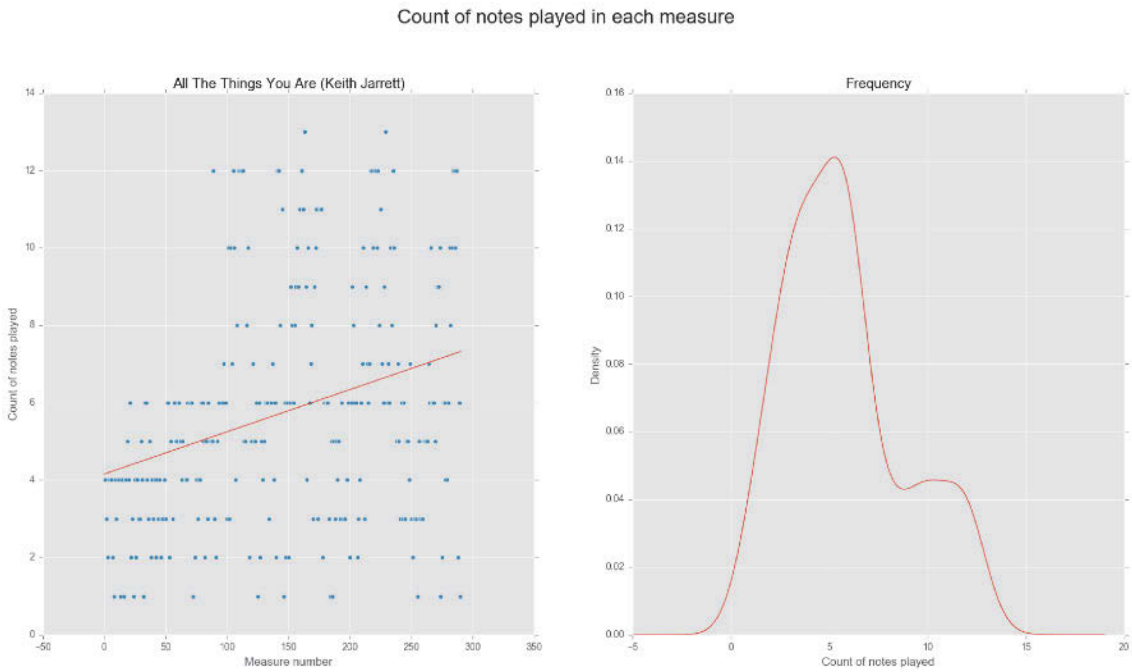
The above figure again demonstrates how overall tempo affects playing rate and notes per measure in the way seen above. Someday My Prince Will Come has a bpm of 148 and, typically of slower tempos, sees more variation in the number of notes per measure. Of interest here is also is Jarrett's tendency to play six notes or twelve notes in a measure rather than eight or sixteen notes. However, Someday My Prince Will Come has a ¾ time signature which equates to six eighth-note durations or twelve sixteenth-note durations per measure, suggesting that Jarrett uses rhythm in similar way regardless of time signature.

Although the rate of playing is less constant at slower tempos, and more undulating at slower tempos, some of the visualisations above, particularly the slower tempo Someday My Prince Will Come, If I Were A Bell and the *Tokyo 96* version of Autumn Leaves, indicate that over time, Keith Jarrett solos do in fact get busier. Anecdotally at least, it often seems that case that, over the course of long jazz solos, improvisors will start out by playing fewer notes and explore different ideas, gradually building them up into frantic passages. Jazz musicians often talk about how busy a soloist is, and discuss this in the context of jazz musicians overplaying.

The structure of this data allows metrics to be placed around this. Figure 5.18 below again plots the number of notes per measure (here seen as a scatter plot), but also tracks the change over time in the average amount of notes being played per measure, indicated by the plotted red line.

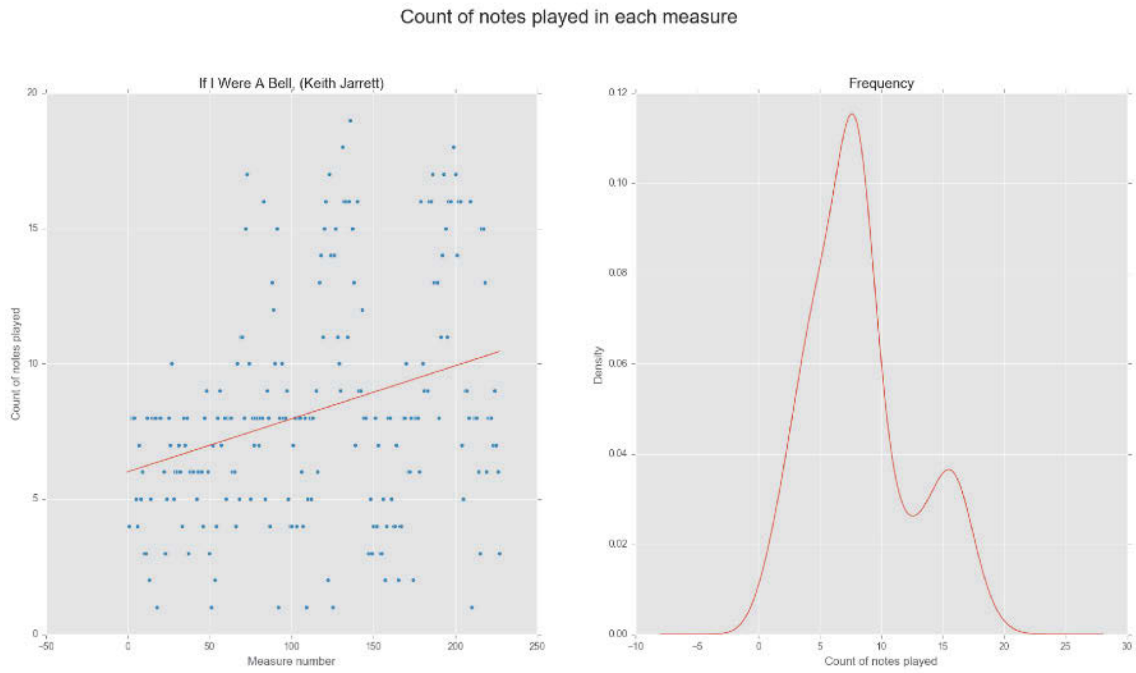
Figure 5.18 shows that in the All The Things You Are solo, there is a tendency to play less notes at the early parts of the solo. Over the course of this solo, Jarrett starts out by playing four notes notes in a measure on average, and then gradually increases this to eight notes per measure on average. On the right side of the Figure I have also included a frequency distribution of the counts of notes per measure.

Figure 5.18.   Count of notes played in each measure



A similar example of this increasing busyness can be seen in <u>If I Were A Bell</u>, in Figure 5.19. Overall, there is a tendency to play eight notes in a bar, however at the beginning of the solo less notes are played on average, and at the end of the solo there are more notes being played on average (increasing over time from five to ten notes on average)

Figure 5.19.   Count of notes played in each measure



However this trend does not hold across the entire dataset. On this metric, Groovin High is the outlier to all the other solos, and this can be seen in Figure 5.20. There is again a clear tendency for Jarrett to play eight notes per measure. However over time, the average declines slightly. While there is certainly not enough data to interrogate why this is the case, it could be due to Groovin High being somewhat different to the other standards. It is a bebop standard, whereas the others are staples from the American Songbook. Additionally, Jarrett has also noted that the playing on this recording is very different from his usual soloing style, being much lighter and in the the bebop style. This is a good example of a question that is simply not possible to pose in a more traditional approach to analysis, and one that could be answered with the addition or more data.

Figure 5.20. Count of notes played in each measure

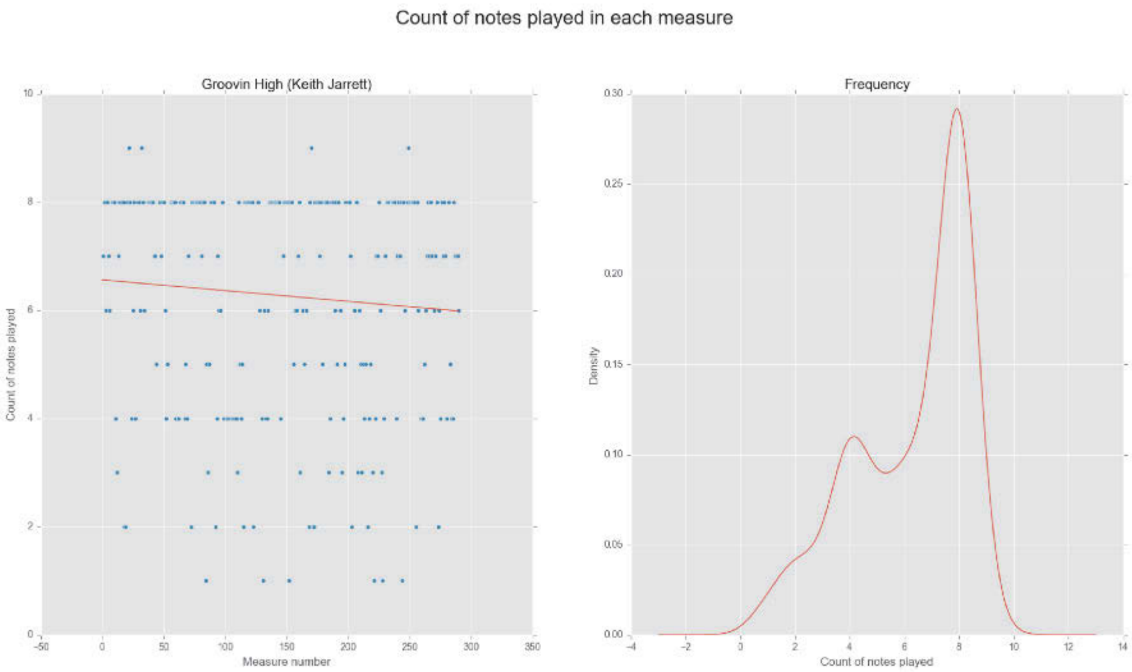


Table 5.5 below provides some further information regarding the number of notes played per measure in each solo. Included is the average amount of notes played, the median amount of notes played and the standard deviation (being an average of how far the count of notes in any given measure is away from the average notes played in a measure).

Table 5.5 Average and median notes per measure and standard deviation, grouped by title

Title	Average	Median	Standard Deviation
All The Things You Are	6.371025	7.0	2.183142
Autumn Leaves	6.283063	6.0	2.300387
Days Of Wine And Roses	8.815068	8.5	4.467512
Groovin High	6.270677	7.0	2.083639
If I Were A Bell	8.246575	8.0	4.186416
In Love In Vain	9.221374	8.0	5.049433
My Funny Valentine	10.064220	10.0	3.895095
Someday My Prince Will Come	5.749117	5.0	3.051094

Stella By Starlight	8.436709	8.0	3.757682

Overall, this presents a picture of consistency in terms of how Keith Jarrett improvises. More moderate and slower tempos have higher average and median numbers of notes played in each measure. Moderate tempos also have a higher standard deviations indicating the tendency to move constantly between eighth notes and sixteenth notes. The slowest tempo, from My Funny Valentine at 122 bpm, has high a high median and average (suggesting more use of smaller subdivisions in a bar) but a lower standard deviation which suggests that, once a tempo is slow enough, Jarrett will start to utilise small subdivisions for most of the time (and in this solo he plays mostly sixteenth notes).

Exploring other factors that may influence the number of notes Jarrett plays does not yield much of value. When correlated to other available metadata such as year composed, and recording location where solos recorded no patterns could be seen.

A comparison of the number of notes played in a measure to chord type can be seen in Table 5.6. There is tendency of many improvisors when they are first starting out, to play more notes on chord types that might be considered easier. As to be expected however, there is not trend around this that can be seen.

Table 5.6. Average amount of notes played in a measure, grouped by chord type and title

Chord type	Title	Average notes in measure
dim7	All The Things You Are	7.222222
dom	All The Things You Are	6.179487
maj7	All The Things You Are	6.306931
min7	All The Things You Are	6.564103

min7b5	All The Things You Are	6.500000
dom	Groovin high	6.395604
maj7	Groovin high	5.616438
min7	Groovin high	6.647059
min7b5	Groovin high	6.588235
dom	Stella By Starlight	8.135593
maj7	Stella By Starlight	9.142857
min7	Stella By Starlight	8.000000
min7b5	Stella By Starlight	8.600000
dim7	Someday My Prince Will Come	6.520000
dom	Someday My Prince Will Come	6.198198
maj7	Someday My Prince Will Come	5.375000
min7	Someday My Prince Will Come	5.219780

This dataset has so far shown that jazz improvisation can be characterised by the use of certain rhythmic subdivisions affected by tempo, a limited range, the playing of notes related to tones used in underlying chords, and a slight tendency to play more notes over time. However this is only at a very high level, with the intent of understanding some general characteristics of Keith Jarrett's improvisational style. In order to interrogate the data at a deeper level the structure of melodic phrases requires further examination.

It is straightforward to transform the base dataset into a second one that groups notes together into phrases. Table 5.7 displays the first three records of the transformed data. Each record denotes a single phrase and also contains additional data pertaining the phrase such as its location, number of notes, range, number of steps and leaps etc. all of which have been derived from the base dataset.

Table 5.7. Three sample records of phrases found in the dataset

Phrase midi numbers	[74, 72, 69]	[72, 69, 74, 72]	[72, 69, 76, 74]
Title	Days Of Wine And Roses	Days Of Wine And Roses	Days Of Wine And Roses
Phrase durations	[0.5, 1.0, 1.0]	[0.5, 1.0, 1.0, 1.0]	[0.5, 1.0, 1.0, 1.0]
Performer name	Keith Jarrett	Keith Jarrett	Keith Jarrett
Performer collection	Keith Jarrett At The Blue Note, The Complete Recordings (Vol. 3)	Keith Jarrett At The Blue Note, The Complete Recordings (Vol. 3)	Keith Jarrett At The Blue Note, The Complete Recordings (Vol. 3)
Number of notes in phrase	3	4	4
Measure in which phrase begins	1	2	3
Measure in which phrase ends	1	2	3
Measure location in which phrase begins	1.5	0.5	0.5
Measure location in which phrase ends	3	3	3
Range of phrase in semitones	5	5	7
Number of pitches used in phrase	3	3	4
Highest pitch used in phrase	74	74	76
Lowest pitch used in phrase	69	69	69
Different pitches used in phrase	[72, 74, 69]	[72, 74, 69]	[72, 74, 76, 69]
Different durations used in phrase	[0.5, 1.0]	[0.5, 1.0]	[0.5, 1.0]
Longest duration used in phrase	1	1	1
Shortest duration used in phrase	0.5	0.5	0.5
Number of different durations used in phrase	2	2	2
Phrase midi numbers transposed to start on middle C	[60, 58, 55]	[60, 57, 62, 60]	[60, 57, 64, 62]
Distances between subsequent phrase notes	[-2, -3]	[-3, 5, -2]	[-3, 7, -2]

Number of positive steps or leaps in phrase	0	1	1
Number of negative steps or leaps in phrase	2	2	2
Number of step movements (by tones or semitones) in phrase	1	1	1
Number of leap movements (by minor thirds or above) in phrase	1	2	2

One of the motivations informing this case study was to explore why there appears to be no repetition in Jarrett's phrases. With the dataset transformed, it now becomes possible to pose this question. Recalling that a melodic phrase is here defined as one more subsequent notes that have no rests between them, Table 5.8 below displays those melodic phrases (here denoted with the midi numbers) which occur more than once in the data set, and the number of times they occur (duration has been ignored).

Table 5.8 Most commonly occurring phrases described by midi number sequence, ignoring rhythm

Sequence of midi numbers	**Number of times phrase occurs**
[67]	7
[70]	5
[62]	3
[72]	3
[82]	2
[79, 77]	2
[68]	2
[74]	2
[67, 65]	2
[74, 72]	2
[70, 68]	2

[86, 84]	2

The highest occurring phrase in the dataset appears only seven times and consists of only one note (the midi number 67 or G4). And though this result technically meets the definition of a phrase in the way I have defined it, it makes little sense to think of this as a distinct melody. Even the idea of a two-note phrase is problematic, and the dataset shows that there are only five examples of two notes phrases which each occur twice.

This suggests that, at the level of melodic phrases, the improvisations of Keith Jarrett's have no repetition. If this group of solos is to be regarded as a representative sample, it could be inferred that Jarrett has the ability to produce endless melodic variation. As such, it would also then make little sense to seek to discuss Jarrett's improvisation within a framework focusing on the use of certain "licks" or melodies that typify his playing which often happens in jazz analysis.

It is also possible to explore the relationship between the melodic phrases and the solos in which they are played. Table 5.9 below shows, for each solo, the total count of phrases that are found, and the percentage of any given phrase that would be expected to appear in a single measure in that solo.

Table 5.9. Count of phrases in each solo, and percentage of phrase in each measure

Performer collection	**Title**	**Count of measures in solo**	**Count of phrases**	**Average percentage of phrase per measure**
Keith Jarrett At The Blue Note, The Complete Recordings (Vol. 3)	Days Of Wine And Roses	163	51	0.312883
Standards Live	Stella By Starlight	161	78	0.484472

Standards, Vol. 1	All The Things You Are	290	121	0.417241
Standards, Vol. 2	In Love In Vain	131	47	0.358779
Still Live	Autumn Leaves	276	137	0.496377
Still Live	My Funny Valentine	111	79	0.711712
Tokyo 96	Autumn Leaves	171	55	0.321637
Up For It	If I Were A Bell	227	91	0.400881
Up For It	Someday My Prince Will Come	290	83	0.286207
Whisper Not	Groovin High	290	71	0.244828

The average-percentage-of-phrase-per-measure metric in the above table captures how much of a phrase can be found, on average, within single measure. This means, for example, that on average, 49% of a phrase will occur in each measure in the *Still Live* version of Autumn Leaves, or alternatively, two measures are required on average to accommodate a single phrase in this solo. In the *Tokyo 96* version of Autumn Leaves, the 32% suggests that, on average, three measure are required to accommodate a single phrase in the solo.

The difference in Autumn Leaves phrase lengths might suggest that Jarrett's more recent solos have longer melodic phrase lengths. However, when considering the other solos in the dataset at least, this does not seem the case. In fact, the dataset shows that phrase length has no bearing on the solo in which it is played. In some solos, phrases will take place over four measures, in other solos phrases will take place of two measures, or three measures. This suggests that phrase length is a mechanism through which Jarrett creates variation in his playing. Furthermore, both short and long phrases occur regardless of tempo and time signature.

Figure 5.21 displays the different phrase lengths used across the entire dataset. The bulk of the phrases are less than 40 notes in lengths, however there are a substantial number of outliers that can be see in the data.

Figure 5.21. Different phrase lengths across all solos

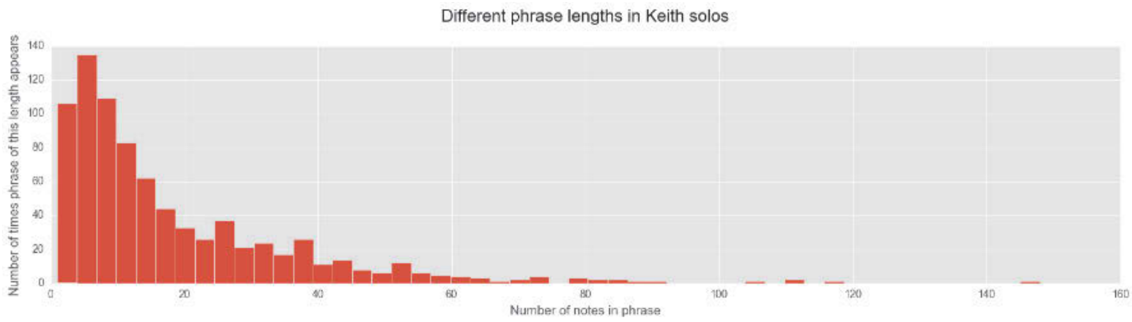


Table 5.10 provides a more nuanced view of the 813 melodic phrases found in the dataset. It shows that phrases have an average of 18 notes each, and range from one note in length to 148 notes in length.

Table 5.10. General characteristics of phrase length across solos

Number of phrases	813.000000
Average number of notes per phrase	18.055351
Standard deviation of phrases	18.370118
Minimum number of notes appearing in a phrase in	1.000000
First Quartile	6.000000
Second Quartile	12.000000
Third Quartile	25.000000
Fourth Quartile	148.000000

The high average phrase length seen in table 5.10 above, along with the high standard deviation is driven predominantly by the outlier melodic phrase lengths (i.e. those melodic phrases with more than fifty notes in length). The majority of phrases, however, are markedly shorter, and 75% of all melodic phrase lengths do not exceed 25 notes in length.

Melodic phrases with 12 notes or less account for almost half of all the examples and these can be seen in Table 5.11. The most common phrase count are four-note melodic phrases in the dataset. Later in the chapter, when examining four-note microphrases it will be seen that four note structures are a critical building block for Jarrett's improvisations.

Table 5.11. Short phrase lengths in the dataset

Count of notes in phrase	**Count of occurrences across entire dataset**
1	33
2	43
3	30
4	57
5	39
6	39
7	43
8	33
9	33
12	31

The ten top melodic phrase lengths are all unique, ranging between 55 notes and 148 notes. The longest phrase in the entire dataset is located in measure 95, of the solo, In Love In Vain, and is rendered below, in full, in Figure 5.22.

Figure 5.22 Phrase excerpt

Title	In Love In Vain
Performer collection	Standards, Vol. 2
Measure in which phrase begins	2.5
Measure location in which phrase begins	95

When examining the individual solos and melodic phrase lengths, different melodic phrase length profiles being to emerge. <u>All The Things You Are</u> and <u>Groovin High</u> have similar upbeat tempos (at 247 bpm and 289 bpm respectively), yet markedly different phrase length profiles, which can be seen in Figures 5.23 and 5.24.

Figure 5.23 Different phrase lengths across all solos

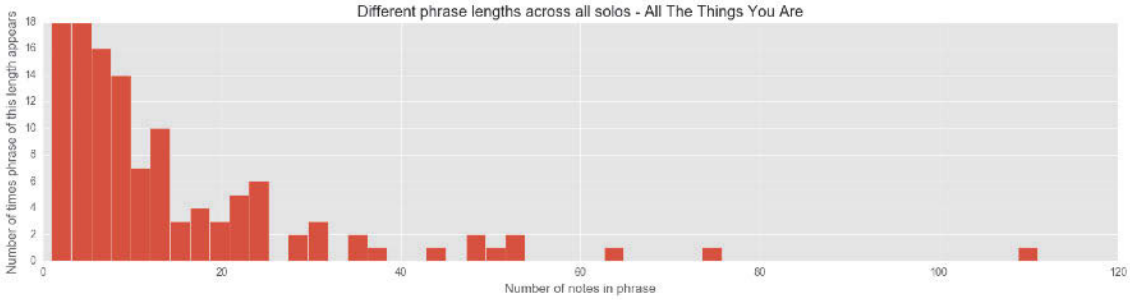
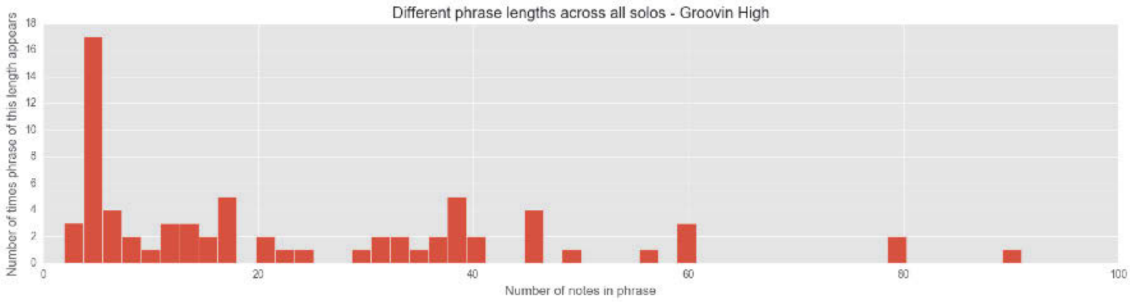


Figure 5.24 Different phrase lengths across all solos



It may again be that this difference is be related to the subtleties of genre. Groovin High is a bebop standard, as opposed to All The Things You Are which is a common standard seen in the American songbook and, with access to more data, it would be possible to test this theory.

Figures 5.25 and 5.26 show the phrase length profiles of the Stella By Starlight and Someday My Prince Will Come (with 151 bpm and 148 bpm respectively) which again, are significantly different.

Figure 5.25 Different phrase lengths across all solos

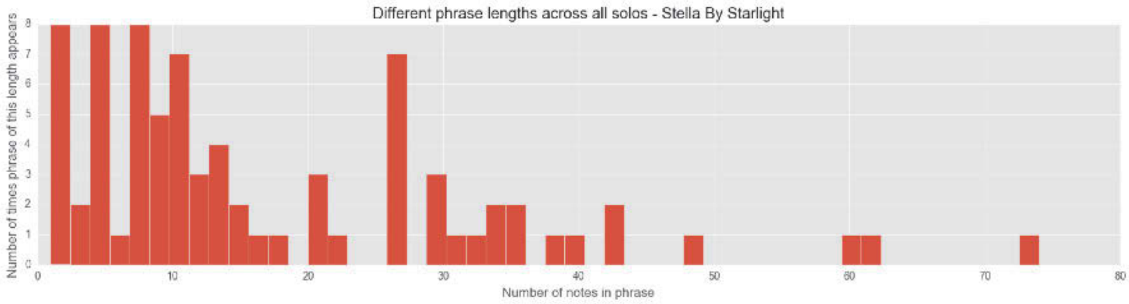
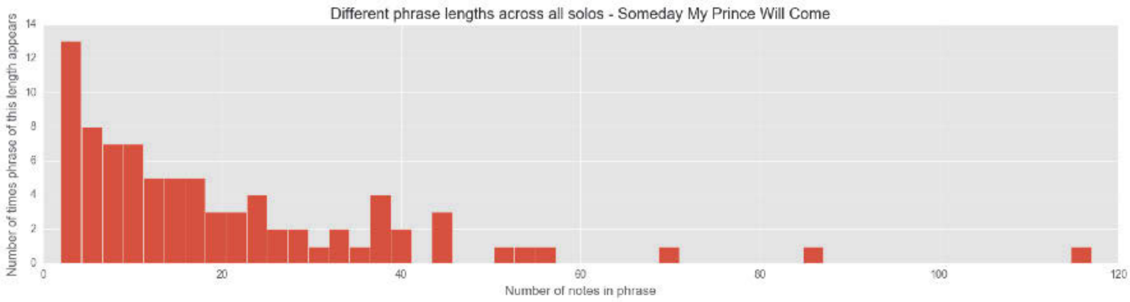


Figure 5.26 Different phrase lengths across all solos



Across the the four different profiles, there is a small similarity can be drawn between Someday My Prince Will Come and All The Things You Are. Interestingly these compositions were written two years apart (in 1937 and 1939) and the songs have highly structured, flowing, melodies, which may suggest that phrase length is related by the phrasing of the melody of the song. This dataset does not include the melodies of the jazz standards under consideration, but with that additional data the question would become straightforward to explore.

Overall however, it is clear that Jarrett's ability to produce different phrase lengths is a principal way through which variation can be created in the course of an improvisation. While the choice of rhythmic subdivisions and note range is characterised by severe limitation, and the amount of notes being played is heavily influenced by tempo, the phrase length appears to be independent of any other factors and a principal way by which repetition is avoided.

Figure 5.27 below provides a visualisation of all the phrase lengths across all the solos. The number of notes in the phrase is plotted on the y-axis and the measure that the phrase commences on the x-axis. Overall this reveals a tendency toward balance where short phrases are contrasted with long phrases.

Figure 5.27. Number of notes in phrase vs. commencing measure

Table 5.12 provides additional details for phrases over 80 notes in length, the very longest phrases in the dataset, including both the measure and measure location in which they begin.

Table 5.12 Phrases over 80 notes in length and commencing measure

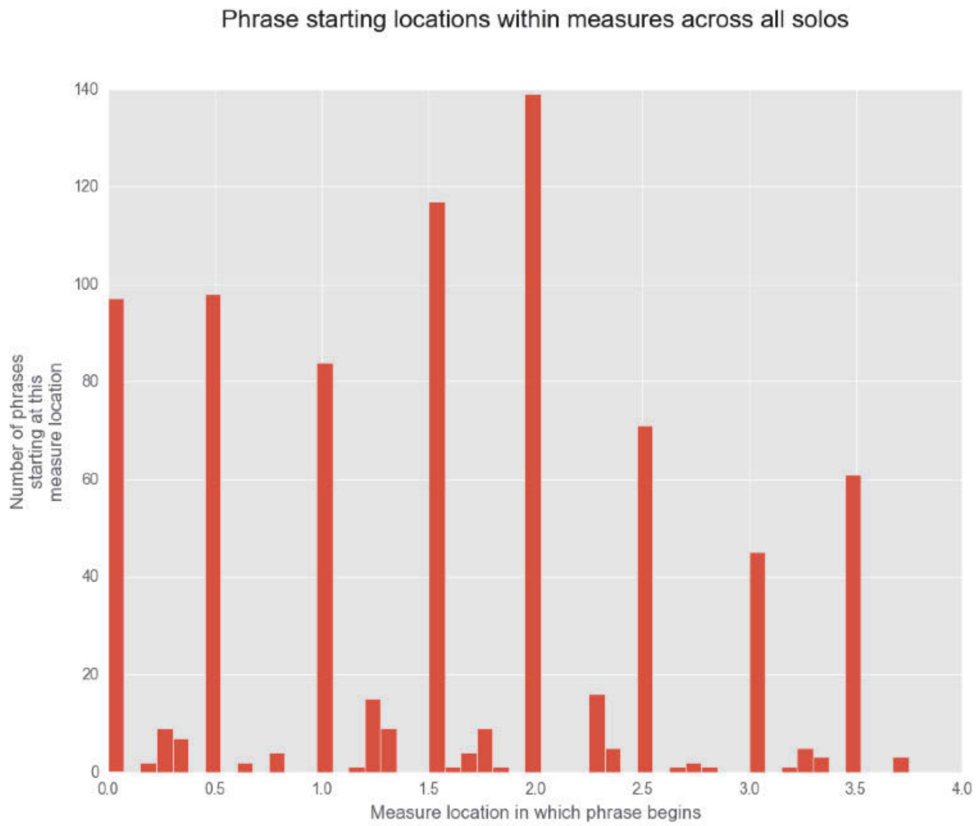
Performer collection	Title	Measure in which phrase begins	Measure location in which phrase begins
Keith Jarrett At The Blue Note, The Complete Recordings (Vol. 3)	Days Of Wine And Roses	43	3.500000
Keith Jarrett At The Blue Note, The Complete Recordings (Vol. 3)	Days Of Wine And Roses	69	2.000000
Keith Jarrett At The Blue Note, The Complete Recordings (Vol. 3)	Days Of Wine And Roses	142	1.000000
Standards, Vol. 1	All The Things You Are	158	2.500000
Standards, Vol. 2	In Love In Vain	38	1.333333
Standards, Vol. 2	In Love In Vain	95	1.500000
Still Live	My Funny Valentine	103	0.000000
Up For It	If I Were A Bell	199	0.000000
Up For It	Someday My Prince Will Come	153	1.000000
Up For It	Someday My Prince Will Come	216	2.250000
Whisper Not	Groovin High	72	3.000000

The data demonstrates Jarrett's slight tendency to play longer phrases in mid-tempo solos rather than the uptempo solos. This appears related to the fact that Jarrett is more likely to use smaller subdivisions of the beat at mid-tempos, such as sixteenth and even thirty-second notes. This, in turn, tends to increase the overall note count in a given melodic phrase.

Overall however, there are no strong patterns to be discerned here. Jarrett commences long melodic phrases on all parts of the measure, and at varying locations within a given solo, which helps to inform the way in which phrase structure can confound repetition.

Although longer phrases (i.e. those over eighty note in length) seem to show no overt tendencies in terms of where they start in the measure, it does raise the question of starting and ending points of melodic phrases, and it is possible to observe these in the dataset. Figure 5.28 shows the starting locations within the measure for all phrases across the entire dataset.

Figure 5.28[4]. Phrase starting locations within measures across all solos

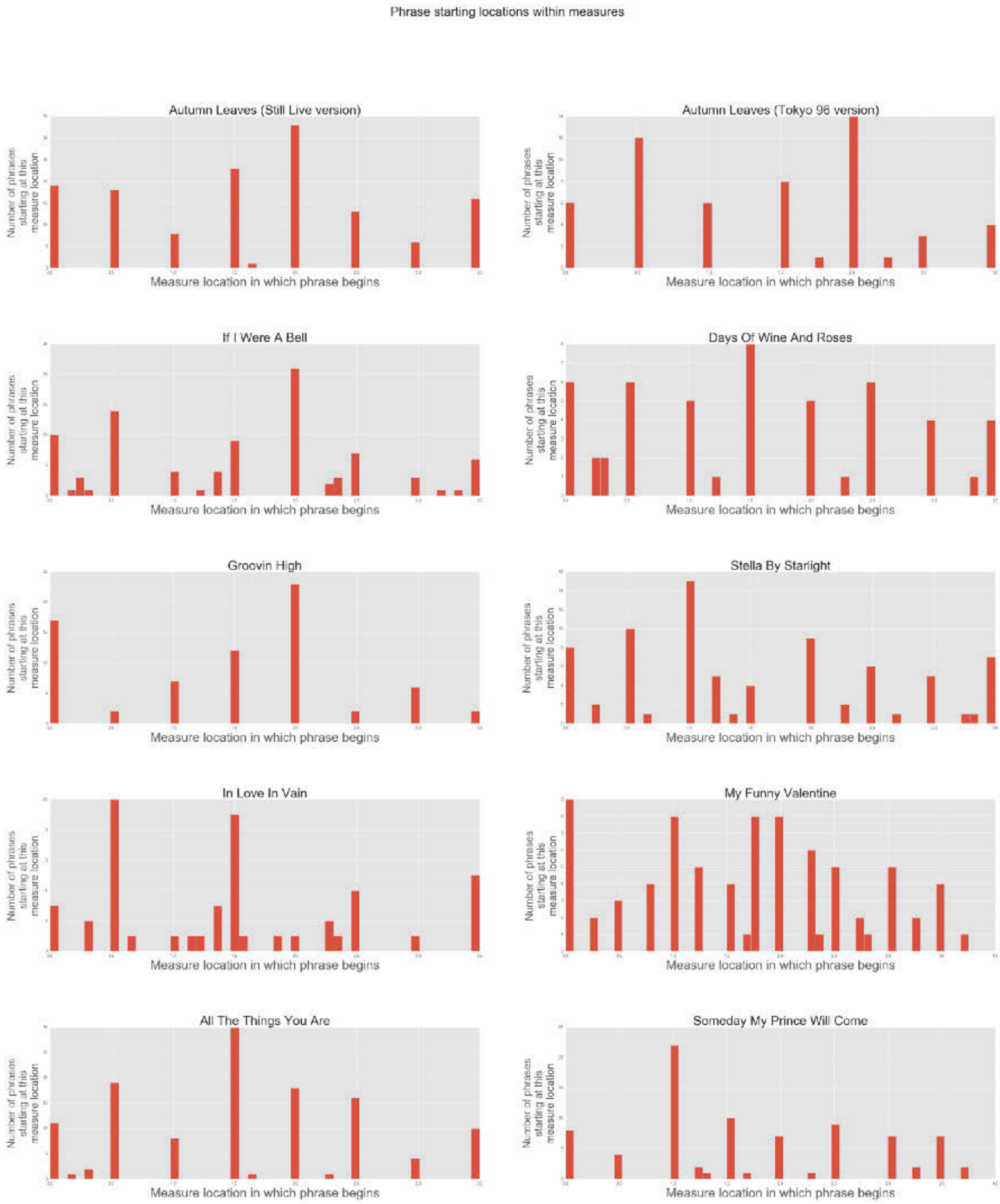


The figure shows an overwhelming tendency to start improvised phrases on eighth note subdivisions. Note also, the higher tendency to start a phrase at position 2.0 in the bar (or the third beat in the bar).

[4] Note that the x-axis on figure 30 starts at 0.0. This position should be equated to the first beat of the bar. Musical time is counted from 1 onwards rather than 0, meaning a bar of 4/4 starts from 1 and finishes at the beginning of 5, or 1st beat of the next bar. Numerically however, the bar starts at 0 and completes at 4, which is the convention that has been adopted here.

Figure 5.29 displays this data again, this time at the level of each individual solo, showing the location of where phrases commence.

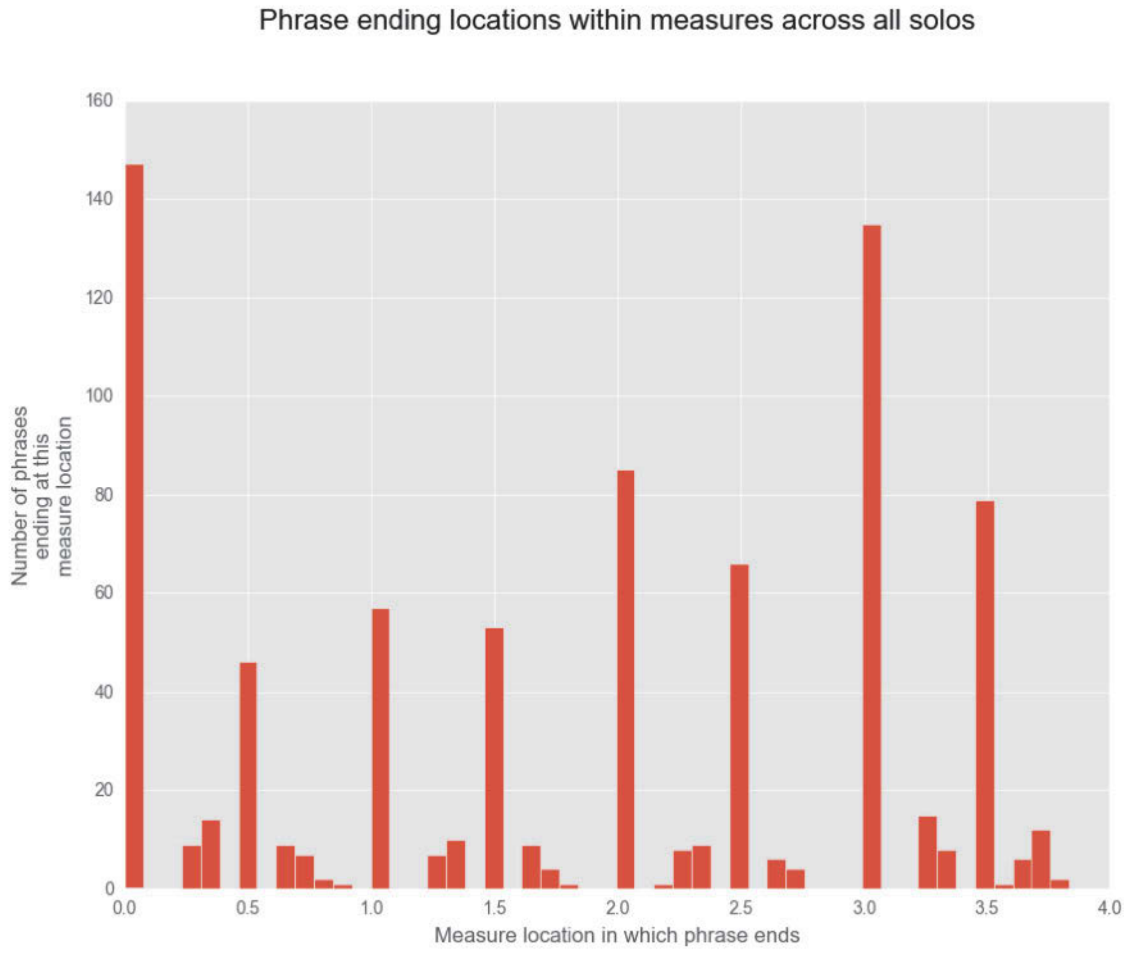
Figure 5.29. Phrase starting locations within measures in each solo



The influence of tempo can again be seen at work here. In the higher tempo solos, such as Groovin High and All the Things You Are, the phrase starting points are far more limited. In the slower tempo phrases, such as My Funny Valentine, there are far more starting positions in the measure at which phrases begin.

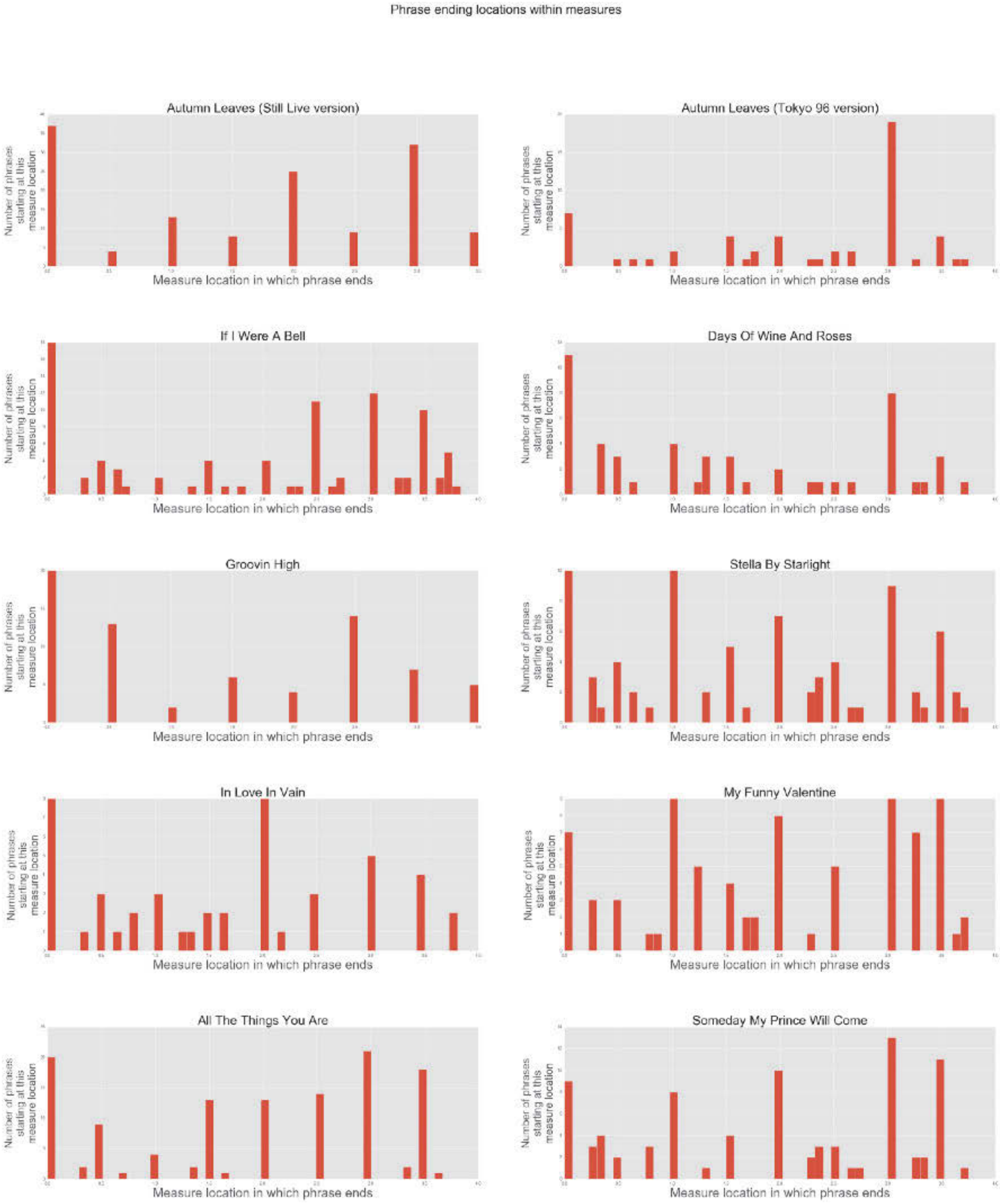
It is also possible to explore the end points of phrases, with a view to understanding if they behave in the same manner, and Figure 5.30 shows the ending locations of all phrases in the dataset.

Figure 5.30. Phrase ending location within measures across all solos



Overall, the locations of phrase endings seems to be more varied than phrase beginnings. Figure 5.31 provides provides the breakdown all phrase ending positions for each of the solos, demonstrating that phase endings tends to be more varied at the solo level. The positioning of phrases, similar to what was discovered about their length, appears to be a principle way through which Jarrett avoids repetition in his playing.

Figure 5.31. Phrase ending location within measures across for each solo



I have not yet explored the internal structure of melodic phrases, and will turn to this now. Figures 5.32 and 5.33 display two typical phrases from the dataset. The first is taken from the mid-tempo solo Days Of Wine And Roses, and the other is from the up-tempo solo Groovin High. The phrases have both marked similarities and differences. Figure 5.32 shows far more rhythmic variety, with the use of eighth notes, triplets, and sixteenth notes. Figure 5.33, though of similar length, is made up almost exclusively of eighth notes. Tempo is again important here, with the first excerpt at a slower bpm having far more rhythmic subdivisions than the second excerpt.

Figure 5.32. Phrase excerpt

Title	Days Of Wine And Roses
Performer collection	
Measure in which phrase begins	
Measure location in which phrase begins	43

Figure 5.33. Phrase excerpt

Title	Groovin High
Performer collection	Whisper not
Measure in which phrase begins	
Measure location in which phrase begins	72

Despite the differences, these phrases still seem somehow similar, and its these similarities that appear to be typical features of Jarrett's overall style. Both phrases can be characterised by an overt use of stepwise movement (i.e subsequent notes being no more than a one tone away in pitch distance). Additionally, when leaps are used, they tend to be in thirds, and follow seventh chord patterns (for example the notes E4, G4, Bb4 and D5 in Figure 5.32, measure five and the notes G4, Bb4, D5, and F5 in the Figure 5.33, measure seven). Finally, although the range of the phrases is limited, they both use all 12 available notes of the octave multiple times, problematising the notion that these phrases could be discussed in terms of the use of particular scales.

So far I have presented Jarrett's ability to avoid repetition at the level of the melodic phrase. However it is also possible to explore other notions of repetition. It is possible to examine if, in the course of playing a particular melodic phrase, particular notes themselves are repeated. The phrases in Figure 5.32 and Figure 5.33 above suggest that Jarrett is comfortable moving between all notes of the octave and repeated notes in phrases are more unlikely, and it is possible to test this.

As an example of what this type of exploration might look like, Figure 5.34 shows a phrase taken from the *Tokyo 96* version of Autumn Leaves. Here, some notes appear only once (such as the note D4) and other notes appear numerous times, such as Eb6, E6 and Eb4. In contrast, in Figure 5.35, the phrase from My Funny Valentine shows a four note phrase in which all notes are appear only once, and none are repeated.

Figure 5.34. Phrase excerpt

Title	Autumn Leaves
Performer collection	Tokyo 96
Measure in which phrase begins	21
Measure location in which phrase begins	0.5

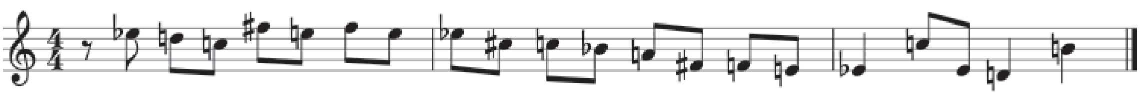


Figure 5.35. Phrase excerpt

Title	My Funny Valentine
Performer collection	Still Live
Measure in which phrase begins	22

Measure location in which phrase begins	3

Examining repetition in this way shows that, on average, 66% of pitches used in a phrase are unique (in that they appear only once in the phrase) and the median percentage of unique pitches is 63%. This means that when Jarrett improvises a phrase in any solo, it can be expected that around 66% of the notes in the phrase will only appear once, and the rest will appear two or more times.

It is also possible to correlate the number of phrases in the dataset with how much of their content is unique (i.e how much of their content is non-repeated notes), and this can be seen in Figure 5.36. A fairly large outlier can be seen in this figure, indicating that in 162 phrases, each pitch will only appear once.

Figure 5.36. Percentage of unique musical frequencies used in phrase in solos

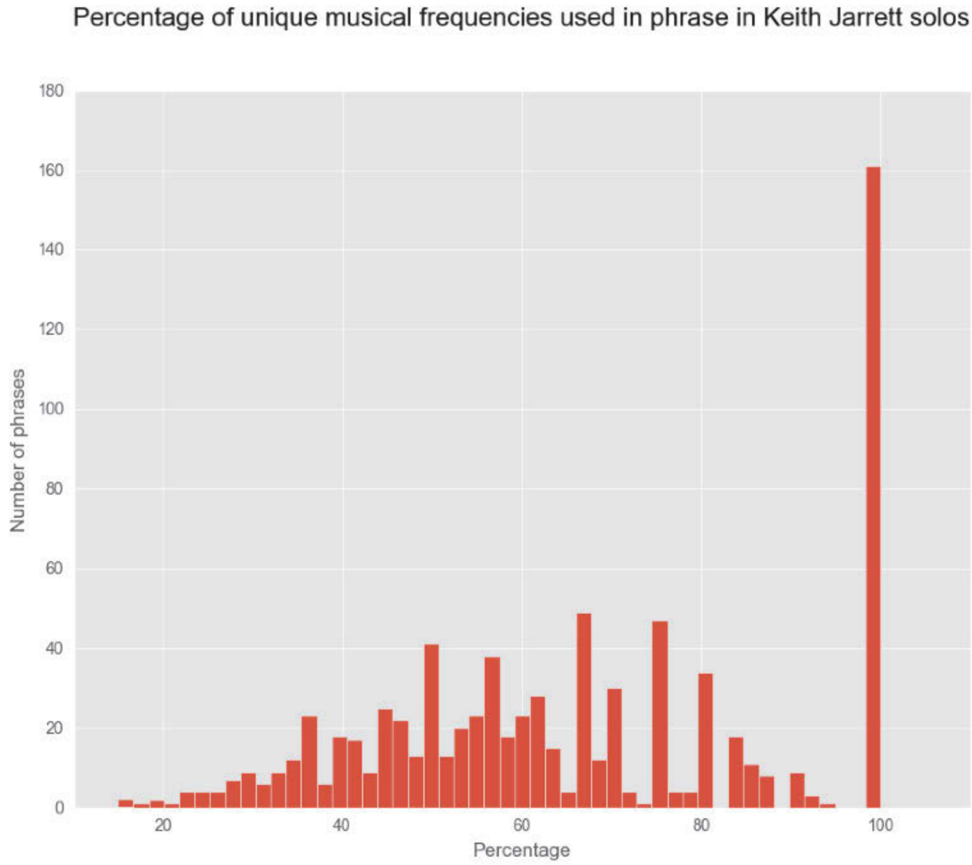


Figure 5.37 explores this idea further by narrowing the criteria to only consider those phrases which have no repeated notes (being the 162 phrases with 100% uniqueness seen in Figure 5.36). When these are examined it shows that uniqueness is mostly related to the length of the phrase: the longer a phrase is, the more likely it seems that particular notes will be repeated.

The boxplot visualisation provided Figure 5.37 provides more information about the length of phrases that have 100% uniqueness. The purple line in the graph below represents the median, showing that these phrases are predominantly only three notes in length. It can also be seen that the middle 50% of all phrases with 100% uniqueness (seen in the orange rectangle) are between two and four notes in length.

Figure 5.37. Count of notes in phrase where all pitches are unique

Yet there are outliers that can be seen in the data too, which are indicated by the '+' symbol on the above figure. This shows that although most phrases with no repeated notes are short, there are exceptions. In particular, there is one phrase in the dataset comprised of 12 notes, all of which are unique in pitch. This is shown in Figure 5.38.

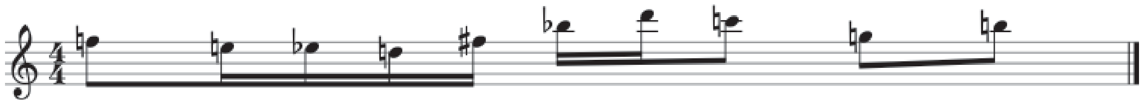
Figure 5.38. Phrase excerpt

Title	Autumn Leaves
Performer collection	Still Live
Measure in which phrase begins	96
Measure location in which phrase begins	2

Figure 38 also highlights the problem of trying to understand Jarrett's improvisational style by appealing to scales. There is a substantial amount of chromatic movement happening here (notes moving the distance of semitones) but there are also constant shifts in tones and minor thirds. The above passage also takes place over A min7b5 followed by and D dominant chord and although a number of notes in the phrase relate to both chords, it could be argued that they just as easily relate to other chords. Much of what is happening in melodic phrase above might be better explained by examining how Jarrett uses voice-leading and handles the preparation and resolution of different notes in different harmonic contexts, which I will examine later in the case study. A second outlier, having a phrase length of ten, can be seen in Figure 5.39.

Figure 5.39. Phrase excerpt

Title	My Funny Valentine
Performer collection	Still Live
Measure in which phrase begins	69
Measure location in which phrase begins	0



The above ten note phrase is played over two different chords, which split the bar, the II-V progression being D minor 7b5 and G dominant 7. Again, this highlights how conceptualising Jarrett in terms of appealing to scales is a daunting task: the E note and F# note are problematic in terms of the A minor 7b5, as is the presence of the eleventh (C6) on the G dominant 7.

Of particular interest in the above example is the four notes that appear, being D5, F#5, Bb5, D6, C6, from the fourth note onwards into the phrase. These four notes can be considered as a D dominant 7#5 chord which could suggest that

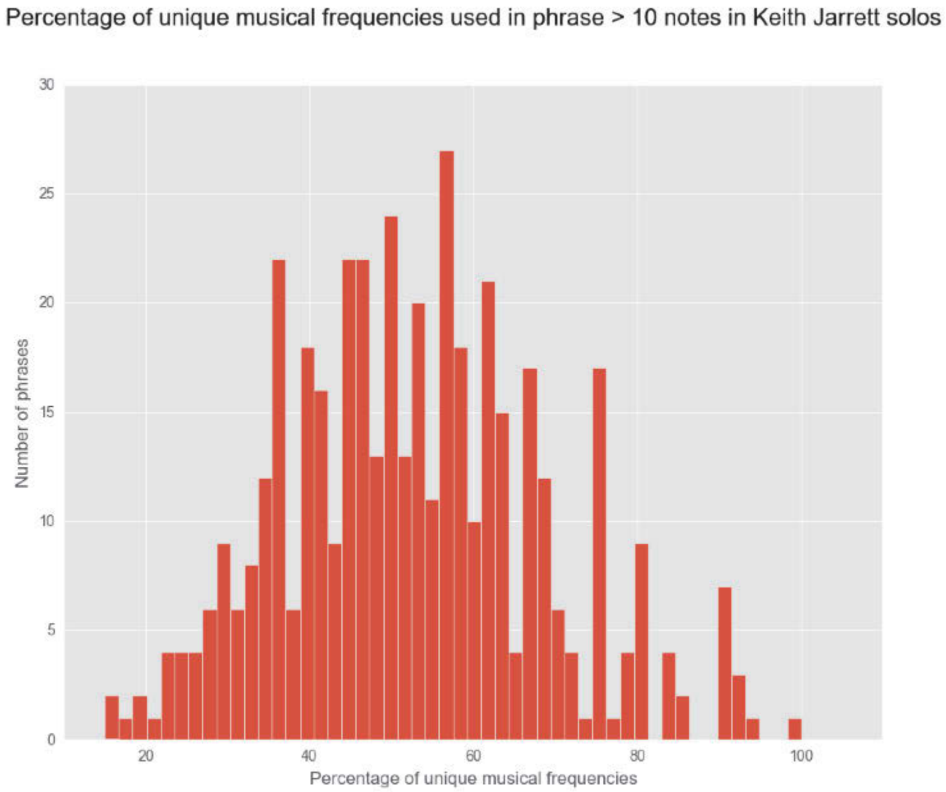
Jarrett is improvising over a reharmonisation, being D minor 7 b5 in the first beat of the bar, then over a D dominant 7#5 chord, and then, in the last beat of the bar, improvising over a G dominant 7.

I would argue here that Jarrett is certainly not setting out to consciously overlay a complicated reharmonisation during the course of an improvised phrase. Also, the appearance of the F# note above, can be seen across Jarrett's playing: II minor chords are very often interpreted as II dominant chords. I will explore this problem in the context of harmony and voice leading later in the chapter as it quickly becomes extremely complicated to view this as a problem of melody.

As a final word on the above example, it highlights how amenable this dataset is to exploratory analysis. In seeking to explore phrases with no repeated pitches, I have discovered a four note pattern which could infer that a D dominant 7#5 is being used in a D minor 7b5 G dominant 7 progression. This can lead to the question of where else this might be happening in the dataset. It is straightforward to extract all the instances of a II minor 7b5 - V dominant 7 progression, and examine any appearance of a super imposed D dominant 7#5. It is also easy to examine if this is limited to certain keys, certain tempos, and even certain geolocations or years in which the solos were played.

Most phrases in this dataset have a degree of repetition. If they did not, Jarrett's playing would most likely resemble a series of 12 tone rows. Further, it is evident from Figure 5.37 above that phrases without repeated pitches tend to have a far shorter length. Figure 5.40 below shows only those phrases that have ten or more notes in length, displaying the percentage of unique pitches in them. To be expected, longer phrases, use repeated pitches in varying degrees.

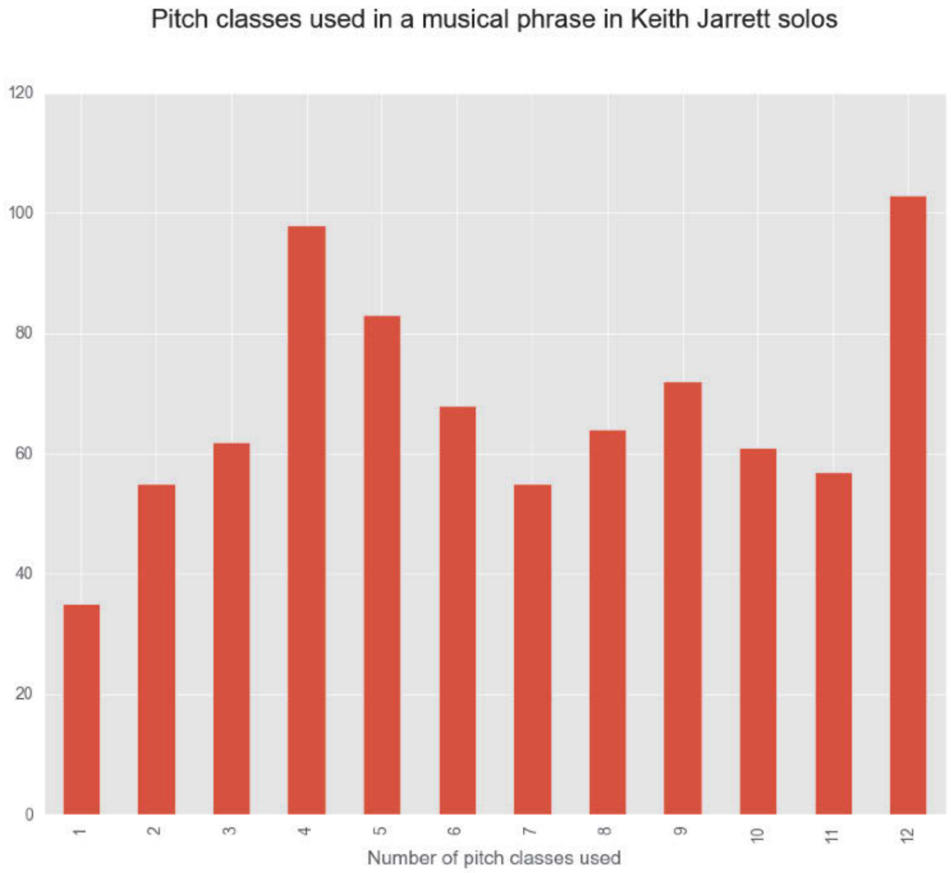
Figure 5.40. Percentage of unique musical frequencies in phrases greater than ten notes



Another way to approach this problem would be to explore whether, regardless of the exact pitches that might be repeated, if all of the pitch classes of the octave tend to appear in any given phrase. The dataset tells us that most phrases have some repetition of particular pitches, yet despite this it appears that many phrases will still use all 12 pitch classes.

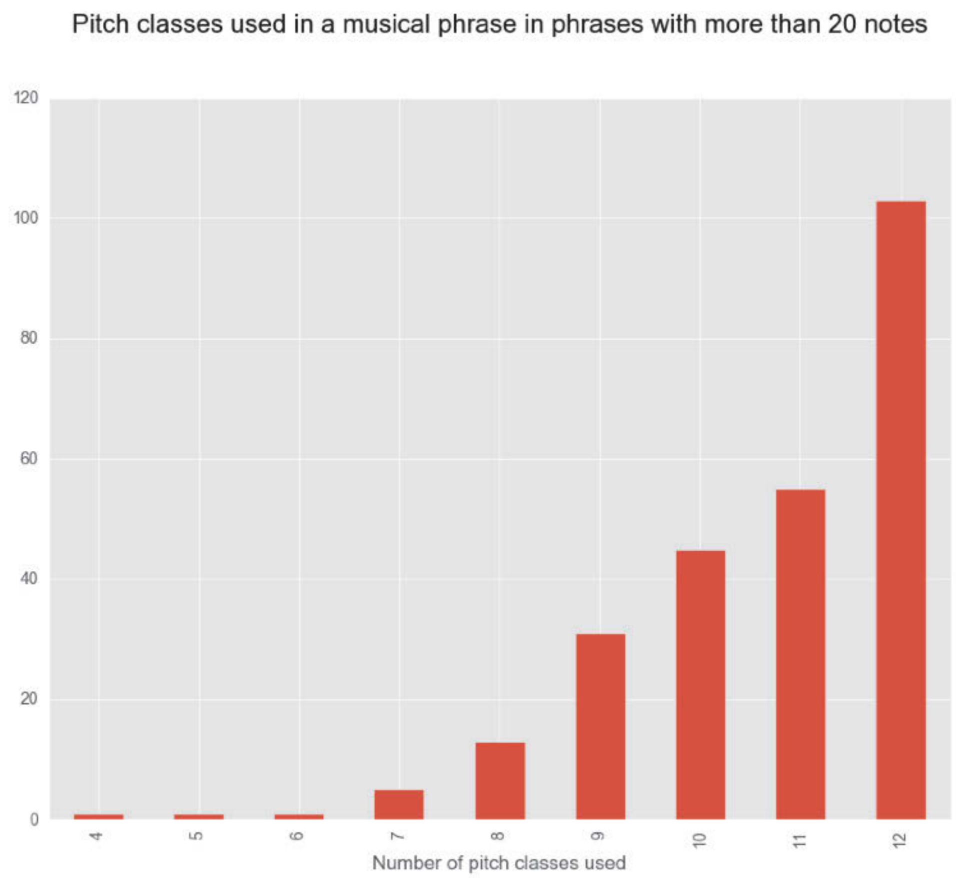
Figure 5.41 provides an initial visualisation of this idea. It shows that over 100 phrases in the data set use all 12 pitch classes (being represented by the column in Figure 5.41 on the far right). On the far left, it can be seen almost forty phrases use only 1 pitch class.

Figure 5.41.   Pitch classes used in melodic phrases in all solos



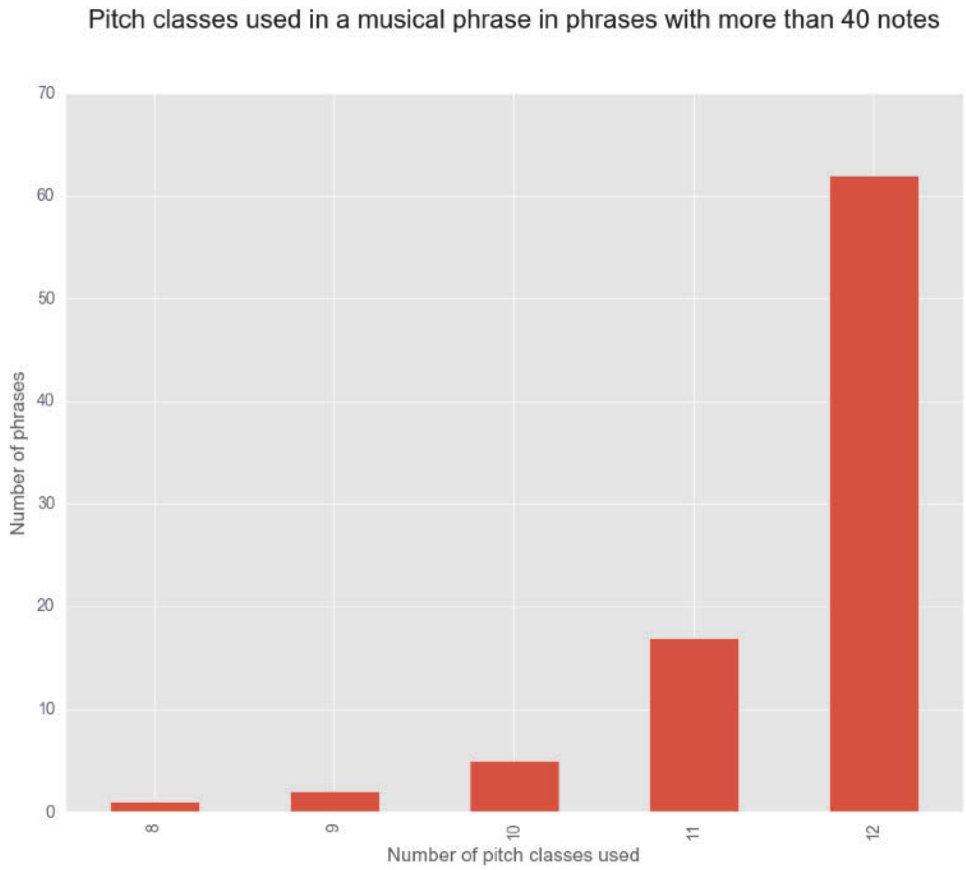
Intuitively, it would seem the use of all pitch classes would be more likely as the phrase becomes longer. Figure 5.42 explores this idea, presenting the same information as Figure 5.41 by filtering the data so only phrases greater than 20 notes in length can be considered.

Figure 5.42. Pitch classes used in melodic phrases in phrases with more than 20 notes



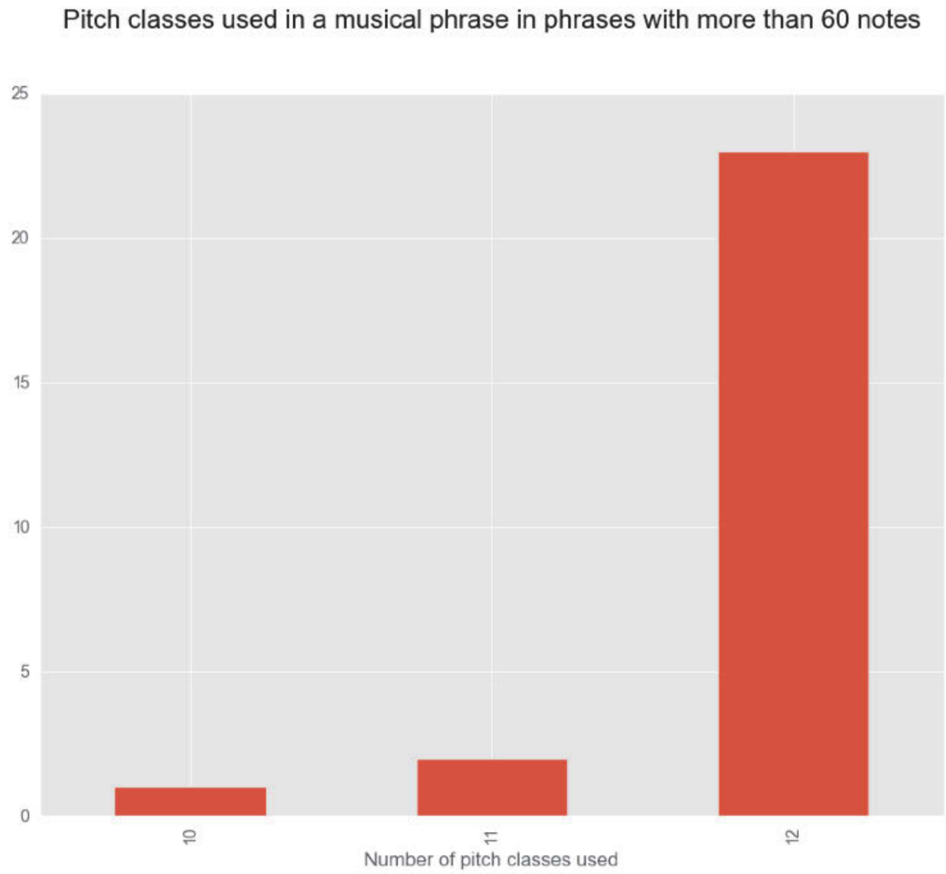
This shows that the more notes in a phrase, the greater the tendency to use all pitch classes. Of all the phrases that are in the dataset which greater than 20 notes in length, over a third will use all pitch classes. Figure 5.43 limits phrases lengths again, this time to those over 40 notes in length.

Figure 5.43. Pitch classes used in melodic phrases in phrases with more than 40 notes



This shows that, in the majority of phrases that have 40 notes or more, all pitch classes are used. To be expected, when the data is filtered to consider only phrases with 60 notes or more, the vast majority use all pitch classes. This can be seen in Figure 5.44.

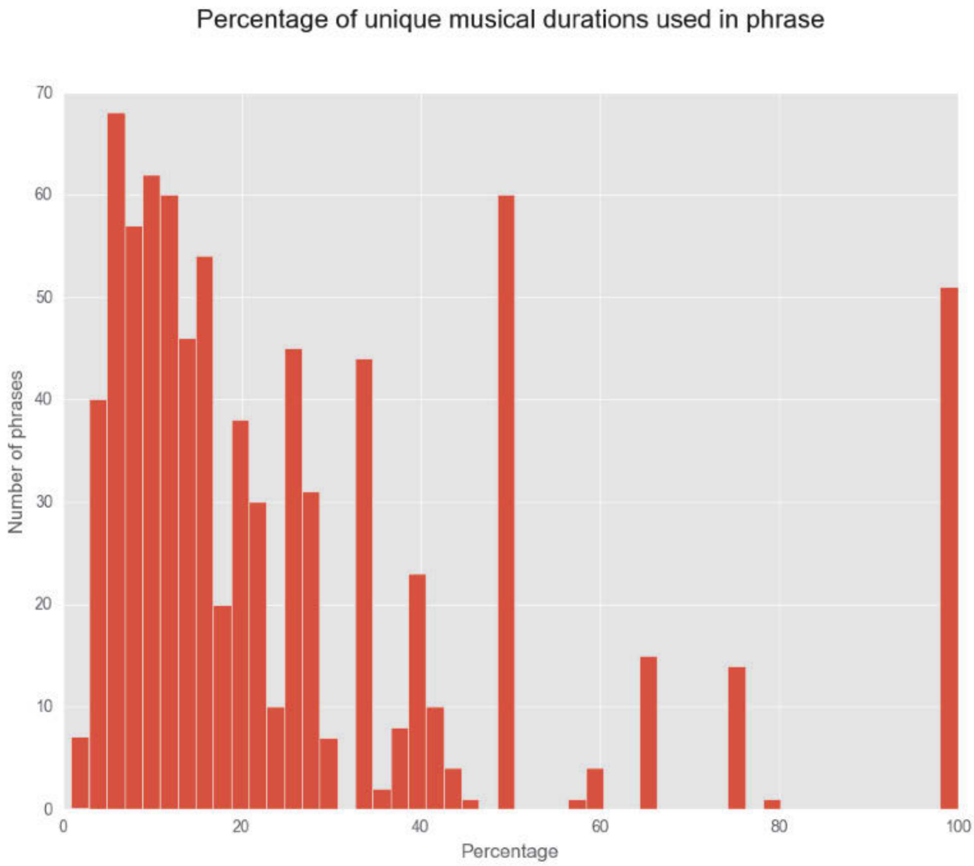
Figure 5.44. Pitch classes used in melodic phrases in phrases with more than 60 notes



Again, this highlights the problem of locating Jarrett's improvisations in any kind of framework related to scales. Jarrett appears to be modulating rapidly through different superimposed tonalities and utilising fairly traditional voice-leading techniques to do this. This idea will be explored further when examining harmony and voice leading later in the chapter.

Turning to the way in which note durations are used in phases, Figure 5.45 correlates the number of phrases with the percentage of unique rhythmic durations used.

Figure 5.45. Percentage of unique musical durations used in phrase



In contrast to how pitch and pitch class are used in the phrases, Figure 5.45 shows that the use of different note durations is severely limited, with most phrases being below 30% in note duration uniqueness. This means that, if a phrase were to contain ten notes, only 3 different note duration types would be employed. Furthermore, there are only 52 phrases across the entire corpus (and the majority of these are very small in length) in which there are only unique durations. Thus, while pitch class choice can be highly varied, especially for long phrases, rhythmic choice is consistently severely limited, and more so as the phrase becomes longer.

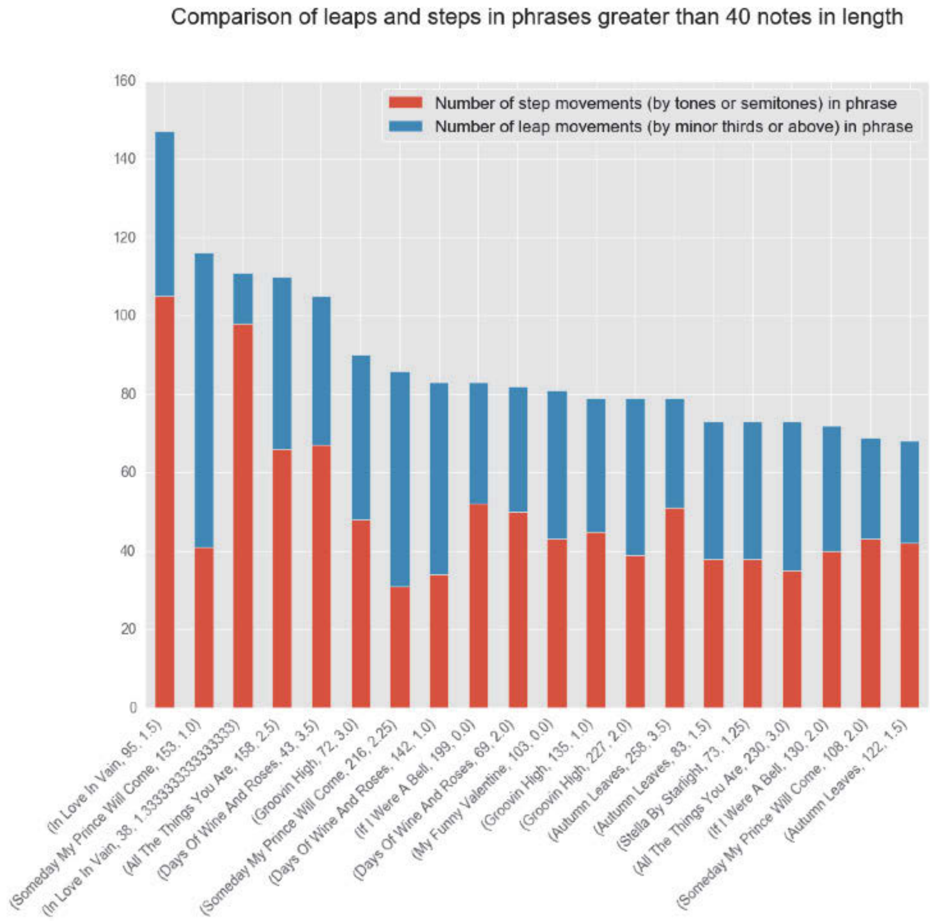
I want to return to the phrase from Figure 5.22 (here recreated in Figure 5.46 for convenience), to examine its structure more closely. It is clear that the notes are not being chosen in a random way and there is a balance between upward and down movement, and stepwise movement and leaps.

Figure 5.46. Phrase excerpt

Title	Days Of Wine And Roses
Performer collection	Keith Jarrett At The Blue Note, The Complete Recordings (Vol. 3)
Measure in which phrase begins	43
Measure location in which phrase begins	3.5

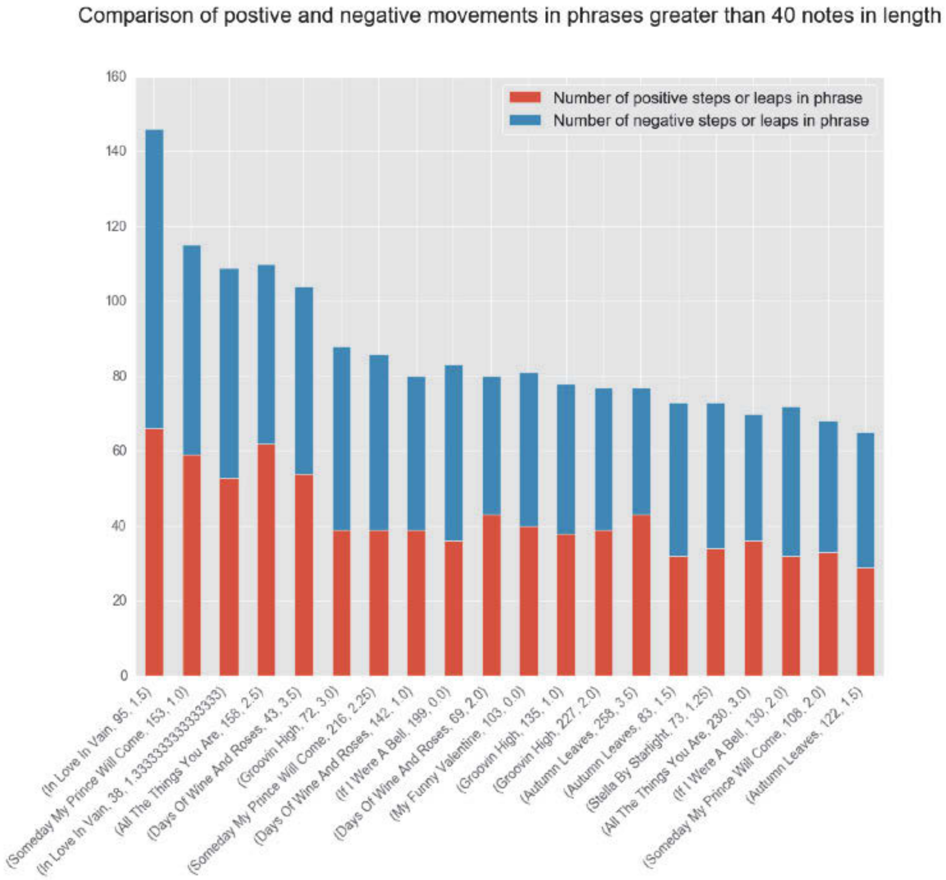
It is possible to place metrics around this, and examine the dataset in terms of how the phrases tend to be contoured. Figure 5.47 below displays all those phrases in the dataset that are greater than 65 notes in length, examining the number of step movements (being movements of a tone or less) in the phrase, as opposed to the number of leap movements (being movements greater than a tone in distance).

Figure 5.47. Comparison of leaps and steps in phrases greater than 40 notes in length



This figure shows that the kind of behaviour seen in Figure 5.46 is widespread throughout the data. In the majority of phrases, there is a constant interplay between step movements and leap movements. Figure 5.48 again displays phrases over 65 notes in length, and this time shows the number of positive movements (being either ascending steps or ascending leaps) opposed to the number of negative movements (being either descending steps or leaps).

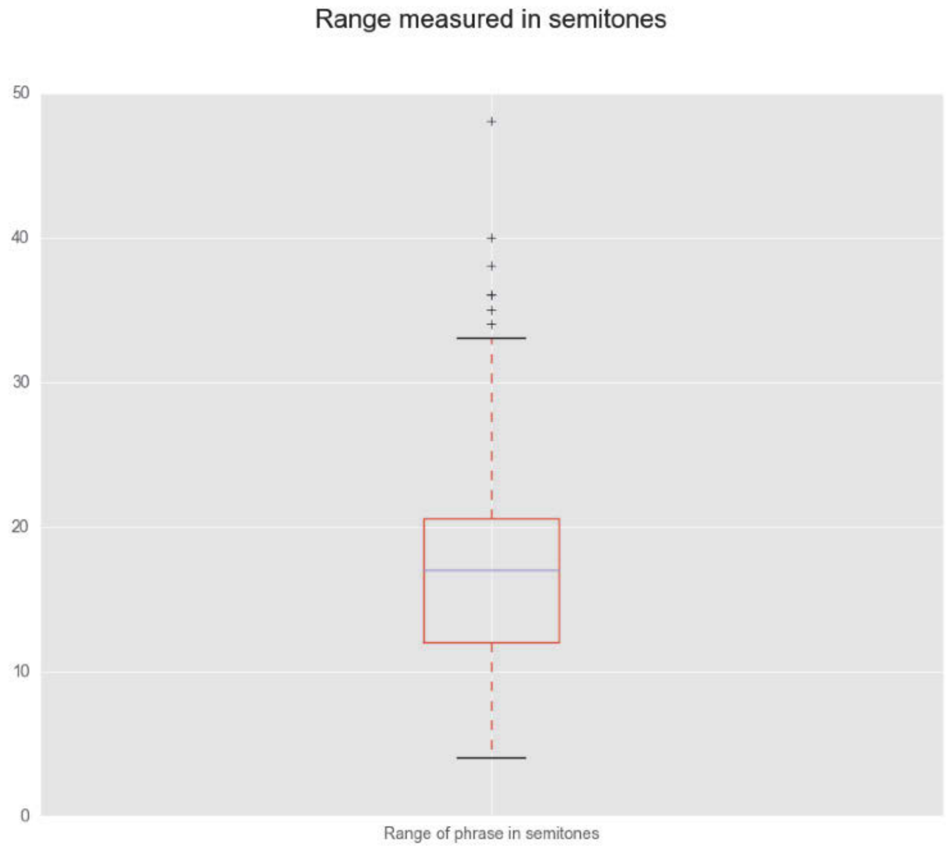
Figure 5.48. Comparison of positive and negative movements in phrases greater than 40 notes in length



Overall, examining the contours of the phrases in this way allows a picture to emerge whereby these phrases are characterised, above all, by balance: upward movement is moderated by downward movement; leaps are moderated by steps.

It is also possible to examine the phrase contour in terms of how range is utilised. The average pitch range found in phrases across the data set is 17.2 semitones, (and the median is 17). This means that, on average, the lowest note of a phrase will only be an octave and half below the highest note. Figure 5.49 displays a boxplot showing how range is used across all the phrases in the dataset. This shows that, in the majority of phrases there is a range between 13 semitones and 21 semitones.

Figure 5.49. Range measured in semitones



Examining phrases at such a high level can provide powerful metrics which can be applied to any set of transcriptions. In the case of Jarrett, it highlights that that the phrase is something that is characterised above all by balance: balance of upward and downward movement, of steps and leaps, different starting points and ending points in the bar. It is this balance which allows the phrase range to remain limited (in that phrases are constantly changing direction). These metrics can also allow us to place concrete measurements in place to assist our understanding of jazz improvisation works, and can be used as points of measurable differentiation into other datasets.

In order to ask more specific questions that go beyond the size and shape of phrases, it is necessary to start examining the patterns that exist within the melodic phases themselves. To do this, I will explore how microphrases (being a partial section of a phrase) are structured in this dataset. This will provide a far more granular understanding of how phrases are being constructed out of underlying building blocks.

In order to explore microphrases in this dataset, it has again been transformed. The transformation works by extracting all microphrases from a larger melodic phrase. For example, if a melodic phrase has seven notes (denoted as the notes n1, n2, n3, n4, n5, n6, n7), and all three note microphrases are to be extracted, the resulting microphrases would be [n1, n2, n3], [n2, n3, n4], [n3, n4, n5], [n4, n5, n6] and [n5, n6, n7]. Note that a three-note microphrase requires at least three notes in order to be considered for a transformation into a three-note microphrase. This means that during the transformation to create three-note microphrases, phrases of length two or less cannot be considered. It is possible to explore any length microphrase in the dataset, however the lengths that will be considered in this case study will be those between two notes and eight notes.

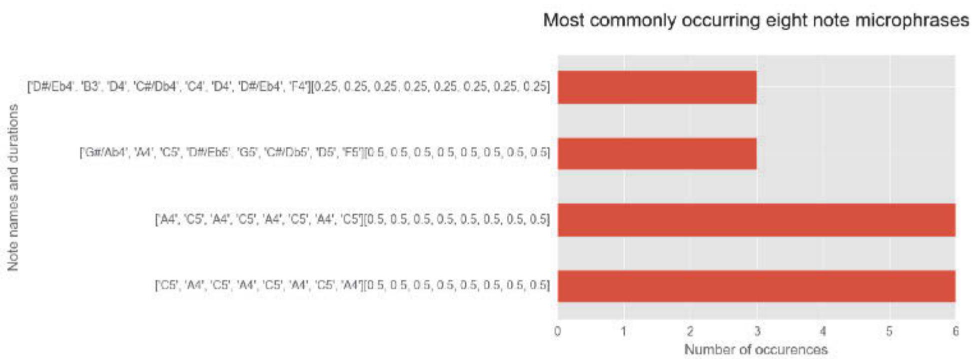
Once the dataset has been transformed, it is possible to count the resulting instances of microphrases, and this can be seen in Table 5.13 below. The table shows that it is possible to construct 9809 eight-note microphrases, through to 13866 two-note microphrases.

Table 5.13. Count of different length microphrases that can be constructed from the dataset

Count of 8 note microphrase	9809
Count of 7 note microphrase	10381
Count of 6 note microphrase	10992
Count of 5 note microphrase	11642
Count of 4 note microphrase	12349
Count of 3 note microphrase	13086
Count of 2 note microphrase	13866

Figure 5.50 displays the four most commonly occurring eight note microphrases. In the process of limiting the melodic phrases via microphrases some (albeit limited) repetition begins to emerge. The figure shows that an identical (in terms of both duration and pitch) eight note length microphrase can be seen in the dataset six times, suggesting that Jarrett, at least on this level, does repeat himself.

Figure 5.50. Most commonly occurring eight-note microphrases



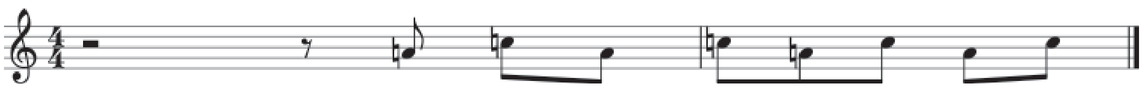
However, on closer inspection of this particular example, it appears to be an outlier in the dataset. The structure that occurs six times does so only because the microphrase is drawn from a larger melodic phrase which is comprised of only two notes. Figures 5.51 and 5.52 show both the examples, being two eight-note microphrases from the Days Of Wine And Roses solo. Note that it would be straight forward to guard against such outliers by filtering out any eight-note microphrases in which there are only two pitch classes.

Figure 5.51. Phrase excerpt

Title	Days Of Wine And Roses
Performer collection	Keith Jarrett At The Blue Note, The Complete Recordings (Vol. 3)
Measure in which microphrase begins	117
Measure location in which microphrase begins	3

Figure 5.52. Phrase excerpt

Title	Days Of Wine And Roses
Performer collection	Keith Jarrett At The Blue Note, The Complete Recordings (Vol. 3)
Measure in which microphrase begins	117
Measure location in which microphrase begins	2.5



Figures 5.53, 5.54 and 5.55 display examples of microphrases which are far more typical in terms of how Jarrett tends to construct melody based on what was seen in the previous section. These examples are really the first substantial instances of repetition that can be seen across the entire dataset. All of the examples are taken from one solo, In Love In Vain. They are played at different parts of the measure, but share the same pitches and durations.

Figure 5.53. Phrase excerpt

Title	In Love In Vain
Performer collection	Standards, Vol. 2
Measure in which microphrase begins	16
Measure location in which microphrase begins	2.5

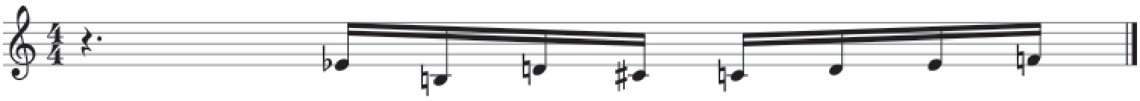


Figure 5.54. Phrase excerpt

Title	In Love In Vain
Performer collection	Standards, Vol. 2
Measure in which microphrase begins	48
Measure location in which microphrase begins	2.25

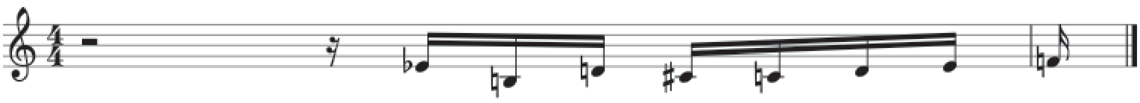
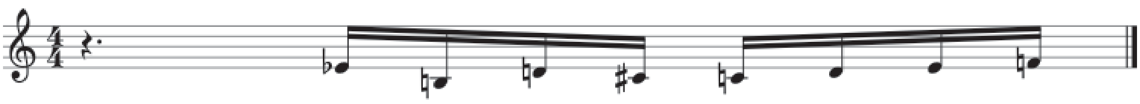


Figure 5.55. Phrase excerpt

Title	In Love In Vain
Performer collection	Standards, Vol. 2
Measure in which microphrase begins	95
Measure location in which microphrase begins	1.5



Of particular interest here is that all three instances of this eight-note microphrase are played over different underlying chords. This suggests that for Jarrett, while there might be a clear correlation between notes in the melody and the notes found in all the chords within a given a jazz standard, the particular underlying chord that he is soloing over at any given time is not as important.

There are three other identical eight-note microphrases that can be seen in the Groovin High solo in Figures 5.56, 5.57 and 5.58.

Figure 5.56. Phrase excerpt

Title	Groovin High
Performer collection	Whisper Not
Measure in which microphrase begins	93
Measure location in which microphrase begins	2

Figure 5.57. Phrase excerpt

Title	Groovin High
Performer collection	Whisper Not
Measure in which microphrase begins	185
Measure location in which microphrase begins	2

Figure 5.58. Phrase excerpt

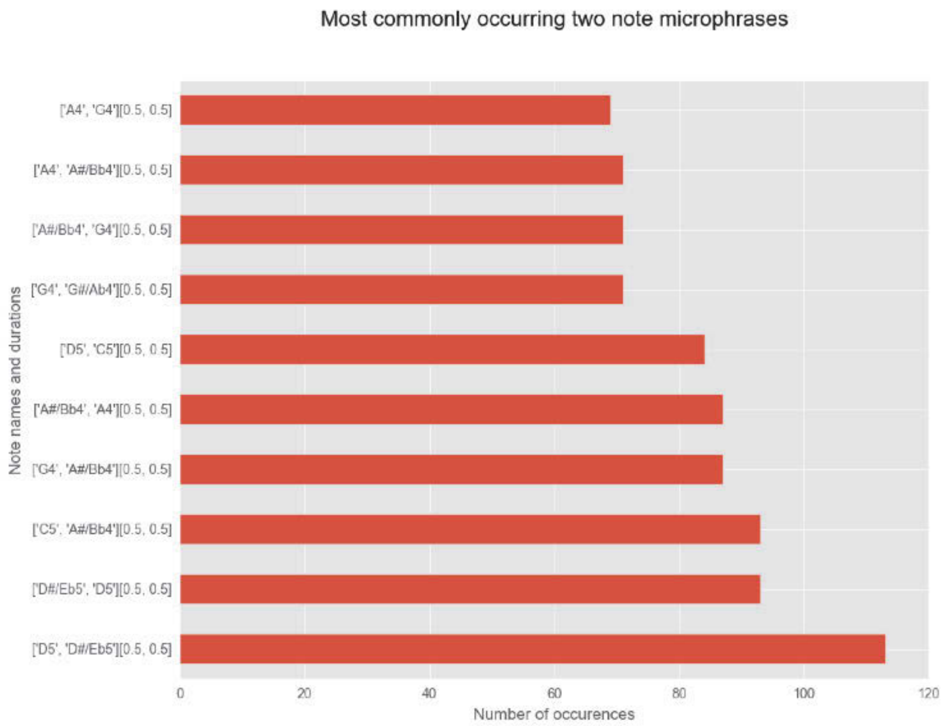
Title	Groovin High
Performer collection	Whisper Not
Measure in which microphrase begins	227
Measure location in which microphrase begins	2

Like the earlier examples, these eight-note microphrases do not occur on the same underlying chords or chord types. This again shows why it is so difficult to locate Jarrett's playing in more typical analytical frameworks often used for jazz.

The phrases that improvisors tend to play most often (or "licks") are usually categorised in the context of specific underlying chord progressions. However, it does not make sense to apply this idea to Jarrett. The repetition that is occurring seems driven not by the specific underlying harmony, but rather the harmony across the whole solo.

At the level of eight-note microphrases, it would still be problematic to characterise this data through any substantial notion of repetition. In order to find further instances of repetition at work, far smaller microphrases need to be examined. Figure 5.59 below shows the ten most commonly occurring two-note microphrases, accounting for both pitch and duration. A comparison from these could be drawn to how an n-gram might function in linguistics: as a basic building block from which larger structures can be derived. When viewed in this way, it appears to behave accordingly. Considering two-note microphrases, a much higher sense of repetition starts to emerge in the dataset.

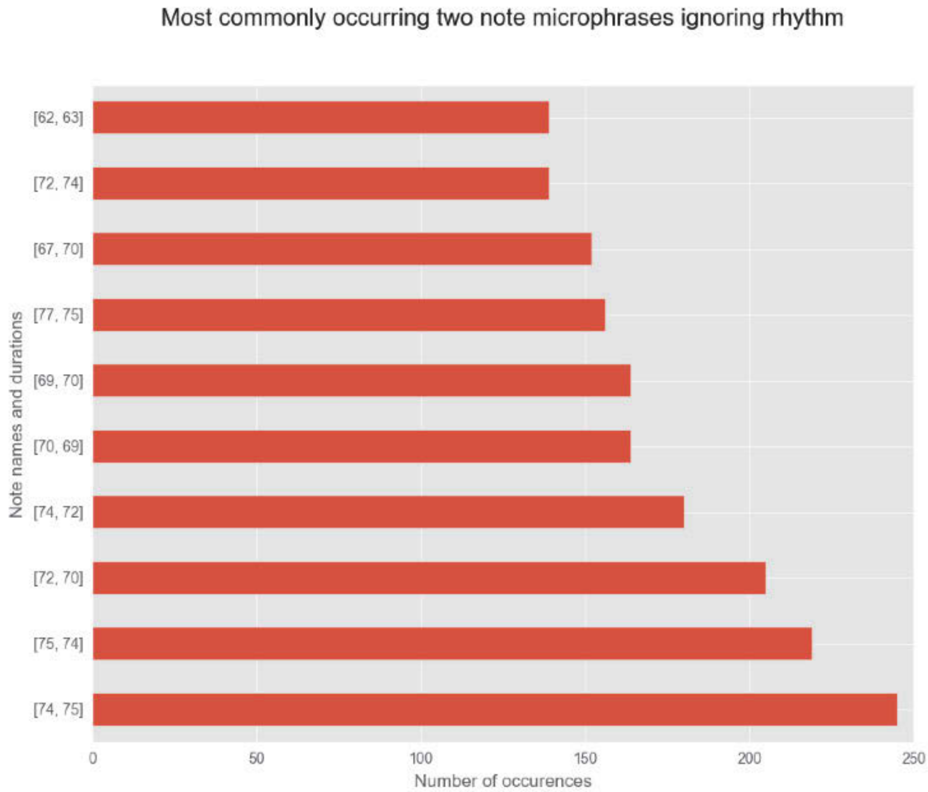
Figure 5.59. Most commonly occurring two-note microphrases



This Figure shows that the most common two-note microphrases (occurring more than 110 times) consists of two eighth-notes. The first is D5, and this is followed by Eb5. The second most commonly occurring microphrase again consists of two eighth-notes, the first being Eb5 and the second being D5. Note that the predominant tonality used across the dataset contains two flats (being Bb major or its relative G minor), and from the point of view of harmony and voice leading in these tonalities, this movement is typical.

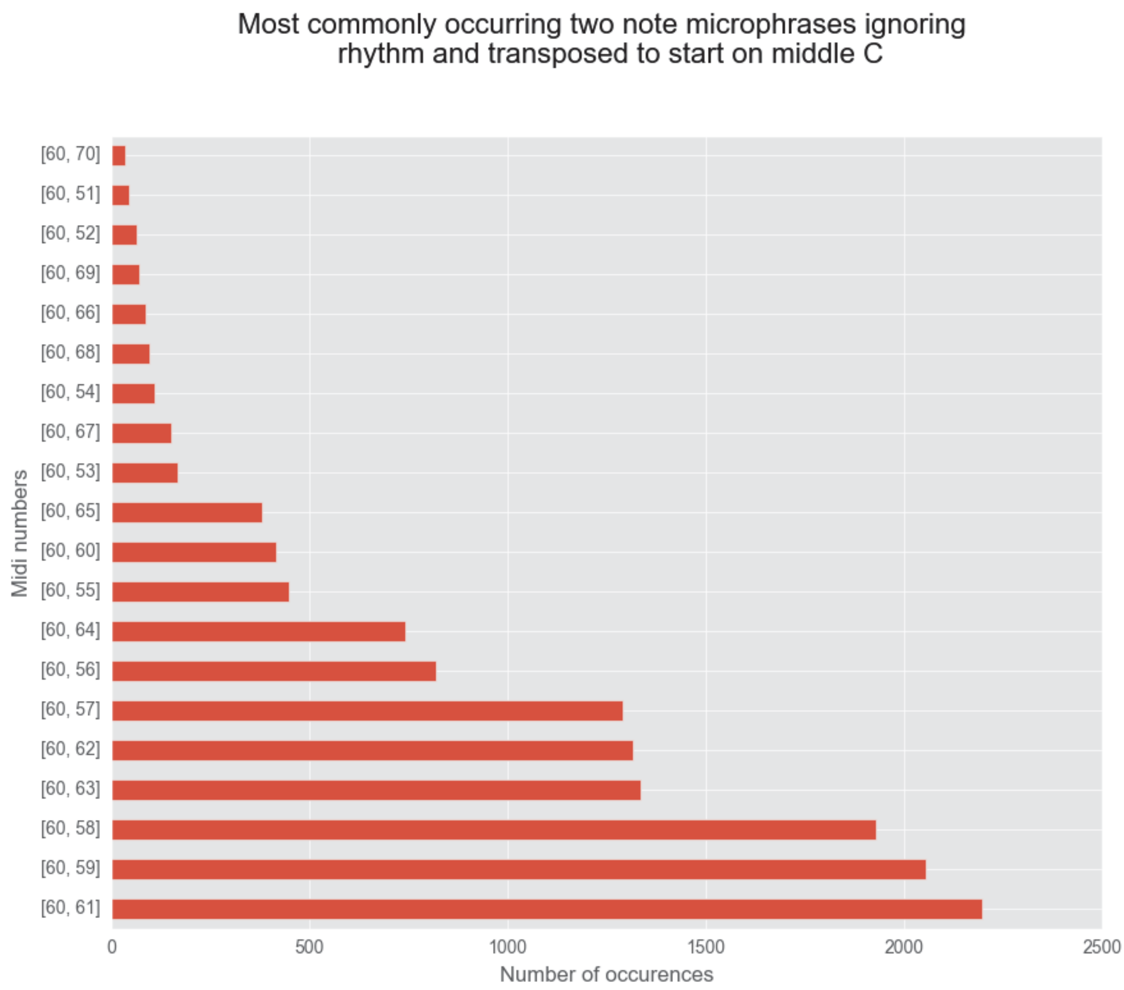
When ignoring rhythmic duration of notes, the instances of two-note microphrases start to grow significantly. Figure 5.60 lists the midi numbers of the most commonly occurring two-note microphrases, revealing that that [74, 75], (being the microphrase of D5 and Eb5) can be found almost 250 times in the dataset.

Figure 5.60. Most commonly occurring two-note microphrases ignoring rhythm



It is also straightforward to transpose all the 13866 two-note microphrases so they each start on middle C. Doing this will allow them to be easily compared regardless of the particular pitch from which they start. This can be seen in Figure 5.61, and shows that the microphrase [60,61], or the movement up a semitone between notes (here between C4 (middle C) and Db4) occurs over 2000 times.

Figure 5.61. Most commonly occurring two-note microphrases ignoring rhythm and transposed to start on middle C



When the two-note microphrases are adjusted to allow for any duration and are transposed to start on middle C, a picture emerges of a large amount of repetitive structure. The above suggests that when Jarrett improvises, after playing any given note, he is over 200 times more likely to then play a note 1 semitone higher (seen above as [60,61]), as opposed to a note that is 10 semitones lower (seen above as [60, 70]). Viewed in this way, it can possible to assign probabilities to the note Jarrett will play next in the course of an improvisation.

Tables 5.14 through 5.19 display the top five combinations of two-note microphrases through to seven-note microphrases. It is clear that the longer microphrases become, the less repetition can be seen. Once a microphrase reaches a length of six notes, it will occur less than ten times across the entire dataset. However, in smaller microphrases, even those up with a length of up to four notes, there is still substantial repetition that can be found.

Table 5.14. Top five two-note microphrases with note names and no rhythm

Sequence of note names in phrase	Count of occurrences
['D5', 'D#/Eb5']	245
['D#/Eb5', 'D5']	219
['C5', 'A#/Bb4']	205
['D5', 'C5']	180
['A#/Bb4', 'A4']	164

Table 5.15. Top five three-note microphrases with note names and no rhythm

Sequence of note names in phrase	Count of occurrences
['D#/Eb5', 'D5', 'C5']	68
['D5', 'D#/Eb5', 'F5']	65
['C5', 'D5', 'D#/Eb5']	61
['C5', 'A#/Bb4', 'A4']	60
['F5', 'D#/Eb5', 'D5']	58

Table 5.16. Top five four-note microphrases with note names and no rhythm

Sequence of note names in phrase	Count of occurrences
['F5', 'D#/Eb5', 'D5', 'C5']	28
['D5', 'D#/Eb5', 'F5', 'G5']	26
['C5', 'D5', 'D#/Eb5', 'F5']	26

['F5', 'E5', 'D#/Eb5', 'D5']	25
['A4', 'A#/Bb4', 'C5', 'D5']	22

Table 5.17. Top five five-note microphrases with note names and no rhythm

Sequence of note names in phrase	Count of occurrences
['C5', 'D5', 'D#/Eb5', 'F5', 'G5']	14
['A#/Bb4', 'B4', 'C5', 'C#/Db5', 'D5']	13
['D5', 'D#/Eb5', 'F5', 'G5', 'G#/Ab5']	12
['D#/Eb5', 'F5', 'D#/Eb5', 'D5', 'C5']	11
['G4', 'A4', 'A#/Bb4', 'B4', 'C5']	11

Table 5.18. Top five six-note microphrases with note names and no rhythm

Sequence of note names in phrase	Count of occurrences
['D5', 'D#/Eb5', 'D5', 'D#/Eb5', 'D5', 'D#/Eb5']	9
['D#/Eb5', 'D5', 'D#/Eb5', 'D5', 'D#/Eb5', 'D5']	8
['A4', 'C5', 'A4', 'C5', 'A4', 'C5']	7
['D5', 'D#/Eb5', 'F5', 'D#/Eb5', 'D5', 'C5']	7
['C5', 'A4', 'C5', 'A4', 'C5', 'A4']	7

Table 5.19 Top five seven-note microphrases with note names and no rhythm

Sequence of note names in phrase	Count of occurrences
['D#/Eb5', 'D5', 'D#/Eb5', 'D5', 'D#/Eb5', 'D5', 'D#/Eb5']	8
['C5', 'A4', 'C5', 'A4', 'C5', 'A4', 'C5']	7
['D5', 'D#/Eb5', 'D5', 'D#/Eb5', 'D5', 'D#/Eb5', 'D5']	7

['A4', 'C5', 'A4', 'C5', 'A4', 'C5', 'A4']	6
['C4', 'D4', 'D#/Eb4', 'F4', 'G4', 'A#/Bb4', 'D5']	4

A final set of visualisations is provided below which displays details of four-note through eight-note microphrases, that have been transposed to start on middle C, again where duration is not taken into account. With the data viewed in this way, repetition becomes one of the defining characteristics of the dataset. The four-note microphrase, [60-59-58-57] (being the notes C4, B3, A#4/Bb3, A3) appears over 200 times in the dataset. The second most highly occurring four-note microphrase is 60-61-62-63 (being the notes C4, C#4/Db4, D4, D#4/Eb4) and appears over 150 times in the dataset. Even when considering longer six-note microphrases, such as [60-59-58-57-56-55] (being the notes C4, B3, A#4/Bb3, A3, G#4/Ab3, G3) this structure appears more than 25 times in the dataset. These visualisations are detailed in Figures 5.62 through 5.66.

Figure 5.62. Most commonly occurring four-note microphrases ignoring rhythm and transposed to start on middle C

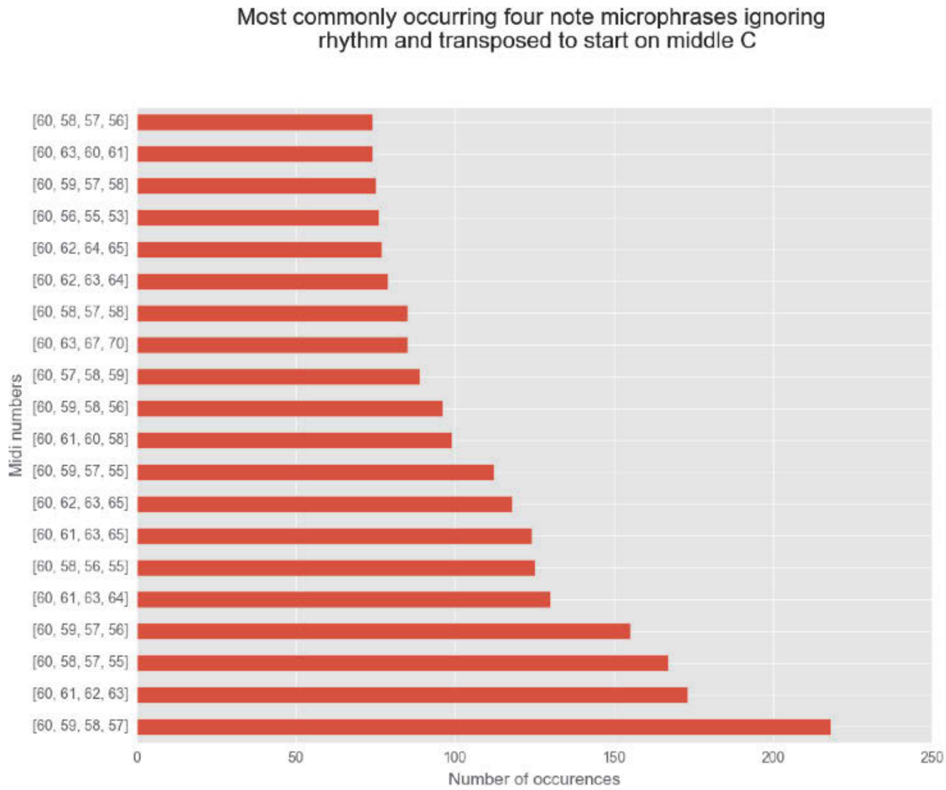


Figure 5.63. Most commonly occurring five-note microphrases ignoring rhythm and transposed to start on middle C

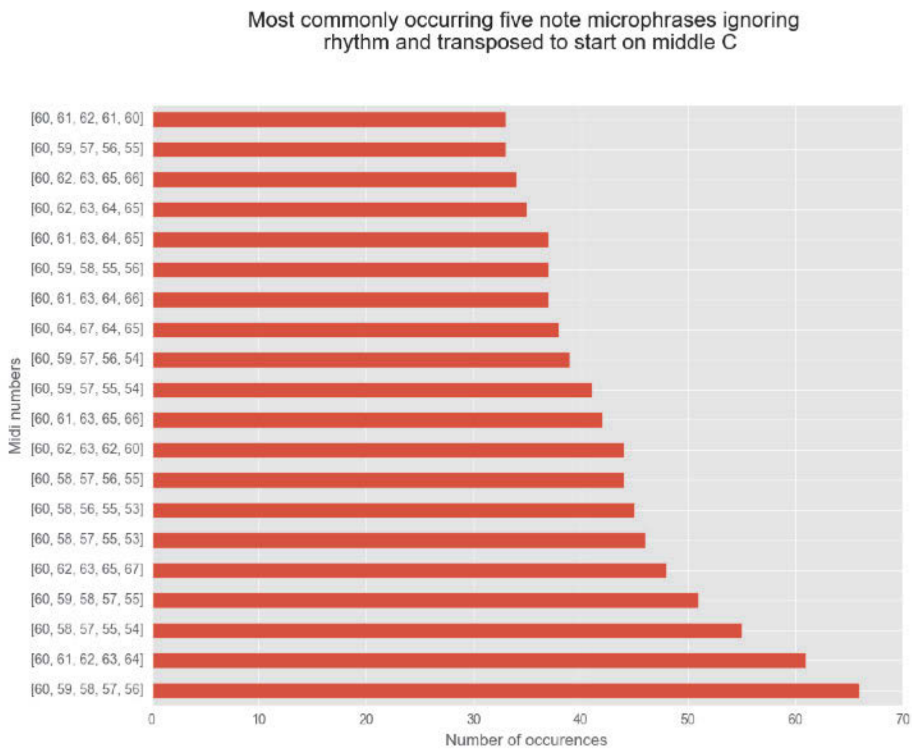


Figure 5.64. Most commonly occurring six-note microphrases ignoring rhythm and transposed to start on middle C

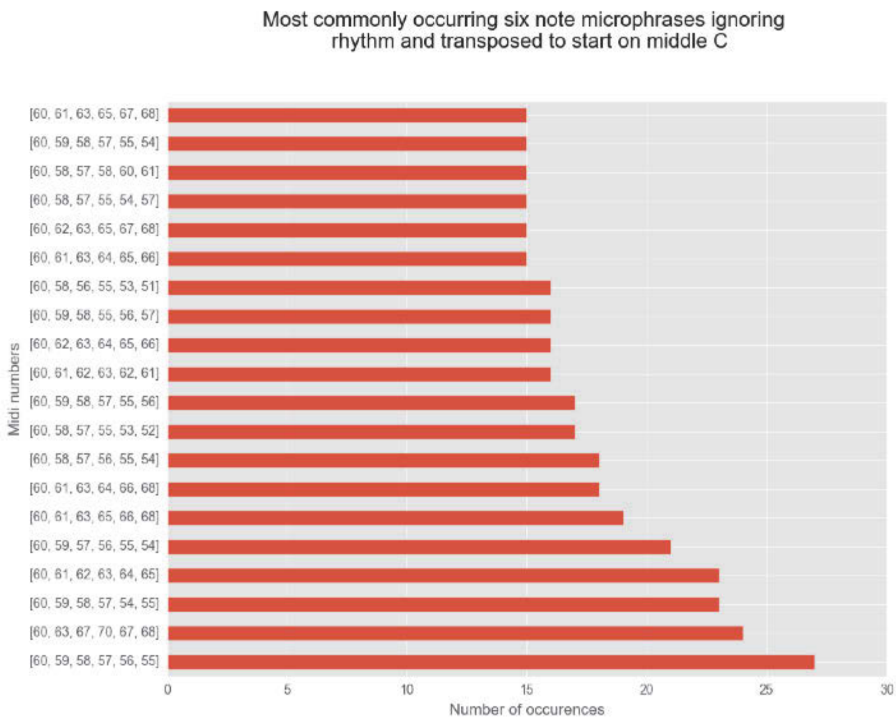


Figure 5.65. Most commonly occurring seven-note microphrases ignoring rhythm and transposed to start on middle C

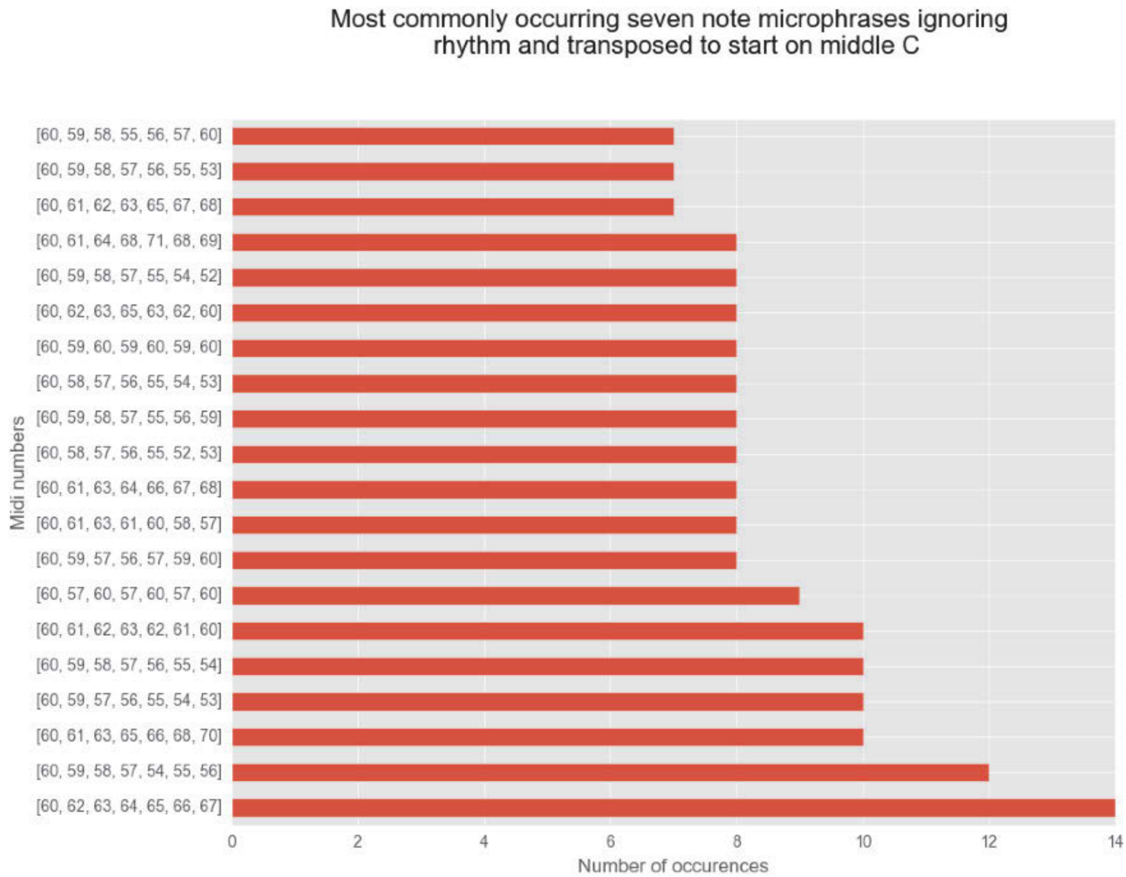
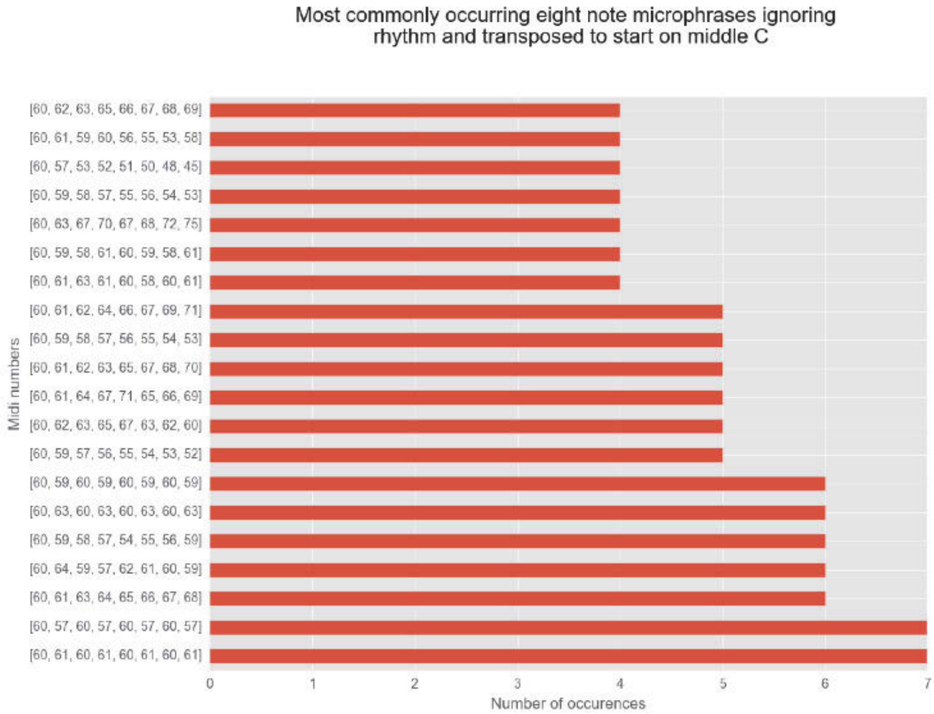


Figure 5.66. Most commonly occurring eight-note microphrases ignoring rhythm and transposed to start on middle C



The following section will explore the behaviour of two-note and four-note microphrases in order to demonstrate how they function as critical building blocks of larger phrases. I will also provide additional visualisations that can be used to explore the microphrases, and can be applied to microphrases of any lengths. The choice to explore only microphrases of two and four notes is due the what is apparent in the dataset: for Keith Jarrett, structures of this length are the critical building blocks of his overall improvisational style.

It was seen above that when this dataset was transformed, 13,866 two-note microphrases could be found (taking into account both pitch and duration). Once duplicates of are removed 3,424 unique two-note microphrases remain.

If the 3,424 two-note microphrases are transposed so they start on middle C, it can be seen that there are 733 two-note microphrases once duplicates are removed (still here accounting for duration). It is possible to also ignore the rhythmic duration of notes, and transpose them all to commence on middle C. Removing duplicates, this leaves 33 two-note microphrases in the data set. This means that, (when ignoring rhythmic duration and allowing for transposition) that each two-note microphrases appears approximately 420 times in other parts of the dataset.

Figures 5.67 through 5.76 provide some typical, highly occurring two-note microphrases that appear across the dataset. I have included their locations in the solos, the location in the measure, as well as the underlying chord from the jazz standard. Four out of the ten examples show typical voice leading movements toward chord tones (being Figures 5.67, 5.69, 5.70 and 5.73). In Figure 5.74, the C#5 note that leads to D5, while not being a resolution to a chord tone of C dominant 7, is a resolution of a chord tone of G minor, and the phrase in question uses more notes related those of G minor 7 (suggesting simple voice leading applied to chordal reharmonisation). Some arpeggiated patterns can be seen below too (in Figures 5.71, 5.75 and 5.76) and in these cases, the arpeggiated patterns go on to resolve to chord tones of the underlying chords.

Figure 5.67.   Phrase excerpt

Title	Stella By Starlight
Performer collection	Standards Live
Measure in which microphrase begins	15
Measure location in which microphrase begins	2
Underlying chord	A minor 7b5

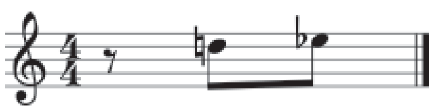


Figure 5.68. Phrase excerpt

Title	Stella By Starlight
Performer collection	Standards Live
Measure in which microphrase begins	66
Measure location in which microphrase begins	4
Underlying chord	A dominant 7

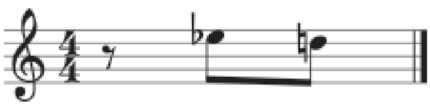


Figure 5.69. Phrase excerpt

Title	Days Of Wine And Roses
Performer collection	Keith Jarrett At The Blue Note, The Complete Recordings (Vol. 3)
Measure in which microphrase begins	16
Measure location in which microphrase begins	3
Underlying chord	E minor 7b5

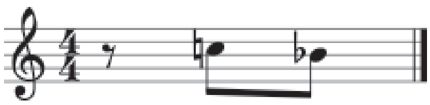


Figure 5.70. Phrase excerpt

Title	Days Of Wine And Roses
Performer collection	Keith Jarrett At The Blue Note, The Complete Recordings (Vol. 3)
Measure in which microphrase begins	17
Measure location in which microphrase begins	1
Underlying chord	D minor 7

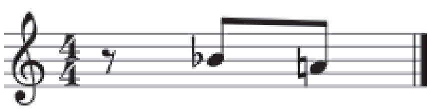


Figure 5.71. Phrase excerpt

Title	Days Of Wine And Roses
Performer collection	Keith Jarrett At The Blue Note, The Complete Recordings (Vol. 3)
Measure in which microphrase begins	16
Measure location in which microphrase begins	3
Underlying chord	A dominant 7 - D minor 7

Figure 5.72. Phrase excerpt

Title	Autumn Leaves
Performer collection	Still Live
Measure in which microphrase begins	242
Measure location in which microphrase begins	3
Underlying chord	G minor

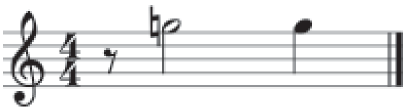


Figure 5.73. Phrase excerpt

Title	If I Were A Bell
Performer collection	Up For It
Measure in which microphrase begins	135
Measure location in which microphrase begins	1
Underlying chord	Ab major 7

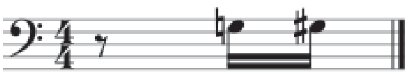


Figure 5.74. Phrase excerpt

Title	In Love In Vain
Performer collection	Standards, Vol. 2
Measure in which microphrase begins	45
Measure location in which microphrase begins	3.266
Underlying chord	C dominant 7

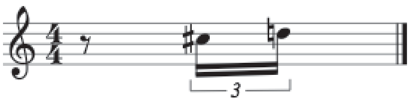
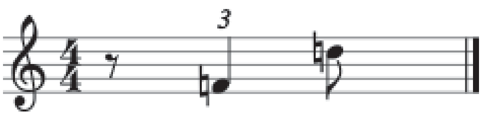


Figure 5.75. Phrase excerpt

Title	My Funny Valentine
Performer collection	Still Live
Measure in which microphrase begins	40 - 41
Measure location in which microphrase begins	3.5
Underlying chord	C min - C minor (major 7)

Figure 5.76. Phrase excerpt

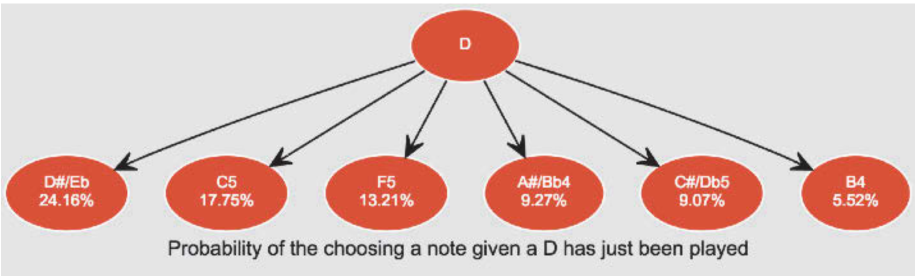
Title	Stella By Starlight
Performer collection	Standards Live
Measure in which microphrase begins	3 - 4
Measure location in which microphrase begins	3.66667
Underlying chord	C minor 7 - F dominant 7



These examples highlight a dichotomy that underlies Jarrett's improvisational style which I will take up further in the section on voice leading: the constant interplay between simple melodic movements favouring step movement or chordal arpeggiation that appears complex is due to many chords being superimposed on the underlying harmony, which makes the melodic movements seem more complicated and seemingly chromatic in nature.

While the music score view above can be used to visualise dataset results, there are certainly other ways to capture this information. Figure 5.77 provides an alternate view, demonstrating how Jarrett's melodic phrases and microphrases can be seen in terms of a decision-tree. The visualisations captures, when a given note is played, the probability of the next note occurring.

Figure 5.77. Decision tree for the probability of choosing a note given a D5 has just been played



This particular example shows that, if a D5 note is played, there is a only a 5.52% chance that the note following this will be the note B4. It is far more likely that the next note will be a D#/Eb5. Being able to access this kind of information about these improvisations is powerful, as it can be used as a measurement to ascertain which notes are more correct than others in the context of an overarching dataset and derive rules from which to create melody.

The purpose of including this visualisation is to give a sense of what is possible with regard to microphrase exploration. However, the decision tree visualisation is difficult to scale: the above figure only contains the top six probabilities above, and soon becomes unwieldy. Things become more complicated for longer phrase lengths too. In the following chapter I will explore a similar visualisation (called a sunburst partition) that I will use in a web based music score search and retrieval engine to track commonly occurring melodic sequences of variable lengths.

An alternate visualisation of two-microphrase occurrences can be seen in Figure 5.78 which allows more results to be viewed in a different way. It shows, in order, the most likely 20 things that will happen after a D5 is played. Playing a D6 note after a D5 note, for example, is extremely rare, but possible. The fact that this it so rare can become something worthy of exploration as it is only used in highly specific situations.

Figure 5.78. All possible outcomes following the note D5

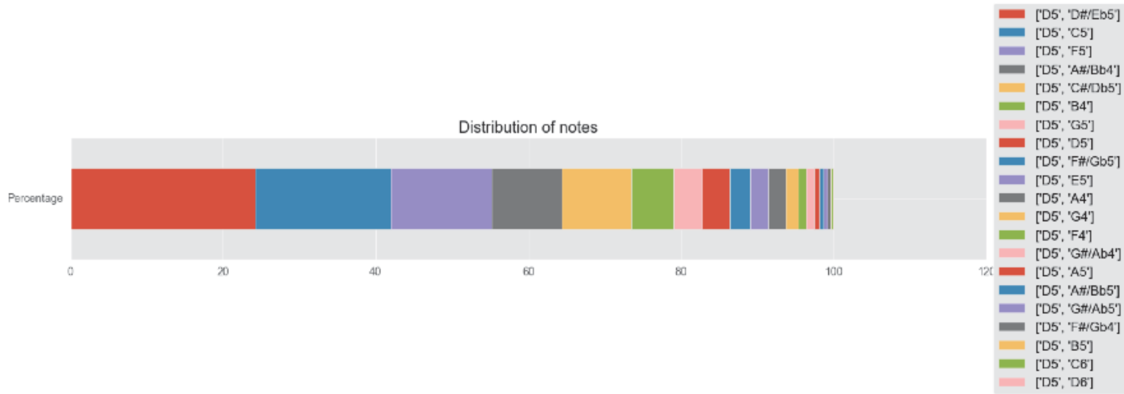
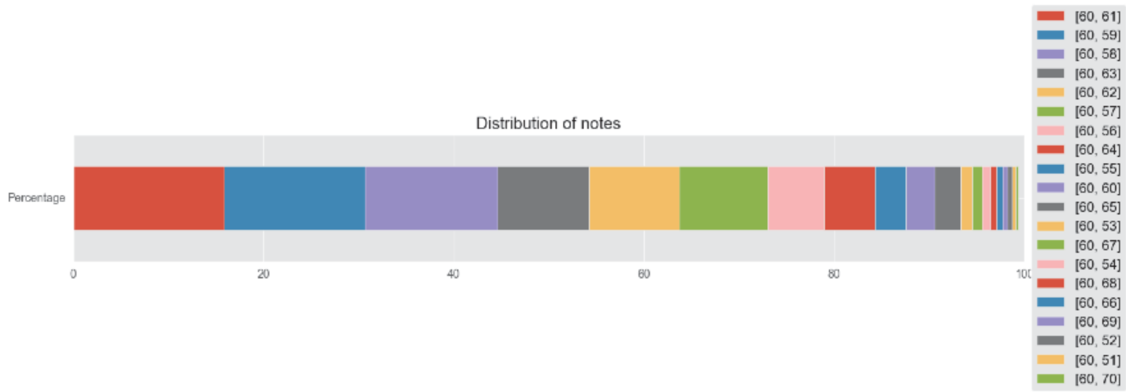


Figure 5.79 shows another similar visualisation of commonly occurring two-note microphrases, but this time transposing them all so that they start on middle C. A similar relationship can be seen to Figure 5.78, indicating that for Jarrett, the particular starting note is less important than the distance between the two notes that are being played. Put another way, for Jarrett, the key he plays in does not affect his note choice.

Figure 5.79. All possible outcomes following the note C4 (middle C)



While exploring two-note microphrases is a powerful way to explore the basic building blocks of improvisational style of Jarrett, the appearance of patterns which could be regarded as melodies does not come through at this level. In contrast to this, it is possible to explore longer, four-note microphrases to shed further light on Jarrett's playing. When microphrases are five notes in length or more, finding repetition becomes problematic, yet Jarrett does use many four-note microphrases throughout all the solos.

In the dataset, there 12,349 four-note microphrases. Taking into account both pitch and duration, these do not appear to support an argument of repetition. Table 5.20 shows that the most commonly occurring four-note microphrase only occurs 28 times (this being the note sequence F5, Eb5, D5, C5).

Table 5.20. Midi number counts with and without durations

Number of times the midi number and duration sequence of "[77, 75, 74, 72][0.5, 0.5, 0.5, 0.5]" occurs	14
Number of times the midi number sequence of "[77, 75, 74, 72]" occurs	28

Figures 5.80 and 5.81 provide different examples of places the most common four-note microphrases appear. Similar to what was found earlier, both examples

occur on different underlying chords (being Eb major 7 and Ab major 7) but exhibit the kinds of voice leading characteristics seen earlier, (the tendency for notes to resolve toward chord tones). If a harmony-centric view of the D5 in Figure 5.81 was taken, it could be interpreted as being the distinctive sharp eleventh of Ab Lydian. However the Ab major chord is occurring in the context of other chords belonging to C minor. Instead, the playing of the D5 indicates what was found earlier, that Jarrett is heavily influenced by the notes both of the underlying chords and the underlying tonality.

Figure 5.80. Phrase excerpt

Title	Someday My Prince Will Come
Performer collection	Up For It
Measure in which microphrase begins	276
Measure location in which microphrase begins	2
Underlying chord	Eb major 7

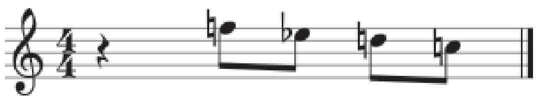
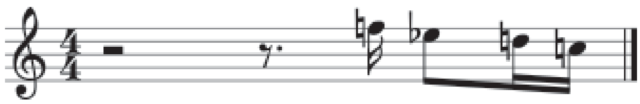


Figure 5.81. Phrase excerpt

Title	My Funny Valentine
Performer collection	Still Live
Measure in which microphrase begins	44
Measure location in which microphrase begins	0.75
Underlying chord	Ab major 7



The way in which improvised notes are related to the underlying groups of chords is becoming increasingly important. Table 5.21 provides further examples of this four-note microphrase along with the underlying tonality of the jazz standard. It shows that in five out of seven cases, the microphrase appears in the tonality with two flats (being Bb major or its relative G minor).

Table 5.21. Count of microphrases with the midi sequence “77, 75, 74, 72”

Performer collection	**Title**	**Count of microphrases in the corpus with the midi number sequence “77,75,74,72”**	**Predominant tonality of jazz standard**
Standards Live	Stella By Starlight	5	Bb major
Standards, Vol. 2	In Love In Vain	1	Bb major
Still Live	Autumn Leaves	2	G minor
	My Funny Valentine	2	C minor
Tokyo 96	Autumn Leaves	3	G minor
Up For It	Someday My Prince Will Come	11	Bb Major
Whisper Not	Groovin High	4	Eb Major

If all of the four-note microphrases are transposed so they start middle C, the extent of their repetition starts to become apparent. Table 5.22 details these, and shows that the most commonly occurring four-note microphrase consists of four ascending semitones, followed by the second most common, consisting of four descending semitones.

Table 5.22. Most commonly occurring four-note microphrases ignoring rhythm and transposed to start on middle C

Sequence of midi numbers	Number of times this sequence occurs across the entire dataset
[60, 59, 58, 57]	218
[60, 61, 62, 63]	173
[60, 58, 57, 55]	167
[60, 59, 57, 56]	155
[60, 61, 63, 64]	130
[60, 58, 56, 55]	125
[60, 61, 63, 65]	124
[60, 62, 63, 65]	118
[60, 59, 57, 55]	112
[60, 61, 60, 58]	99

Figures 5.82 through 5.91 show some instances where these microphrases appear in the dataset. Note that I have kept the transposition in place to make comparison more convenient.

Figure 5.82. Phrase excerpt

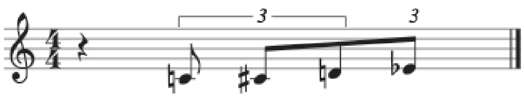
Title	Days Of Wine And Roses
Performer collection	Keith Jarrett At The Blue Note, The Complete Recordings (Vol. 3)
Measure in which microphrase begins	70
Measure location in which microphrase begins	2

Figure 5.83. Phrase excerpt

Title	My Funny Valentine
Performer collection	Still Live
Measure in which microphrase begins	45
Measure location in which microphrase begins	2

Figure 5.84. Phrase excerpt

Title	If I Were A Bell
Performer collection	Up For It
Measure in which microphrase begins	142
Measure location in which microphrase begins	1



Working through the most common four-note microphrases, it can be seen that some basic variations on stepwise patterns soon start to emerge: four-note microphrases are still often characterised by movement in semitones, but other figures start to emerge using tones. Several examples of this can be seen in figures 5.85 through 5.91 (note they have all been transcribed to start on middle C).

Figure 5.85. Phrase excerpt

Title	Days Of Wine And Roses
Performer collection	Keith Jarrett At The Blue Note, The Complete Recordings (Vol. 3)
Measure in which microphrase begins	4
Measure location in which microphrase begins	0.5

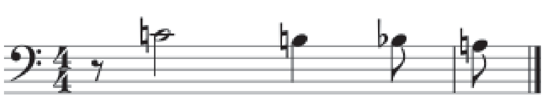


Figure 5.86. Phrase excerpt

Title	Days Of Wine And Roses
Performer collection	Keith Jarrett At The Blue Note, The Complete Recordings (Vol. 3)
Measure in which microphrase begins	12
Measure location in which microphrase begins	1.25

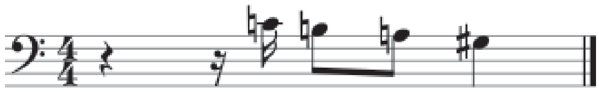


Figure 5.87. Phrase excerpt

Title	Days Of Wine And Roses
Performer collection	Keith Jarrett At The Blue Note, The Complete Recordings (Vol. 3)
Measure in which microphrase begins	80
Measure location in which microphrase begins	2.5

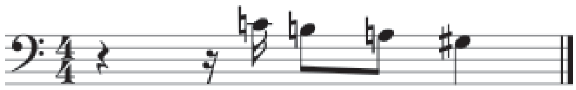


Figure 5.88. Phrase excerpt

Title	Days Of Wine And Roses
Performer collection	Keith Jarrett At The Blue Note, The Complete Recordings (Vol. 3)
Measure in which microphrase begins	26
Measure location in which microphrase begins	1.5

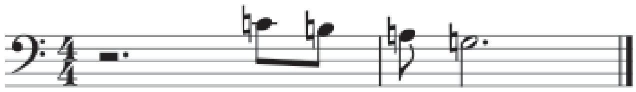


Figure 5.89. Phrase excerpt

Title	Days Of Wine And Roses
Performer collection	Keith Jarrett At The Blue Note, The Complete Recordings (Vol. 3)
Measure in which microphrase begins	80
Measure location in which microphrase begins	2.5

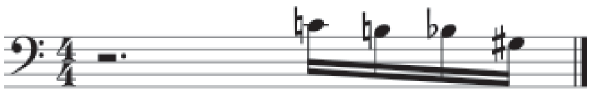


Figure 5.90. Phrase excerpt

Title	My Funny Valentine
Performer collection	Still Live
Measure in which microphrase begins	103
Measure location in which microphrase begins	0

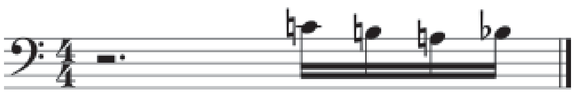
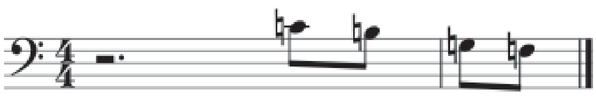


Figure 5.91. Phrase excerpt

Title	All The Things You Are
Performer collection	Standards, Vol. 1
Measure in which microphrase begins	241
Measure location in which microphrase begins	2.5



While it does not make sense to build a repository of typical Jarrett “licks” (as it would soon become too exhaustive) it is possible to apply such an approach to four-note microphrases, in order to obtain the distinct four note patterns that are regularly seen and can be used to characterise Jarrett’s style. The examples above also reveal a tendency shared in common with the first two notes of a four note microphrase (being a descending semitone) and another way of exploring this would be to only examine those four-note microphrases that start with the movement of a downwards semitone down, and then exploring what tends to happens next. Examples of this can be seen in Figure 5.92 through 5.94. Like the two-note microphrases, these are typified by and interplay between stepwise movement and movement around chord tones.

Figure 5.92. Phrase excerpt

Title	In Love In Vain
Performer collection	Standards, Vol. 2
Measure in which microphrase begins	93
Measure location in which microphrase begins	3.5

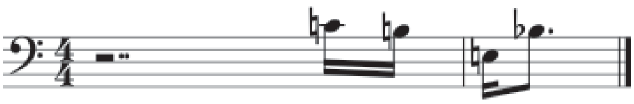


Figure 5.93. Phrase excerpt

Title	In Love In Vain
Performer collection	Standards, Vol. 2
Measure in which microphrase begins	62
Measure location in which microphrase begins	1.5

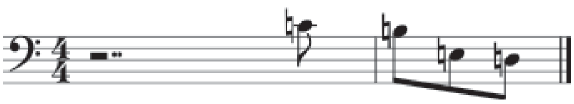
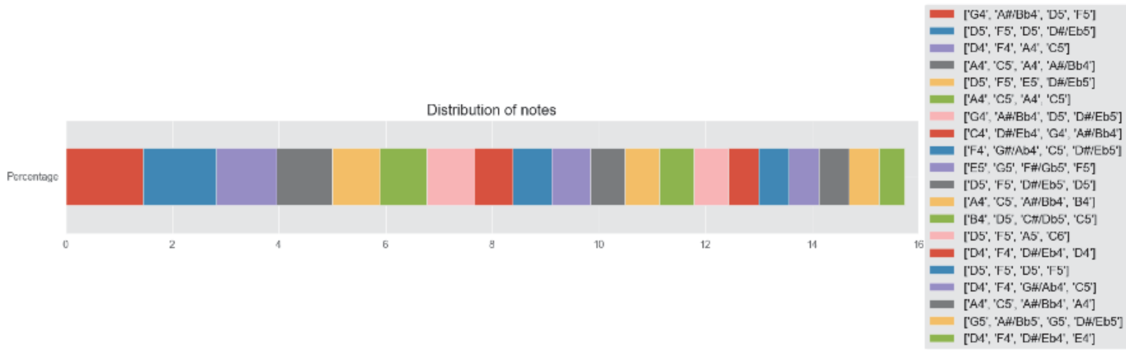


Figure 5.94. Phrase excerpt

Title	If I Were A Bell
Performer collection	Up For It
Measure in which microphrase begins	199
Measure location in which microphrase begins	0

Though using a visualisation such as a decision tree to explore four-note microphrases would become unwieldy, the alternative distribution-style visualisation, used earlier, can be utilised as a powerful way of understanding the different four-note microphrases that Jarrett uses. In Figure 5.95, I have applied a criteria of a four-note microphrase having any starting note and then moving upwards a minor third (such as G4 leading to an Bb4) before having any two additional notes.

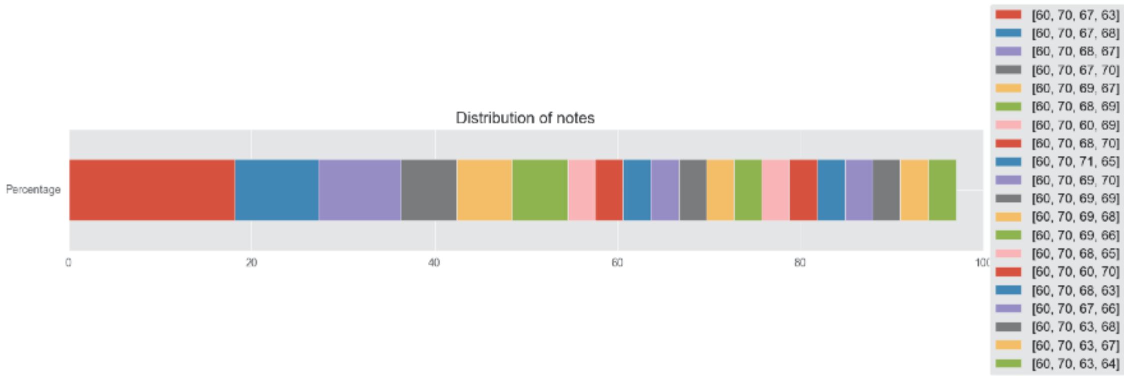
Figure 5.95. All possible outcomes following the note sequence G4-Bb4



The figure above reveals that there are two likely things that can occur once Jarrett plays an interval of an ascending minor third, regardless of the note he starts from. The first is that the next note will be a movement up a major third, (such as G going to Bb going to D) and the second is a movement back down a minor third (such as G going to Bb going to G again). This also suggests that if Jarrett plays two notes that are a leap rather than a step, he is then far more likely to undertake another leap movement rather than a step movement. Thus, the playing of intervals greater than a tone, tends to be followed by further intervals greater than a tone.

A second distribution visualisation can be seen below. This time I have examined four-note microphrases which start with the interval of an ascending flattened seventh, and have transposed all examples to start on middle C. It shows a clear tendency for flat seventh intervals to followed by quite limited possibilities.

Figure 5.95. All possible outcomes following the note sequence C4-Eb4 after transposition of all four-note microphrases to middle C



Examining the dataset in this way soon starts to become highly exploratory: there are very specific uses of steps and leaps starts employed by Jarrett in this dataset (for example, the C5 - A#/Bb5 flattened seventh movement in this dataset will be followed only by possible six possible choices, being G5, G#5/Ab5, C5, C#6/Db6, A5 or D#5/Eb5). Although the phrases overall have very little repetition, depending on the notes just played there appears to be a limited set of possibilities that can come next.

Exploring the data in this fashion soon becomes problematic however. It would be possible to explore four-note microphrases that start with a flattened seventh and correlate this to what the underlying chords are, or to particular solos, or tempos: the data is completely amenable to any of these questions. But the case study quickly becomes unwieldy. The purpose here is to provide a sense of what is possible.

While exploring how microphrases operate across the dataset, I have alluded to notions of harmony and voice leading, and will examine this in the next section. In this dataset, every note or rest contains an underlying chord and chord type. Note that there is a simplification going on here: the chord and chord types have been taken from chord changes typically seen in the lead sheets of the jazz standards, and does not reflect any implicit re-harmonisations that might be inferred from the notes of the solo. Yet even with such a high level of simplification, there is still a great deal of insight than can be gained.

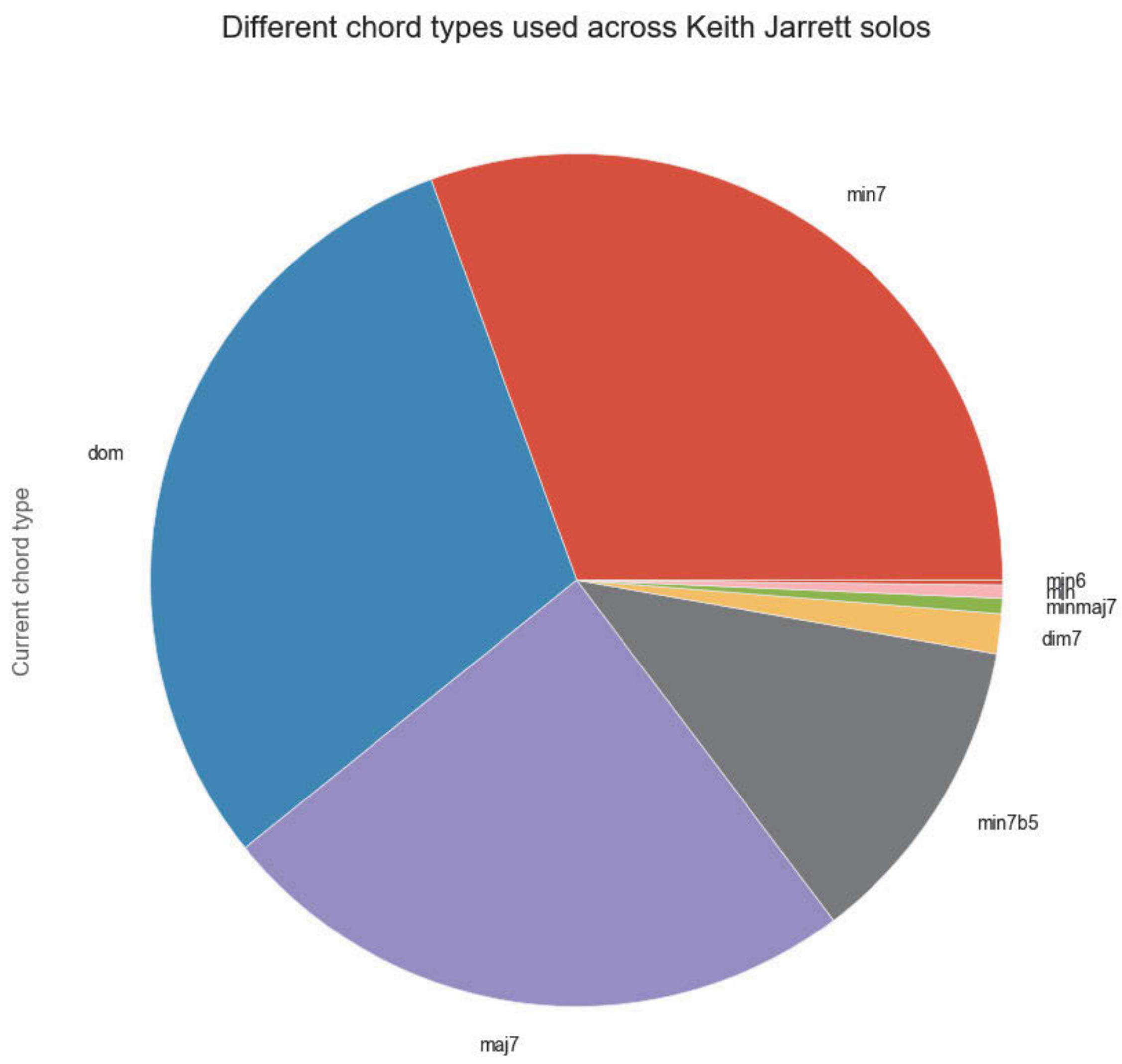
To explore harmony and voice leading, I have included a “harmonic degree” attribute for all the records in the datasets that are notes (which defaults to null if the record is a rest). The list of all the possible harmonic degrees can be seen below in Table 5.23. Using these labels makes it possible to explore the relationship between any given note in an improvised phrase, and the underlying chord.

Table 5.23. Names of harmonic degrees with an example on the root note C

Chord Root	**Example Melody Note**	**Harmonic degree**
C	C	Root
C	C#/Db	Flat ninth
C	D	Ninth
C	D#/Db	Sharp ninth
C	E	Third
C	F	Eleventh
C	F#/Gb	Sharp eleventh
C	A	Thirteenth
C	A#/Bb	Flat seventh
C	B	Major seventh

Figure 5.97 shows the distribution of all the different chord types found in this dataset, indicating that they are overwhelmingly dominant seventh, minor seventh and major seventh. It also shows a handful of other chord types which are far more rare.

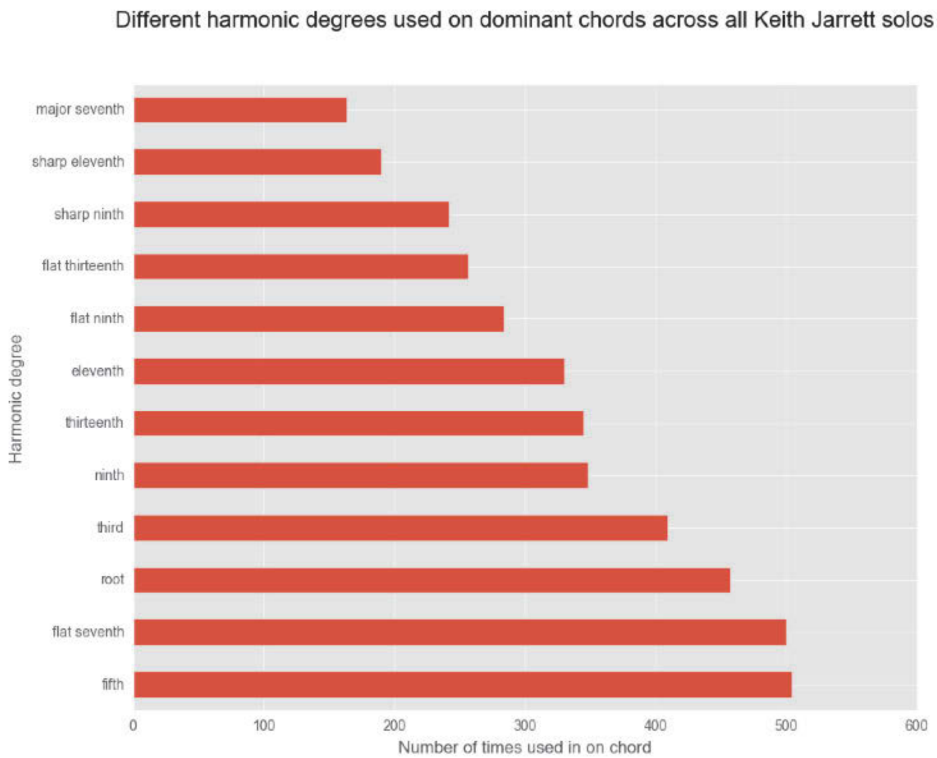
Figure 5.97. Different chord types used across all solos



The high counts of minor 7, minor 7b5, dominant 7 and major 7 can be related to the extremely common chord progression that appears all throughout the dataset and more generally in jazz: the II-V-I progression (with II as a minor 7b5 when I is minor, and II as a minor 7 when I is a major 7). The dataset can be easily be interrogated to build a compendium of typical progressions, however given that the solos all come from jazz standards whose structure is predictable (being based around II-V-I movements and chord movements generally moving in fourths) I will not do this. Instead, I would like to examine how different harmonic degrees are treated in the solos. Even when ignoring notions of chord superimposition, there are still very strong patterns that can be found in the way voice-leading operates works in these solos.

Figure 5.98 shows all the harmonic degrees used on dominant seventh chords in these solos. What is immediately apparent is that all the pitch classes of the octave are in used.

Figure 5.98. Different harmonic degrees used on dominant chords across all solos



Some of these results are to be expected: there is a predominance of thirds and sevenths (the so called tritone, which when played as an interval provides the dominant seventh's distinctive sound). Further, the top four occurring harmonic degrees are the notes from the dominant seventh chord: the fifth, flat seventh, root, and third.

However, in terms of jazz improvisation, this is unusual. When improvising over the dominant seventh, jazz musicians will often favour altered chord extensions such as the flat-thirteenth, sharp-eleventh, sharp-ninth and flat-ninth. When jazz is discussed in terms of scales, these upper chordal extensions are often presented in the context of the altered scale (being a mode of the melodic minor scale) which can be used when playing over dominant chords. However these choices of harmonic degrees are used less than others such as the fifth and eleventh.

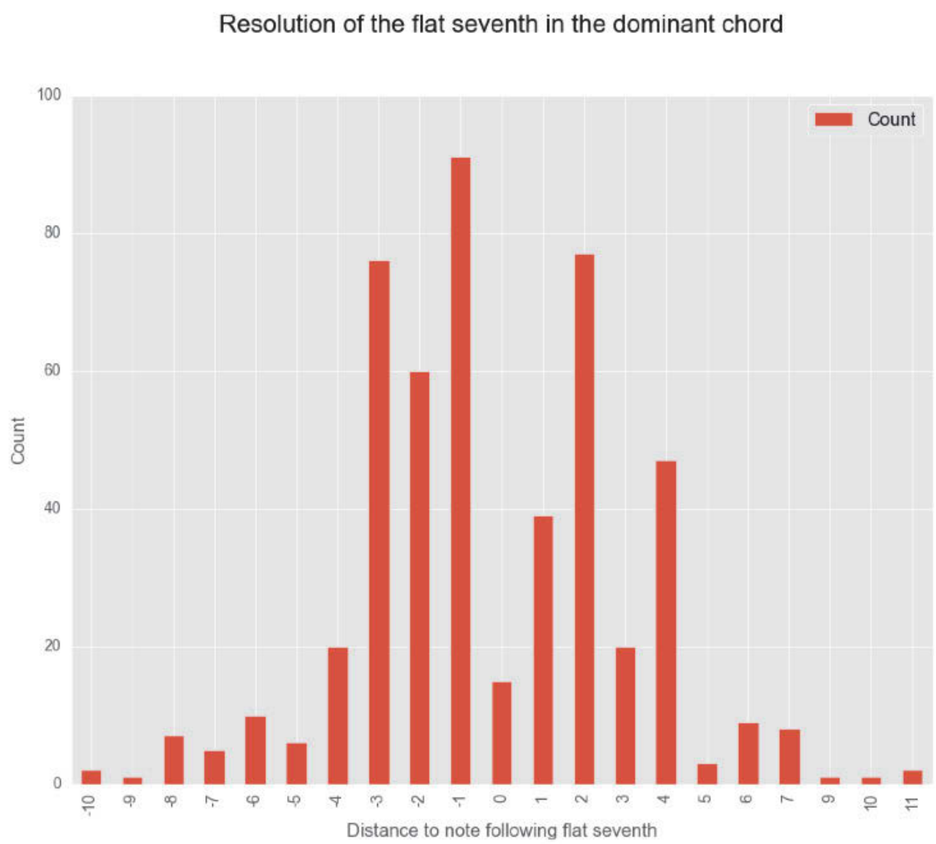
The heavy use of the fifth is also unexpected. Though it is outside the scope of this dataset, Jarrett almost never plays unaltered left-hand dominant chords in the trio while improvising (or more generally) so the appearance of the natural fifth seems unusual.

What is more surprising however, is the amount of times a major seventh occurs (for example, playing a B note while soloing on a C dominant chord). While this is the most least likely harmonic degree to be played over a dominant chord, it occurs over 150 times across the solos. Whereas the dominant seventh appears as a note choice 12% of the time on a dominant chord, the major 7th is not far behind, at 4%.

To further interrogate how these notes are used, they can be examined in terms of voice-leading: the ways in which harmonic degrees are both prepared and resolved. As it is straightforward in this dataset to examine the distance between one note and the next, the dataset can be grouped by both chord type and distance.

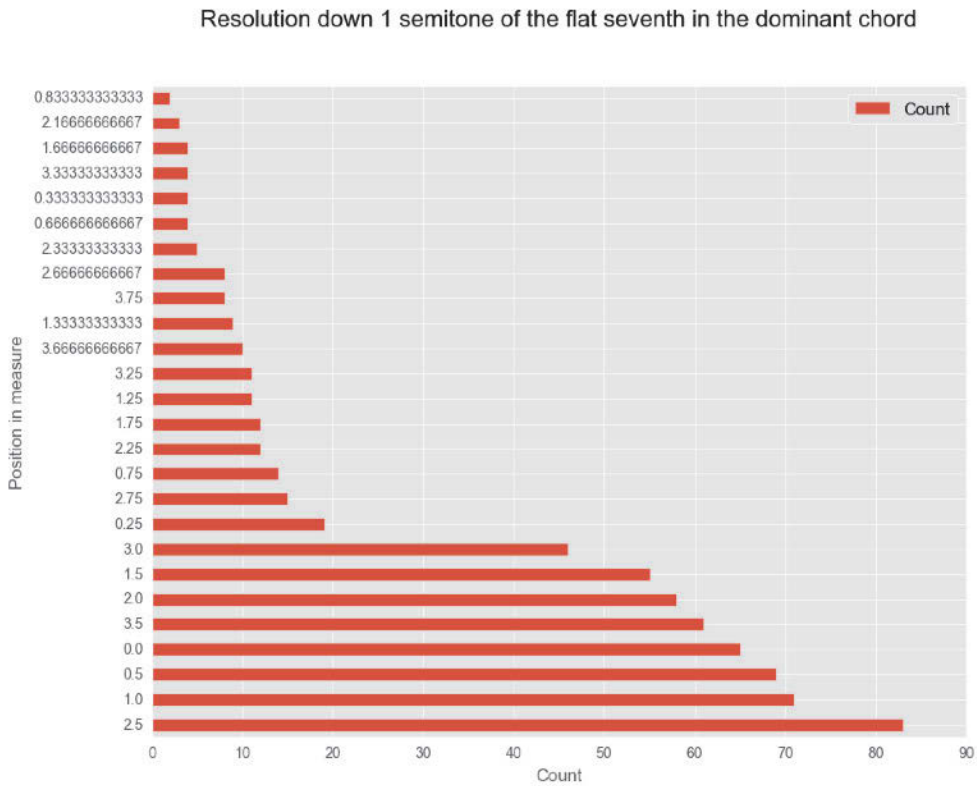
The flat-seventh is the most commonly used harmonic degree used on the dominant seventh chord in this dataset. Figure 5.99 shows the different ways it is resolved by calculating the distance to the following note, and counting the instances of these distances. While there are a number of outliers here, it can be seen that when Jarrett plays a flat-seventh on a dominant chord, he will most likely resolve this in one of four ways. The most common way is by moving down one semitone (for example a Bb note on a C dominant-seventh resolving to an A note). This is followed by going up a tone, then down a minor third, and then down a tone.

Figure 5.99.   Resolution of the flat-seventh in the dominant chord



It is possible to delve deeper into this example of the flat-seventh to examine where in the measure it is used. The flat seventh will most often occur at position 2.5 (of halfway through the third beat of the bar). Again, it can be seen that, though there are outliers, there are clear tendencies in the rhythmic placement of the flat seventh in Jarrett's solos.

Figure 5.100. Resolution down one semitone of the flat-seventh in the dominant chord



Examining the dataset in this way again highlights that what can be asked here are very specific and exploratory questions. It is possible to filter the dataset to explore how one harmonic degree operates in one type of chord, and then explore specific ways in which this dealt with. Highly marginal use cases can also be explored (for example, the resolution of a flat-seventh down nine semitones (for example a Bb4 note on a C dominant chord resolving to C#/Db4). This makes it possible to move away from music being defined in a rules based framework, and allow this metadata to be examined in the context of highly specific and contextual questions. Here, it becomes possible to ask about the rarest ways that Jarrett will use the flat-seventh when soloing. Again, answering these questions can lead to findings that are specific to particular solos, tempos or time-signatures. As an example, once I start to examine the most common harmonic degree (the flat-seventh) and the most common part the measure it is used on (2.5 being half way through the third beat of the bar), I can then focus in further to examine the particular solos where this is occurring. Table 5.24 shows a sample of the instances where this situation occurs.

Table 5.24. The use of the flat-seventh on beat 2.5 on a dominant chord

Title	Performer collection	Current measure
Autumn Leaves	Still Live	123
Autumn Leaves	Tokyo 96	18
Autumn Leaves	Tokyo 96	38
If I Were A Bell	Up For It	193
Someday My Prince Will Come	Up For It	1
Someday My Prince Will Come	Up For It	105
Someday My Prince Will Come	Up For It	269
Stella By Starlight	Standards Live	132

The dataset also shows that the major-seventh is used in a very specific way on the dominant seventh. Table 5.25 below lists the first ten examples of where this occurs (here ordered by title) and shows that in 80% of cases when a major-seventh is used the note before it is one semitone higher (for example a B5 on a C dominant seventh chord being preceded by a C6). While there are more possibilities of what could occur when a flat-seventh is used, when a major-seventh appears, it will be prepared and resolved in far more limited ways.

Table 5.25. Examples of major seventh being used on a dominant chord

Title	Performer collection	Current measure	Distance to previous midi number
All The Things You Are	Standards, Vol. 1	229	1
All The Things You Are	Standards, Vol. 1	276	1
Autumn Leaves	Still Live	29	1
Autumn Leaves	Still Live	41	1
Autumn Leaves	Still Live	47	-2
Autumn Leaves	Still Live	125	1
Autumn Leaves	Still Live	133	1
Autumn Leaves	Still Live	149	2
Autumn Leaves	Still Live	189	1
Autumn Leaves	Still Live	193	1

When looking across all the examples, it shows that when the major seventh is used, it will be prepared the same way 78% of the time (from one semitone above). There are only six other ways it could be prepared, and three of these have a less than 2% chance of occurring. The counts of all the preparation can be seen in Table 5.26.

Table 5.26. Preparation of the major seventh on a dominant chord

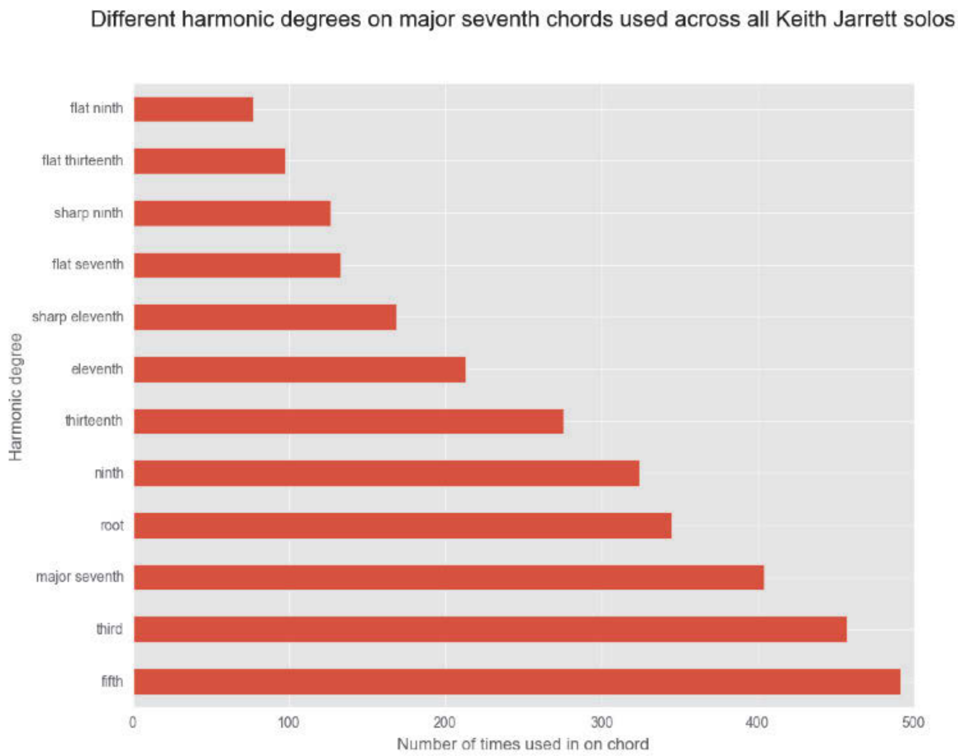
Distance to previous note	Count of times this is found
1	39
-1	4
4	2
-4	1
-3	1
-2	1
3	1
2	1

What starts to become evident when examining the dataset in this fashion, is the very high level of predicability and probability that can be assigned to Keith Jarrett's note choice. Though each melodic phrase may be unique, the inner workings of the phrases are highly structured and dependent on predictable notions of voice-leading. This appears to an extent in the microphrases, but becomes far more clear when examining harmonic degrees are used. It also turns out that working through complicated reharmonisation analysis is not even needed to uncover this structure. It may be that the use of a major-seventh on a dominant seventh chord is indicative another super-imposed harmony (such as the B note being part of a G dominant seventh structure being super imposed on a C dominant seventh) however, the voice leading underpinning the B note will behave similarly regardless of any chordal superimposition.

These highly specific and nuanced questions again highlight the difficulty of having the case study structured in this way. It would be easy to undertaken an entire chapter to explore the way Jarrett solos just across dominant seventh chords, and examine all the harmonic degrees along with the specific way in which they are dealt with. The downsize of having the ability to ask such specific questions of the dataset is that it challenges the suitability of framework in which questions are asked.

Turning to the major seventh chord, Figure 5.101 shows the different harmonic degrees that are used during Jarrett's solos.

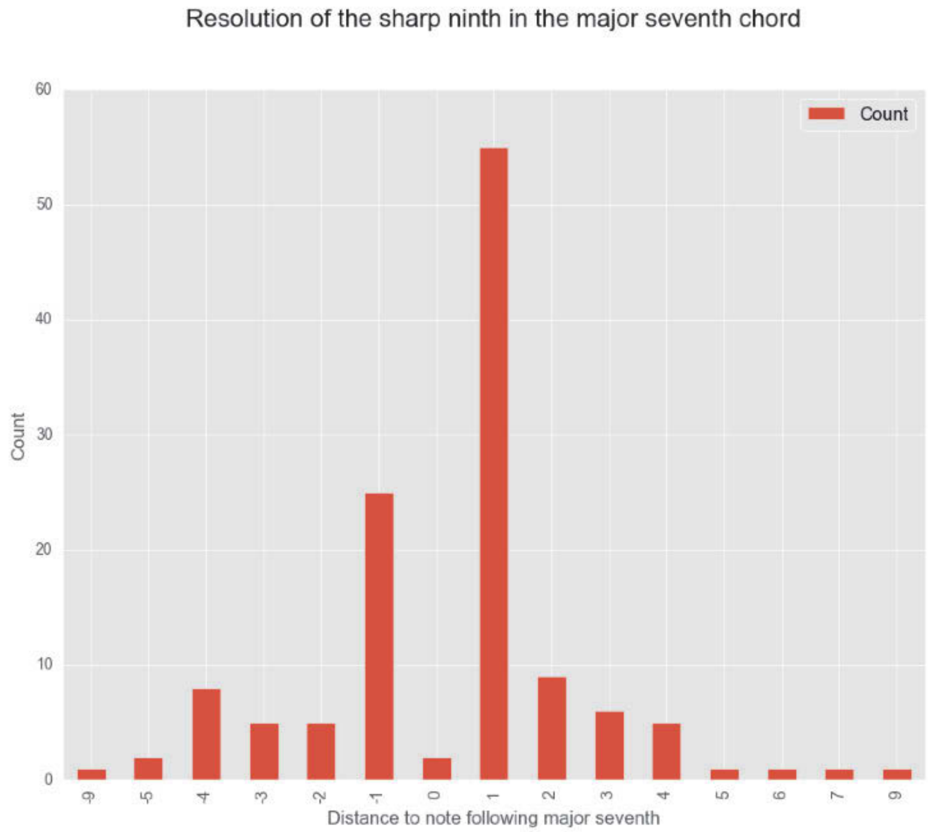
Figure 5.101. Different harmonic degrees used on major seventh chords across all solos



Similarly to the dominant-seventh, the chord tones are still the most common notes, being the fifth, third, major seventh and root. Again, it is the fifth that is the most prominent. All twelve pitch classes are still in use and the order of commonly occurring harmonic degrees is very similar to the dominant chord.

Just like the appearance of the major seventh in the dominant chord, the use of the sharp-ninth on a major seventh chord (for example, playing a D#/Eb note on a C major seventh chord) is unusual, and one of least most occurring harmonic degree choices. Figure 5.102 shows the different ways that the sharp-ninth is resolved.

Figure 5.102. Resolution of the sharp ninth in the major seventh chord



The figure shows that, by far, the use of the sharp-ninth on a major seventh chord will be resolved by moving a semitone upwards (for example an D#5/Eb5 note resolving to a E5 note on a C major seventh chord). This means that if Jarrett plays a sharp-ninth, he is half as likely to resolve this down a semitone, rather than up a semitone. The chances of resolving it any other way are less than five times as likely. Additionally, five of the resolutions are outliers, each occurring only once in the entire dataset.

Table 5.27 provides further information about the use the sharp-ninth on a major-seventh chord. It can also be seen that for the case of the sharp-ninth (and this is the first instance I have come across this in the dataset) key appears to be a factor. If a sharp ninth occurs on on a major chord, it is far more likely to be an

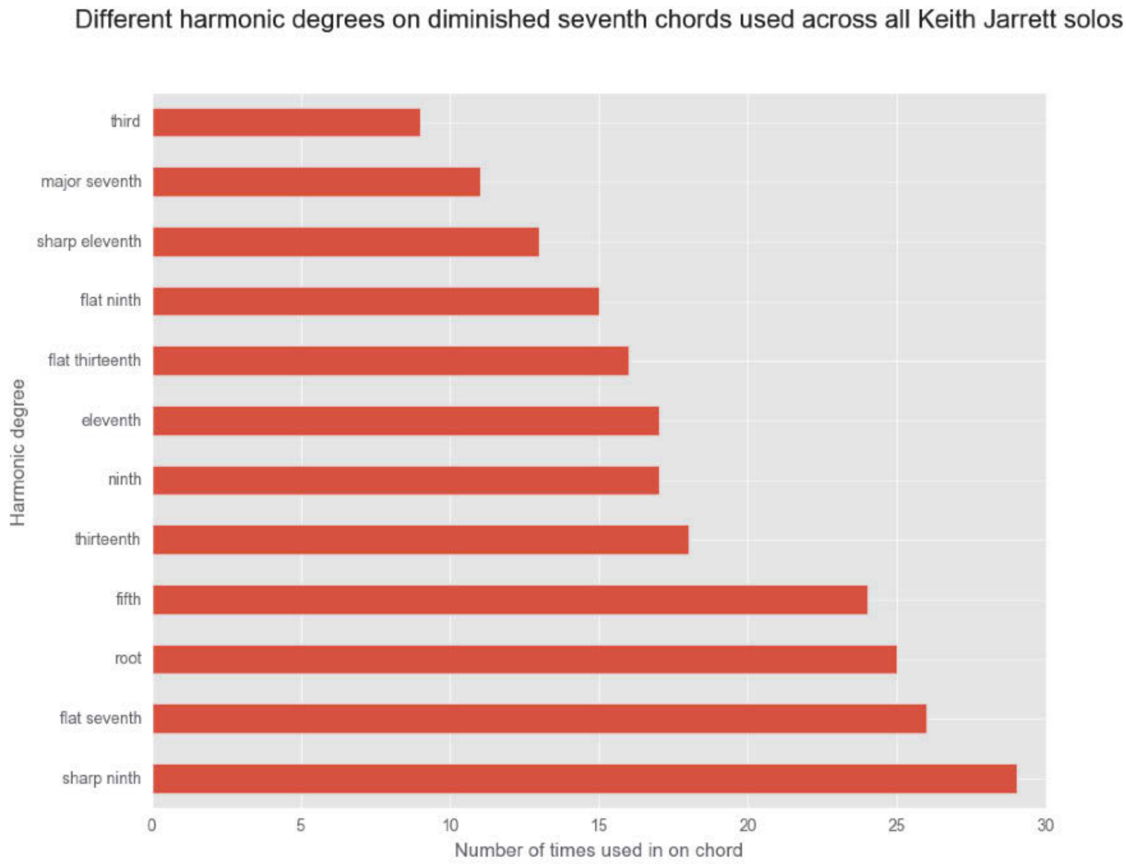
Eb major seventh chord. Also, the preparations of the sharp-ninth on an Eb major-seventh are limited to only three semitones above, or one semitone above.

Table 5.27. Examples of the sharp ninth being used on a major seventh chord

Title	**Performer collection**	**Current measure**	**Distance to previous midi number**	**Current chord root**
Autumn Leaves	Still Live	112	3	D#/Eb
Autumn Leaves	Still Live	144	3	D#/Eb
Autumn Leaves	Still Live	183	1	D#/Eb
Groovin High	Whisper Not	20	1	D#/Eb
Groovin High	Whisper Not	211	4	D#/Eb
Days Of Wine And Roses	Keith Jarrett At The Blue Note, The Complete Recordings (Vol. 3)	101	-6	F
Someday My Prince Will Come	Up For It	68	3	D#/Eb
Someday My Prince Will Come	Up For It	132	3	D#/Eb

It is clear from Figure 5.97 that the diminished chord appears far less times across this dataset. However it still possible to view it in terms of the harmonic degrees that are used, and how they are prepared and resolved. Figure 5.103 shows the counts of the different types of harmonic degrees that are used with the diminished chord.

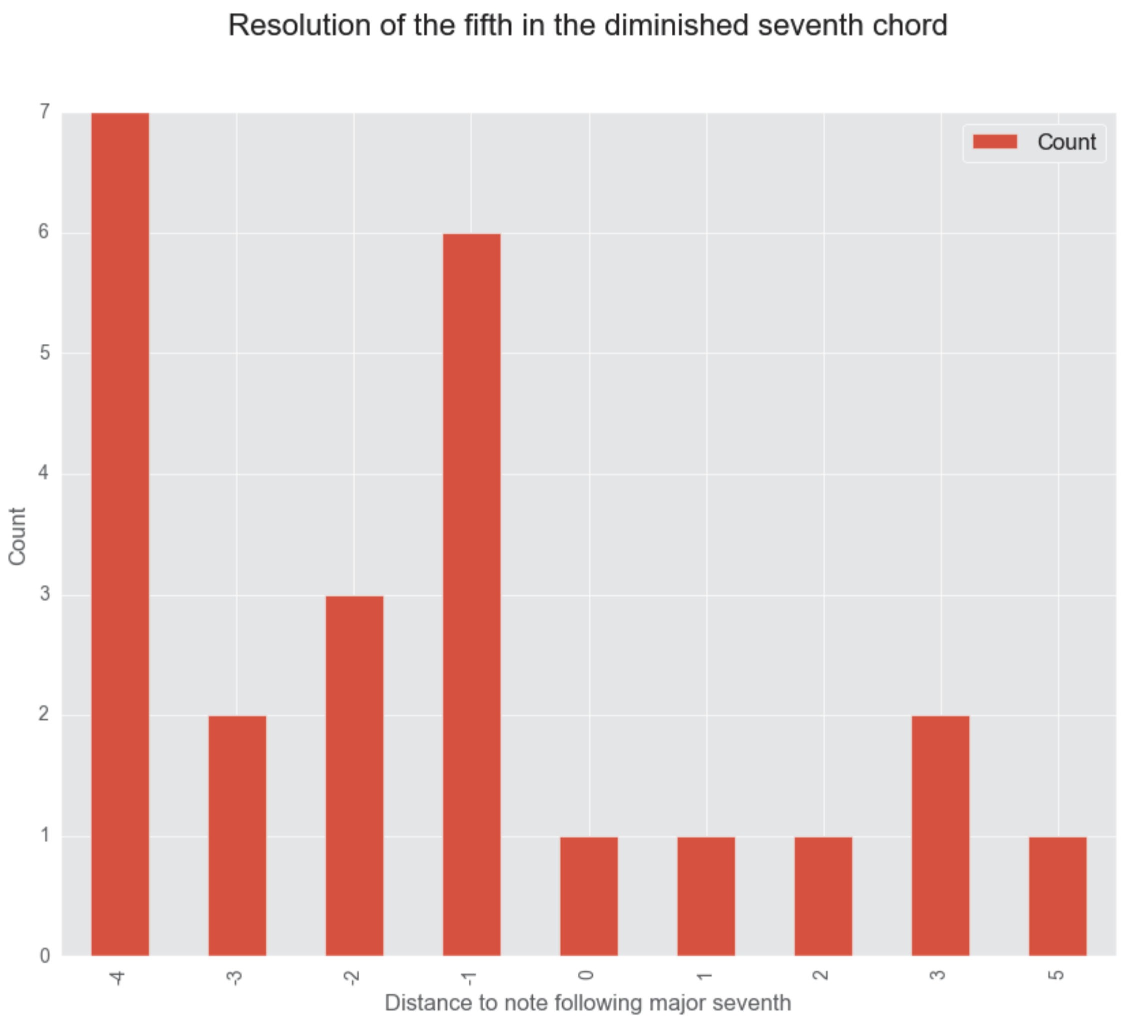
Figure 5.103. Different harmonic degrees used across all solos



There are similar trends to the other chord types that can be seen. Five of the four highest occurring harmonic degrees are chord tones from the diminished seventh. The fifth is again important (and the importance of the fifth across all chord types is unexpectedly characteristic of Jarrett's soloing). And again, all pitch classes are used, the most least occurring being the third (for example an E5 note played on a C diminished-seventh chord)

Playing the harmonic degree of a fifth on diminished-seventh chord is a marginal choice. To examine this further, Figure 5.104 shows the different ways the fifth is resolved when used on a diminished-seventh chord.

Figure 5.104. Resolution of the fifth in the diminished seventh chord



Both of the most common resolutions move toward diminished-seventh chord tones (the most common being down four semitones to the minor third of the diminished chord, or down one semitone to the sharp eleventh of flat fifth of the chord. The preparation of the fifth in the diminished-seventh chord sees more variability (however it should be noted the sample, being only fifth notes used on diminished seventh chord) is very small.

Table 5.28. Examples of the fifth being used on a diminished seventh chord

Title	Performer collection	Current measure	Distance to previous midi number	Current chord root
All The Things You Are	Standards, Vol. 1	34	5	B
Someday My Prince Will Come	Up For It	15	1	C#/Db
Someday My Prince Will Come	Up For It	75	-3	C#/Db
Someday My Prince Will Come	Up For It	111	2	C#/Db
Someday My Prince Will Come	Up For It	221	3	E
Someday My Prince Will Come	Up For It	239	3	C#/Db
Someday My Prince Will Come	Up For It	285	-7	E

It would be problematic to say that key is a factor here, as there are simply not enough examples to inform a representative sample, but it again highlights that although Jarrett appears endlessly inventive at the phrase level, when harmonic degrees and voice leading are considered, there are only a limited number of choices he will make at any given time in the solo.

With the sheer volume of the data that might be explored when the data is structured in this way, things quickly become problematic. There are however, alternate ways to present information regarding harmonic degrees. One of these is to cross-tabulate the harmonic degrees with any other aspect of the data in order to produce heat-map style visualisations, showing when certain harmonic degrees tend to appear. Tables 5.29, 5.30 and 5.31 show a cross-tabulation of harmonic degrees with measure position, for dominant-seventh, diminished seventh, and minor seventh chords. The numbers in the table refer to the counts of when different harmonic degrees occur at certain positions in the measure. If a count is denoted as 'NaN' (meaning not a number) it should be regarded as zero.

Table 5.29. Cross tabulation of harmonic degrees and position in the measure in which they are used on the dominant seventh chord

	11th	5th	b9th	b7th	b13th	M7	9th	Root	#11th	#9th	3rd	13th
0.00	3.0	5.0	6.0	9.0	8.0	6.0	4.0	6.0	6.0	5.0	7.0	6.0
0.25	3.0	9.0	10.0	9.0	5.0	2.0	12.0	4.0	2.0	4.0	2.0	10.0
0.50	11.0	8.0	7.0	7.0	4.0	5.0	8.0	7.0	2.0	6.0	3.0	8.0
0.75	6.0	9.0	4.0	11.0	5.0	6.0	9.0	6.0	4.0	5.0	7.0	6.0
1.00	9.0	6.0	5.0	8.0	3.0	2.0	5.0	10.0	6.0	3.0	7.0	9.0
1.25	7.0	9.0	2.0	9.0	6.0	7.0	6.0	11.0	4.0	NaN	4.0	10.0
1.50	10.0	9.0	3.0	12.0	6.0	6.0	7.0	5.0	4.0	5.0	4.0	10.0
1.75	9.0	9.0	7.0	6.0	6.0	1.0	9.0	10.0	5.0	6.0	9.0	4.0
2.00	5.0	3.0	5.0	7.0	7.0	2.0	5.0	8.0	4.0	4.0	15.0	8.0
2.25	6.0	9.0	8.0	8.0	4.0	2.0	5.0	5.0	4.0	10.0	10.0	6.0
2.50	7.0	11.0	4.0	8.0	5.0	2.0	6.0	9.0	4.0	6.0	12.0	5.0
2.75	7.0	8.0	9.0	8.0	4.0	7.0	6.0	9.0	3.0	6.0	4.0	7.0
3.00	NaN	6.0	1.0	5.0	4.0	1.0	4.0	8.0	3.0	4.0	6.0	7.0
3.25	2.0	6.0	8.0	3.0	3.0	2.0	3.0	8.0	1.0	3.0	6.0	4.0
3.50	2.0	1.0	4.0	8.0	6.0	3.0	3.0	7.0	4.0	1.0	8.0	4.0
3.75	2.0	7.0	5.0	7.0	1.0	2.0	3.0	11.0				

Table 5.30. Cross tabulation of harmonic degrees and position in the measure in which they are used on the diminished seventh chord

	11th	**5th**	**b9th**	**b7th**	**b13th**	**M7**	**9th**	**Root**	**#11th**	**#9th**	**3rd**	**13th**
0.00	NaN	1.0	NaN	1.0	NaN	NaN	3.0	1.0	NaN	NaN	NaN	NaN
0.25	NaN	2.0	NaN	1.0	1.0	NaN	NaN	1.0	NaN	1.0	NaN	NaN
0.50	2.0	NaN	1.0	NaN	NaN	1.0	NaN	NaN	1.0	2.0	NaN	1.0
0.75	1.0	1.0	NaN	NaN	2.0	NaN	NaN	3.0	NaN	NaN	1.0	NaN
1.00	NaN	NaN	NaN	1.0	1.0	NaN	NaN	NaN	NaN	NaN	NaN	2.0
1.25	NaN	NaN	2.0	1.0	1.0	1.0	NaN	NaN	1.0	NaN	NaN	NaN
1.50	1.0	NaN	NaN	NaN	1.0	NaN	2.0	NaN	NaN	1.0	NaN	2.0
1.75	NaN	1.0	NaN	1.0	NaN	1.0	NaN	1.0	NaN	2.0	1.0	NaN
2.00	NaN	NaN	1.0	1.0	NaN	1.0	1.0	NaN	NaN	1.0	NaN	3.0
2.25	1.0	3.0	1.0	NaN	NaN	NaN	NaN	2.0	NaN	NaN	NaN	NaN
2.50	1.0	1.0	NaN	1.0	NaN	NaN	1.0	NaN	NaN	2.0	NaN	NaN
2.75	NaN	1.0	1.0	NaN	NaN	1.0	NaN	1.0	2.0	NaN	1.0	NaN
3.00	NaN	NaN	NaN	1.0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
3.25	1.0	NaN	NaN	NaN	NaN	NaN	NaN	NaN				

Table 5.31. Cross tabulation of harmonic degrees and position in the measure in which they are used on the minor seventh chord

	11th	5th	b9th	b7th	b13th	M7	9th	Root	#11th	#9th	3rd	13th
0.00000	8.0	8.0	NaN	8.0	2.0	5.0	3.0	10.0	3.0	8.0	2.0	5.0
0.25000	7.0	7.0	5.0	4.0	2.0	2.0	6.0	6.0	1.0	6.0	9.0	5.0
0.50000	3.0	8.0	4.0	7.0	8.0	4.0	11.0	6.0	4.0	13.0	2.0	2.0
0.75000	9.0	7.0	3.0	6.0	5.0	1.0	15.0	10.0	3.0	5.0	3.0	6.0
1.00000	11.0	14.0	5.0	5.0	5.0	1.0	7.0	9.0	2.0	8.0	6.0	5.0
1.25000	10.0	7.0	6.0	10.0	5.0	2.0	10.0	6.0	1.0	14.0	3.0	7.0
1.50000	5.0	22.0	6.0	14.0	4.0	2.0	10.0	12.0	1.0	7.0	3.0	4.0
1.75000	6.0	14.0	2.0	15.0	5.0	8.0	13.0	6.0	2.0	14.0	2.0	7.0
2.00000	10.0	13.0	4.0	7.0	6.0	8.0	10.0	10.0	5.0	12.0	NaN	1.0
2.25000	12.0	9.0	1.0	8.0	5.0	5.0	12.0	14.0	5.0	7.0	7.0	3.0
2.50000	6.0	10.0	9.0	9.0	3.0	4.0	11.0	4.0	3.0	17.0	3.0	5.0
2.75000	10.0	9.0	5.0	10.0	5.0	9.0	9.0	9.0	4.0	9.0	1.0	3.0
3.00000	8.0	7.0	5.0	4.0	5.0	1.0	9.0	4.0	4.0	6.0	3.0	7.0
3.00625	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	1.0	NaN	NaN
3.25000	10.0	3.0	4.0	2.0	4.0	2.0	5.0	9.0	4.0	8.0	6.0	3.0
3.50000	5.0	8.0	3.0	4.0	3.0	1.0	8.0	5.0	5.0	12.0	5.0	3.0
3.75000	3.0	6.0	6.0	7.0	2.0	6.0	5.0	5.0				

This approach shows that there are strong tendencies to play certain harmonic degrees rather than others, and this can be related to the part of the bar Jarrett is playing in. For example, it is always more likely that Jarrett will play a sharp-ninth rather than a sharp-eleventh in a minor seventh chord. But the dataset shows that the extent of this unlikeliness changes during the course of the measure. At certain times in the measure, Jarrett is fourteen times more likely to do this, at others times he is less than twice as likely. It is even possible to model this change in likelihood over time as a mathematical function, and then compare this with other other harmonic degree likelihood functions in either this chord type, or various various others, and then filter the data to explore correlation.

For the purposes of understanding Keith Jarrett, this analysis has allowed deep and far reaching questions to be posed, and answered them in an evidenced based way. It has demonstrated that Jarrett's improvisations, though seemingly endlessly inventive, are high structured and to an extent, highly predictable when viewed in a framework of harmony and voice leading. It has provided a wide range of metrics that can easily be applied other solos of Jarrett, or any other improvising musician. Although this analysis has been created with an audience of musicians in mind, it is possible to explore the same dataset using methods found in other fields, such as statistics and machine learning. In this way, analysis can be customised to the user at hand, with the metadata also being decoupled from this user customisation.

But at same time, a challenge remains. Easily being able to interrogate music metadata can quickly lead to very large explorations, that are simply not suited to a written text format. As such, it becomes increasingly important to find a more viable framework through which to undertake this work.

Chapter 6
Conclusion

The intent of this dissertation has been to take a different approach toward music analysis. More specifically, it has asked the question: can adopting a search and retrieval based approach to music score metadata change the way music theory and analysis is practiced?

In answering this, I have set out to reframe problems found in music theory as problems of information search and retrieval, and also untangle the complicated relationship between the music score and music itself. I have also set out to explicitly reimagine the music score as music score metadata to highlight that, however it is that we may understand music, our understanding is often confounded by the practical difficulties of retrieving and searching through information.

Adopting this approach was partly a response to the difficulties that can be seen in more traditional approaches to music theory and analysis. Examining these approaches reveals a complicated history (explored in chapter one) in which music and the music score are often conflated in vague and confusing ways. It is also a history where notions of truth are little more than localised phenomena, rarely evidenced based, and not enduring. And at the same time, it is a history that often utilises a heavily pseudo-scientific language, prescribing rigid rules for music construction and practice.

The field of Music Information Retrieval (MIR), with its emphasis on information search and retrieval of music metadata, provides a powerful response to these issues. Though the field is not yet heavily engaged with developing models of music analysis for musicians, by explicitly viewing music and music information as discrete entities, it manages to sidestep so many of complex problems that inform music analysis models, and locates investigation into clearly defined bodies of information that can be susceptible to scientific methods. Yet there are challenges here too: there are currently very few tools that can be used to explore the questions of music analysis in MIR, and much of the research of this field is not appropriate for non-technical end users.

The analysis chapter provided a tangible example of what music analysis could look like if it were to leverage off an information retrieval paradigm. While the chosen corpus could have been any set of music scores, here I chose to focus on a corpus that featured transcribed jazz improvisations. I did this partly because jazz improvisation represents such profound challenges to more traditional approaches of music analysis (the difficulty it poses for music analysis was explored in chapter three). Further, music scores in jazz are often extremely minimal in their information content. Yet by utilising a larger corpus of ten jazz improvisations, deep insights could still be derived.

My other motivation for choosing jazz was that it allowed me to explore the improvisational style of Keith Jarrett. Jarrett has had almost no analysis carried out on his work, and existing models of analysis are difficult to use because of the apparent lack of repetition in his improvised melodic phrases. The analysis allowed deep insights and a new understanding of the nature of Jarrett's playing, and provided set a measurable benchmarks from which to compare other musicians.

The limitations encountered in this thesis have arisen predominantly because music data (be it score related or otherwise) now finds itself dispersed across such varied disciplines and requires such different domains of expertise. Exploring music has become a deeply mathematical and computational problem, but remains just as richly a problem far outside these disciplines. Although this certainly has been the case for many decades, it is becoming more pronounced and shows no signs of abating. Paradoxically, this leads to limitations in the dissertation itself, requiring to be written for multiple audiences and domains, and having to skirt specialised nomenclature. It is this limitation that I am seeking to address in the future work coming from this dissertation.

In the preface to his text on score based analysis, Nicholas Cook claimed that “there is something fascinating about the idea of analysing music” (Cook, 1987, p. 1). This is certainly true but for me this statement misses something fundamental: because although analysing music may be fascinating, it is also highly inconvenient and usually limited to small volumes of information. This dissertation has set out to show that, by being able to easily find and transform the kind of metadata found on music scores, we can endlessly challenge and transform our understanding of music.

Future Work

In this dissertation I have set out to re-imagine the music score as a site of scaleable metadata that is highly suited to exploratory data transformation, search and retrieval, and analysis, to find new ways to undertake music analysis and understand music theory. In applying such an approach in the last chapter, far reaching insights could be seen, both in terms of the improvisational style of Keith Jarrett as well as how this kind of analytical framework might be used to undertake jazz analysis.

But such an approach is not without its challenges. While the methods presented in the previous chapter can enable a deep analysis into ten Keith Jarrett solos, together they represent only a very small fraction of Jarrett's recorded output, which limits the breadth of analysis that might be carried out. Further, there are many other solos from comparable jazz improvisors that could be used to shed light on jazz improvisation. And while the approach taken is certainly scalable to an extent, things can soon become unwieldy.

The approach taken in the previous chapter assumes a reasonably high level knowledge of data tools and methods. This means that, although this might be able to provide a deep examination of music analysis in an evidenced based way, its practical applications for many musicians will be limited. This chapter will address this by developing an alternative framework that can respond to these challenges. I will leveraging off existing approaches that utilise music metadata (though not from music scores) to show how these features could influence the creation of a search and retrieval framework for music metadata in which music score metadata is used. The chapter will outline some of the key features that typify such applications before presenting a software project as an example of future work in this space, in order to demonstrate what a music score search and retrieval framework might look like.

Popular examples of applications that heavily utilise music metadata include Spotify, iTunes, Google Play, Pandora and Shazam. These applications utilise metadata in order to allow users to easily navigate information about audio files. Though they each have a slightly different emphasis, they share a number of key features, that could be leveraged off in order to build a music score focused metadata application.

The first of these features is the ubiquitous availability of music metadata. Google Music, for example is estimated to hold over 40 million songs (Morris & Powers, 2013, p. 108), and Spotify and iTunes hold comparable collections. Each of these audio files has been tagged with extensive metadata holding many different attributes, facilitating the ease of search and retrieval (for example, title, duration, genre, sub-genre, album name, musician names etc).

Applications such as Spotify and iTunes also allow users to access this metadata in a way that is completely decoupled with the user interface, through the use of an application programming interface (API). This makes it possible use the data to build more software. The Music Score Metadata Builder software I presented in Chapter 4 has been integrated with the iTunes Data API to easily obtain information about music which can then be combined with the data taken from a music score. The Shazam application, which allows users to identify song names from audio recording, also leverages off the Spotify and iTunes data API to allow users to go and download these songs are the application identifies it.

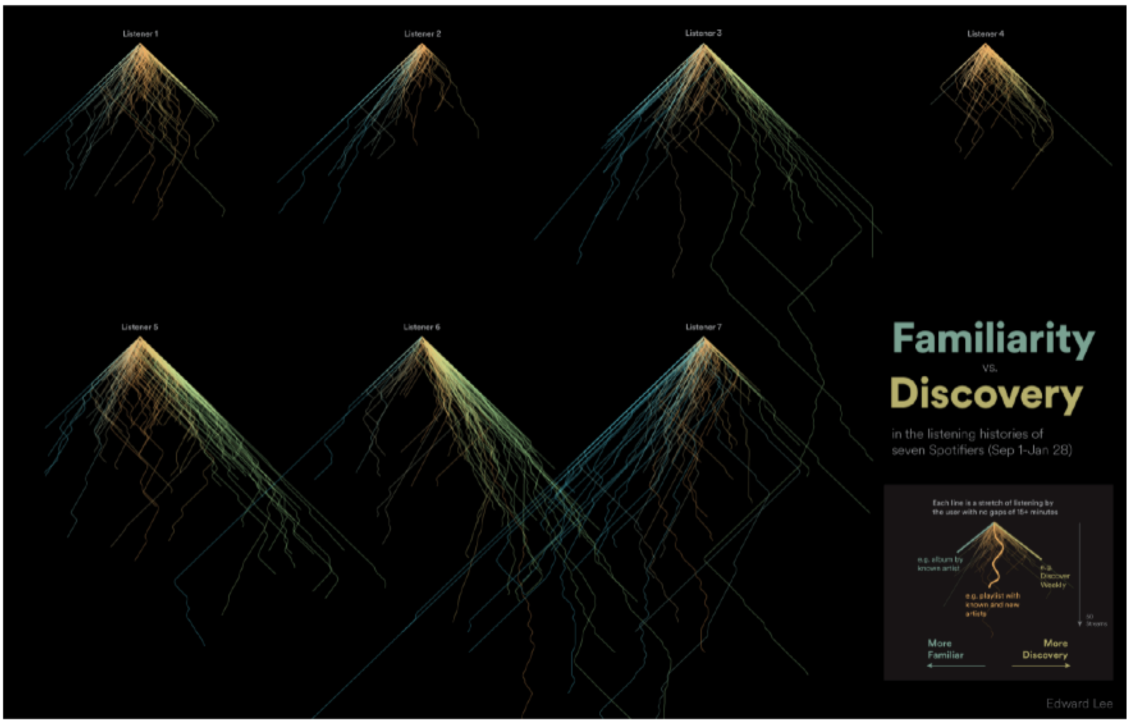
Though there is a vast amount of music metadata held by these applications, the metadata itself is of course markedly different from the kind of metadata found on the music score. Online repositories holding music score information are also far more limited. The most extensive of these is the IMSLP/Pretrucci Music Library, which holds music scores for over 400,000 pieces of music (just 1% of what is estimated to be available on Google Play), or around eight million pages of music. Currently data is not available through a API, and much of it still to be yet to be digitised into machine readable data format such as MusicXML. However problems of data availability in music score metadata have made strong advances in the last decade, particularly in automated transcriptions and optical music recognition.

The applications such as Spotify, iTunes and Google Play are also optimised to allow for ease of user searching. Users can easily input words, or even partial words of song names, artist names, and results are dynamically returned almost instantly to the user. More advanced filtering can also be accessed as users can search for music of particular genres or timeframes. These applications also have internal filtering processes (such as removing songs with explicit language or titles from returned results). Third party applications can also be built off the Data API which allow far more nuanced searching,

Currently this functionality is limited when it comes to music score metadata. An example discussed in Chapter 3 was PeachNote, but this limited to searching for melodies over a small corpus. Yet this is a problem that can be solved. If music score metadata can be transformed into appropriately structured data, it can become straightforward to filter the data on such things as melody, range, harmonic structure and voicing, tempo, dynamic markings, instrumental combinations, and time signatures.

Having access to well structured metadata can also facilitate the building of powerful data visualisations from which to explore music. An example of this can be seen in Figure 6.1 below. In this example, listening data was extracted from seven Spotify employees in order to visualise whether the music they were listening to was predominantly new music (i.e music they had not heard before and were 'discovering') or familiar music. The length of the line from the top of each listener structure below was an indication of how long they had played the song for. The slope of the line (whether it veered to the right of left) is an indication of whether this is new, 'discovered' song, or a familiar song, based on their listener history (Lee, para 5, 2017).

Figure 6.1. Spotify discovery visualisation



If an application can be created that has access to well structured music score metadata, it would become straightforward to build visualisations to interactively explore different aspects of music, such as melodic sequences a, chord progressions, instrumental combinations. Being able to visualise music data in different ways can also facilitate the learning of different types of music (i.e. for musicians seeking to learn about orchestration who cannot read music).

Applications such as Spotify, iTunes and Google Play also seek to curate the listening experience by applying machine learning and recommender system algorithms to usage patterns of their applications. The ability to predict user preference is also a key research question of the Music Information Retrieval community. An example of this is in action can be seen in Spotify's weekly "Discovery" playlist, that aims to understand explores user playing habits and comes up with a weekly list suited to user tastes. Speaking of its success, Ogle notes:

> We now have more technology than ever before to ensure that if you're the smallest, strangest musician in the world, doing something that only 20 people in the world will dig, we can now find those 20 people and connect the dots between the artist and listeners.
> https://qz.com/571007/the-magic-that-makes-spotifys-discover weekly-playlists-so-damn-good/(2018)

It is certainly possible to do this for well structured music score metadata, and this is perhaps the most exciting possibility of a music score search and retrieval application. It paves the way for music theory that is crowd sourced and can evolve over time in line the user's taste. Customising results to the individual tastes of the user and using recommendation algorithms to promote explorations based on similar users, has the potential to replace the need for the expert curator seen in much traditional analysis.

At the heart of existing music metadata applications is the ability to allow users to interact with audio. Searching for a song, or browsing a genre, aims to create a listening experience.

Though it is not currently possible to link specific parts of music sore information through existing data API, web technologies have evolved in the last decade to allow the building of sophisticated online synthesisers which would allow users to play music score examples, and allow muting or changing instruments in them to explore the different sonic textures.

The question I am left with after considering the kinds of features that exist in existing music metadata applications is: why do these types of applications not exist for the exploration of music scores, so that musicians can explore music structure and practice in a similar way?

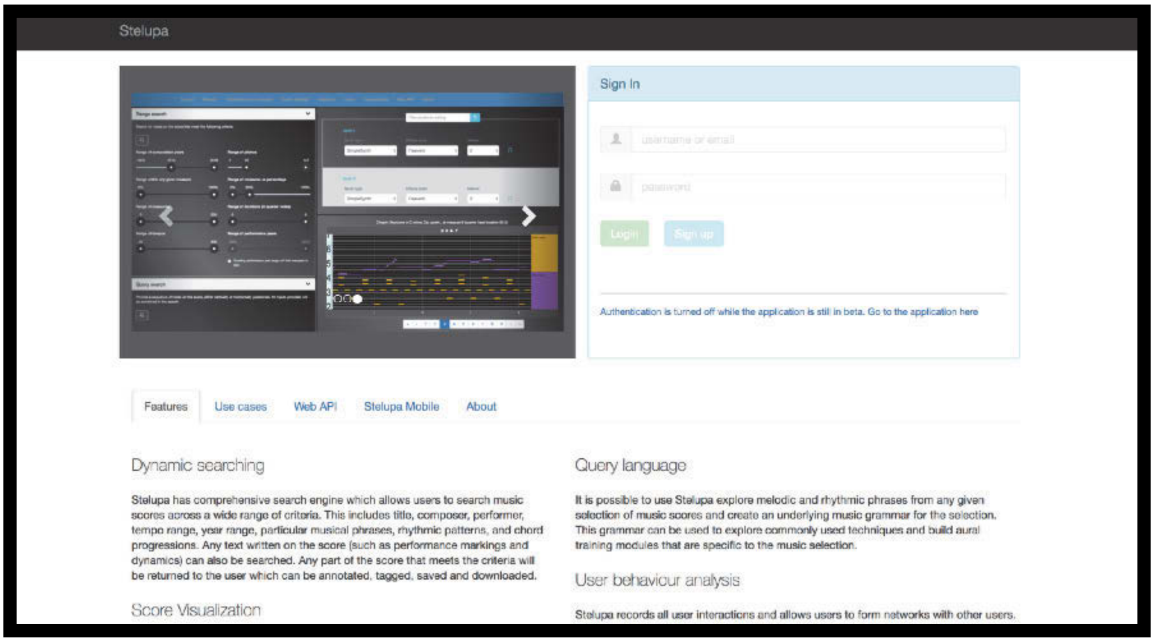
In responding to this problem, and as part of this dissertation, I have created Stelupa, a search and retrieval engine that utilises metadata taken from the music scores. The following section will present this in the form of a proof of concept application that has been structured as an open source project. In building this software, I have leveraged off a number of widely used web technologies and libraries such as Node.js, React, Electron, Tone.js, Django and Postgres, that are well suited to data rich applications in which there is high level user interaction.

Stelupa is an open source music score metadata search engine that uses the data structures of the analysis chapter and extends it to an interactive environment.

The application can be explored can be viewed at www.stelupa.com, and I have outlined a number of its core features below.

A screenshot of the landing page for the application can seen below in Figure 6.2.

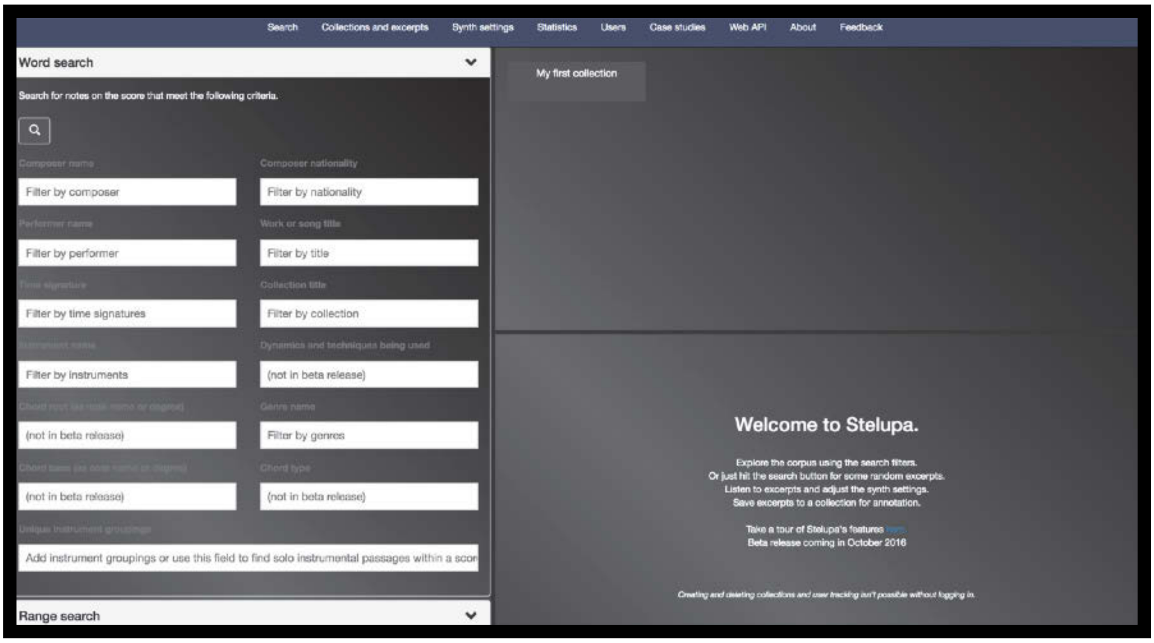
Figure 6.2. Stelupa landing page



This application allows users to intuitively search, retrieve, and categorise excerpts from music scores based on a wide range of criteria. It has been designed for users who have varying amounts of domain specific knowledge to undertake music analysis and does not require knowledge of coding or data transformation. Jazz musicians can use the application to build repositories of licks; orchestrators are able explore explore instrumental combinations found on large scores, and musicologists could use it to hold examples of dissonance found in music of different periods and locations. It implements many of the features found in other music metadata applications, such as dynamic searching using multiple filters, different visualisations and user behaviour analysis. The applications features are also detailed on the application landing page.

After users log into the application, they encounter the main search page seen in Figure 6.3 below. There are three widows on the right hand side in a scrollable pane, that together provide comprehensive search capabilities. The top right hand side holds user curated collections, and the bottom right hand side holds a pane that returns results from searches. All windows in the application are resizable and moveable depending on user preference.

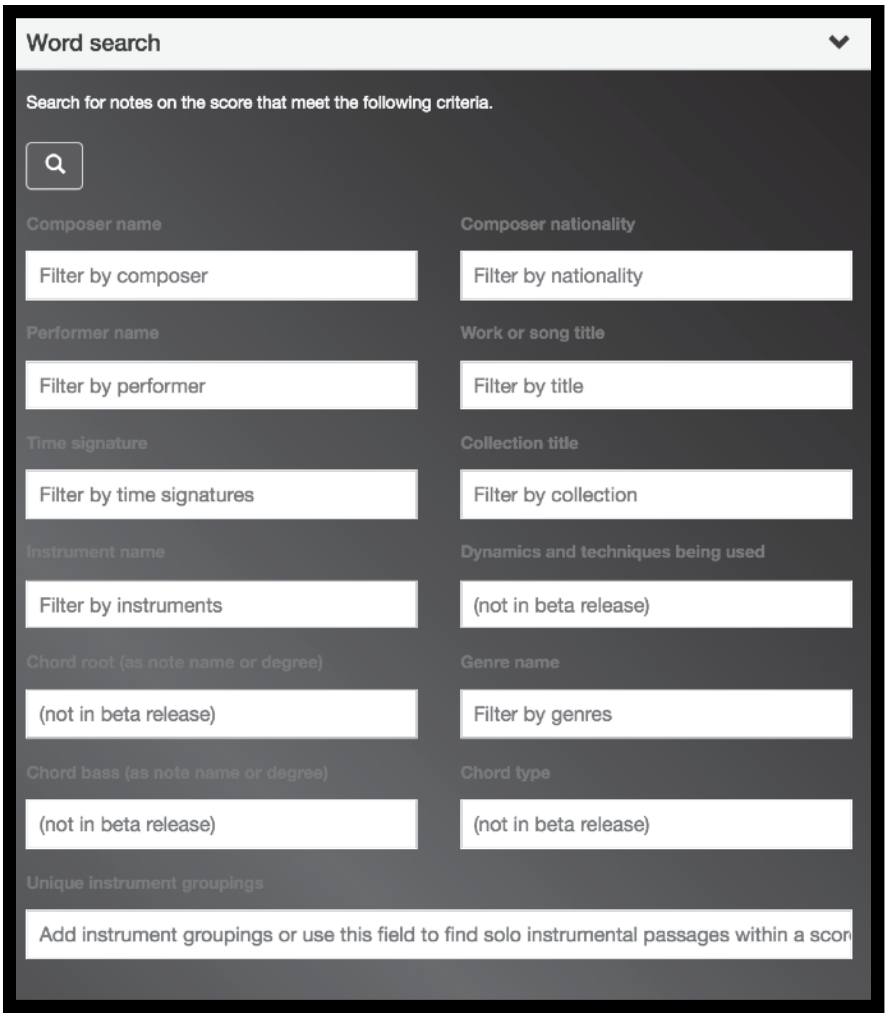
Figure 6.3. Main search page



The application has search capabilities for finding words and ranges, as well as its own built in query language. Figure 6.4 provides a screenshot of different word based filters currently available, such as composer, nationality, performer, time signature, instrument, instrumental grouping etc. All notes and rests in the underlying data structure (which is the same as that used in the methodology and analysis chapters) have been encoded with this information in order to allow any note that meets this criteria to be returned. To allow some context for the returned result (in that it makes little sense to only return one note from a score that meets a certain criteria), the notes and rests both before and after the found results are also returned (it is possible to change the amount of contextual results to either side of the result in the application settings).

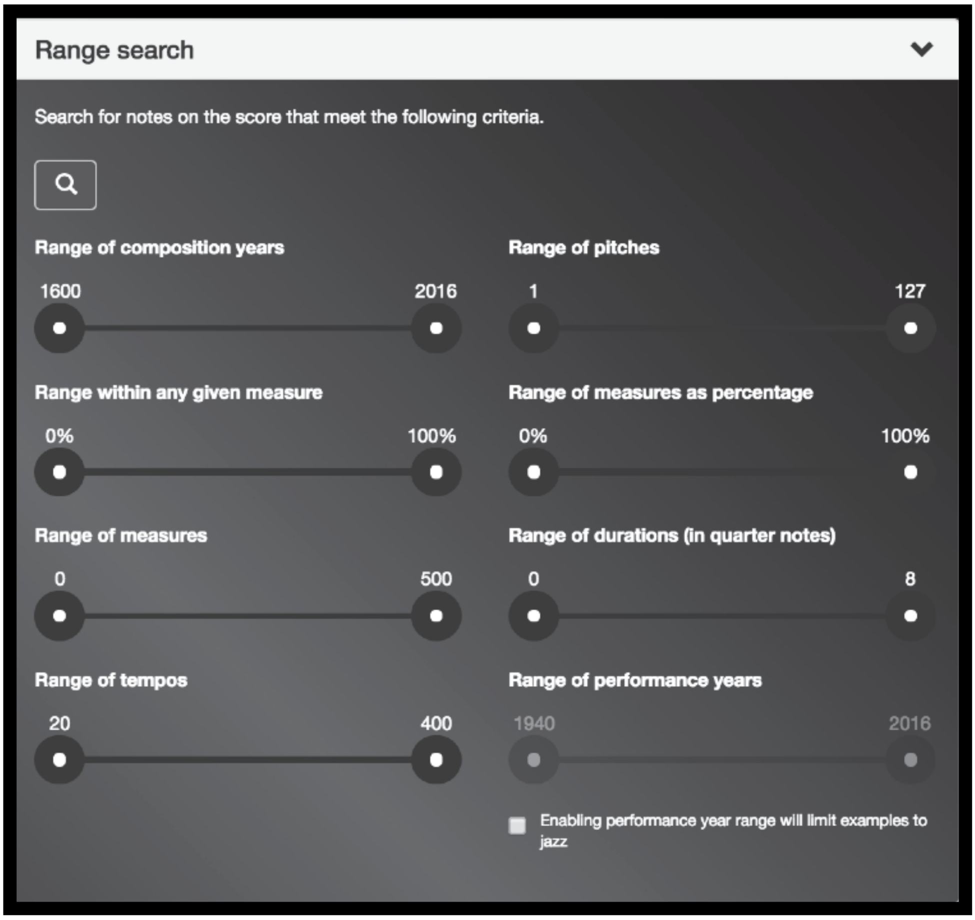
Any search term that is inputted will act as a filters on the data. Each single input allow either/or searching and and using multiple search fields means that results must meet the criteria of the multiple filters. As an example, it is possible to choose Mahler and Bach in the composer name field, which would limit all results to any works by these composers. Adding an additional search input, such as choosing nationality of Austrian, will restricting the results to be either Mahler and Bach, and Austrian. (the semantics of this search would be ((“Composer: Bach” OR “Composer: Mahler”) AND “Nationality” : “Austrian”).

Figure 6.4. Word filtering metadata



As well as searching for words in the metadata, it is possible to search across multiple range criteria. Range can be limited to criteria such as composed or performed year, pitch range, measure range etc, and this can be seen in Figure 6.5 below. All range filters are applied cumulatively to specific notes or rests, along with other contextual records.

Figure 6.5 Range filtering metadata



To analyse music it is often important to search for very specific structures, such as particular melodies and harmonic voices. In accomodating this, the application has a built in query language that allows for searching of specific note sequences and chordal structures. The query language also accommodates relative note distances and searching for structures that are spread across multiple instruments. This is shown in Figure 6.6.

Figure 6.6. More nuanced searching

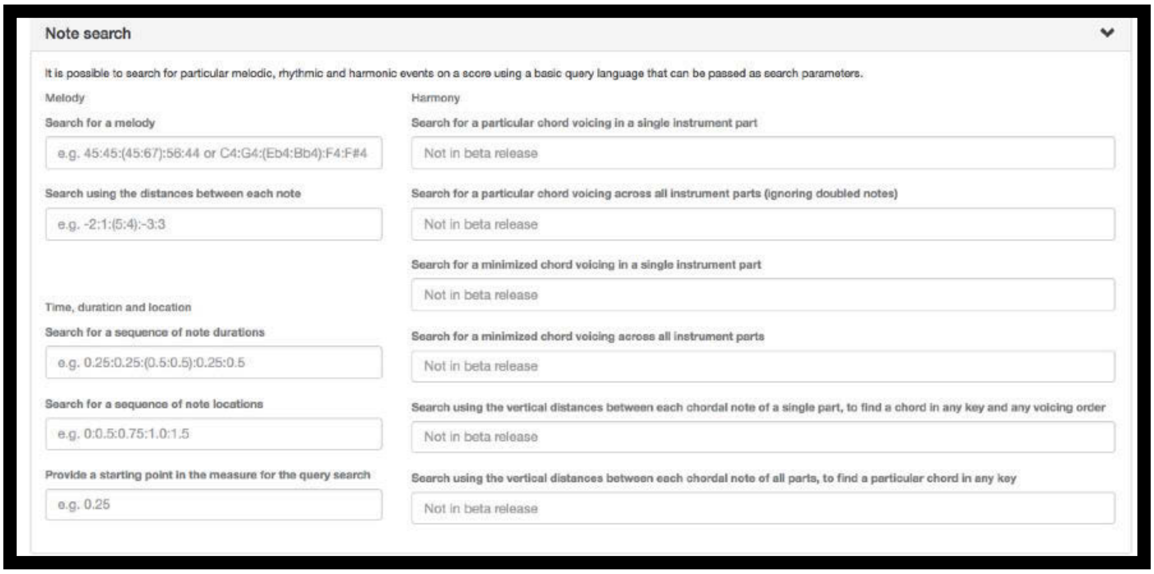
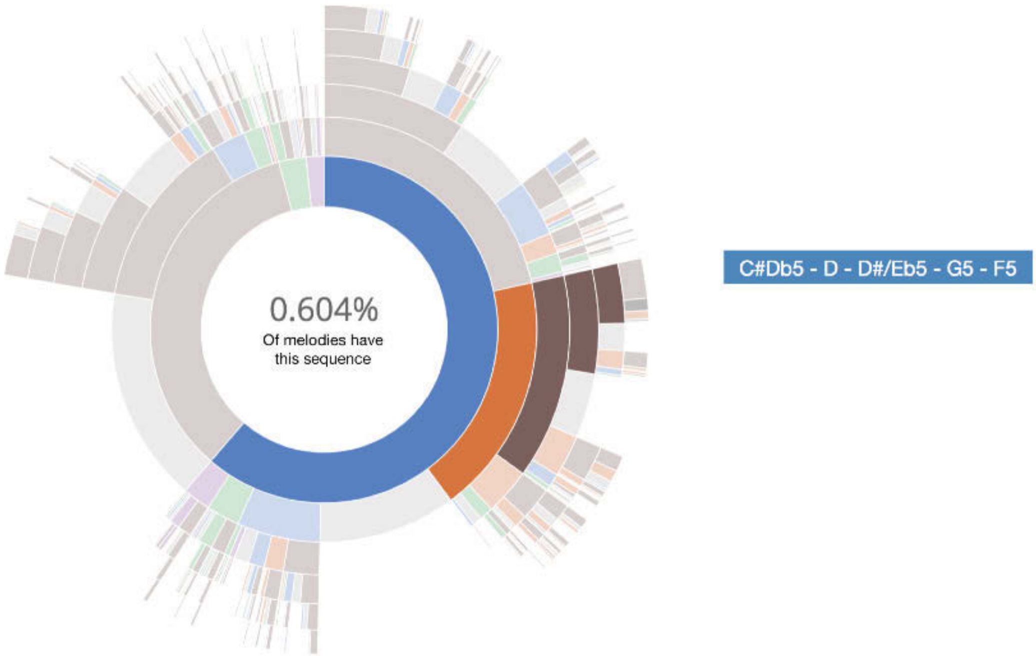


Figure 6.7 shows an alternate sunburst partition search view possible in the application, that allows users to see how melodic structures are distributed across a corpus (or a corpus filtered by user chosen criteria). As the user moves the mouse over the visualisation, a percentage of the number of melodies in the corpus that have this pattern is returned.

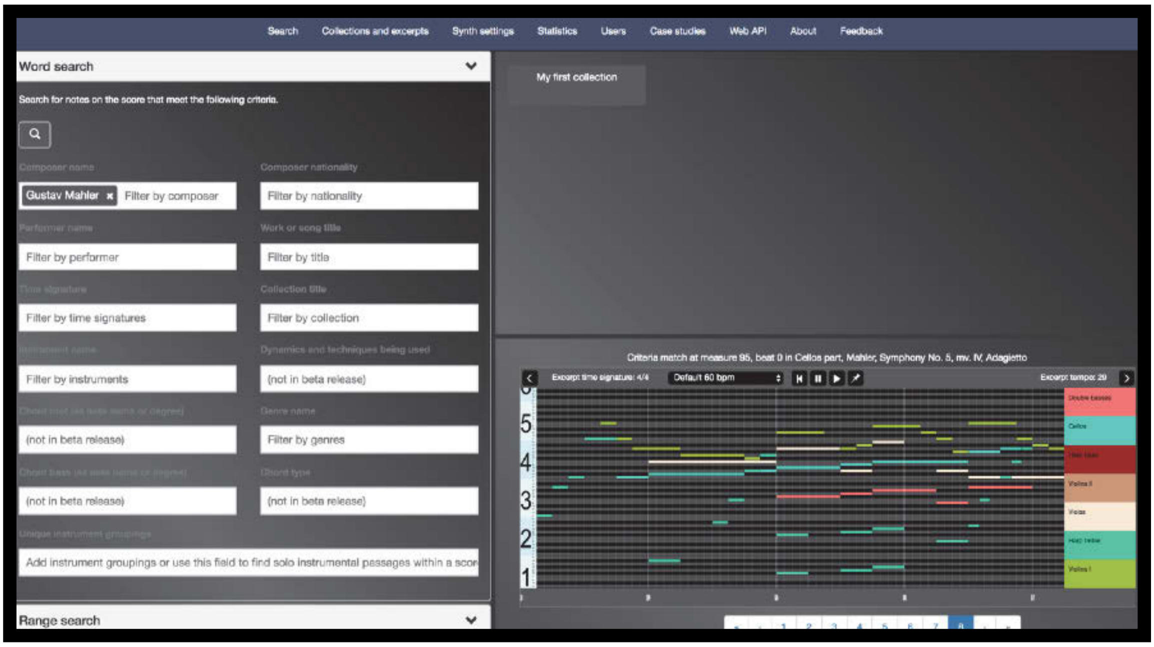
Figure 6.7. Phrase sequence searching using a sunburst partition



In some cases, undertaking some searches (such as searching for the note middle C) will return a large number search results. In such a scenario, the application limits to the results to ten randomised instances that meet this criteria.

Figure 6.8 below displays a returned excerpt, seen in the bottom right hand pane. By default, results are returned in piano roll format (though experimentation is being done in the JavaScript libraries VexFlow and D3 with a view to rendering music notation visualisations also). The query below shows an example from Mahler (a note from a Cello part in the Symphony No. 5 Adagio). The result, along with other contextual results (being the notes and rests around this note) have also been returned. At the bottom of the excerpt a pagination bar can be seen, showing that it is possible to move between ten returned excerpts. The search criteria here was that Mahler was the composer, and these are the first ten results that have been returned.

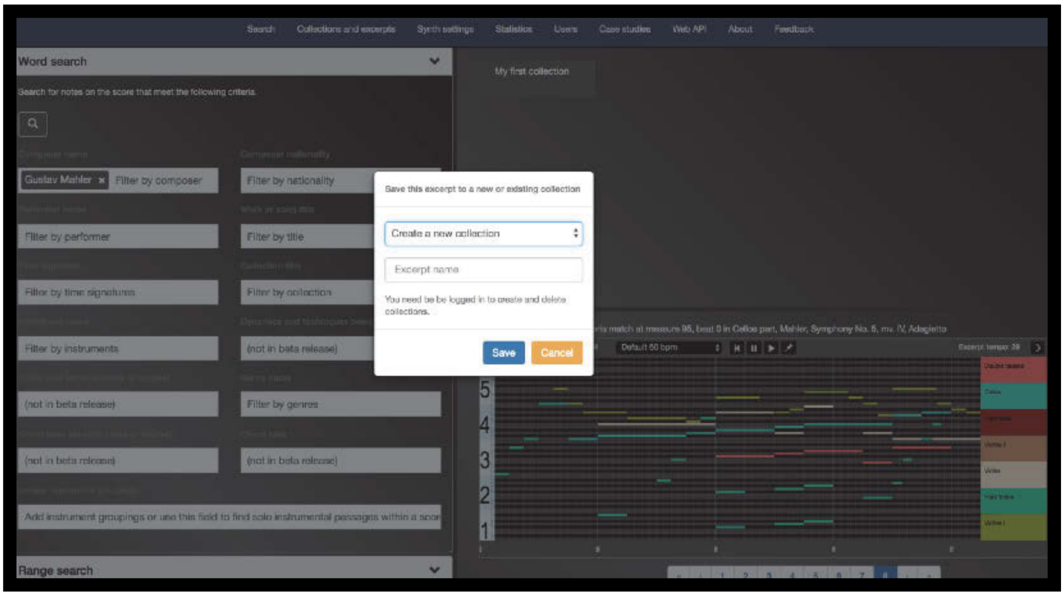
Figure 6.8. Piano-roll visualisation to render results



The application uses visualisation of the piano roll as a primary view due to its ease of navigation and existing popularity in music software. It has been built with custom SVG in the browser (meaning that no images need to be rendered) making it very fast to return to the user. The numbers listed on on the side of the visualisation indicates the octave of the note, and each instrument that has been returned is given a particular colour (which is configurable in the application settings). Once a visualisation is returned, it is also possible to highlight part of the piano roll and include that as part of search criteria for the application.

A critical feature in music metadata applications is the ability to “pin” or tag results of interest. This application allows users to click the pin icon on the expert toolbar, should they be interested in it. Figure 6.9 provides a screenshot of what happens once the user presses the pin icon.

Figure 6.9. Pinning a result



The user will be prompted to choose a collection (which will hold a list of excerpts), either by creating one or using existing one. The user can then provide a name of the pinned example, which can be seen in Figure 6.10.

Figure 6.10. Building collections

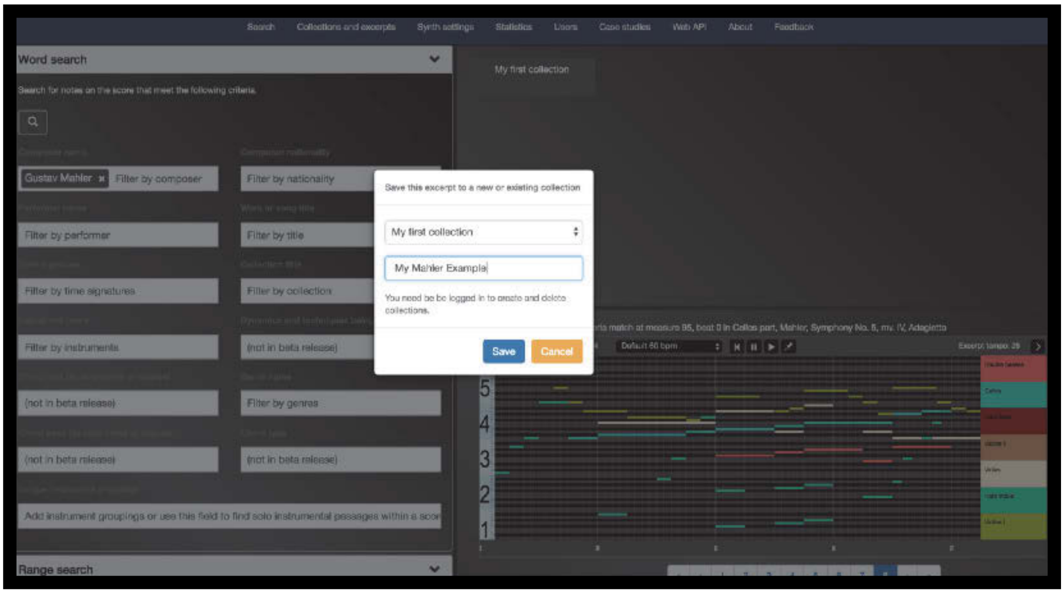
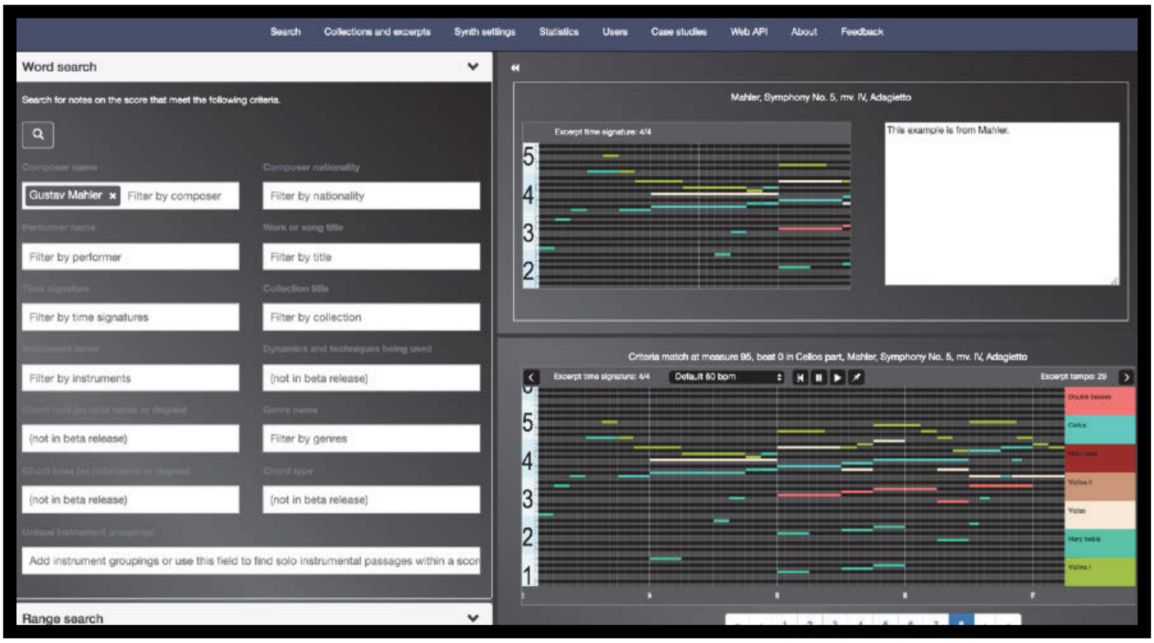


Figure 11 shows that the resulting example has now become part of a user defined collection that can now be annotated. For example, a musicologist might use the search criteria to locate several solo oboe passages in symphonies by composers in different periods in order to explore changes in how this instrument is scored over time. Having these in a clear collection allows easy navigation and annotation. For the Keith Jarrett case study of the previous chapter, the application could be used to find specific four note microphrases across different solos, and tag these. An example of notes being made for a particular example can be seen below in Figure 6.11.

Figure 6.11. Annotating a pinned excerpt

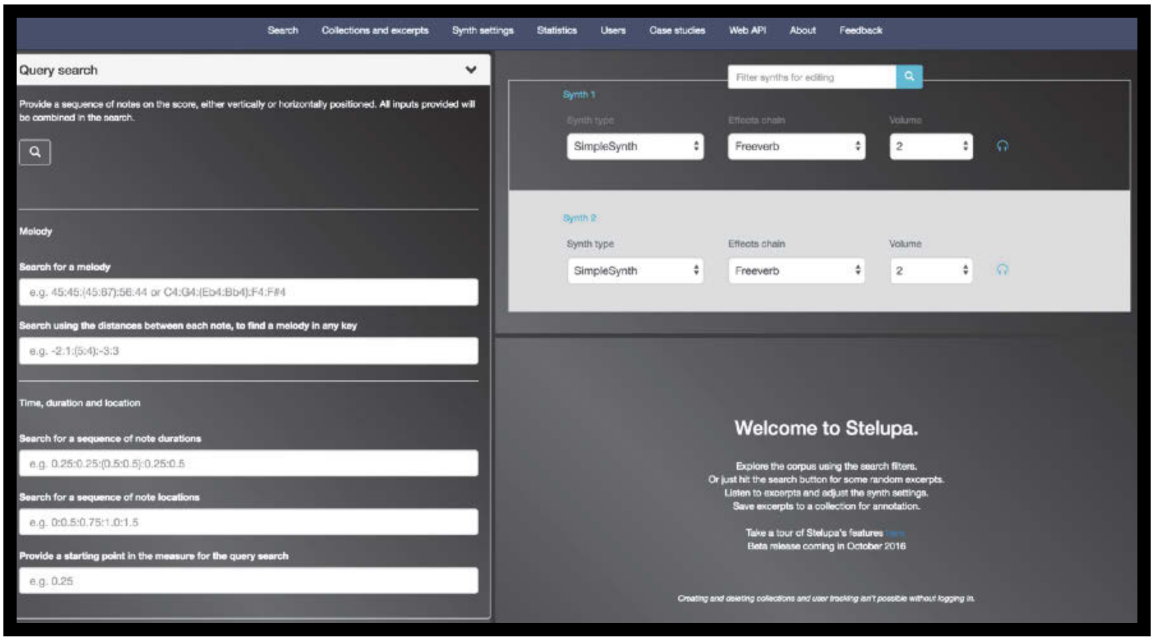


In keeping with the importance of relating the metadata to sound, the application also has functionality to allow users to interact directly with audio. The toolbar on the top of the excerpt provides play pause, and rewind buttons similar to a music player, and allows users to choose different tempos for playback. A multi-timbral synthesiser, built in JavaScript, provides excerpt playback, and is currently limited to eight voices, all of which can be manipulated in terms of sound and effects chain.

Admittedly, the use of audio in this application is currently very limited. However, the rapid development of streaming audio technologies seen in metadata applications suggests that this is a solvable problem. Ideally the application should allow playing audio recordings as well as synthesised audio, allow multiple speeds and pitch change of these, and allow the streaming of high quality orchestral sample libraries to explore music.

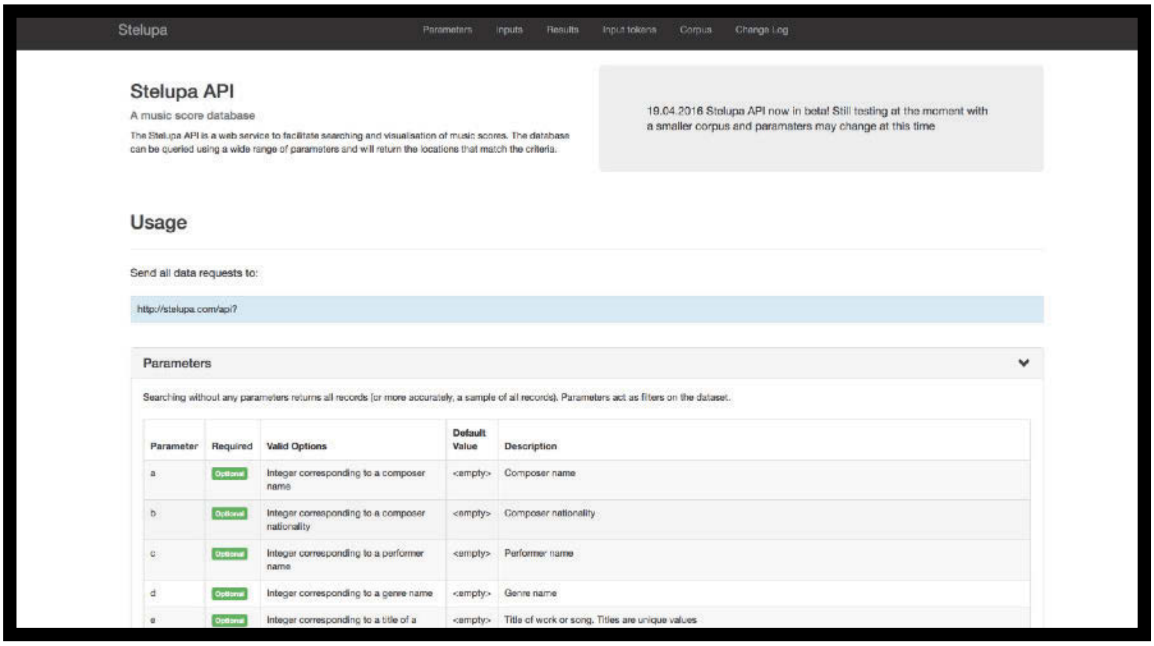
The application’s synth can be accessed by clicking the synth settings menu items on the top and then it will appear in the top right hand side pane. It can can be seen in Figure 6.12. below.

Figure 6.12. Built in Javascript synth



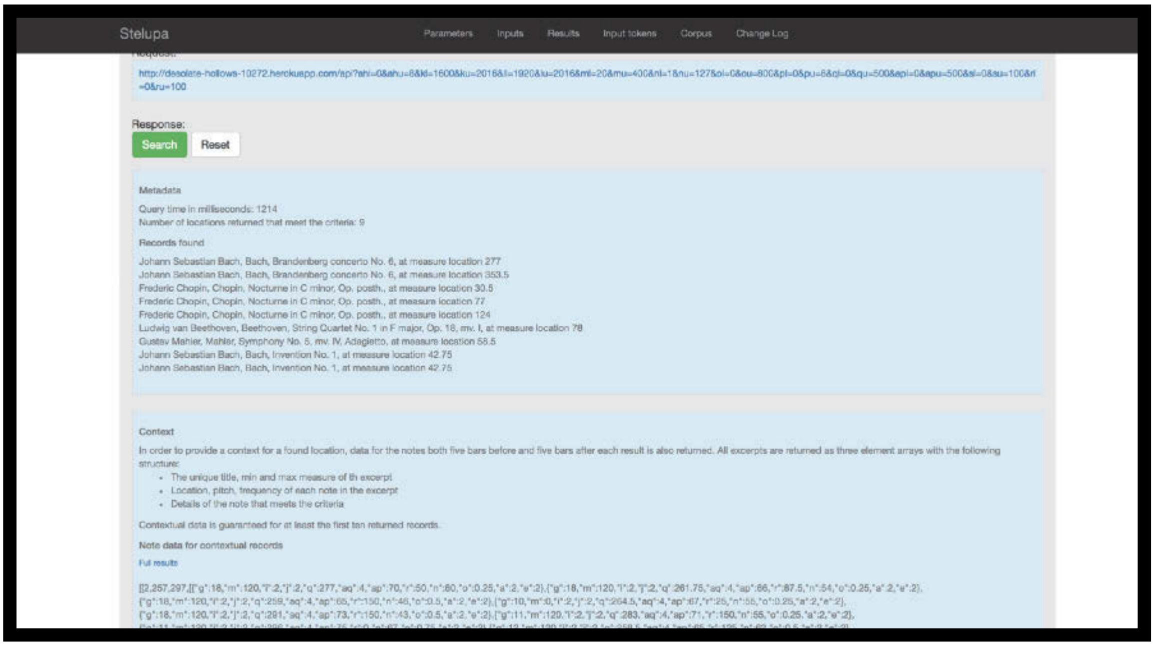
Having access to the raw data that informs music metadata applications (such as the Spotify data API) is a powerful mechanism with which to allow third party to applications to explore data without being limited to any given user. This software has also been built to accomodate this. Figure 6.13 shows a screenshot of the Stelupa Data API that provides comprehensive search functionality, but rather than returning visualisations, just returns the raw data.

Figure 6.13 Stelupa Data API



The API allows the data to easily be exported into other applications for exploration. To undertake the Keith Jarrett case study, the API was used to obtain raw music score metadata for the Keith Jarrett solos which was then imported into Jupyter Notebook where the analysis was carried out. Figure 6.14 shows returned records coming straight from data API. At the bottom of the screen an extract from the raw data of the first returned record can be seen. Users can also click on the Full results link which downloads this in JSON format, and CSV format will be supported in the future.

Figure 6.14. Searching for data in the data API



Stelupa is only one example of what might be possible in terms of a search and retrieval framework for music score metadata. I have created this software as an open source project whose code base is intended to be extended as group effort. Its scope goes beyond this dissertation and there is much more functionality that can be built into it. For example, there is currently no user preference and recommendation functionality built into this software, which would allow users to find others with similar tastes and explore interactive music scores in a collaborative fashion.

Appendix 1
Keith Jarrett Transcriptions

Title	Performer collection	Date recorded	Composer collection	Date composed	Quarter beats per minute	Tonality	Number of records
All The Things You Are	Standards, Vol. 1	1983	Very Warm For May	1939	247	Ab major	2027
Autumn Leaves	Still Live	1986	Les Portes De La Nuit	1945	251	G minor	1826
Autumn Leaves	Tokyo 96	1996	Les Portes De La Nuit	1945	224	G minor	1243
Days Of Wine And Roses	Keith Jarrett At The Blue Note, The Complete R...	1994	Days Of Wine And Roses	1962	160	F major	1424
Groovin High	Whisper Not	1999	Shaw Nuff	1945	289	Eb major	1811
If I Were A Bell	Up For It	2002	Guys And Dolls	1950	167	Ab major	1982
In Love In Vain	Standards, Vol. 2	1983	Centennial Summer	1946	147	Bb major	1280
My Funny Valentine	Still Live	1986	Babes In Arms	1937	122	C minor	1254
Someday My Prince Will Come	Up For It	2002	Snow white and the seven dwarfs	1937	148	Bb major	1815
Stella By Starlight	Standards Live	1983	The Uninvited	1944	151	Bb major	1512

All The Things You Are
Composed by Jerome Kern
Keith Jarrett piano solo (Standards, Volume 1, 1983)

[Production note: Content removed due to copyright restrictions.]

Autumn Leaves
Composed by Joseph Kosma
Keith Jarrett piano solo (Still Live, 1986)

[Production note: Content removed due to copyright restrictions.]

Autumn Leaves
Composed by Joseph Kosma
Keith Jarrett piano solo (Tokyo 96, 1996)

[Production note: Content removed due to copyright restrictions.]

Days Of Wine And Roses
Composed by Henry Mancini
Keith Jarrett piano solo (Keith Jarrett At The Blue Note, Vol. 3, 1983)

[Production note: Content removed due to copyright restrictions.]

Groovin High
Composed by Dizzy Gillespie
Keith Jarrett piano solo (Whisper Not, 1999)

[Production note: Content removed due to copyright restrictions.]

If I Were A Bell
Composed by Frank Loesser
Keith Jarrett piano solo (Up For It, 2002)

[Production note: Content removed due to copyright restrictions.]

In Love In Vain
Composed by Jerome Kern
Keith Jarrett piano solo (Standards Vol. 2, 1983)

[Production note: Content removed due to copyright restrictions.]

My Funny Valentine
Composed by Richard Rogers
Keith Jarrett piano solo (Still Live, 1986)

[Production note: Content removed due to copyright restrictions.]

Someday My Prince Will Come
Composed by Jerome Kern
Keith Jarrett piano solo (Standards, Volume 1, 1983)

[Production note: Content removed due to copyright restrictions.]

Stella By Starlight (tempo: 151 bpm,)
Composed by Victor Young
Keith Jarrett piano solo (Standards Live, 1983)
Tempo: 151 bpm
Solo start time in recording: 5'11

[Production note: Content removed due to copyright restrictions.]

Appendix 2
Notes for software related to this dissertation: Music Metadata Builder, Jupyter Analysis Notebooks and Stelupa

All accompanying software is hosted on my github account at: https://github.com/jgab3103/

Music MetaData Builder
This repository contains the code used to convert MusicXML into a JSON format suited for data analysis, and allows merging of this data with other metadata (such as look up data from the iTunes API).

Further details at: https://github.com/jgab3103/musicXML2MusicJSON

Jupyter Analysis Notebooks
These notebooks contain all code relating to the analysis chapter. Also hosted here is the prepared datasets used in the analysis.

Further details of software at: https://github.com/jgab3103/Phd-Jupyter-Notebooks
Further details of data used at: https://github.com/jgab3103/Phd-Data

Stelupa
This is a full stack web application that provides a multimodal environment to search music score metadata and has both polyphonic examples (for example Mahler, Bach) and jazz examples (the Keith Jarrett solos used in this dissertation).

A youtube walk through exploring an earlier version of the software (built in Angular.js and MongoDB) can also be viewed at: https://www.youtube.com/watch?v=P9xebSuW9ys&t=97s

Further details at: https://github.com/jgab3103/stelupa-1.1

Bibliography

Antila, C., & Cumming, J. (2014). The Viz Framework: Analyzing Counterpoint in Large Datasets, *International Society of Music Information Retrieval, Conference Proceedings*.

Atcherson, W. T., (1972). Symposium on Seventeenth-Century Music Theory: England. *Journal of Music Theory*, *16*(1/2), 6. http://doi.org/10.2307/843323

Baggi, D. L. (1974). *Realization of the Unfigured Bass by Digital Computer*.

Baker, N. K. (1977). The Aesthetic Theories of Heinrich Christoph Koch. *International Review of the Aesthetics and Sociology of Music*, *8*(2), 183. http://doi.org/10.2307/836886

Balkwill, L.-L., & Thompson, W. F. (1999). A Cross-Cultural Investigation of the Perception of Emotion in Music: Psychophysical and Cultural Cues. *Music Perception: An Interdisciplinary Journal*, *17*(1), 43–64. http://doi.org/10.2307/40285811

Baker, D. (2005). *Jazz Improvisation (Revised): A Comprehensive Method for All Musicians*. Alfred music.

Barker, A. (1984). *Greek musical writings*. Cambridge: Cambridge University Press.

Barlow, H., & Morgenstern, S. (1948). *A Dictionary of Musical Themes*. New York: Crown Publishers.

Bas de Haas, W., Magalhaes, J. P., ten Heggeler, D., Bekenkamp, G., & Ruizendaal, T. (2012). Chordify: Chord transcription for the masses. *International Society of Music Information Retrieval, Conference Proceedings*.

Batlle, E., & Cano, P. (2000). Automatic Segmentation for Music Classification using Competitive Hidden Markov Models. *International Society of Music Information Retrieval, Conference Proceedings*.

Beard, D., & Gloag, K. (2005). *Musicology: the key concepts*. London: Routledge.

Bello, J., Guiliano, M., & Sandler, M. (2000). Techniques for Automatic Music Transcription. *International Society of Music Information Retrieval, Conference Proceedings*.

Bello, J. P. (2007). Audio-Based Cover Song Retrieval Using Approximate Chord Sequences: Testing Shifts, Gaps, Swaps and Beats. *International Society of Music Information Retrieval, Conference Proceedings*.

Bendor, D., & Sandler, M. (2000). Time Domain Extraction of Vibrato from Monophonic Instruments. *International Society of Music Information Retrieval, Conference Proceedings*.

Bent, I. (1996). *Music theory in the age of Romanticism*. Cambridge: Cambridge University Press.

Bergeron, K., & Bohlman, P. V. (1992). *Disciplining music: musicology and its canons*. Chicago: University of Chicago Press.

Blake, J. (2016). *Improvising optimal experience: Flow theory in the Keith Jarrett Trio* (Doctoral dissertation, The University of North Carolina at Chapel Hill).

Blume, G. (2003). Blurred affinities: tracing the influence of North Indian classical music in Keith Jarrett's solo piano improvisations. *Popular Music*, *22*(2), 117-142.

Briginshaw, S. B. (2012). A neo-riemannian approach to jazz analysis. *Nota Bene: Canadian Undergraduate Journal of Musicology*, *5*(1), 57.

de Bruin, L. (2015). Theory and practice in idea generation and creativity in Jazz improvisation. *Australian Journal of Music Education*, (2), 91-106.

Busse, W. G. (2002). Toward objective measurement and evaluation of jazz piano performance via MIDI-based groove quantize templates. *Music Perception: An Interdisciplinary Journal*, *19*(3), 443-461.

Bohak, C., & Marolt, M. (2009). Calculating Similarity of Folk Song Variants with Melody Based Features. *International Society of Music Information Retrieval, Conference Proceedings*.

Bonardi, A. (2000). IR for Contemporary Music: What the Musicologist Needs. (Abstract of invited talk) *International Society of Music Information Retrieval, Conference Proceedings*.

Bostock, M., "D3: Data-Driven Documents", Date viewed: 12 Jan 2018, d3js.org/.

Capuzzo, G. (2006).The Nature of the Guitar: An Intersection of Jazz Theory and Neo-Riemannian Theory. *Music Theory Online*, *12*, 2.

Caplin, W. E. (2000). *A theory of formal functions for the instrumental music of Haydn, Mozart and Beethoven*. New York: Oxford University Press.

Campion, T., & Wilson, C. R. (2002). *A new way of making fowre parts in counterpoint by Thomas Campion and rules how to compose by Giovanni Coprario, edited by Christopher R. Wilson*. Aldershot, Hants, England: Ashgate.

Carr, I. (1992). *Keith Jarrett: The man and his music*. Da Capo Press.

Chai, W., & Vercoe, B. (2000). Using User Models in Music Information Retrieval Systems. *International Society of Music Information Retrieval, Conference Proceedings*.

Christensen, T., Damschroder, D., & Williams, D. R. (1992). Music Theory from Zarlino to Schenker: A Bibliography and Guide. *Notes*, *48*(4), 1306. http://doi.org/10.2307/942150

Christensen, T. S. (2002). *The Cambridge history of Western music theory*. Cambridge: Cambridge University Press.

Christensen, T., (2004), *Rameau and Musical Thought in the Enlightenment,* Cambridge: Cambridge University Press.

Christensen, T., (2016), *The Works of Music Theory: Selected Essays*, Routledge.

Clark, S., ed. (1997), The Letters of C.P.E Bach, Clarendon Press, UK

Clausen, M., Engelbrecht, R., Meyer, D., & Schitz, J. (2000). PROMS: A Web-based Tool for Searching in Polyphonic Music. *International Society of Music Information Retrieval, Conference Proceedings*.

Cliff, D., & Freeburn, H. (2000). Exploration of Point-Distribution Models for Similarity-based Classification and Indexing of Polyphonic Music. *International Society of Music Information Retrieval, Conference Proceedings*.

Cornelis, O., Leman, M., Moelants, D. (2009) Exploring African Tone Scales, *International Society of Music Information Retrieval, Conference Proceedings*.

Cowart, G. (1989). *French musical thought: 1600-1800*. Ann Arbor u.a., UMI Research Press

Cook, N. (1987). *A guide to musical analysis*. New York: G. Braziller.

Cope, D. (1991). *Computers and musical style*. Madison, WI: A-R Editions.

Crochemore, M., Iliopolous, C., Pinzon, Y., & Rytter, W. (2000). Finding Motifs with Gaps. *International Society of Music Information Retrieval, Conference Proceedings*.

Dahlhaus, C. (1987). *Schoenberg and the new music: essays*. Cambridge: Cambridge University Press.

Davis, M., (1985), “Miles Davis: 'Coltrane was a very greedy man. Bird was, too. He was a big hog’: a classic interview from the vaults”. Date viewed: 12 Nov

2017, https://www.theguardian.com/music/2012/nov/06/miles-davis-interview-rocks-backpages

Dean, T. *New structures in jazz and improvised music since 1960*. Open University, 1992.

Descartes René. (ed. 1968). *Musicae compendium*. New York: Broude.

Dixon, S., Gouyon, F., & Widmer, G. (2003). Towards Characterisation of Music via Rhythmic Patterns. *International Society of Music Information Retrieval, Conference Proceedings*.

Django Website, "Django: The Web framework for perfectionists with deadlines", Date viewed: 10 Dec 2018, www.djangoproject.com/.

The Last Date 1964, audio recording, EMARCY Records.

Doornbusch, P. (2010). Gerhard Nierhaus: Algorithmic Composition: Paradigms of Automated Music Generation. *Computer Music Journal*, *34*(3), 70–74. http://doi.org/10.1162/comj_r_00008

Doraisamy, S., & Ruger, S. M. (2001). An Approach Towards A Polyphonic Music Retrieval System. *International Society of Music Information Retrieval, Conference Proceedings*.

Duckles, V. H., Reed, I., & Keller, M. A. (1997). *Music reference and research materials: an annotated bibliography*. New York: Schirmer Books.

Dudeque, N. (2005). *Music theory and analysis in the writings of Arnold Schoenberg (1874-1951)*. Aldershot, Hants, England: Ashgate.

Dunsby, J. (1983). *A Hitchhikers Guide to Semiotic Music Analysis*, Music Analysis, 1983

Eddington, A., (1927) Gifford Lectures (online). Date viewed: Aug 21, 2016, http://www-history.mcs.st-and.ac.uk/extras/eddington_gifford.html

Elsdon, P. (2008). Style and the Improvised in Keith Jarrett's Solo Concerts. *Jazz Perspectives*, *2*(1), 51-67.

Evans, B., (1980), Breakfast With Bill Evans, Date viewed: 3 Jan 2018, *https://www.allaboutjazz.com/breakfast-with-bill-evans-bill-evans-by-bob-kenselaar.php?page=1/2018*

Feisst, S. (2011). *Schoenberg's new world: the American years*. New York: Oxford University Press.

Forte, A. (1998). *The atonal music of Anton Webern*. New Haven: Yale University Press.

Node Website, “Node.js”, Date viewed: 10 Mar 2018, nodejs.org/en/

Foote, J. A., (2000): Retrieving Orchestral Music by Long Term Structure. *International Society of Music Information Retrieval, Conference Proceedings*.

Fujinaga, I., & Riley, J. (2002). Digital Image Capture of Musical Scores. *International Society of Music Information Retrieval, Conference Proceedings*.

Futrelle, J., & Downie, J. S. (2003). Interdisciplinary Research Issues in Music Information Retrieval: ISMIR. *Journal of New Music Research*, *32*(2), 121–131. http://doi.org/10.1076/jnmr.32.2.121.16740

Fux, J. J., Mann, A., & Edmunds, J. (1965). *The study of counterpoint from Johann Joseph Fux's Gradus ad parnassum*. New York: W.W. Norton.

Flexer, A. (2007). A Closer Look on Artist Filters for Musical Genre Classification. *International Society of Music Information Retrieval, Conference Proceedings*.

Ganseman, J., Scheunders, P. and D'haes, W. Using XML-Formatted Scores in Real-Time Applications. In *Proceedings of ISMIR 2009 10th International Conference on Music Information Retrieval* (Kobe, Japan, October 26-30, 2009), pp. 663-668.

Gao, B., Dellandrea, E., & Chen, L. (2013). Sparse Music Decomposition onto a Midi Dictionary driven by Statistical Music Knowledge. *International Society of Music Information Retrieval, Conference Proceedings*.

Gibson, S. (2005). *Aristoxenus of Tarentum and the birth of musicology*. New York: Routledge.

Giger, A., & Mathiesen, T. J. (2003). *Music in the mirror: Reflections on the history of music theory and literature for the twenty-first century*. Lincoln, Neb.: University of Nebraska Press.

Girdlestone, C. (1969). *Jean-Philippe Rameau: his life and work*. New York: Dover Publications.

Gjerdingen, R. O. (1988). *A classic turn of phrase: music and the psychology of convention*. Philadelphia: University of Pennsylvania Press.

Good, M. D. (2000), MusicXML. *Structuring Music through Markup Language Designs and Architectures*, 187–192. http://doi.org/10.4018/978-1-4666-2497-9.ch009

Hartley, R. V. L. (1928). Transmission of Information 1. *Bell System Technical Journal*, *7*(3), 535–563. http://doi.org/10.1002/j.1538-7305.1928.tb01236.x

Helmholtz, H. von, & Ellis, A. J. (1954). *On the sensations of tone as a physiological basis for the theory of music*. New York: Dover Publications.

Herissone, R. (2000). *Music theory in seventeenth-century England*. New York: Oxford University Press.

Heyer, D. J. (2012). Applying Schenkerian Theory to Mainstream Jazz: A Justification for an Orthodox Approach. *Music Theory Online*, *18*(3).

Hickey, M. (2009). Can improvisation be 'taught'?: A call for free improvisation in our schools. *International Journal of Music Education*, *27*(4), 285-299.

Hiller, L., & Bean, C. (1966). Information Theory Analyses of Four Sonata Expositions. *Journal of Music Theory*, *10*(1), 96. http://doi.org/10.2307/843300

Hodson, R. (2007). Interaction, improvisation, and interplay in jazz. Routledge.

Hoos, H., Renz, K., & Gorg, M. (2001). GUIDO/MIR — an Experimental Musical Information Retrieval System based on GUIDO Music Notation. *International Society of Music Information Retrieval, Conference Proceedings*.

Huron, D., (2000), Perceptual and Cognitive Applications in Music Information Retrieval, *International Society of Music Information Retrieval, Conference Proceedings*.

Iverson, E., (2009), Interview with Keith Jarrett, https://ethaniverson.com/interviews/interview-with-keith-jarrett/ viewed 9 Sep 2017

Izmirli, O. (2009) Tonal-atonal classification of Music Audio using Diffusion Maps. *International Society of Music Information Retrieval, Conference Proceedings*.

Jarrett, K., (2014), Keith Jarrett: Interview and Speech at NEA Jazz Masters Awards 2014, https://www.youtube.com/watch?v=fDbOKHOuy9M/2018, viewed on 4 Jan 2018

Jiang, N., & Muller, M. (2013). Automated Methods for analyzing Music Recordings in Sonata Form. *International Society of Music Information Retrieval, Conference Proceedings*.

"JavaScript." Mozilla Developer Network, Date viewed: 19 Jan 2018, developer.mozilla.org/en-US/docs/Web/JavaScript

Joost-Gaugier, C. L. (2009). *Pythagoras and Renaissance Europe: Finding heaven*. Cambridge: Cambridge University Press.

Keirnan, F. (2000). Score-based style recognition using artificial neural networks. *International Society of Music Information Retrieval, Conference Proceedings*.

Kerman, J. (1985). *Contemplating music: challenges to musicology*. Cambridge, MA: Harvard University Press.

Keil, C. M. (1966). Motion and feeling through music. *Journal of Aesthetics and Art Criticism*, 337-349.

Kirlin, P., & Jensen, D. (2011). Probabilistic Modelling of Heirachical Music Analysis. *International Society of Music Information Retrieval, Conference Proceedings*.

Kraehenbuehl, D., & Coons, E. (1959). Information as a Measure of the Experience of Music. *The Journal of Aesthetics and Art Criticism*, *17*(4), 510. http://doi.org/10.2307/428224

Lange, A., (1984), Interview with Keith Jarrett, Date viewed: 3 Jan 2018, http://downbeat.com/archives/artist/keith-jarrett.

LaRue, J. (1966). *Aspects of medieval and Renaissance music: a birthday offering to Gustave Reese*. New York: W.W. Norton.

Lee, E., (2017), “Familiarity vs. Discovery” in “10 Data Art Projects by Spotify Analysts & Designers”, *https://insights.spotify.com/es/2016/09/29/10-data-art-projects/* , viewed 6 Jun 2017

Lee, W., & Chen, L. P. (2000). Efficient Multi-Feature Index Structures for Music Data Retrieval . *International Society of Music Information Retrieval, Conference Proceedings*.

Lemstrom, K., & Perttu, S. (2000). SEMEX - An efficient Music Retrieval Prototyp. *International Society of Music Information Retrieval, Conference Proceedings*.

Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music*. Cambridge, MA: MIT Press.

Levitin, D. J. (2010). Why music moves us. *Nature*, *464*(7290), 834–835. http://doi.org/10.1038/464834a

Larson, S. (1998). Schenkerian analysis of modern jazz: questions about method. *Music Theory Spectrum*, *20*(2), 209-241.

Luttmann, S. (2009). *Paul Hindemith a research and information guide*. New York, NY: Routledge.

Mardirossian, A., & Chew, E. (2007). Visualising Music: Tonal Progressions and Distributions. *International Society of Music Information Retrieval, Conference Proceedings*.

Martin, H., (1996), Jazz Theory: An Overview, Annual Review of Jazz Studies 1-17, Newark NJ.

Martin, R., (2015), *At 70, Keith Jarrett Is Learning How To Bottle Inspiration*, Date viewed: 22 Oct 2017, https://www.npr.org/2015/05/10/404975326/at-70-keith-jarrett-is-learning-how-to-bottle-inspiration,

Matt Ogle talks Discover Weekly and Spotify's personalised evolution. Retrieved August 21, 2016, from http://musically.com/2016/03/21/matt-ogle-discover-weekly-spotify/

Mazzola, G., & Cherlin, P. B. (2008). *Flow, gesture, and spaces in free jazz: Towards a theory of collaboration*. Springer Science & Business Media.

Mersenne, M., & Williams, R. F. (1972). *Marin Mersenne: an edited translation of the fourth treatise of the Harmonie universelle*. Ann Arbor, MI: University Microfilms.

Meyer, L. B. (1957). Meaning in Music and Information Theory. *The Journal of Aesthetics and Art Criticism*, *15*(4), 412. http://doi.org/10.2307/427154

Monelle, R. (1992). *Linguistics and semiotics in music*. Chur, Switzerland: Harwood Academic.

Moreno, J. (1999). Body'n'Soul?: Voice and Movement in Keith Jarrett's Pianism. *The Musical Quarterly*, *83*(1), 75-92.

Morris, J.W, & Powers, D. (2013). Control, Curation, and Musical Experience in Streaming Music Services, *Creative Industries Journal, 8* (2)*, 106-122*

Mugglestone, E., & Adler, G. (1981). Guido Adler's "The Scope, Method, and Aim of Musicology" (1885): An English Translation with an Historico-Analytical

Commentary. *Yearbook for Traditional Music*, *13*, 1. http://doi.org/10.2307/768355

MuseScore Website ,"Create, play and print beautiful sheet music.", Date viewed: 10 Jan 2018, musescore.org/en.

"10 Top Companies that Used Node.Js in Production (Examples)." *Netguru Blog on DevOps*, Date viewed: 17 Dec 2017, www.netguru.co/blog/top-companies-used-nodejs-production

Nattiez, J. J. (1990). *Music and discourse: toward a semiology of music*. Princeton, NJ: Princeton University Press.

Nettl, B. (1983). *The study of ethnomusicology: twenty-nine issues and concepts*. Urbana: University of Illinois Press.

Ogle, M., (2016) 11 Awesome Spotify Tips and Tricks you're probably not using', viewed 1 April 2018, https://www.lifehacker.com.au/2016/01/11-more-awesome-spotify-tips-and-tricks-youre-probably-not-using/

Oxford English Dictionary. Retrieved August 20, 2016, from http://www.oed.com/

Pachet, F., & Zils, A. (2001). Evolving automatically high level Music descriptors from Acoustic Signals. *International Society of Music Information Retrieval, Conference Proceedings*.

Panken, T., (2014), *For Keith Jarrett's 69th Birthday, Full Interviews From 2000, 2001, and 2008, plus an 2008 Interview with Manfred Eicher,* Date viewed: 18 Jun 2017, https://tedpanken.wordpress.com/tag/keith-jarrett/

Page, T., (2009). *Motivic Strategies in Improvisations by Keith Jarrett and Brad Mehldau*, Masters Thesis, Sibelius Academy, Finland

Peeling, P. H., Cemgil, A. T., & Godsill, S. J. (2007, September). A Probabilistic Framework for Matching Music Representations. In *ISMIR* (pp. 267-272).

Piaget, J., (ed. 2016), *Structuralism*, Psychology Press 1 Edition

"Seventh String Software - the home of Transcribe!" Date viewed: 10 Aug 2015, www.seventhstring.com/.

Pickens, J. (2000). A Comparison of Language Modeling and Probabilistic Text Information Retrieval Approaches to Monophonic Music Retrieval. *International Society of Music Information Retrieval, Conference Proceedings*.

Pinkerton, R. C. (1956). Information Theory and Melody. *Sci Am Scientific American*, *194*(2), 77–87. http://doi.org/10.1038/scientificamerican0256-77

"Project Jupyter." Date viewed: 27 Jan 2018, jupyter.org/.

Quintilianus, A., & Mathiesen, T. J. (1983). *On music, in three books*. New Haven: Yale University Press.

Quinn, I. (2000). Highpoints: A Study of Melodic Peaks Zohar Eitan Melodic Similarity: Concepts, Procedures, and Applications Walter B. Hewlett Eleanor

Selfridge-Field. *Music Theory Spectrum*, *22*(2), 236–245. http://doi.org/10.2307/745962

Raphael, C. (2001). Automated Rhythm Transcription. *International Society of Music Information Retrieval, Conference Proceedings*.

Rizzo, D., Ponce de Leone, P., Perez-Sancho, C., & Pertusa, A. (2006). A Pattern Recognition Approach for Melody Track Selection in MIDI Files. *International Society of Music Information Retrieval, Conference Proceedings*.

Rosenthal, T., (1996), Keith Jarrett Interview: The "insanity" of doing more than one (musical) thing, http://tedrosenthal.com/tr-kj.htm, viewed 3 Jan 2018

Ruwet, N., & Everist, M. (1987). Methods of Analysis in Musicology. *Music Analysis*, *6*(1/2), 3. http://doi.org/10.2307/854214

Sacks, O. (2007). *Musicophilia: tales of music and the brain*. New York: Alfred A. Knopf.

Sandvold, V., Aussenac, T., Celma, O., & Herrera, P. (2006). Good Vibrations: Music Discovery through Personal Musical Concepts. *International Society of Music Information Retrieval, Conference Proceedings*.

Schoenberg, A., & Stein, L. (1975). *Style and idea: selected writings of Arnold Schoenberg*. New York: St. Martins Press.

Schoenberg, A. (1978). *Theory of harmony*. Berkeley: University of California Press.

Schopenhauer, A., Norman, J., Welchman, A., & Janaway, C. (2010). *The world as will and representation.* Cambridge: Cambridge University Press.

Schuller, G. (1958). Sonny Rollins and the challenge of thematic improvisation. *The Jazz Review*, *1*(1), 6-21.

Schueller, H. M. (1955). Immanuel Kant and the Aesthetics of Music. *The Journal of Aesthetics and Art Criticism*, *14*(2), 218. http://doi.org/10.2307/425860

Shannon, E, A. (1948). Mathematical Theory of Communication. http://doi.org/10.1109/9780470544242.ch1

Shirali, S. A. (2013). Marin Mersenne, 1588–1648. *Resonance Reson*, *18*(3), 226–240. http://doi.org/10.1007/s12045-013-0034-2

Smith, J., Burgoyne, J. A., & Fujinaga, I. (2011). Design and Creation of a Large Scale Database of Structural Annotations. *International Society of Music Information Retrieval, Conference Proceedings*.

Smith, G. E., (1983) Homer, Gregory, and Bill Evans?: the theory of formulaic composition in the context of jazz piano improvisation, Ann Arbor, University Microfilms

Steege, B. (2012). *Helmholtz and the modern listener*. Cambridge: Cambridge University Press.

Strange, P. (2003). *Keith Jarrett's up-tempo jazz trio playing: Transcription and analysis of performances of "Just In Time",* Doctoral Thesis, University of Miami.

Terefenko, D. (2004). Keith Jarrett's transformation of standard tunes.

Strunk, S. (2005). Notes on Harmony in Wayne Shorter's Compositions, 1964-67. *Journal of Music Theory, 49*(2), 301-332. Retrieved from http://www.jstor.org/stable/27639402

Sudnow, D., & Dreyfus, H. L. (2001). Ways of the hand: A rewritten account.

Sweetman, B. (1999). *The failure of modernism: the Cartesian legacy and contemporary pluralism*. Mishawaka, IN: American Maritain Association.

Tagg, P. (1982). Analysing popular music: theory, method and practice. *Popular Music*, *2*, 37. http://doi.org/10.1017/s0261143000001227

Tagg, P., & Brackett, D. (1998). Interpreting Popular Music. *American Music*, *16*(2), 224. http://doi.org/10.2307/3052568

The Jazzomat Research Project. Retrieved August 21, 2016, from http://jazzomat.hfm-weimar.de/workshop2014/workshop.html

Viro, V. (2011). Peachnote: Music Score Search and Analysis Platform. *International Society of Music Information Retrieval, Conference Proceedings*.

Walser, R. (1993). Out of notes: Signification, interpretation, and the problem of Miles Davis. *The Musical Quarterly*, *77*(2), 343-365.

Wang, E. J. (2011). *Mistuning the world: a cultural history of tuning and temperament in the seventeenth century*.

West, M. L. (1992). *Ancient Greek music*. Oxford: Clarendon Press.

Worthen, J. (1992). Linguistics and Semiotics in Music. http://doi.org/10.4324/9781315076942

Weigl, D., & Gaustavino, C. (2011). User Studies in the Music Information Retrieval Literature. *International Society of Music Information Retrieval, Conference Proceedings*.

Wiil, U. K. (2005). *Computer music modeling and retrieval: Second International Symposium, CMMR 2004, Esbjerg, Denmark, May 26-29, 2004: revised papers*. Berlin: Springer.

Youngblood, J.E., (1958), Style as Information, Journal of Music Theory, April 1958, pp. 24-35.

Zhang, B., Kreitz, G., Isaksson, M., Ubillos, J., Urdaneta, G., Pouwelse, J. A., & Epema, D. (2013). Understanding user behavior in Spotify. *2013 Proceedings IEEE INFOCOM*. http://doi.org/10.1109/infcom.2013.6566767